UNIVERSITY OF CALIFORNIA, SAN DIEGO

Essays in Development Economics

A dissertation submitted in partial satisfaction of the requirements for the degree Doctor of Philosophy

in

Economics

by

Ketki Sheth

Committee in charge:

Professor Karthik Muralidharan, Chair
Professor Craig McIntosh, Co-Chair
Professor Gordon Dahl
Professor Paul Niehaus
Professor Joshua Graff Zivin

2014

The Dissertation of Ketki Sheth is approved, and it is acceptable in quality and form for publication on microfilm and electronically:

Co-Chair

Chair

University of California, San Diego

2014

TABLE OF CONTENTS

Signature Page ………………………………………………………………………………………. iii

Table of Contents ……………………………………………………………………………………. iv

List of Figures ……………………………………………………………………………………….. v

List of Tables ………………………………………………………………………………………... vi

Acknowledgements ………………………………………………………………………….…..... ix

Vita …………………………………………………………………………………………………... x

Abstract of the Dissertation ………………………………………………………………………….. xi

Chapter 1: Bridging Education Gender Gaps in Developing Countries: The Role of Female Teachers ….………………………………………………………………………………………….. 1

Appendix ……………………………………………………………………………... 29

References …………………………………………………….. 35

Chapter 2: Evaluating Health-Seeking Behavior, Utilization of Care, and Health Risk: Evidence from a Community Based Insurance Model in India ………………………………………………... 38

Appendix ……………………………………………………………………………... 84

References …………………………………………………….. 89

Chapter 3: The Distributional Consequences of Micro Health Insurance: Can a Pro-Poor Program Prove to be Regressive? ……………………………………………………………………………... 90

Appendix ……………………………………………………………………………... 141

References ………………………………………………………………………………. 150

LIST OF FIGURES

Figure 2.1: Illness by Months Enrolled …………………………………... 60

Figure 2.1: Health Care Consumption by Months Enrolled …………………………………………. 60

Figure 3.1: Yearly Health Care Expenditure and Income Rank (Control Households) ……………... 121

Figure 3.2: Yearly Health Care Expenditure Concentration Index (Control Households) ………….. 121

Figure 3.3: Yearly Health Care Expenditure Concentration Index by Enrollment Status (Control Households) ………………………………………………………………………………………. 122

Figure 3.4: Yearly Health Care Expenditure Concentration Index by Enrollment Status …………… 122

Figure 3.5: Yearly Health Care Expenditure Concentration Index of Early Insured Households (Treatment Households) ………………………………………………………………………….. 123

LIST OF TABLES

Table 1.1: Summary Statistics by Gender …………………………………………………………... 21

Table 1.2: Learning Gaps by Gender and Grade …………………………………………………….. 22

Table 1.3: Characteristics of Classrooms Assigned to Female Teachers ……………………………. 23

Table 1.4: Gender Differentials in Learning Trajectories from Lower to Higher Grades …………... 24

Table 1.5: Impact of Female Teachers on the Learning Gains of Female Students ……................... 25

Table 1.6: Heterogeneous Effects on Test Score Gains of Girls by Teacher Characteristics and Teacher Gender …………………………….. 26

Table 1.7: Impact of Female Teachers on the Learning Gains of Female Students by Subject 27

Table 1.8: Impact of Female Teachers on the Attendance of Female Students ……………………... 28

Table A.1.1: Estimating Sample by Year and Grade ………………………………………………... 32

Table A.1.2: Entering Cohorts by Gender …………………………………………………………… 33

Table A.1.3: Attrition by Gender Matching …………………………………………………………. 34

Table 2.1: Enrollment and Claims Summary Statistics ……………………………………………... 61

Table 2.2: Summary Statistics and Balance on Demographic Characteristics ………………………. 62

Table 2.3: Endline Survey – Health Incidence and Expense ……………………………………….... 63

Table 2.4: SHG Monthly Surveys – Illness, Care, and Expenditure ………………………………… 64

Table 2.5: Individual Health Behavior Conditional On Illness ……………………………………... 65

Table 2.6: Assets and Loans ………………………………………………………………………… 66

Table 2.7: Financial Vulnerability …………………………………………………………………… 67

Table 2.8: Psuedo-Baseline Health Status …………………………………………………………… 68

Table 2.9: Difference-in-Difference using Pseudo-Baseline ………………………………………... 69

Table 2.10: Endline Survey Non-response ………………………………………………………….. 70

Table 2.11: Reasons for 2012 October Endline Attrition (by fraction) ……………………………... 71

Table 2.12: Attrition in 2012 Endline Survey by February 2011 SHG Health Status ………………. 72

Table 2.13: SHG Monthly Survey Health Responses by Endline Response Status ……................... 73

Table 2.14: SHG Survey Response Rates and Demographics ………………………………………. 74

Table 2.15: Length of Exposure to Insurance Offer …………………………………………….……… 75

Table 2.16: Variation on Time of Village Enrollment on Health Variables ………………………… 76

Table 2.17: Enrolled Members in Control versus Treatment Villages ………………………………. 77

Table 2.18: Endline Survey – Treatment Effect on the Treated ……………………………………... 78

Table 2.19: SHG Monthly Survey – Treatment Effect on the Treated ……………………………..… 79

Table 2.20: Enrolled Members Versus Unenrolled Members ………………………………………. 80

Table 2.21: Indirect Effects of Insurance ……………………………………………………….…. 81

Table 2.22: Endline Survey – Village Doctor Visits ……………………...………………………… 82

Table 2.23: SHG Monthly Survey – Village Doctor Visits …………………………………………. 83

Table A.2.1: Health Illness Summary Statistics …………………………………………………….. 84

Table A.2.2: SHG Monthly Summary Statistics …………………………………………………….. 85

Table A.2.3: Summary Statistics on Assets …………………………………………………………. 86

Table A.2.4: Summary Statistics of SHG Finances (Internal and External) ………………………… 87

Table A.2.5: Distribution of Visits by Insurance Doctor to Villages ………………………………... 88

Table 3.1: Demography Summary Statistics ………………………………………………………… 124

Table 3.2: Enrollment Summary Statistics ………………………………………………………….. 125

Table 3.3: Enrolled Houses ………………………………………………………………………….. 126

Table 3.4: Claim Summary Statistics ……………………………………………………………….. 127

Table 3.5: Health Summary Statistics ……………………………………………………………….. 128

Table 3.6: Duration Between Enrollment and First Claim …………………………………………... 129

Table 3.7: Poverty and Health Care Utilization, Year Recall ……………………………………….. 130

Table 3.8: Poverty and Health Care Expenditure ……………………………………………………. 131

Table 3.9: Poverty and Health Care Consumption …………………………………………………... 132

Table 3.10: Demand for MHI …………………………………………………………………... 133

Table 3.11: Poverty and Health Care Expenditure by Enrollment Status 134

Table 3.12: Poverty and Health Care Expenditure by Enrollment Status, Concentration Index ……. 135

Table 3.13: Heterogeneous Effect of Insurance by Poverty Status …………………………………. 136

Table 3.14: Poverty and Health Care Expenditure by Insurance Status …………………………….. 137

Table 3.15: Poverty and Health Care Expenditure by Insurance Status, Concentration Index ……… 138

Table 3.16: Poverty and Health Care Expenditure of Early Enrollers ………………………………. 139

Table 3.17: Enrollment Patterns (Early Enrollers) ………………………………………………….. 140

Table A.3.1: Poverty and Health Care Consumption ……………………………………………….. 146

Table A.3.2: Poverty and Health Care Utilization by Enrollment Status ……………………………. 147

Table A.3.3: Heterogeneous Effect of Insurance by Poverty Status ………………………………… 148

Table A.3.4: Poverty and Health Care Utilization by Insurance Status …... 149

ACKNOWLEDGEMENTS

I would like to thank Craig McIntosh and Karthik Muralidharan, my primary advisors, for their insightful advice and constant support. I would also like to thank my committee members, Gordon Dahl, Paul Niehaus, and Joshua Graff Zivin, for their invaluable guidance. I also thank the staff at Chaitanya, particularly Kalpana Pant, for their overall support and enthusiasm for my dissertation project. I would also like to acknowledge the Microinsurance Innovation Facility for their support of my dissertation project.

Chapter 1 is currently being prepared for submission for publication of the material. Muralidharan, Karthik; Sheth, Ketki. Bridging Education Gender Gaps in Developing Countries: The Role of Female Teachers. The dissertation author was the primary investigator and author of this material.

VITA

2004	Bachelor of Arts, Economics and Psychology, UC Berkeley
2005 – 2006	Research Assistant, Peterson Institute for International Economics
2006 – 2008	Indicorps Fellowship, Chaitanya (Maharashtra, India)
2011	Masters of Arts, Economics, UC San Diego
2009 – 2013	Research Assistant, Professor Karthik Muralidharan (UC San Diego)
2010	World Bank Consultant, South Asian Human Development Sector
2011	World Bank Consultant, Regional Study on Education Quality
2011 – 2014	Principal Investigator, ILO/EUDN Grant
2014	Doctor of Philosophy, Economics, UC San Diego

PUBLICATIONS

"Gender Discrimination in the Family" (with Prashant Bharadwaj and Gordon Dahl), in *Family Economics*, ed. by Esther Redmount, ABC-Clio Publishers, forthcoming.

"US China Trade Disputes: Rising Tides, Rising Stakes" (with Gary Hufbauer and Yee Wong). Institute for International Economics, Washington D.C., 2006.

FIELDS OF STUDY

Major Field: Economics (Development and Labor Economics)
Professor Craig McIntosh and Karthik Muralidharan, Advisors

ABSTRACT OF THE DISSERTATION

Essays in Development Economics

by

Ketki Sheth

Doctor of Philosophy in Economics

University of California, San Diego, 2014

Professor Karthik Muralidharan, Chair
Professor Craig McIntosh, Co-Chair

The following dissertation evaluates methods to improve the delivery of education and health care in low income countries.

In “The Education Gender Gap in Developing Countries: The Role of Female Teachers”, joint with Karthik Muralidharan, we add to the limited empirical literature on whether female teachers improve girls’ education outcomes in developing countries, a policy frequently advocated. Using a difference-in-difference estimate with fixed effects, we find that teachers are relatively more effective at teaching students of their own gender. However, female teachers are more effective overall, resulting in improvement of girls’ test scores by .036 standard deviations per year and a lack of adverse effects for boys.

In “The Distributional Consequences of Micro Health Insurance: Can a Pro-Poor Program Prove to be Regressive?”, I estimate heterogeneous effects of heath care consumption by poverty status and the related redistribution of premiums. Understanding these effects can inform optimal design of MHI contracts to maximize benefits and reduce unintended adverse effects. I document that poorer households consume significantly less health care at baseline, suggesting MHI may unintentionally lead to poorer households subsidizing wealthier households. But strikingly, twenty months after the introduction of MHI, there is no significant relationship between health care consumption and income among enrolled households. Thus, even though *ex-ante* health care consumption suggests MHI will result in regressive premium redistribution, *ex-post* behavior suggests the poor will not subsidize wealthier households.

The next two chapters discuss community based micro health insurance (MHI). MHI, insurance targeted at low income populations, has been an increasingly popular policy, though empirical evidence of its effectiveness has been limited. In “Evaluating Health-Seeking Behavior, Utilization of Care, and Health Risk: Evidence from a Community Based Insurance Model in India”, I assess the extent to which MHI reduces vulnerability and increases access to health care. I exploit a staggered expansion of MHI in which the villages offered the contract were randomly selected. I fail to find support for increased health-seeking behavior, but find suggestive evidence of reduced health shocks. This suggests the potential of MHI to improve the poor’s health and implies strengthened financial sustainability of MHI programs.

Chapter 1: Bridging Education Gender Gaps in Developing Countries:

The Role of Female Teachers

Abstract: Recruiting female teachers is frequently suggested as a policy option for improving girls' education outcomes in developing countries, but there is surprisingly little evidence on the effectiveness of such a policy. We study gender gaps in learning outcomes, and the effectiveness of female teachers in reducing these gaps using a large, representative, annual panel data set on learning outcomes in rural public schools in the Indian state of Andhra Pradesh. We report six main results in this paper. (1) We find a small but significant negative trend in girls' test scores in both math (0.02σ/year) and language (0.01σ/year) as they progress through the public primary school system; (2) Using five years of panel data, school-grade and student gender by grade fixed effects, we find that both male and female teachers are more effective at teaching students of their own gender; (3) However, female teachers are more effective overall, resulting in girls' test scores improving by an additional 0.036σ in years when they are taught by a female teacher, with no adverse effects on boys when they are taught by female teachers; (4) The overall gains from having a female teacher are mainly attributable to their greater effectiveness at improving math test scores than male teachers (especially for girls); (5) We find no effect of having a same-gender teacher on student attendance, suggesting that the mechanism for the impact on learning outcomes is not on the extensive margin of increased school participation, but on the intensive margin of more effective classroom interactions; (6) Finally, the increasing probability of having a male teacher in higher grades can account for around 10-20% of the negative trend we find in girls' test scores as they move to higher grades.

1. Introduction

Reducing gender gaps in education attainment has been an important priority for international education policy, and is explicitly listed as one of the United Nations Millennium Development Goals (MDGs). This commitment has been reflected in the policies of many developing countries, and substantial progress has been made in the past decade in reducing gender barriers in primary school enrollment. One key policy that is credited with increasing girls' education is the increased recruitment of female teachers (UNESCO 2012, Herz and Sperling 2004, UN 2012). UNICEF has documented the practice in a variety of countries, including Bangladesh, India, Liberia, Nepal, and Yemen, and the United Nations Task Force for achieving the MDGs has advocated hiring more female teachers as an effective policy mechanism for reaching the goal of universal primary education of girls (UNDG 2010, Rehman 2008, Slavin 2006).

While the idea that hiring more female teachers can bridge gender gaps is widely prevalent among policy makers, there is very little empirical evidence on testing this hypothesis in developing countries. In this paper, we study the causal impact of having a female teacher on the learning gains of female students, using one of the richest datasets on primary education in a developing country. The dataset features annual longitudinal data on student learning measured through independent assessments of learning conducted over five years across a representative sample of 500 rural schools and over 90,000 students in the Indian state of Andhra Pradesh. The data also includes detailed information on teacher characteristics and on their assignments to specific classrooms in each year.

The combination of panel data and variation in the gender of teachers and students allows us to estimate the causal impact of matching teacher and student gender in a *value-added* framework. Identification concerns are addressed by showing that our causal estimates of gender matching do not change under an increasingly restrictive set of specifications including school, school-grade, and student gender by grade fixed effects. We also show that there is no correlation between the probability of being assigned a female teacher and either the fraction of female students in the class or their mean test scores at the start of the year. Further, our estimation sample is restricted to schools that only have one section per grade, which precludes the possibility that students may be tracked across sections and that female teachers may be assigned to different sections based on unobservables.

We report six main findings in this paper. First, we find a small but significant negative trend in girls' test scores in both math (0.02σ/year) and language (0.01σ/year) as they advance through the five grades of primary school.[1] Girls have significantly higher test scores in language and equal test scores in math relative to boys at the end of grade one, but score almost on par with boys in language and significantly worse in math by the end of grade five. These results are consistent with evidence of gender gaps in test scores (particularly in math) documented in both high and low income countries (Fryer and Levitt 2010, Bharadwaj et al. 2012), and suggest that the growing gender gaps documented at later ages in both these papers probably reflect a cumulative effect of a trend that starts as early as primary school.

Second, using five years of panel data and school-grade and student gender by grade fixed effects, we find that teachers are .034σ/year more effective in teaching students of their own gender *relative* to teachers of the opposite gender. In other words, female teachers are .034σ/year more effective at reducing the gender gap in achievement than male teachers. Since female teachers differ from male teachers on several characteristics that may be correlated with teacher quality, we test the robustness of the 'gender-match' result by including interactions between student gender and each of the teacher characteristics on which female and male teachers differ, and find that our estimates above are essentially unchanged.

The result above is a difference-in-difference estimate that compares the relative advantage of female teachers in teaching girls rather than boys with the relative disadvantage of male teachers in teaching girls rather than boys, and is *symmetric by construction*. However, the overall effectiveness of a teacher is also determined by his or her effectiveness at teaching students of the opposite gender. Our

[1] As we discuss later, this estimate is based on the sample of test takers in public schools, and cannot account for the biases that may occur due to differential migration to private schools and differential absence on the day of the test by gender.

third result speaks to this issue and we find that female teachers in our setting are more effective overall than male teachers. We find that girls who have a female teacher in a given year have .036σ higher annual test score gains than if she had a male teacher. However, boys perform similarly regardless of the gender of their teacher. Thus, girls are likely to benefit from a policy of hiring more female teachers, and overall educational performance is likely to increase due to the lack of any offsetting effect on boys.

Fourth, these effects differ by subject. In particular, female teachers are more effective at teaching math relative to language when compared to male teachers. While girls continue to fare better with female teachers relative to male teachers in both language and math, the effect is greater in math relative to language. Boys though fare a little worse with female teachers (relative to male teachers) in language, and experience no differential effect of teacher gender in math. Together, these results suggest that the overall gains from hiring female teachers come mainly from improving mean math test scores relative to male teachers (positive for girls, no effect for boys) than from language (positive for girls, negative for boys, and no overall effect).

Fifth, we also study the impacts of a teacher-student gender match on student attendance, and find no evidence that teachers are more effective at raising the attendance for students of the same gender. This suggests that the likely mechanism for the 'matching' effect on test scores is not on the extensive margin of increased student-teacher contact time, but rather on the intensive margin of more effective classroom interactions.

Finally, we document that female teachers are more likely to teach in earlier grades. Combined with the results above, we estimate that around 10-20% of the trend of increasing gender gaps in test scores over time can be attributed to the reduction in the probability of girls being taught by female teachers as they advance to higher grades. Since teachers in higher grades are more likely to be male across several countries (NCES 2011, UNESCO 2010), our results suggest that an important channel for growing gender gaps in achievement (especially in math) could be the greater likelihood of having male teachers in higher grades.

While there have been several studies on the impact of shared gender between teachers and students on learning outcomes in developed country contexts, there is surprisingly little well-identified evidence on this question from developing countries. In the US and UK, studies have shown improved test scores, teacher perception, student performance, and engagement of girls when taught by a female teacher in schools, with magnitudes of test score impacts similar to those found in our paper (Dee 2007, Dee 2005, Nixon and Robinson 1999, Ehrenberg et al. 1995, Ouazad and Page 2012). However, other studies conducted in both the US and in European countries have failed to find such an effect (Holmund and Sund 2008, Carrington, Tymms and Merrell 2008, Lahelma 2006, Winters et al. 2013, Marsh et al. 2008, Driessen 2007, Neugebauer et al. 2011). In higher education institutions in the US, female professors have been found to have small effects on female students' course selection, achievement, and major choice (Bettinger et al. 2004, Carrell et al. 2010, Hoffmann and Oreopoulos 2009).[2]

However, the question of the role of female teachers in reducing gender gaps is much more salient in developing country contexts, where gender gaps in school enrollment and attainment are much larger and increased recruitment of female teachers is actively advocated (OECD 2010, Hausmann et al. 2012, Muralidharan and Prakash 2013, Bharadwaj et al. 2012). The only related paper in a developing country setting is Rawal and Kingdon (2010), who use test score data on 2nd and 4th grade students in the Indian states of Bihar and Uttar Pradesh, and find a positive impact on educational achievement for girls taught by female teachers, but find no similar effect for boys.

In addition to providing well-identified estimates of the impact of matching teacher and student gender on learning outcomes in a developing country (where the literature is very sparse), our dataset

[2] Analogous to gender, studies in the United States have also looked at the effect of sharing the ethnicity of a teacher and have generally found positive effects on such educational outcomes as drop outs, pass rates, and grades at the community college level, and teacher perceptions and student achievement in school going children (Dee 2004, Dee 2005, Farlie et al. 2011). We find no similar effect on other important dimensions in the Indian context, particularly disadvantaged castes and minority religions. We do not focus on caste and religion because the fraction of teachers and students in the relevant categories are small (typically less than 20%) and as a result the fraction of 'matches' are usually less than 5% (and often much smaller), which makes the estimates less stable to the series of robustness checks that we use in this paper to ensure that the estimates of the 'match' are well identified.

allows us to make advances relative to both the developed and developing country literatures on this subject. First, while several existing papers in this literature (especially those looking at college-level outcomes) use grades or test scores assigned by the students' own teachers, the test scores used in this paper are based on independent assessments and grading. This limits the concern that the measured effects of gender matching may reflect more generous grading by teachers towards students who share their own gender and allows us to be confident that the effects we measure reflect genuine impacts on learning.

Second and more important, the majority of papers in the global literature on this question (including Dee 2007 and Rawal and Kingon 2010) use student fixed effects and variation in the gender of teachers across different subjects to identify the impact of the gender match on learning, but they are based on comparing *levels* of test scores as opposed to value added. Thus, it can be difficult to interpret the magnitudes of the estimated effects without knowing the gender composition of the teachers in that subject in previous grades.[3] Our use of five years of annual panel data on test scores allows us to estimate the impact of a gender match on the *value-added in the year that the match occurred*, which has a much clearer interpretation relative to the standard in the literature. Finally, we observe students at a younger and more formative age than most of the literature, when the role of sharing gender may be especially important. This is also the age that is most relevant to policy for reducing education gender gaps in developing countries since the majority of students do not complete more than eight years of school education.

The remainder of this paper is organized as follows: Section 2 describes the dataset and presents summary statistics on students and teachers; Section 3 lays out the estimation and identification strategies; Section 4 presents the main results, and section 5 concludes.

[3] Thus, if this approach finds that a girl in eighth grade who has a female language teacher and a male math teacher does better in language, the interpretation of the point estimate is confounded by the possibility that the girl is also more likely to have had female language teachers in earlier grades (especially if teacher gender is correlated with subjects taught across grades, which is likely to be true).

2. Context and Dataset

India has the largest primary schooling system in the world, catering to over 200 million children. As in other developing countries, education policy in India has placed a priority on reducing gender disparities in education, and both the Five Year Plans and Sarva Shiksha Abhiyan (SSA), the flagship national program for universal primary education, have called for an increase in recruiting female teachers as a policy for increasing girls' education. SSA requires that 50% of new teachers recruited be women, and the 11th Five Year Plan suggested that it be increased to 75% (Government of India 2008). These calls for increased female teachers reflect a belief that through such mechanisms as role model effects, increased safety, reduced prejudices, and greater identification and empathy, female teachers are arguably more effective in increasing girls' achievement in primary school relative to their male counterparts (Ehrenberg et al. 1995, Stacki 2002, Dee 2005).

This paper uses data from the Indian state of Andhra Pradesh (AP), which is the 5th most populous state in India, with a population of over 80 million (70% rural). The data was collected as part of the Andhra Pradesh Randomized Evaluation Studies (AP RESt), a series of experimental studies designed to evaluate the impact of various input and incentive-based interventions on improving education outcomes in AP.[4] The project collected detailed panel data over five years (covering the school years 2005-06 to 2009-10) on students, teachers, and households in a representative sample of 500 government-run primary schools (grades 1 through 5) across 5 districts in Andhra Pradesh. The dataset includes annual student learning outcomes as measured by independently conducted and graded tests in language (Telugu) and math (conducted initially at the start of the 2005-06 school year as a baseline, and subsequently at the end of each school year), basic data on student and teacher demographics, and household socio-economic data for a subset of households. The test scores are normalized within each year-grade-subject combination and all analysis is conducted in terms of normalized test scores, with magnitudes being reported in standard deviations.

---

[4] These interventions are described in Muralidharan and Sundararaman (2011).

The Appendix provides further details on the dataset, including sample size and attrition between years. There is some differential attrition in the sample over time by gender (where attrition is defined as the fraction of students who had taken a test at the end of year 'n-1', but did not take a test at the end of year 'n'), with female students more likely to be in the test-taking sample (around 3% each year). However, this attrition over time is not a first-order concern for this paper because it is highly unlikely that the additional 3% of female students who appear for the test each year (relative to boys) would have test scores that are *differentially affected* by having a female teacher. This is further supported by Table 8 and by Table A.1.3, where we show that having the same gender as the teacher does not have any impact on either student attendance on a typical school day or on student presence in the end of year test.

Table 1 - Panel A, presents descriptive statistics on students who have at least one recorded test score and data on gender in the dataset.[5] Girls comprise 51% of the sample of public-school students in our sample. This does not imply that more girls are going to school than boys since it is likely that more boys are attending private schools (Pratham 2012). However, it does illustrate that on average, girls are well represented in public primary schools and in our sample. The girls in the sample come from modestly better off socioeconomic backgrounds than the boys, and have parents who are slightly more educated and affluent. These differences probably reflect two dimensions of selection into the sample – better off households are more likely to send girls to school, and better off households are more likely to send boys to private schools. However, the magnitudes of these differences are quite small (often in the range of 0-2 percentage points), and the statistical significance reflects the very large sample size. Since the household surveys were completed for only 70% of the sample of students for whom we have test score data, our main specifications do not include household controls.[6]

---

[5] Less than 3 percent of students with test scores have no recorded gender.

[6] While there are a few observable differences between the boys and girls in the sample, including these in the estimation will only matter if there are differential interactions between these household characteristics and teacher gender across boys and girls. We verify that our results are robust to the inclusion of household characteristics, but prefer to not include household characteristics in our main estimating equations because doing so reduces the sample size by 30% and it is possible that the

Table 1 - Panel B, presents summary statistics for the teachers in our analysis. Female teachers comprise 46% of the total teacher body, but are less experienced, less likely to have completed high school or a masters degree, and less likely to hold a head-teacher position. Not surprisingly, their mean salaries are also lower. They also comprise a much greater share of the contract teacher work-force than that of regular civil-service teachers. Since teacher characteristics vary systematically by gender, we will report our key results on the impact of matching teacher and student gender, both with and without controls for these additional teacher characteristics. We will also conduct robustness checks of our main results on the effects of a teacher-student 'gender match' on learning outcomes, by including interactions of student gender with each of the teacher characteristics that are different across male and female teachers.

Table 2 - Panel A presents summary statistics on gender differences in test scores by grade. We see that girls score as well as boys in math and score 0.05σ *higher* on language in grade 1. However, there is a steady decline in girls' test scores in both math and language as they advance through higher grades, and by the last two years of primary school (grades 4 and 5) we see that girls' initial advantage in language scores has declined and they do significantly worse than boys on math (by around .1σ). Table 2 - Panel B quantifies the annual decline in girls' relative scores by including an interaction term between student gender and grade in a standard value-added specification. We find evidence of a growing education gender gap among test takers in public primary school, with a mean decline of 0.02σ/year in math scores and 0.01σ/year decline in language scores for girls relative to boys. Since the data includes nine different cohorts of students (see Appendix), we also include cohort fixed effects, and see that the estimates of the gender gaps and of the trend in the gender gap across grades are unchanged. Similarly, the results are also robust to including school fixed effects.

---

remaining sample may have some non-random attrition. Results with household controls are available on request.

One caveat to the interpretation of the above numbers is that they are based on a representative sample of test-taking students in public schools. Relative to the gender gap in the universe of primary-age school children, our estimate may be biased downwards if higher-scoring boys are differentially more likely to attend private schools. Conversely, they may be biased upwards if lower-scoring boys are more likely to be absent on the day of testing. While we cannot estimate these, it is more likely that we under-estimate the gender gap, because boys aged 7-10 in rural AP are around 10 percentage points more likely to be enrolled in a private school during this period (45% versus 35% in 2010 - Pratham 2010), whereas girls in public schools are only 3% more likely to be present on the day of testing (Table A.1.3).

In spite of these caveats, this documentation of gender gaps in a representative sample of public schools in rural AP is a useful contribution to the literature on gender gaps in test scores in developing countries, because there are very few longitudinal data sets on student test scores in low-income settings, and no other paper that we are aware of is able to document these gaps with *cohort fixed effects.* Further, the literature on gender gaps in test scores mostly relies on samples of students who take tests in schools, and therefore has the same limitations we discus above.

3. Estimation and Identification

Our main estimating equation takes the form:

$$(1) \qquad E_{itjk} = \alpha + \gamma E_{it-1j-1k} + \beta_1 (F * g)_{itjk} + \beta_2 g_{itjk} + \beta_3 F_{itjk} + \boldsymbol{\delta T_{itjk}} + \mu_{itjk}$$

where E_{itjk} are student educational outcomes (test scores and attendance) for student *i*, in year *t*, grade *j*, and school *k* respectively. g_{itjk} is an indicator for whether the student is a girl, F_{itjk} is an indicator for whether the student's current teacher is female, and $F*g_{itjk}$ is an indicator for whether a girl student shares her teacher's gender in the current year. $\boldsymbol{T}_{itjk}$ is a vector of additional teacher characteristics, and μ_{itjk} is a stochastic error term. The inclusion of the lagged test score on the right-hand side of (1) allows us to estimate the impact of contemporaneous inputs in a standard *value-added* framework. Since all test scores are normalized by grade and subject, the estimated coefficients can be directly interpreted as the correlation between the covariate and annual gains in normalized test scores.[7] When studying attendance we do not include the lagged attendance of the previous year.

The above estimating equation allows us to calculate the marginal impact of changing each component of the feasible student-teacher gender combinations relative to boys taught by male teachers (the omitted category).

The first coefficient of interest in this paper is β_1, which indicates the extent to teachers are relatively more effective at teaching to their own gender compared to teachers of the opposite gender. Since the indicator variable is based on the interaction of dummies for teacher and student gender, the coefficient is a 'difference in difference' estimate of the impact of female teachers when teaching girls rather than boys *relative* to their male counterparts teaching girls rather than boys. The coefficient on the interaction term therefore reflects the sum of the relative advantage of female teachers when teaching girls (rather than boys) and the relative disadvantage of male teachers when teaching girls (rather than boys). (i.e., β_1 = *(female teachers teaching girls – female teachers teaching boys) – (male teacher teaching girls – male teachers teaching boys)*).

A more intuitive way of understanding this is to note that β_1 represents the relative effectiveness of female teachers (compared to male teachers) in reducing the test score gap between girls and boys. By construction, this is symmetric and equivalent to the relative effectiveness of male teachers teaching boys compared to girls *relative* to female teachers teaching boys compared to girls. It is important to highlight that a positive β_1 does not necessarily imply that both boys and girls have better outcomes when sharing their teacher's gender. For example, a positive β_1 could co-exist with a situation where all students are better off with female (or male) teachers because the general effectiveness of female (or male) is considerably higher (even for students of the opposite gender).

[7] In the case of grade 1 where there is no lagged score (since there was no testing prior to enrolling in school), we set the normalized lagged score to zero. Our results on the impact of 'gender matching' on test score gains are unchanged if we drop grade 1 from the analysis.

β_2 is the difference in test score gains of girls taught by male teachers relative to boys taught by male teachers (i.e., *male teachers teaching girls – male teachers teaching boys)*. β_3 is the difference in test score gains of boys taught by female teachers relative to when taught by male teachers (i.e., *female teachers teaching boys – male teachers teaching boys)*. Thus, β_3 estimates the extent to which boys perform differently when they are taught by a female teacher relative to a male teacher.

Starting with the omitted category (of male teachers teaching boys), adding combinations of $\beta_1, \beta_2,$ and β_3 allow us to measure other marginal effects of interest. Analogous to β_3 for boys, testing if $\beta_1 + \beta_3 > 0$ provides a formal test of whether girls gain by being paired with female teachers relative to male teachers. The derivation is below:

$$(2)\ \textit{Female teachers teaching girls} - \textit{Male teachers teaching girls} > 0$$

$$\Rightarrow (\alpha + \beta_1 + \beta_2 + \beta_3) - (\alpha + \beta_2) > 0 \Rightarrow \beta_1 + \beta_3 > 0$$

As highlighted earlier, it is possible that female teachers are relatively more effective at teaching girls than boys compared to male teachers (a positive β_1), but that female teachers are overall less effective (a negative β_3), resulting in girls being better off with male teachers despite the loss in gains from not sharing their teacher's gender ($\beta_1 + \beta_3 < 0$).

Additionally, if we value both boys' and girls' educational achievement equally, then we would be interested in knowing whether the positive gain for girls taught by female teachers outweighs any adverse effects from mismatching boys to being taught by female teachers (i.e., (*potential gain to girls + potential loss to boys).* The formal test for this is $\lambda_g\beta_1 + \beta_3 > 0$ where λ_g is the proportion of girls in schools. The derivation is below:

$$(3)\ \lambda_g * \textit{potential gain to girls} + (1 - \lambda_g) * \textit{potential loss to boys} > 0$$

$$\Rightarrow \lambda_g(\beta_1 + \beta_3) + \left(1 - \lambda_g\right)(\beta_3) > 0 \Rightarrow \lambda_g\beta_1 + \beta_3 > 0$$

Thus, if the effect of female teachers on boys was negative, but their effect on girls was positive, we would find that: $\beta_3 < 0$ *and* $\beta_1 + \beta_3 > 0$. The test outlined in Eq (3) can also be interpreted as the overall effectiveness of female teachers relative to male teachers. Intuitively the impact of replacing a male teacher in a classroom with a female teacher is equal to the sum of the impact of the female teacher on all students (β_3), and the additional gains to female students from matching with a female teacher (β_1), weighted by the fraction of female students in the classroom (λ_g).

The main identification challenge in interpreting these coefficients causally is that teachers are not randomly assigned to schools, and it is possible that schools with more female teachers are in areas with greater overall girls' education levels and steeper learning trajectories. Thus, it is possible that girls would perform well in these schools regardless of their teacher's gender. In such a case, the estimate of β_1 could be confounded by omitted variables correlated with both the probability of having a female teacher and steeper learning trajectories for girls. We address this concern by augmenting (1) with school fixed effects, and thereby estimating the impact of a gender-match on value-added *relative* to the schools' average effectiveness at improving value-added.

A further concern could be that teachers are not assigned randomly to grades within schools, and a similar omitted variable concern would apply if female teachers are differentially assigned to grades in which girls are more likely to show greater test score gains (for instance, if female teachers are more likely to be assigned to younger grades and if girls outperform boys in earlier grades). To address this concern, we include school grade fixed effects, which controls for the average performance in a given *grade* in the school (instead of the overall performance of the school). Finally, to account for potentially differential trajectories of learning in different grades *by gender*, we also include grade fixed effects by student gender to estimate the parameters of interest by comparing educational outcomes relative to girls' and boys' average learning trajectories in each grade. Our preferred specification therefore includes both school-grade fixed effects and grade fixed effects by gender to address this concern.[8]

[8] Since the data are drawn from schools that were exposed to various experimentally-assigned programs, all estimates include dummy variables indicating the treatments assigned to the school. This turns out to not matter in practice because our main specifications of interest use school-fixed effects, which makes the treatment status of the school irrelevant for identification purposes.

A final concern could be that if grades in a school have multiple sections, then the assignment of teachers to sections within grades could be based on omitted variables such as a greater probability of assigning female teachers to sections that have girls with a greater likelihood of improving test scores. However, this is not an important factor in our setting because schools typically have fewer teachers than grades, and the typical teaching arrangement is one of multi-grade teaching (where the same teacher simultaneously teaches multiple grades) and so there are only few cases where there are multiple sections per grade with different teachers assigned to different sections. We drop all such cases (6% of observations) where there are multiple teachers per grade.

Note that our identification strategy does not require teacher gender to switch in a given school grade over time, and neither does it require teacher gender to switch within a cohort over time (across different grades).[9] Rather, the inclusion of school-grade and gender grade fixed effects implies that the identifying variation is coming from the differential effectiveness of teachers (by gender) at teaching girls versus boys relative to (a) the mean value added experienced by students in that school and grade over the five years of data, and (b) the mean value added for girls relative to boys in that grade across all schools in the sample.

3.1. Testing the Identifying Assumptions

Table 3 shows the correlation between various classroom characteristics and the probability of the classroom having a female teacher. We see that there is no significant correlation between having a female teacher and the fraction of girls in the classroom or with the average test scores of incoming cohorts for either gender. Female teachers *are* more likely to be assigned to younger grades. But once school-grade fixed effects are included, this is no longer an issue for average female teacher effects, and it

[9] We avoid using a student fixed effects estimate because the identifying variation in a specification with student fixed effects would come from changes in teacher gender in different grades. However, as we see in 3.1, girls have higher value-added in lower grades, and female teachers are more likely to be assigned to lower grades. This would therefore create an upward bias in the 'matching estimate'.

continues to be case that there is no significant correlation between having a female teacher in the class and either the fraction of female students or the test scores of the incoming cohort (columns 5 and 6).

However, we see in Table 4 that girls do have a slightly more concave learning trajectory than boys. We estimate a standard value-added model that controls for lagged test scores (as in Eqn. 1), but allow for an interaction between student gender and grade, and we see that female students have lower value-added in higher grades. Since female teachers are more likely to be assigned to lower grades, the inclusion of school-grade fixed effects (i.e., the average test score gain in a grade within a school over the five years across both student genders) does not address the possible spurious correlation from female teachers being more likely to be assigned to grades where female students fall behind boys at a lower rate. Therefore, we also include grade fixed effects by student gender in our main specifications to control for average value-added test scores in each grade *by student gender*. Thus, the parameters of interest in Eq. (1) are identified relative to the average learning trajectory for girls in the same grade (student gender grade fixed effects) and relative to the average learning trajectory in the same school for that grade (school-grade fixed effects).

We also verify that there is no significant difference between classrooms taught by male and female teachers on any of the household socio-economic variables listed in Table 1 (tables available on request), but we focus our attention on the test-scores of incoming cohorts as the most useful summary statistic of previous inputs into education to test balance on, because the sample size with the household survey is 30% smaller than that of just the test scores.

4. Results

The main results of the paper (from the estimation of Equation 1) are presented in Table 5, which pools the results across subjects (results separated by subject are in Table 7). The columns show increasingly restrictive identification assumptions with school fixed effects (Column 2), school-grade fixed effects (Column 3), and both of these with grade fixed effects by student gender (Column 4 and 5). Column 6 expands the preferred specification in Column 5 with the inclusion of teacher covariates to differentiate between a pure "gender effect" versus effects driven by teacher characteristics correlated with teacher gender. Thus, the estimates in column 5 are relevant to the policy question: "What will happen if we replace a male teacher with a female teacher whose characteristics are the same as those of the average female teacher?" On the other hand, the estimates in column 6 answer the question: "What will happen if we just switch a teacher's gender from male to female holding other observable characteristics constant?" While our main results are remarkably stable and robust under the various specifications, our discussion below will use the estimates in columns 5 and 6, unless mentioned otherwise.

Averaged across subjects, we see that teachers are .034σ/year more effective in teaching to their own gender relative to a student of the opposite gender compared to teachers of the other gender. In other words, female teachers are .034σ/year more effective in reducing the gender gap between girls and boys relative to male teachers. We find no negative effect on boys from being taught by female teachers relative to male teachers (β_3 is close to zero). We estimate that girls gain an extra .036σ/year when taught by female teachers instead of male teachers ($\beta_1 + \beta_3$), and that there is a statistically significant net increase in annual test score gains of .019σ/year from replacing a male teacher with a female one ($\lambda_g * \beta_1 + \beta_3$). However, once we control for teacher characteristics, this net welfare effect drops to .013σ/year, suggesting that characteristics correlated with female teachers may partly contribute to female teachers being more effective overall.

This discussion points to an important caveat to the interpretation of these results. Since female teachers are systematically different from their male counterparts (Table 1 - Panel B), it is possible that the β_1 estimated in (1) reflects not just the effect of female students matching with female teachers, but the effect of female students matching with teacher *characteristics* that are systematically more commonly found in female teachers. We address this concern in Table 6, where we show a series of regressions where we follow the specification in (1), but include teacher characteristics and the interaction of this characteristic with student gender. These include teacher demographic characteristics that may be correlated with teaching effectiveness (such as education, training, contractual status, seniority, and salary) as well as teaching conditions (multi-grade teaching) and measures of teacher effort (absence). Doing so allows us to test the extent to which the positive β_1 found in Table 5 reflects a 'gender' match as opposed to other characteristics of female teachers that differentially effect girl students.

Table 6 reports the key results without controlling for other teacher characteristics. The main result is that the estimates of β_1 are remarkably robust to including the student interactions with teacher characteristics that vary by teacher gender.[10] In all cases, the estimate of the gain to a female student from switching to a female teacher $(\beta_1 + \beta_3)$ is positive and significant (ranging from 0.03 to 0.04 σ/year), and so is the estimate of the overall gain to a classroom $(\lambda_g * \beta_1 + \beta_3)$ from having a female instead of a male teacher (ranging from 0.015 to 0.025 σ/year). The results in Panel B show that the figures are even more consistent (and always significant) when controls for other teacher characteristics are included. The range of the magnitudes is much tighter with $(\beta_1 + \beta_3)$ mostly being 0.031σ/year and $(\lambda_g * \beta_1 + \beta_3)$ *always* being 0.015σ/year.

Table 7 breaks down the results by subject (Panels A and B). Comparing $(\beta_1 + \beta_3)$ across subjects suggests that the gains to girls from having a female teacher are higher in math. Finally, comparing the total social gains of shifting from a male to a female teacher $(\lambda_g * \beta_1 + \beta_3)$ across subjects, we see that the gains in math are significantly larger than those in language (not shown). Further, once we control for teacher characteristics, all the gains in Column 6 of Table 5 can be attributed to the better performance of female teachers in math (where female teachers do much better with girls and no worse with boys) with the net effects in language being close to zero (positive for girls and negative for boys).

---

[10] In the interest of space, we only show these results for characteristics that are significantly different across teacher gender (see Table 1 - Panel B). The estimate of β_1 is unchanged and significant for interactions with other teacher characteristics (such as religion and caste) as well.

We also study the impact of a teacher-student gender match on student attendance. We find no significant effect of a gender-match on student attendance (Table 8). We do find that female teachers are slightly more effective at increasing attendance overall (by around 0.6 percent), but there is no differential impact by student gender. This result is interesting because the rhetoric of hiring female teachers is often based on the belief that having female teachers increases the safety and comfort of girls in school, and that their presence therefore encourages girls to attend school. Our results suggest however, that the mechanism for the positive impact of a gender match on test scores is less likely to be due to effects on the extensive margin of school participation, but more due to the increased effectiveness of classroom transactions between teachers and students.

Of course, this result could be reflecting a scenario where total primary school enrollment for both boys and girls is over 98% (Pratham 2012) and the role of female teachers in increasing attendance of female students may be more limited in such a setting. Nevertheless, our results suggest that even after achieving gender parity in school enrollment, there may be continued benefits to a policy of preferred hiring of female teachers due to their greater overall effectiveness in improving learning outcomes, and specifically due to their effectiveness in reducing gender gaps in test scores.

Finally, we calculate what proportion of the growing gender gap calculated in Table 2 can be attributed to girls being less likely to have a female teacher as they advance through primary school. Regressing the probability of a female teacher on the grade taught (with school fixed effects), we find that there is a 4 percentage point reduction in the probability of a student having a female teacher at each higher grade. Multiplying the reduced probability of a female teacher by the cost to girls of not having a female teacher in a given year ($\beta_1 + \beta_3$), and dividing this by the total annual increase in the test score gender gap (estimated in Table 2), we estimate that the reduced likelihood of female teachers in higher grades accounts for 9% of the annual growth in the gender gap in math and 21% in language (the fraction of the growing gender gap in language that is accounted for by this channel is higher than in math because the absolute magnitude of the annual growth in the gender gap is lower in language). Using estimates without school fixed effects, these figures would be 8% and 15% respectively (because the overall trend in the gender gap is slightly larger without school fixed effects - see Table 2).

5. Conclusion

We study gender gaps in primary school learning outcomes in a low-income setting using one of the richest datasets on primary education in a developing country. We find that at the start of primary school, girls have a slight advantage in the local language (approximately .05σ) and are at par in math. However, girls lose this advantage in both language (by 0.01σ/year) and in math (by 0.02σ/year) as they progress through the schooling system.

While these trends likely reflect a broad set of household, school, and social factors, one specific school-level policy that has been posited as a promising channel for mitigating these trends is the greater use of female teachers in low-income settings. This is a policy that has been widely recommended and adopted, but there has been very little well-identified evidence to support this claim. In this paper, we present some of the first well-identified empirical tests of this hypothesis in a low-income setting, using an extremely rich data set collected annually over five years in the Indian state of Andhra Pradesh.

Our results suggest that female (and male) teachers are relatively more effective when teaching to their own gender, that learning for girls increases when they are taught by female teachers relative to male teachers, and that boys do not suffer adverse effects when taught by female teachers relative to male teachers, even when controlling for teacher observables. These results differ across subjects, and the value to girls of having a teacher of the same gender is greater in math than in language. One possible explanation for this could be that boys and girls face different stereotypes in math and language and that shared teacher gender matters more in areas with negative stereotypes, such as a stereotype that girls are less good at math).

From a policy perspective, our estimates suggest that expanding the hiring of female teachers - both at the margin of the current patterns of hiring (assuming that the marginal female teacher hired has the same characteristics as the average female teacher), and also when holding other characteristics constant, would improve overall learning outcomes and be especially useful as a tool for bridging gender gaps in learning trajectories over time. While we find evidence to suggest that the mechanism of impact is through more effective classroom interactions (as opposed to increased teacher-student contact time), our data does not allow us to explore the further granularity of the specific mechanisms through which shared gender may influence learning (such as role model effects, greater empathy, and closer identification between teachers and students of the same gender). Decomposing the reduced form effects further could help in crafting more nuanced policies to capture these positive gains without having adverse effects on either gender.

Chapter 1 is currently being prepared for submission for publication of the material. Muralidharan, Karthik; Sheth, Ketki. Bridging Education Gender Gaps in Developing Countries: The Role of Female Teachers. The dissertation author was the primary investigator and author of this material.

Table 1.1: Summary Statistics by Gender

	Obs	Mean	Male	Female	Female - Male
	Panel A: Students				
	Obs	Mean	Male	Female	Female - Male
Female	94599	0.509			
Literate Father	66511	0.592	0.582	0.600	0.0185***
Literate Mother	66827	0.439	0.429	0.449	0.0199***
Proper House	66851	0.311	0.306	0.315	0.00981***
Has Toilet	66974	0.289	0.284	0.294	0.0106***
	Panel B: Teachers				
	Obs	Mean	Male	Female	Female - Male
Female	2680	0.457			
Head Teacher	2680	0.288	0.377	0.182	-0.195***
Regular Teacher	2680	0.503	0.497	0.511	0.0141
Contract Teacher	2680	0.188	0.116	0.273	0.157***
Completed Education: 12th Pass	2680	0.931	0.962	0.893	-0.0696***
Completed Education: Masters	2680	0.226	0.270	0.174	-0.0964***
Has Teacher Training	2661	0.833	0.909	0.743	-0.166***
Native to Village	2679	0.234	0.175	0.304	0.128***
Married	2676	0.810	0.845	0.769	-0.0762***
Active in Union	2674	0.183	0.276	0.074	-0.202***
Salary (monthly)	2674	9560	10697	8209	-2487.5***
Age	2660	36.905	39.542	33.750	-5.791***
Years Experience	2285	12.953	14.465	11.076	-3.389***
Teacher Absence	2666	0.191	0.197	0.184	-0.0135**
Multigrade Classroom	2680	0.458	0.475	0.437	-0.0386**
Classroom Enrollment	2680	23.225	22.869	23.647	0.778

Notes: (1) All variables are binary indicators, except for salary which ranges from 300 to 38400 (with a standard deviation of 5776), age which ranges from 12 to 58 (with a standard deviation of 9.76), and years of experience which ranges from 1 to 42 (with a standard deviation of 7.94). (2) Significance levels are as follows: *10%, **5%, and ***1%.

Table 1.2: Learning Gaps by Gender and Grade

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Panel A: Gender Differentials in Test Scores by Grade									
	Dependent Variables: Normalized Test Score (Within Grade)								
	Pooled Across Subjects			Math			Telugu		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Female (Grade 1)	0.0279**	0.0279**	0.0207**	0.00238	0.00237	-0.00377	0.0531***	0.0531***	0.0448***
	(0.0122)	(0.0122)	(0.00925)	(0.0135)	(0.0135)	(0.0101)	(0.0127)	(0.0127)	(0.0100)
No. of observations	66660	66660	66660	33187	33187	33187	33473	33473	33473
Female (Grade 2)	0.00526	0.00507	0.00580	-0.0271**	-0.0273**	-0.0241***	0.0376***	0.0374***	0.0356***
	(0.0114)	(0.0114)	(0.00828)	(0.0117)	(0.0117)	(0.00881)	(0.0122)	(0.0121)	(0.00889)
No. of observations	70953	70953	70953	35453	35453	35453	35500	35500	35500
Female (Grade 3)	-0.0217*	-0.0217*	-0.0225***	-0.0569***	-0.0570***	-0.0572***	0.0136	0.0135	0.0122
	(0.0118)	(0.0117)	(0.00813)	(0.0120)	(0.0119)	(0.00863)	(0.0128)	(0.0127)	(0.00894)
No. of observations	74715	74715	74715	37349	37349	37349	37366	37366	37366
Female (Grade 4)	-0.0442***	-0.0444***	-0.0375***	-0.0956***	-0.0957***	-0.0876***	0.00709	0.00698	0.0126
	(0.0120)	(0.0120)	(0.00770)	(0.0122)	(0.0121)	(0.00815)	(0.0130)	(0.0130)	(0.00864)
No. of observations	79972	79972	79972	39973	39973	39973	39999	39999	39999
Female (Grade 5)	-0.0262**	-0.0263**	-0.0206***	-0.0749***	-0.0750***	-0.0669***	0.0225*	0.0224*	0.0256***
	(0.0115)	(0.0115)	(0.00738)	(0.0123)	(0.0123)	(0.00771)	(0.0123)	(0.0123)	(0.00846)
No. of observations	85572	85572	85572	42777	42777	42777	42795	42795	42795
Panel B: Trends in Gender Differentials in Test Scores from Lower to Higher Grades									
Female	0.0311**	0.0311**	0.0271***	0.0115	0.0115	0.00814	0.0506***	0.0505***	0.0458***
	(0.0132)	(0.0132)	(0.00993)	(0.0142)	(0.0141)	(0.0106)	(0.0139)	(0.0139)	(0.0107)
Female*Grade	-0.0144***	-0.0144***	-0.0126***	-0.0207***	-0.0207***	-0.0189***	-0.00805*	-0.00803*	-0.00631**
	(0.00383)	(0.00382)	(0.00281)	(0.00410)	(0.00409)	(0.00298)	(0.00410)	(0.00410)	(0.00308)
No. of Observations	377872	377872	377872	188739	188739	188739	189133	189133	189133
Cohort Fixed Effects	No	Yes	Yes	No	Yes	Yes	No	Yes	Yes
School Fixed Effects	No	No	Yes	No	No	Yes	No	No	Yes

Notes: (1) Standard errors (in parentheses) are clustered at the school level for OLS regressions not including school fixed effects, and are clustered at the student level for OLS regressions including school fixed effects. (2) Significance levels are as follows: *10%, **5%, and ***1%.

Table 1.3: Characteristics of Classrooms Assigned to Female Teachers

	Dependent Variable: Classroom Has a Female Teacher					
	(1)	(2)	(3)	(4)	(5)	(6)
Proportion of Female Students	-0.000673	-0.0121	0.00109	-0.0103	0.00523	-0.00914
	(0.0338)	(0.0323)	(0.0209)	(0.0204)	(0.0205)	(0.0204)
Grade 1	0.0156	0.0641***	0.0243**	0.0658***		
	(0.0163)	(0.0153)	(0.0121)	(0.0116)		
Grade 2	0.0228	0.0491***	0.0278**	0.0460***		
	(0.0150)	(0.0141)	(0.0121)	(0.0116)		
Grade 4	-0.0671***	-0.0398***	-0.0676***	-0.0358***		
	(0.0150)	(0.0146)	(0.0123)	(0.0118)		
Grade 5	-0.140***	-0.0629***	-0.134***	-0.0539***		
	(0.0170)	(0.0163)	(0.0121)	(0.0119)		
Test Score of Incoming Cohort of Male Students	-0.0142	-0.0111	-0.00149	-0.000472	-0.00906	-0.00380
	(0.0228)	(0.0214)	(0.0134)	(0.0129)	(0.0136)	(0.0134)
Test Score of Incoming Cohort of Female Students	0.0189	0.00683	-0.00698	-0.00476	0.00172	0.00571
	(0.0191)	(0.0188)	(0.0119)	(0.0113)	(0.0124)	(0.0119)
Number of Observations	10974	9641	10974	9641	10974	9641
Teacher Characteristics	No	Yes	No	Yes	No	Yes
School Fixed Effects	No	No	Yes	Yes	No	No
School*Grade Fixed Effects	No	No	No	No	Yes	Yes
Boys' Test Score = Girls' Test Score (p-value)	0.3168	0.5708	0.7932	0.8298	0.6117	0.6483

Notes: (1) "Teacher Characteristics" are salary, age, experience, teacher absence, class enrollment size and indicators for caste, teacher status, education, training, native to school location, marital status, union status, and a multigrade class. (2) Standard errors (in parentheses) are clustered at the school level for OLS regressions not including fixed effects, and are clustered at the student level for OLS regressions including fixed effects. (3) Significance levels are as follows: *10%, **5%, and ***1%.

Table 1.4: Gender Differentials in Learning Trajectories from Lower to Higher Grades

	Dependent Variable: Normalized Test Scores					
	(1)	(2)	(3)	(4)	(5)	(6)
Female	0.0251**	0.0255***	0.00464	0.00667	0.0464***	0.0453***
	(0.0120)	(0.00923)	(0.0132)	(0.0102)	(0.0125)	(0.00980)
Female*Grade	-0.00624*	-0.00725***	-0.00830**	-0.0106***	-0.00563*	-0.00562**
	(0.00322)	(0.00252)	(0.00368)	(0.00281)	(0.00339)	(0.00271)
No. of Observations	304410	304410	151785	151785	152625	152625
School Fixed Effects	No	Yes	No	Yes	No	Yes

Notes: (1) Regressions include student's previous year's test score as an independent variable. (2) Standard errors (in parentheses) are clustered at the school level for OLS regressions not including school fixed effects, and are clustered at the student level for OLS regressions including school fixed effects . (3) Significance levels are as follows: *10%, **5%, and ***1%.

Table 1.5: Impact of Female Teachers on the Learning Gains of Female Students

	Dependent Variable: Normalized Test Scores					
	(1)	(2)	(3)	(4)	(5)	(6)
(β_1) Female Student * Female Teacher	0.0383***	0.0362***	0.0354***	0.0350***	0.0343***	0.0347***
	(0.00997)	(0.00788)	(0.00753)	(0.00792)	(0.00757)	(0.00804)
(β_2) Female Student	-0.0120*	-0.0140***	-0.0126**			
	(0.00676)	(0.00522)	(0.00498)			
(β_3) Female Teacher	-0.0154	-0.00344	0.000700	0.00212	0.00132	-0.00305
	(0.0188)	(0.00629)	(0.00697)	(0.00634)	(0.00699)	(0.00805)
$\beta_1 + \beta_3$	0.023	0.033	0.036	0.037	0.036	0.032
F-statistic (H_0: $\beta_1 + \beta_3 = 0$)	1.575	30.113***	29.585***	37.954***	28.615***	16.722***
λ_g*$\beta_1 + \beta_3$	0.004	0.015	0.019	0.020	0.019	0.015
F-statistic (H_0: λ_g*$\beta_1 + \beta_3 = 0$)	0.054	10.194***	10.913***	17.643***	10.944***	4.625**
Number of Observations	268548	268548	268548	268548	268548	235022
Teacher Characteristics	No	No	No	No	No	Yes
School Fixed Effects	No	Yes	No	Yes	No	No
School*Grade Fixed Effects	No	No	Yes	No	Yes	Yes
Grade Fixed Effects by Student Gender	No	No	No	Yes	Yes	Yes

Notes: (1) Regressions include student's previous year's test score as an independent variable. (2) "Teacher Characteristics" are salary, age, experience, teacher absence, class enrollment size and indicators for caste, teacher status, education, training, native to school location, marital status, union status, and a multigrade class. (3) Standard errors (in parantheses) are clustered at the school level for OLS regressions not including fixed effects, and are clustered at the student level for OLS regressions including fixed effects. (4) Significance levels are as follows: *10%, **5%, and ***1%.

Table 1.6: Heterogeneous Effects on Test Score Gains of Girls by Teacher Characteristics and Teacher Gender

	Dependent Variable: Normalized Test Scores						
	Panel A: Excludes Additional Teacher Correlates						
Teacher Characteristic:	Head Teacher	Contract Teacher	Completed 12th	Teacher training	Native to Village	Absence	MG
(β_1) Female Student * Female Teacher	0.0312***	0.0342***	0.0351***	0.0346***	0.0338***	0.0352***	0.0335***
	(0.00772)	(0.00765)	(0.00758)	(0.00764)	(0.00759)	(0.00761)	(0.00758)
(β_3) Female Teacher	-0.00228	0.00243	-0.000149	0.000203	-0.00156	-0.00350	0.000449
	(0.00708)	(0.00708)	(0.00700)	(0.00709)	(0.00702)	(0.00704)	(0.00698)
(δ_1) Female Student *Characteristic	-0.0179**	0.000603	0.0260	-0.00151	0.00786	-0.00428	-0.0295***
	(0.00846)	(0.0122)	(0.0188)	(0.0127)	(0.0100)	(0.0178)	(0.00770)
(δ_3) Teacher Characteristic	-0.0185**	-0.00902	-0.0423***	0.0102	0.00785	-0.0666***	-0.0162**
	(0.00743)	(0.0105)	(0.0158)	(0.0109)	(0.00837)	(0.0142)	(0.00767)
$\beta_1 + \beta_3$	0.029	0.037	0.035	0.035	0.032	0.032	0.034
F-statistic (H_0: $\beta_1 + \beta_3 = 0$)	18.248***	29.589***	27.669***	26.730***	23.338***	22.437***	26.115***
$\lambda_g*\beta_1 + \beta_3$	0.014	0.020	0.018	0.018	0.016	0.014	0.017
F-statistic (H_0: $\lambda_g*\beta_1 + \beta_3 = 0$)	5.586**	11.894***	9.762***	9.567***	7.547***	6.361**	9.532***
Number of Observations	268548	268548	268548	267475	268482	264581	268264

Notes: (1) Regressors include student's lagged normalized test score, school*grade fixed effects and grade fixed effects by student gender (Specification from Column 5 and 6 of Table 5). (2) "Teacher Characteristics" are salary, age, experience, teacher absence, class enrollment size and indicators for caste, teacher status, education, training, native to school location, marital status, union status, and a multigrade class. (3) Standard errors (in parantheses) are clustered at the school level for OLS regressions not including fixed effects, and are clustered at the student level for OLS regressions including fixed effects. (4) Significance levels are as follows: *10%, **5%, and ***1%.

Table 1.7: Impact of Female Teachers on the Lea1.rning Gains of Female Students by Subject

	Dependent Variable: Normalized Test Scores					
	(1)	(2)	(3)	(4)	(5)	(6)
	Panel A: Math					
(β_1) Female Student * Female Teacher	0.0338***	0.0335***	0.0336***	0.0312***	0.0315***	0.0302***
	(0.0111)	(0.00876)	(0.00843)	(0.00880)	(0.00847)	(0.00901)
(β_2) Female Student	-0.0374***	-0.0428***	-0.0408***			
	(0.00729)	(0.00575)	(0.00553)			
(β_3) Female Teacher	-0.0139	0.00240	0.00806	0.00916	0.00917	0.00924
	(0.0209)	(0.00701)	(0.00785)	(0.00706)	(0.00786)	(0.00909)
$\beta_1 + \beta_3$	0.020	0.036	0.042	0.040	0.041	0.039
F-statistic (H_0: $\beta_1 + \beta_3 = 0$)	0.928	28.781***	30.823***	35.745***	29.364***	20.412***
λ_g*$\beta_1 + \beta_3$	0.003	0.019	0.025	0.025	0.025	0.025
F-statistic (H_0: λ_g*$\beta_1 + \beta_3 = 0$)	0.027	13.635***	15.392***	22.120***	15.464***	10.257***
Number of Observations	133907	133907	133907	133907	133907	117205
	Panel B: Language (Telugu)					
(β_1) Female Student * Female Teacher	0.0429***	0.0393***	0.0373***	0.0385***	0.0364***	0.0392***
	(0.0104)	(0.00851)	(0.00819)	(0.00856)	(0.00824)	(0.00875)
(β_2) Female Student	0.00971	0.0104*	0.0113**			
	(0.00724)	(0.00563)	(0.00539)			
(β_3) Female Teacher	-0.0174	-0.00858	-0.00531	-0.00389	-0.00485	-0.0140
	(0.0182)	(0.00687)	(0.00762)	(0.00692)	(0.00764)	(0.00880)
$\beta_1 + \beta_3$	0.026	0.031	0.032	0.035	0.032	0.025
F-statistic (H_0: $\beta_1 + \beta_3 = 0$)	2.130	22.243***	19.361***	27.818***	18.819***	8.821***
λ_g*$\beta_1 + \beta_3$	0.005	0.011	0.014	0.016	0.014	0.006
F-statistic (H_0: λ_g*$\beta_1 + \beta_3 = 0$)	0.069	4.894**	4.830**	9.099***	4.848**	0.642
Number of Observations	134641	134641	134641	134641	134641	117817
Teacher Characteristics	No	No	No	No	No	Yes
School Fixed Effects	No	Yes	No	Yes	No	No
School*Grade Fixed Effects	No	No	Yes	No	Yes	Yes
Grade Fixed Effects by Student Gender	No	No	No	Yes	Yes	Yes

Notes: (1) Regressions include student's previous year's test score as an independent variable. (2) "Teacher Characteristics" are salary, age, experience, teacher absence, class enrollment size and indicators for caste, teacher status, education, training, native to school location, marital status, union status, and a multigrade class. (3) Standard errors (in parantheses) are clustered at the school level for OLS regressions not including fixed effects, and are clustered at the student level for OLS regressions including fixed effects. (4) Significance levels are as follows: *10%, **5%, and ***1%.

Table 1.8: Impact of Female Teachers on the Attendance of Female Students

	Dependent Variable: Student Attendance					
	(1)	(2)	(3)	(4)	(5)	(6)
(β_1) Female Student * Female Teacher	-0.00266	-0.00380	-0.00346	-0.00326	-0.00288	-0.00178
	(0.00349)	(0.00306)	(0.00307)	(0.00308)	(0.00310)	(0.00333)
(β_2) Female Student	0.00759***	0.00740***	0.00679***			
	(0.00264)	(0.00207)	(0.00207)			
(β_3) Female Teacher	0.000113	-0.00372	0.00752***	0.00375	0.00720**	0.00587*
	(0.00461)	(0.00247)	(0.00287)	(0.00250)	(0.00288)	(0.00336)
Number of Observations	148791	148791	148791	148791	148791	129890
Male Student with Male Teacher Mean	0.777	0.777	0.777	0.777	0.777	0.777
$\beta_1 + \beta_3$	-0.003	-0.008	0.004	0.000	0.004	0.004
F-statistic (H_0: $\beta_1 + \beta_3 = 0$)	0.346	10.317***	2.172	0.043	2.438	1.600
$\lambda_g*\beta_1 + \beta_3$	-0.002	-0.011	0.012	0.004	0.012	0.010
F-statistic (H_0: $\lambda_g*\beta_1 + \beta_3 = 0$)	0.087	9.142***	6.020**	1.268	5.949**	3.065*
Teacher Characteristics	No	No	No	No	No	Yes
School Fixed Effects	No	Yes	No	Yes	No	No
School*Grade Fixed Effects	No	No	Yes	No	Yes	Yes
Grade Fixed Effects by Student Gender	No	No	No	Yes	Yes	Yes

Notes: (1) Attendance is calculated as the average of the indicator of whether the student was present or not on the day of 2 to 6 visits per year. (2) "Teacher Characteristics" are salary, age, experience, teacher absence, class enrollment size and indicators for caste, teacher status, education, training, native to school location, marital status, union status, and a multigrade class. (3) Standard errors (in parentheses) are clustered at the school level for OLS regressions not including fixed effects, and are clustered at the student level for OLS regressions including fixed effects. (4) Significance levels are as follows: *10%, **5%, and ***1%.

Chapter 1 Appendix: Data and Attrition

The data used in this paper were collected over 5 school years from 2005-06 to 2009-10 from a representative sample of government-run rural primary schools in the Indian state of Andhra Pradesh (AP). Since primary school consists of grades 1 to 5, a total of nine cohorts of students are present in our data (with the oldest cohort being in grade 5 in Year 1 of the project, and the youngest cohort being in grade 1 in Year 5). Table A.1.1 shows the number of student observations by grade and year in our core estimation sample.

For students in grades 2 through 5, the estimating sample includes only those who have a test score in the current grade/year *and* have a test score from the previous grade/year (which is needed to estimate value-added). For grade 1 students, we include all those who have a test score in Grade 1 and set the normalized lagged test score to zero since there is no previous test (the estimates in Tables 5-8 are unchanged even if we exclude Grade 1). For grades 2-5, field teams conducted two rounds of testing at the end of each year (the first test covered competencies from the previous year, while the second test covered current year competencies). Since student attendance rates are around 70% having two rounds of testing helps considerably with reducing attrition from the sample.[1] However, there is only one round of testing at the end of grade 1 (since there are no previous grade competencies to be covered). Thus, the grade 2 sample in any year is smaller than the other grades.[2]

This sample is further limited to observations for which we have student gender data (97%) and for specifications that include the teacher characteristics, the sample is restricted to cases where teacher interviews were conducted (which is 88% of the sample conditional on having student test data and student and teacher gender data for the year).

[1] Student scores are first normalized with respect to each test and then averaged across the two tests, and so we have a valid normalized test score for any student who took at least one of the two tests.
[2] Grade 1 has the highest number of missing students in the end-line, but does not require a baseline; and grades 3-5 have the benefit of fewer missing data points since they are less likely to have missing test score data from the previous year (where there would have been 2 rounds of testing).

Moving across a row in Table A.1.1 (over years), we observe a reduction in student observations. This is because the share of private school enrollment is growing considerably in rural Andhra Pradesh (Pratham 2012) and fewer students are entering the public school system over time. Table A.1.2 tests whether entering cohorts over time differ in relative ability by student gender. We find no differences in Grade 1 test scores over time by student gender suggesting that the ability of girls relative to boys is not changing over time for the later entering cohorts. Thus, our estimates of the gender gap or of the impact of students sharing a teacher's gender are unlikely to be affected by the changing cohort sizes and composition over time.

We next review how attrition from the sample will affect our estimates and interpretation of the gender gap and the effect of 'gender matching' in Table A.1.2. Attrition is defined as the fraction of students in a given year who are in the potential estimation sample (which comprises of all students who have a valid test score for the previous year), but are not in the final sample because they were absent from the end of year test (i.e., have no recorded test score for the current year). Grade 1 students are not included in the attrition analysis because they do not have a test-score from the previous year, and we therefore cannot define attrition for grade 1. As mentioned earlier, all the results in Tables 5-8 are robust to excluding grade 1.

From our analysis on student attendance (Table 8), we know that girls are less likely to be absent from school on any given school day. Similarly, we find that girls have lower attrition (of 3%) in the sample used for the value-added calculations (Table A.1.3). But we also see that there is no effect of a student having the same gender as the teacher on the probability of attrition. Thus, our main estimates (presented in Tables 5-7) are unlikely to be biased due to the lower attrition of girls from our estimation sample. Furthermore, the differential attrition by student gender will only change our interpretation of the gender matching effect if the students who attrite are *differentially affected* by shared teacher gender, which is unlikely given the lack of any effect of gender matching on either student attendance (Table 8 -

columns 5 and 6) or on the probability of taking an end of year test conditional on having taken the test at the end of the previous school year (Table A.1.3 - columns 5 and 6).

Table A.1.1: Estimating Sample by Year and Grade

	Year 1	Year 2	Year 3	Year 4	Year 5
Grade 1	14011	13030	11332	11150	9194
Grade 2	10286	8021	8322	6778	6162
Grade 3	11496	10381	10372	9757	8276
Grade 4	14119	11430	10702	11010	9711
Grade 5	15415	14024	11801	11295	10473

Table A.1.2: Entering Cohorts by Gender

Dependent Variable: Normalized Test Score		
	(1)	(2)
Female Student	-0.00135	-0.0000674
	(0.0261)	(0.0210)
Year	-0.00532	-0.00540
	(0.0120)	(0.00480)
Female Student * Year	0.0102	0.00725
	(0.00832)	(0.00665)
Number of Observations	66660	66660
School Fixed Effects	No	Yes

Notes: (1) Sample limited to students in Grade 1. (2) Standard errors (in parentheses) are clustered at the school level for OLS regressions not including fixed effects, and are clustered at the student level for OLS regressions including fixed effects. (3) Significance levels are as follows: *10%, **5%, and ***1%.

Table A.1.3: Attrition by Gender Matching

	Dependent Variable: Indicator of Attrition					
	(1)	(2)	(3)	(4)	(5)	(6)
Female Student * Female Teacher	-0.0067	-0.0062	-0.0087**	-0.0029	-0.0054	-0.0049
	(0.0053)	(0.0045)	(0.0044)	(0.0045)	(0.0044)	(0.0048)
Female Student	-0.0320***	-0.0334***	-0.0308***			
	(0.0038)	(0.0030)	(0.0029)			
Female Teacher	0.0089	0.0156***	0.0047	-0.0055	0.0030	0.0050
	(0.0069)	(0.0037)	(0.0041)	(0.0037)	(0.0041)	(0.0047)
Number of Observations	131585	131585	131585	131585	131585	115592
Male Student Attrition Mean	0.227	0.227	0.227	0.227	0.227	0.227
Female Student Attrition Mean	0.193	0.193	0.193	0.193	0.193	0.193
Teacher Characteristics	No	No	No	No	No	Yes
School Fixed Effects	No	Yes	No	Yes	No	No
School*Grade Fixed Effects	No	No	Yes	No	Yes	Yes
Grade Fixed Effects by Student Gender	No	No	No	Yes	Yes	Yes

Notes: (1) Student Attrition is calculated as an indicator for being absent for the test in a given year and having taken the test the preceding year. (2) Grade 1 students are excluded because they do not have a test score prior to enrollment in school. (3) Year 1 students who drop out of the sample in the first year are excluded. (4) Standard errors (in parentheses) are clustered at the school level for OLS regressions not including fixed effects, and are clustered at the student level for OLS regressions including fixed effects. (5) Significance levels are as follows: *10%, **5%, and ***1%.

References

Bettinger, E. P., and B. T. Long (2005): "Do Faculty Serve as Role Models? The Impact of Instructor Gender on Female Students," *American Economic Review*, 95, 152-157.

Bharadwaj, P., G. D. Giorgi, D. Hansen, and C. Neilson (2012): "The Gender Gap in Mathematics: Evidence from Low and Middle Income Countries," NBER Working Paper 18464.

Carrell, S., M. Page, and J. West (2010): "Sex and Science: How Professor Gender Perpetuates the Gender Gap," *Quarterly Journal of Economics*, 125, 1101-1144.

Carrington, B., P. Tymms, and C. Merrell (2008): "Role Model, School Improvement and the Gender Gap - Do Men Bring out the Best in Boys and Women the Best in Girls?," *British Educational Research Journal*, 34.

Dee, T. (2004): "Teachers, Race and Student Achievement in a Randomized Experiment," *The Review of Economics and Statistics*, 86, 195-210.

— (2005): "A Teacher Like Me: Does Race, Ethnicity, or Gender Matter?," *American Economic Review*, 95, 158-165.

— (2007): "Teachers and the Gender Gaps in Student Achievement," *Journal of Human Resources*, 42, 528-554.

Driessen, G. (2007): "The Feminization of Primary Education: Effects of Teachers' Sex on Pupil Achievement, Attitudes and Behavior," *Review of Education*, 53, 183-203.

Ehrenberg, R. G., D. D. Goldhaber, and D. J. Brewer (1995): "Do Teachers' Race, Gender, and Ethnicity Matter? Evidence from the National Educational Longitudinal Study of 1988.," *Industrial and Labor Relations Review*, 48, 547-561.

Fairlie, R., F. Hoffman, and P. Oreopoulos (2011): "A Community College Instructor Like Me: Race and Ethnicity Interactions in the Classroom ": NBER Working Paper 17381.

Fryer, R. G., and S. D. Levitt (2010): "An Empirical Analysis of the Gender Gap in Mathematics," *American Economic Journal: Applied Economics, American Economic Association*, 2, 210-40.

GoI (2008): "Eleventh Five Year Plan 2007 - 2012," New Delhi: Oxford University Press.

Hausmann, R., L. D. Tyson, and S. Zahidi (2012): "The Global Gender Gap Report 2012," *World Economic Forum*.

Herz, B., and G. B. Sperling (2004): "What Works in Girls' Education Evidence and Policies in the Developing World. ," USA: Council on Foreign Relations.

Hoffman, F., and P. Oreopoulos (2009): "A Professor Like Me: The Influence of Instructor Gender on College Achievement," *Journal of Human Resources*, 44.

Holmlund, H., and K. Sund (2008): "Is the Gender Gap in School Performance Affected by the Sex of the Teacher?," *Labor Economics*, 15.

Lahelma, E. (2000): "Lack of Male Teachers: A Problem for Students or Teachers?," *Pedagogy, Culture and Society*, 8, 173-86.

Marsh, H. W., A. J. Martin, and J. H. S. Chend (2008): "A Multilevel Perspective on Gender in Classroom Motivation and Climate: Potential Benefits of Male Teachers for Boys?," *Journal of Educational Psychology*, 100, 78-95.

Muralidharan, K., and N. Prakash (2013): "Cycling to School: Increasing Secondary School Enrollment for Girls in India," NBER Working Paper 19305.

Muralidharan, K., and V. Sundararaman (2011): "Teacher Performance Pay: Experimental Evidence from India," *Journal of Political Economy*, 119.

NCES (2011): "National Assessment of Educational Progress (Naep) Data Explorer. ," NCES.

Neugebauer, M., M. Helbig, and A. Landmann (2011): "Unmasking the Myth of the Same-Sex Teacher Advantage," *European Sociological Review*, 27, 669-689.

Nixon, L., and M. Robinson (1999): "The Educational Attainment of Young Women: Role Model Effects of Female High School Faculty," *Demography*, 36.

OECD (2010): "Pisa 2009 Results: What Students Know and Can Do -- Student Performance in Reading, Mathematics and Science (Volume 1)".

Ouazad, A., and L. Page (2012): *Students' Perceptions of Teacher Biases: Experimental Economics in Schools*. London, UK: Center for the Economics of Education. London School of Economics.

Pratham (2010): *Annual Status of Education Report*.

— (2012): *Annual Status of Education Report*.

Rawal, S., and G. Kingdon (2010): "Akin to My Teacher: Does Caste, Religious, or Gender Distance between Student and Teacher Matter? Some Evidence from India," London: Institute of Education, University of London.

Rehman, N.-U. (2008): "Yemen Makes Progress in Girls' Education with Unicef - Supported Literacy Programmes", http://www.unicef.org/education/yemen_43424.html.

Slavin, P. (2006): "Liberia Launches Girls' Education National Policy with Support from Unicef" At a glance: Liberia, http://www.unicef.org/education/liberia_33458.html.

Stacki, S. (2002): "Women Teachers Empowered in India: Teacher Training through a Gender Lens," UNICEF.

UN (2012): "We Can End Poverty 2015 Millenium Development Goals."

UNDG (2010): "Thematic Paper on Mdg3: Promote Gender Equality and Empower Women," United Nations Development Group.

UNESCO (2010): "Institute for Statistics."

— (2012): "Enrolment and Gender Trends: Primary Education," UNESCO.

Winters, M. A., R. C. Haight, T. T. Swaim, and K. A. Pickering (2013): "The Effect of Same-Gender Teacher Assignment on Student-Achievement in the Elementary and Secondary Grades: Evidence from Panel Data," *Economics of Education Review*, 34, 69-75.

Chapter 2: Evaluating Health-Seeking Behavior, Utilization of Care, and Health Risk: Evidence from a Community Based Insurance Model in India

Abstract: Providing community based health insurance, CBHI, has been an idea gaining recent attention as a method to reduce vulnerability and increase access to health care in poorer rural populations. This study evaluates a community based health insurance contract by randomizing the insurance offer to women in microfinance Self Help Groups in rural India. I find no support for increased use of health care, and instead find limited suggestive evidence of reduction in health shocks and health care utilization. I also find suggestive evidence that the insurance offer reduces health expenditure and health related debt. This suggests scope for additional indirect benefits of increased health to insured members and assisting in the financial sustainability of CHBI contracts.

1. Introduction

In recent years, there have been increased efforts to reduce vulnerability where formal insurance markets are missing. Though health care has been documented as a significant expenditure in poorer households (Banerjee et al. 2009, Dupas and Robinson 2009), and informal risk pooling shown to be incomplete (Townsend 1994, Morduch 1999, Jalan and Ravallion 1998), health insurance in most developing countries is virtually non-existent, with private prepaid plans being a small fraction of private expenditure on health.[1] Researchers and practitioners have long discussed the lack of insurance markets in poorer rural parts of the world, and the provision of community based health insurance, CBHI, also referred to as micro health insurance (MHI), has been gaining attention as a method to fill this gap.

Similar to the microfinance revolution providing missing credit markets to the poor, CBHI are arguably able to overcome the high loading costs and asymmetric information that have prevented formal insurance markets from serving the poor. Though community based health insurances differ in design subtleties, they also share a variety of common characteristics, such as lowering the price of health care, creating a network of facilities, and having a relatively low upper limit of coverage (Jakab and Krishnan 2003, Morduch 2003, Ekman 2004). CBHI differs from larger insurance companies in that they are often organized in closer connection to the local population, and in recent years, many have attempted to reach remote poorer populations by building upon preexisting organizations, the most common of which are microfinance institutions.

A primary purpose of most of these programs is to both lower health expenditures and improve health care access of those who become insured. However, the extent to which CBHIs successfully achieve these goals is critically dependent on how insured members change their demand for health care in response to the insurance contract. When faced with lower health care costs, the direction and amount of change for health care consumption is ambiguous. To the extent that the insurer cannot observe the

[1] According to WHO Core Health Indicators:
http://apps.who.int/whosis/database/core/core_select_process.cfm

required treatment for the illness and lowers the cost of care, the quantity of health care demanded will increase. Such an increase may be seen as welfare enhancing by increasing access to health care for a population typically seen as underserved, such as the rural poor in a country like India. In theory, an increase in health care demanded could even lead to an increase in out of pocket health care expenditure by members, though this would imply a price elasticity beyond what has been commonly estimated. Members may even respond to being insured by increasing their health care consumption by such a large amount that the insurance contract becomes financially unsustainable and unravels as the cost of insurance becomes higher and higher.

But unlike other goods, the demand for health care is dependent on both cost and health status, which is a function of previously consumed health care. For example, if greater health care is initially purchased due to lower prices, this may lead to a long run increase in health status and reduced amount of needed health care. This dynamic relationship between health and health care could lead to a *decrease* in the overall health care sought even if the price of health care has decreased (Dupas 2011).

The effect of the insurance contract on long run health care consumption depends on which of these opposing effects dominates the change in health care usage.

This paper assesses the causal impact on health incidents and health care utilization when lowering the costs of health care through community based health insurance contracts. I fail to find that households increased their health care utilization and find limited suggestive evidence of a decrease in health incident, care, and expenditure. I also find limited suggestive evidence of lower levels of debt used towards health, suggesting that these households found it less necessary to rely on credit as a coping mechanism for health shocks.

This paper is one of the few studies to find that health care use does not increase after being offered and enrolled in a CBHI program (Jutting 2004, Chankova et al. 2008, Jakab and Krishnan 2004, Wagstaff and Lindelow 2008). Previous studies evaluating the impact of health insurance on changes in health seeking behaviors have primarily used case studies, and identification of a causal link has been problematic. Because most evaluations compare the insured versus uninsured, it is unclear to what extent these results stem from the effect of being insured versus preexisting differences between those who choose to enroll in the insurance and those who do not.

This paper adds to the literature by being the first of my knowledge in providing a causal link between CBHI and health incidence, health care utilization, and financial expenditure through a randomized controlled trial design. As described above, many studies have been stymied by identification and compare users and non users of an insurance program; the randomized controlled trial methodology employed in this paper attempts to overcome this barrier and provide causal estimates without relying on differences between insured status. I review a CBHI scheme in India that shares many of the common features typical of the widespread growth of CBHI in developing countries. Unlike the majority of the studies, I reject the null hypothesis of an increase in the use of health care and find limited financial protection against health expenditures.

The remainder of this paper is organized as follows: Section 2 describes the insurance contract of the CBHI, Section 3 outlines a theoretical model of the effect of insurance on health care incidence and utilization, Section 4 describes the methodology, Section 5 reviews the datasets, Section 6 discusses results, robustness analysis and alternative interpretations, and Section 7 concludes.

2. Overview of the CBHI Contract

In January 2011, Chaitanya, a non-profit microfinance institution (MFI) working on women's empowerment and microfinance in Junnar sub-district of rural Maharashtra, expanded its community based health insurance program, DAN. DAN capitalizes on Chaitanya's pre-existing microfinance Self-Help Groups[2] (SHGs) structure, and the option to purchase the insurance is limited to SHGs in which at least 80 percent of members are willing to purchase DAN (though women can decide who in their family will be included in the coverage). The cost of membership to DAN is Rs. 200 (USD 4) per person per

[2] SHG are groups of 15 – 20 women who voluntarily come together to save and access micro credit from Chaitanya.

year if the household insures 1 or 2 persons, or Rs. 150 (USD 3) per person per year if the household insures 3 or more persons. DAN does not involve a third party insurer, and health claims and operational costs are borne by the premiums collected[3].

The main provisions of the health insurance contract are discounted prices (ranging from 5 to 20 percent) negotiated at private network medical facilities, which include hospitals, medical laboratories, and pharmacies. Additionally, for in-patient treatment, the member will receive 60 percent reimbursement of their medical fees at network private hospitals, and 100 percent reimbursement at government medical facilities, up to a limit of Rs. 15,000 (USD 300)[4]. The product also includes a 24-7 medical help-line, health camps, and monthly visits by a doctor to villages to offer referrals and basic medicines. However, village visits by a doctor were intermittent and only one health camp was implemented during the timeframe of the research study.

Though CBHIs differ in design, DAN shares many of the common characteristics of CBHIs, including implementation through an existing MFI infrastructure, reducing the price of health care, establishing a network of medical facilities, and implementing an upper limit in coverage.

Chaitanya began enrollments into DAN in one area (*Block 1: semi-urban*) of Junnar sub-district in February 2011 and the remaining two areas (*Block 2: more rural*, *Block 3: tribal and rural*) in May 2011. Though enrollments were initially gradual, 61 percent of the 1,311 members[5] offered the contract were enrolled for at least some part of the study. In October 2012, the month in which a majority of the data used in this paper was collected, 47 percent of members were enrolled, and 57 percent were enrolled during the year recall period. Health claims were disbursed to 10 percent of enrolled members, with an average payout of Rs. 3,610 (USD 73) (See Table 2.1).

---

[3] A team of medical doctors, who are able to judge the technical validity of the claim, reviews the reimbursement claims. Afterwards, the claims are sent to a committee composed of local women from the Self Help Groups who decide the final outcome (e.g., should more or less be given).

[4] Specific illnesses may have lower upper limits based on predefined categories of illness types.

[5] The households included in all analysis were those that were present at the start of the research study. Households were considered to be present if at least one SHG meeting was held in the 3 to 4 months preceding the start of enrollments in the area.

3. A Simple Model on Changes in Health Care and Health Incident

When reducing the price of health care, we often assume that the overall quantity of health care consumed will increase. A common concern of insurance is that because it effectively lowers the price of health care for the insured, individuals consume more health care than if they were uninsured. In a developing country context such as India, this may be considered welfare enhancing by increasing access to healthcare. Nevertheless, because households also decide when and what type of health care to access, it may be the case that health improves and overall health care consumption decreases. The dynamics between these two factors, decreasing the costs of assessing health care and the timing and quality of the health care purchased, leads to a theoretically ambiguous response in the change of health care utilization when members become insured.

Consider a household that has the choice of seeking health care immediately or waiting to seek health care in the future depending on the course of the illness. If the household chooses to wait, with a certain probability they will recover on their own and will not have incurred any health cost. Alternatively, the illness may advance over time and require an increased amount of health care. Below I outline a simple two period model in which a household can either 1)seek care immediately when illness is still uncertain and face lower health expenditure with certainty, or 2) wait until the second period where the illness shock will become known, but conditional upon receiving a health shock the health expenditure will be higher.

I assume the household derives utility from two parameters, consumption and health. If the household chooses to purchase health care in period 1, then the household is not in risk of a health shock in period 2, and has the following expected utility (with certainty):

$$(1)\ EU = (1+\beta)u(Y - P * H_1, \bar{H} + H_1),$$

$$s.t. P * H_1 < Y$$

However, if the household chooses not to purchase health care in period 1, they risk a negative health shock in period 2 and have the following the expected utility:

$$(2)\ EU = (1-\pi_s)(1+\beta)u(Y,\overline{H}) + \pi_s[u(Y,\overline{H}) + \beta u(Y - P*H_2, \overline{H} - \theta + H_2)]\ ,$$

$$s.t. P*H_2 < Y$$

where π_s is the expected probability of the health shock in the second period, β is the discount rate for the second period, Y is the household's income endowment, $\overline{H}$ is the household's health endowment, P is the price of health care, and H is the amount of health care required to be purchased, assuming $H_1 < H_2$.

Depending on which equation yields a higher expected utility, the household will either purchase or wait to purchase health care in the first period. Depending on the curvature of the utility and the above parameters, such as the discount factor and the difference in health care required between the periods, one may choose to take the risk of increased health care in the future on the chance of not having to pay any health expenses. Assuming a homogenous society, if the expected utility of Eq (1) is higher, we would expect the population's average health care utilization to be H_1, with an average cost of $P*H_1$. If expected utility of Eq (2) is higher, then average health care utilization would be $\pi_s * H_2$, with average costs being $\pi_s * P * H_2$. It is not obvious whether the lower expected health care consumption will be optimal due to the discount rate. For example, as we imagine households to have higher and higher discount rates, they will be more likely to forego health care in the first period since the potential cost in the second period is valued less in the present period, even if poor health and health care utilization would be lower had they chosen to seek health care earlier.

A health insurance program effectively lowers the price of health care, P. While this is often done through directly lowering the monetary price of health care, it could also include other measures that lower the cost of seeking health care, such as creating a network of health facilities with increased quality or doctor visits which reduce the costs associated with travel. Using the model described above, a decrease in the price of health care could either cause people to seek care earlier (now that the foregone income is lower) or cause people to seek care later (now that the risk to income from waiting has also reduced). Depending on which effect dominates, we could see a rise or fall of health status and health care utilization.

In the above model I assumed a fixed requirement of health care. However, the amount of health care purchased is also a factor in the household's decision making process. Though the potential health burden increases in period 2 if health care if not sought earlier, the household still chooses how much health care to purchase in both periods (i.e., H_1 and H_2 are usually not fixed amounts as depicted in the model above). Thus for any of the given periods, assuming increasing returns to health care, a drop in the price will lead to an increase in the consumption of health care.

Thus, the combination of reducing the price of health care with the dynamic element of when and what type of health care to purchase leads to ambiguity when predicting how health care utilization will change under a health insurance program that lowers the cost of health care.

4. Methodology

Finding a valid comparison group for estimating the effect of CBHI has been elusive due to endogeneity of placement of programs and voluntary enrollment. To overcome this issue, the CBHI program evaluated in this paper randomized the offer of the health insurance. Half of the 43 villages in which Chaitanya was operational were randomly offered the health insurance DAN in the Junnar sub-district of Maharashtra.[6] The randomization was stratified upon three distinct areas (referred to as Block *1, 2, 3*), which become increasingly rural.

The randomization of the insurance offer assists in estimating the casual effect of the health insurance offer in the community. Using the following equation I estimate the effect of the insurance offer on illness, health care utilization, and health expenditures in the past week, month, and year.

$$(3)\ y_{igvt} = \beta_1 TreamentVillage_v + BlockFixedEffects_v + \varepsilon_{igvt}$$

where *y* is the variable of interest (health and financial variables), *TreatmentVillage* is an indicator of whether the household lives in a village that was offered DAN, *BlockFixedEffects* are indicators for whether the household lives in the area upon which the randomization was stratified; subscript *i* indicates

[6] The randomization was originally done for 61 villages. However, in the early stages of the study it was realized that 18 of these villages were not operational and so were dropped from the study. These villages were equally assigned to treatment and control villages.

the household or individual, subscript g indicates the SHG to which the household belongs, subscript *v* indicates the village, and subscript *t* indicates the month of the survey for estimations that use panel data. The randomization of DAN suggests *TreatmentVillage* is less likely to be correlated with the error term, ε_{igvt}, a necessary requirement for the consistent estimate of β_1.

4.1 Robustness Analysis

Upon estimating β_1 from Eq (3), I test for evidence of pre-exisitng differences between treatment and control villages. Pre-existing differences between village types would yield inconsistent estimates of β_1. Additionally, bounding exercises and comparison of responses across surveys are used to test for whether non-response rates raise doubt on the differential representativeness of the sample yielding a spurious estimate of β_1.

In addition to reviewing balance and attrition across treatment status, I run through a series of robustness analysis. One would expect that the treatment effect increases with greater exposure to the insurance contract, as more members become aware and enrolled in the program. Thus, I expand Eq (3) with a time trend interacted with the village's treatment status to test for increasing effects with exposure to the insurance.

$$(4)\ y_{igvt} = \beta_1 TreamentVillage_v + \beta_2 t + \beta_3 TreatmentVillage * t_{vt} + BlockFixedEffects_v + \varepsilon_{igvt}$$

where *t* the number of months since the start of the insurance offer. Since more and more members were enrolled over the course of the study, I hypothesize $\beta_3 > 0$.

The varied timing of enrollments provide additional opportunities for robustness analysis by including household fixed effects and estimating the effect of the treatment by comparing households before and after enrollment into the insurance program.

$$(5)\ y_{igvt} = \delta_1 Enrolled_{igvt} + \alpha_i + \varepsilon_{igvt}$$

where α_i are household fixed effects, and *Enrolled* is an indicator for whether any member in the household is enrolled in the given month. One concern for δ_1 to be consistently estimated is the timing of the enrollment may not be exogenous. For example, we may expect that households choose to become enrolled into the health insurance contract when they foresee health consumption in the near future – biasing δ_1 upwards. Alternatively, it could be the case that households have a health incident and suddenly value the health insurance more because of the increased saliency of health, biasing δ_1 in the opposite direction. In this context, when entire SHGs are enrolling at the same time and enrollment is dependent on a qualified staff from the MFI to enroll the members, such endogeneity of the timing of enrollment is likely to be limited. As an additional robustness check, I also estimate Eq (5) using the timing of the first enrollment in the village (in which case *Enrolled* varies with village and month). Though this is also subject to the same type of endogeneity concerns, it is arguably even less likely that the timing of the village's first enrollment is based on sorting by household into the contract.

At the start of the study, rudimentary pilot surveys were conducted in the SHGs on health parameters. Though these are not at the household level, I employ a difference-in-difference technique with the SHG as the unit of observation.

$$(6)\ \ y_{gvt} = \gamma_1 TreamentVillage_v + \gamma_2 PostBaseline_{gvt} + \gamma_3 TreatmentVillage * PostBaseline_{gvt} + BlockFixedEffects_v + \varepsilon_{gvt}$$

where *PostBaseline* is an indicator for whether the survey source was conducted after the intervention. This estimation is limited to very few indicators for which data was collected in the pilot surveys. Nonetheless, it provides a method to control for pre-existing differences between treatment and control villages. γ_1 is a measure of differences in the two types of villages not related to the insurance, and γ_3 is the consistent estimate of the offer of the insurance contract. One concern is that the pilot surveys were technically implemented after the start of the study, though before most enrollments. Thus, γ_3 may not be the entire effect of the insurance offer, but perhaps isolating effects of actual enrollment into the contract. Likewise, γ_1, can either be interpreted as pre-existing differences between village types or an effect of the insurance offer that is not dependent on actually being insured.

Upon completion of the study, the SHGs in the control villages were also offered the insurance. This provided a natural test for the Treatment Effect on the Treated using a difference-in-difference technique with the enrolled households in the control and treatment villages.

$$(7)\ y_{igvt} = \gamma_1 TreamentVillage_v + \gamma_2 Enrolled_{igvt} + \gamma_3 TreatmentVillage * Enrolled_{igvt} + BlockFixedEffects_v + \varepsilon_{igvt}$$

where *Enrolled* is now an indicator of the households who became in the treatment *and* control villages. Though there may be differences between the type of households who enrolled in the treatment villages almost two years earlier, it is likely that these households are similar. γ_1 is the average difference between unenrolled members in treatment and control villages (which may be pre-existing differences and/or externalities from the insurance program), γ_2 is a measure of the type of households who choose to enroll in the program, and γ_3 is the parameter of interest – the effect of the insurance on those who enrolled (as opposed to the insurance offer).

5. Data

The primary data source is an Endline Survey conducted in October 2012 on a randomly selected subsample of the population, approximately 18 to 21 months after the insurance was introduced. This survey was a detailed questionnaire on household demographics and illnesses in the past week and year. This survey provides a cross section of detailed information at the individual and household level.

In addition to the Endline Survey, short health surveys were conducted during monthly SHG meetings from October 2011 to July 2012. These SHG Monthly Surveys asked basic questions on household's rate of illness and health care utilization since the previous SHG meeting (i.e., a one month recall period). Unlike the Endline Survey, these surveys provide a panel on health status and health care use. However, the survey is limited to the household level (as opposed to the individual) and is dependent on whether the SHG meeting was held in the given month. Additionally, two pilot SHG Surveys were conducted in February and July 2011. These surveys reported illness in the household as a proportion of SHG Members for a one and three month recall period.

Financial activity with the MFI was sourced from Chaitanya's records. Households have two sources for loans – within their own SHGs and across SHGs through the MFI. SHG records were collected from September 2011 to August 2012. These records include all financial activities of the members, both within and across SHGs. However, the quality of data is low and highly dependent on whether the SHG chose to retain and complete the records. MFI records were collected from August 2011 to September 2012. These financial records are of much higher quality, but only include financial activities across SHGs through the MFI (i.e., larger loans).

Enrollment, claims, and the insurance's doctor village visits are accessed from Chaitanya's internal records.

For all estimations, I only assess data collected from those households who were members of the MFI at the start of the insurance offer. This prevents the estimates from being driven by the entry and exit of members, which may be an effect of the insurance offer itself.

5.1 Summary Statistics

Table 2.2 describes the demographics of the households in the research study collected in the Endline Survey. A significant number of households in this area are below the poverty line, belong to castes recognized as disadvantaged by the government, and have at least some household participation as agricultural laborers for employment. The population is approximately 50 percent female, has an average education level of 6^{th} grade, and an average age of 31.

Table 2.2 also tests for balance in these characteristics between control and treatment villages. Treatment households have slightly higher socioeconomic status, which are statistically significant for some variables. Additionally, treatment villages have slightly more females (approximately 2%), though the magnitude of this difference is very small.

6. Empirical Estimations

6.1 Incidence of illness, health care utilization, health expenditure

Table 2.3 estimates Eq (3) using data collected in the Endline Survey. When asked about the previous week's health incidents, I find that treatment villages are more likely to have been sick (statistically significant) and have higher health expenditure. When asked about the previous year, however, I find that treatment areas are 7 to 9 percentage points less likely to have had a major health shock, and have substantially less health expenditure (approximately $USD 100) and debt (approximately USD 50). This is true even if I control for initial health status by include the SHG health variables asked in the pilot SHG surveys as independent variables (even columns).[7]

Table 2.4 echoes the annual recall results in Table 2.3 using the SHG Monthly Surveys panel data. Treatment areas report being 6 percentage points less ill, less likely to seek health care, less likely to have extreme illnesses (as indicated by lower levels of being admitted or prolonged bed rest), and lower health expenditure. However, I do not find any statistically significant difference in prolonged bed rest when recalling the previous month, raising some doubt on the lower levels of illnesses of households in villages offered insurance. Table 2.3 and 3.4 also estimates bounds for the estimates (discussed in the following section). Table A.2.1 and A.2.2 describe summary statistics of the variables used in Table 2.3 and 2.4.

One possible explanation for the differences in the week versus the year and month recall may be the type of illnesses that are being recalled. It is likely that smaller illnesses and minor health care consumption are more likely to be recalled in a week period, but not in longer timeframes (Das et al. 2011). The data collected on the previous year is limited to larger illnesses, which I proxy by asking households whether they suffered a health incident in which a household member was on prolonged bed rest for 5 or more days, admitted to a health care facility, or incurred health expenses which totaled over Rs. 1,000 (USD 20). Though the surveys recalling the previous month do not limit themselves to larger health illnesses, it may be the case that a month's recall period is too long of a timeframe for households

[7] In the remaining tables, I estimate all regressions with the health variables as controls from the SHG Surveys and note if there are differences in estimated parameters when excluding the variables.

to recall smaller and more regular illnesses. This is consistent with the possibility that households in treatment villages are seeking more health care, though for smaller health incidences. As a result, these households have a decrease in health incidents and expenditure for larger illnesses (captured in the month and year recall).

In order to better understand why the insurance led to lower health care utilization, I estimate health care consumption differences in only those members who reported an illness in their household. Because incidence of illness is correlated with the insurance offer, these estimations provide suggestive evidence but should not be viewed as consistent estimates of the insurance product. Table 2.4 suggests that even when ill, treatment villages are less likely to report going to visit a health facility and have lower health expenditure. This is consistent with illnesses in treatment village households being less severe. Table 2.5 observes individual level responses conditional upon being ill and exhibits a similar pattern – even conditional upon being ill, individuals in treatment villages are ill for fewer days and miss less days of school/work. For larger illnesses, they are admitted for shorter durations, have lower health expenditures, and less likely to have borrowed for health. These results continue to be more robust for the year recall as opposed to the week recall. Relative to the year recall, the week recall suggests slightly higher expenditure and visits to health facilities for treatment households, though these results are not statistically different from zero.

6.2 Financial Vulnerability

Being insured against health shocks and lowing the cost of health care would theoretically lead to a decrease in financial vulnerability. One would expect to see a decrease in debt and selling of assets used to finance health. However, this may result in capital and credit being more available for investments leading to a potentially ambiguous result in the overall debt burden. Table 2.6 finds no statistically significant effect on the selling or purchasing of assets or levels of debt (though point estimates on credit use are in the expected direction). Table 2.7 reviews the financial records of the SHG and MFI and finds a lack of statistically significant results. Though treatment villages have a lower level of outstanding debt in the SHGs (Panel A), Column (1) and Column (2) suggest that treatment areas are both saving and borrowing at lower amounts. Furthermore, the lower debt burden is limited only to the smaller loans in SHGs, and I find no impact on the larger loans offered by the MFI. Due to the lack of consistent and statistically significant results in the financial patterns of the insured village, the remainder of the paper focuses only on health outcomes.[8]

6.3 Robustness Analysis for Health Incidences and Health Care

Pre-existing Differences

A potential concern is that pre-existing differences, instead of the offer of the insurance, account for the difference between households in treatment and control villages. Though Table 2.2 suggested that differences along household demographics were minimal, Table 2.8 raises the concern of the two areas differing in health status from the start of the intervention. The health data for Table 2.8 is from the pilot SHG Surveys conducted in February and July of 2011. These surveys recorded the proportion of the SHG that had experienced household illness in the past month, and prolonged bed rest or high health expenditure in the past three months. Unfortunately, this data has relatively low response-rate, a slight imbalance in the response rate by treatment status, and identification only at the SHG level (not at the household level). Also, these surveys were technically conducted after the start of the intervention – though insurance coverage only began in February 2011, and enrollments had only minimally begun in Block 2 and 3 by July 2011. Nonetheless, the results of these initial surveys are disconcerting as they report that SHGs in treatment villages had a lower proportion of illnesses in their households at magnitudes similar to the ITT estimates found in Table 2.3 and 3.4. This holds true in every Block which the randomization was stratified. However, this same pattern does not hold for other the indicators of health, prolonged bed rest and high expenditure, where treatment and control villages are relatively balanced.

[8] The estimations outlined in the robustness analysis for financial outcomes also find varying and statistically insignificant results. These tables are available upon request.

Using the Pilot SHG Monthly Surveys in February and July 2012, I am able to construct a panel for estimating a difference-in-difference of the treatment effect with the SHG as the unit of observation (Eq. 6).[9] As expected, a difference-in-difference estimation suggests no difference (or even an increase) of illness in the treatment villages after the insurance is offered (see Table 2.9). The coefficient on prolonged bed rest continues to suggest a decrease of five percentage points. Neither coefficient is statistically significant, most likely due to a reduction in statistical power from the decrease in the number of observations. These results do give rise to the concern that the Intent to Treat Effect of the insurance found in Table 2.3 and 3.4 are not due to the insurance itself, but perhaps to pre-existing differences or treatment effects on survey response. Though I attempt to control for these initial health variables by including them as independent variables in estimations[10], the low correlation between the measures (especially once treatment status is included), does not result in large differences in the estimated coefficient of the treatment effect.

Survey Non-Response

Survey non-response is also a primary concern in the robustness of the effect of the insurance offer estimated in Table 2.3 and 3.4. The response rate of the October 2012 Endline Survey was 80% in treatment villages and 79% in control villages. As Table 2.10 confirms, there was no statistically significant difference in the response rate by treatment status. Table 2.11 provides the reasons for non-response rates – very few households refused consent and the majority of households not surveyed were due to relocation, which seems unlikely to be a result of the insurance offer. One primary reason for the low-response stemmed from three villages which were experiencing difficulties with the MFI due to high

[9] In the SHG Monthly Surveys, members are directly asked about illness and prolonged bed rest in their household in the past month. The prolonged bed rest in both surveys use a different number of days as the minimum (3 versus 5), and in the latter survey is reconstructed by the past 3 months' surveys to match the recall period. Averages of the SHG are then taken to match the unit of observation in the Pilot SHG Surveys.

[10] In order to not lose observations by including health controls from the February 2011 Pilot Survey, village averages are assigned to those SHGs where the Pilot Survey was not conducted. If no SHG in the village was surveyed, then the respective treatment status average was assigned.

defaults. This made it difficult for surveyors to contact households in these villages, and thus accounts for over 50% of the unknown non-response. In a small number of households, a shorter survey was implemented which asked basic health and expenditure questions, increasing the response rate to 80% in both treatment and control villages.

Table 2.12 (Columns 1 and 2) and Table 2.13 provide suggestive evidence of the health status of non-respondents in the Endline Survey. Table 2.12 uses the February 2011 Pilot SHG Survey to test whether non-respondents in treatment villages differed from those in control villages. Here I find suggestive evidence that non-respondents in treatment villages were more likely to be sick than those in control villages. However, Table 2.13 compares responses in the SHG Monthly Survey by Endline Survey response status and suggests these differences are relatively small. These results suggests that non-respondents in treatment areas were sicker relative to respondents (compared to households in control households), biasing estimated ITT effects downward. However, the magnitude of these differences is small enough making it unlikely to be accounting for the entire estimated ITT in Table 2.3 and 3.4.

Table 2.3 estimates bounds by quintile. Though the attrition rate is too large to confidently provide a range for most variables, the incidence of large health shocks is robust to a 20 percentage point difference in treatment and control household non-respondents, the high end of what Table 2.12 and 3.13 suggest.

Table 2.12 (Column 3 and 4) similarly suggest that non-respondents from treatment villages in the SHG Monthly Surveys are also more likely to be ill in the previous month, though not more likely to have experienced a large health shock in the previous three months. Table 2.14, Panel A, is analogous to Table 2.13, where I compare responses on the 2012 Endline Survey based on respondent status in the SHG Monthly Survey. If anything, non-respondents in treatment areas are even less ill than those in the control areas (relative to respondents). Thus it is unclear whether the ITT coefficient estimated in Table 2.4 is biased upwards or downwards. Panel B tests for balance between non-respondents and respondents using the demographic data collected in the Endline Survey – though results are mostly statistically insignificant, it is suggestive of non-respondents being slightly better off socioeconomically.

Time Variation

Due to the concerns of pre-existing differences and non-response rates in the surveys, I next turn to time variation as a method to test for the robustness of the initial results found in Table 2.3 and 3.4. Table 2.15 estimates whether the length of the exposure to the insurance offer results in increased treatment effects. I find no statistically significant evidence that the effects of the insurance offer is increasing over time. This is surprising since enrollments are varied over the first year of the program, with many enrollments not occurring until at least 6 months after the start of the intervention. This result echoes the concern raised by the Pilot SHG Surveys, suggesting differences existed even before most members are enrolled in the insurance program.

Figure 1 and 2 plot point estimates, θ, from a regression that includes a variable for each month enrolled in the program, using only the sample of households in treatment villages:

$$(8)\ y_{igvt} = \theta_1 Enrolled1Mon_{igvt} + \theta_2 Enrolled2Mon_{igvt} + \ldots + \theta_{18} Enrolled18Mon_{igvt} + MonFixedEffects_t + AreaFixedEffects_v + \varepsilon_{igvt}$$

The figures show steadily decreasing level of illness and utilization of health care facilities as the household is enrolled in the program for longer duration. However, an odd feature of the pattern is that these estimates are relative to before the household was enrolled – it is surprising that the month after enrollment, the household has lower rates of illness and health care consumption. If we believed that the health insurance encouraged seeking care earlier or better care, we should expect that the insurance first causes an increase in health care and then a subsequent drop – however, this initial rise is absent from the data.

Table 2.16 uses the timing of the village and household enrollment, along with member fixed effects, to compare households before and after enrollment. Panel A compares each household before and after the village had its first enrollment (a proxy for the initial exposure to the insurance offer). Village timing of enrollment is more likely determined by constraints of the MFI's human resources than household health status at the time. Panel B compares the household to itself before and after becoming enrolled in the program. Both panels suggest that even when comparing within households, households are less likely to be ill, consume health care, and have high health expenditures after being enrolled in the program. This result is supportive of our initial findings in Table 2.3 and 3.4 and do not rely on methods balance in characteristics and non-response between treatment and control households. Nonetheless, the methodology is open to concerns of endogeneity if the timing of enrollment is a function of health status and decreasing time trends of poor health. For example, an alternative interpretation to Table 2.16 is households and villages become insured when health care needs are high and salient, and thus the estimated treatment effect is spurious.

Enrolled Members in Control Villages: Difference-in-Difference

Upon completion of the research timeframe, the insurance was offered to control villages. This provides potential identification of households who would have enrolled in the program had the insurance been offered to all villages initially. Table 2.17 estimates a difference-in-difference of household demographics and health status among the enrolled in the control versus treatment villages. Though differences are not statistically significant, enrolled households in treatment villages appear to be slightly better off socioeconomically and have lower levels of health incidence. Table 2.18 and 3.19 estimate the difference-in-difference outlined in Eq. 7 and shows no statistically significant Treat Effect of the Treated from the insurance on the enrolled. This may be due to low statistical power, but even the point estimates are of low magnitude and opposite signs of the results found in Table 2.3 and 3.4. I continue to see the coefficient on treatment village remain large and statistically significant, suggesting either very large externalities to households that chose not to enroll or that differences do not stem from the insurance itself. Of course, it may be the case that those enrolled in the control villages are not a comparable group to the enrolled in the treatment village, though it seems unlikely that the TET would not be more similar to estimates found in Table 2.16.

Comparative Methodologies: Comparing Enrolled Members

As a comparison to the most common methodologies in previous papers, I analyze the data as if the insurance offer had not been randomized. I compare those who enrolled in the insurance contract with those who did not, limiting the sample to the villages that were offered the insurance contract. Table 2.20 presents drastically different estimates, suggesting that the type of people who choose to enroll in the insurance differ from those who do not enroll in the very variables in which we are interested in measuring an effect. While Panel A, variables at the household level, suggest that the insurance has no effect, Panel B and C illustrate that at the individual level the insurance contract is correlated with poorer health and higher health care consumption. This highlights the differences between those who choose to enroll and those who do not, and the necessity of a valid comparison group in evaluating CBHI programs when enrollment into the program is optional.

Potential Mechanisms: Indirect Effects

Table 2.21, Panel A, estimates household behavior when ill in villages offered the insurance. One would expect that treatment areas would seek care faster, forego treatment less, and be more likely to recover from illnesses. However, I find no statistically significant effect on any of these variables and even the sign of the point estimate is often opposite of the expected direction. Panel B estimates whether these behaviors are affected by the number of times the village was visited by the insurance's doctor[11]. Again, I find no evidence that there was a positive effect on such behaviors from the insurance or the insurance's doctor's visits.

Whether these estimates provide lower or upper bounds depends on our beliefs of the characteristics of the individuals who are not ill in the villages offered the insurance. For example, if we believe that those who did not suffer a health shock were the type of people who would have had milder shocks requiring less health care, then the estimates provided in Table 2.21 are lower bounds. The

[11] Table A.2.5 summarizes the frequency of village visits by the insurance doctor for out-patient care (OPD Doctor).

assumptions required for the estimates of foregoing treatment and days waited to seek medical care to be lower bounds is that those who did not become ill were the type of people who would be less likely to wait or forego medical care. The opposite assumptions would imply the estimates are upper bounds of the effect of the insurance offer conditional upon illness.

Table 2.22 and 3.23 estimate whether village visits by the insurance's doctor had any discernible effect on health incidents, health consumption, and health expenditure. Though I do not argue that the doctor's visits to villages are exogenous, the correlation can help to understand whether the results found in Table 2.3 and 3.4 are spurious or due to such mechanisms as the visits from the doctor. Though Table 2.22 finds no consistent effect of the doctor on health, and Table 2.23 suggests that the doctor did encourage consumption of health care. I also find a correlation between illness incidence and doctor visits, which may be a function of endogeneity of the doctor visits or that the diagnosis and recall of the illness was altered by the doctor visits.

7. Conclusion

The success and effectiveness of insurance contracts are critically dependent on the health care utilization of its member base. An increase in health care consumption is often an indirect goal of community based insurance providers, though it also raises concerns of the financial stability of the insurance contract. Contrary to the majority of studies evaluating CBHI, I find this insurance contract does not increase health care consumption. Instead, I find limited suggestive evidence of the insurance possibly reducing the consumption and expenditure of health care. Not only is this indicative of improving the health status of the insured, but additionally helps households with their health expenditure beyond the discounted health care received directly through the insurance. Furthermore, a decrease in health care utilization increases the probability of financial sustainability of the program. However, the insurance lowering the need and demand for health care is only one possible interpretation of the findings, as discussed in Section 6, and thus should be considered an upper limit.

In general, the potential of CBHI to increase health status and lower the amount of health care warrants further research. Numerous factors in the design of the CBHI may be responsible for decreasing the barriers of access to health care and potentially reduced health shocks: direct price reductions, network facilities with quality checks, and local doctors being monitored. Further research is required to decipher which of these factors led to a decrease in health shocks and health care utilization and how these can be promoted and integrated into the designs of CBHI programs.

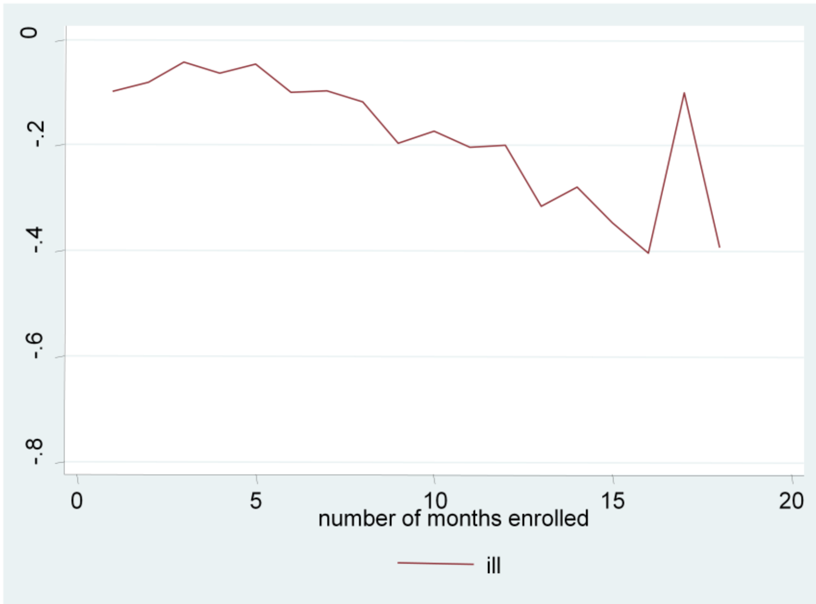

Figure 2.1: Illness by Months Enrolled

Figure 2.2: Health Care Consumption by Months Enrolled

Table 2.1: Enrollment and Claims Summary Statistics

Panel A: Enrollment	
Never Enrolled	512
Enrolled	799
Percent Ever Enrolled	60.9%
Percent Enrolled at Endline Survey	47.0%
Percent Enrolled in Year Recall of Endline Survey	57.2%
Panel B: Claims (Conditional upon Enrollment)	
Percent Filed Claim	9.8%
Average Claim Disbursement	Rs. 253
Conditional upon Receiving Claim	
Average Claim Disbursement	Rs. 3610

Table 2.2: Summary Statistics and Balance on Demographic Characteristics

	Control		Treatment		Treatment - Control	Treatment - Control
	No. SHG Women	Mean	No. SHG Women	Mean		
Household Characteristics:						
Household Size	938	5.801	748	5.698	-0.0643	-0.103
Forward Caste	887	0.297	707	0.633	0.226**	0.336***
Hindu	923	0.963	739	0.922	-0.0391	-0.0406
Above Poverty Line	905	0.012	710	0.029	0.0160**	0.0177**
Subsidized Ration Card	905	0.526	710	0.612	-0.0226	0.0859
Below Poverty Line	905	0.435	710	0.327	-0.00375	-0.108
Stamped Below Poverty Line	905	0.027	710	0.031	0.0104	0.00391
Agricultural Laborer	938	0.701	749	0.635	-0.0189	-0.0659
House Type	931	2.284	744	2.443	0.0683	0.160*
Individual Characteristics:						
Female	5333	0.496	4249	0.513	0.0179*	0.0171*
Age	5357	31.319	4276	31.347	-0.478	0.0288
Education	5176	6.131	4265	6.257	-0.0529	0.126

Observations are limited to households selected and surveyed in the Endline Health Survey and are weighted to be representative of the target population. Source: 2012 October Endline Survey. Statistical significance levels are as follows: *10%, **5%, ***1%.

Table 2.3: Endline Survey - Health Incidence and Expense

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
	Panel A: In the past week, the household experienced									
	no. of sick hh		health expenses							
Treatment	0.167**	0.173**	388.5	327.1						
Village	(0.0757)	(0.0719)	(439.0)	(326.6)						
Bounds by quintile:										
0%	-2.878***	-3.011***	-39198.3***	-40789.6***						
20%	-1.738***	-1.872***	-23689.5***	-25289.4***						
40%	-0.625	-0.736*	-8180.7	-8763.5						
60%	0.488	0.483	6948.3	7372.8						
80%	1.602***	1.672***	22457.2***	23913.6***						
100%	2.676***	2.797***	37139.8***	39471.8***						
Obs	1643	1643	1591	1591						
	Panel B: In the past year, the household experienced									
	Rs. 1000 illness		admitted		prolonged bedrest		health expenses		health loan	
Treatment	-0.0720*	-0.0910***	-0.0783**	-0.0930***	-0.0527	-0.0729*	-5096**	-6709***	-2368***	-2195***
Village	(0.0364)	(0.0335)	(0.0362)	(0.0321)	(0.0400)	(0.0365)	(2273.1)	(2242.5)	(619.2)	(586.5)
Bounds by quintile:										
0%	-0.285***	-0.297***	-0.293***	-0.303***	-0.281***	-0.293***	-128530***	-133596***	-70546***	-72877***
20%	-0.199***	-0.212***	-0.204***	-0.214***	-0.192***	-0.204***	-79107***	-83755***	-43043***	-45621***
40%	-0.114***	-0.123***	-0.118***	-0.125***	-0.106***	-0.115***	-31656	-34256	-16730	-17820
60%	-0.0288	-0.0308	-0.0338	-0.0338	-0.0211	-0.0229	15971	17419	9576	10458
80%	0.0564	0.0612	0.0524	0.0602	0.0652*	0.0711*	63422***	67877***	35889***	38949***
100%	0.139***	0.149***	0.137***	0.150***	0.149***	0.161***	110259***	117761***	61437***	65738***
Obs	1690	1690	1685	1685	1684	1684	1677	1677	1670	1670
Controls	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes

Observations are limited to households selected and surveyed in the Endline Health Survey and are weighted to be representative of the target population. Regressions includes block fixed effects. Standard errors are clustered at the village level. Source: 2012 October Endline Survey. Statistical significance levels are as follows: *10%, **5%, ***1%.

Table 2.4: SHG Monthly Surveys - Illness, Care, and Expenditure

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	In the past month,							
	illness	visiting facility/ dan doctor	visiting health facility	doctor (not village visit)	doctor (including village visit)	admitted	prolonged bedrest	health expenses
Treatment Village	-0.0578**	-0.0368	-0.0749***	-0.0721***	-0.0319	-0.0125**	-0.00630	-737.8**
	(0.0228)	(0.0226)	(0.0175)	(0.0173)	(0.0237)	(0.00537)	(0.00739)	(286.6)
Upper Bound	0.00719	0.0330	-0.0123	-0.0125	0.0362	-0.00736	-0.000133	-625.3*
Lower Bound	-0.204***	-0.188***	-0.222***	-0.218***	-0.185***	-0.0239***	-0.0222***	-762.6**
N (SHG Member Month)	12765	12797	12766	12777	12796	12755	12742	10852
	Conditional upon being ill							
Treatment Village		-0.0562*	-0.0668**	-0.0690**	-0.0531*	-0.0190	0.00209	-1339.9**
		(0.0285)	(0.0324)	(0.0320)	(0.0296)	(0.0148)	(0.0203)	(607.6)
Upper Bound		-0.0398*	-0.0495*	-0.0407	-0.0253	-0.00480	0.0189	-1361.9**
Lower Bound		-0.152**	-0.226**	-0.357***	-0.242***	-0.0609***	-0.0556***	-1389.4**
N (SHG Member Month)		4429	4429	4429	4429	4424	4415	4281

All regressions include health status variables from the February 2011 SHG Pilot Survey as independent control variables. Households are weighted to be equal per month. Regressions include block area fixed effects. Standard error are clustered at the village level. Upper and Lower bounds refer to Lee (2009) bounds, stratified by block areas. Source: SHG Monthly Surveys. Statistical significance levels are as follows: *10%, **5%, ***1%.

Table 2.5: Individual Health Behavior Conditional On Illness

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	Panel A: Week Recall						
	Conditional upon being sick in the past week,						
	was sick	poor health today	days ill	days missed	health facility	total health expenses	
Treatment Village	0.0277**	-0.0611	-0.388*	-0.0720	0.0119	183.7	
	(0.0136)	(0.0491)	(0.203)	(0.197)	(0.0108)	(280.0)	
Obs (Individuals)	9651	1212	1213	1174	1221	1228	
	Panel B: Year Recall						
	Conditional upon having experienced a health shock in the previous year,						
	experienced a health shock	days missed	doctor	total days admitted	health expenses	loan for health	borrowed
Treatment Village	0.00162	-5.757***	0.00850	-1.116*	-7186.6	-0.0243	-1970.1*
	(0.0100)	(1.548)	(0.00518)	(0.605)	(4292.1)	(0.0337)	(1009.2)
Obs (Individuals)	9664	878	901	904	889	795	904

Observations are limited to households selected and surveyed in the Endline Health Survey and are weighted to be representative of the target population. All regressions include health status variables from the February 2011 SHG Pilot Survey as independent control variables. A health shock is defined as having had a health incident with high expenditure, being admitted, or being on prolonged bed rest. Regressions include block fixed effects. Standard errors are clustered at the village level. Source: 2012 October Endline Survey. Statistical significance levels are as follows: *10%, **5%, ***1%.

Table 2.6: Assets and Loans

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	In the past year,			*Of current outstanding loans,*			
	mortgaged assets	sold assets	purchased assets	number of loans	has health loan	has home/ business improvement loan	total loan outstanding
Treatment Village	-0.00579	-0.0244	-0.0270	-0.0802	-0.0218	0.0261	-6027.3
	(0.0245)	(0.0257)	(0.0244)	(0.0746)	(0.0187)	(0.0394)	(8542.4)
Obs (SHG Members)	1657	1650	1654	1529	1516	1516	1520

Observations are limited to households selected and surveyed in the Endline Health Survey and are weighted to be representative of the target population. All regressions include health status variables from the February 2011 SHG Pilot Survey as independent control variables. Regressions include block area fixed effects. Standard errors are clustered at the village level. Source: 2012 October Endline Survey. Statistical significance levels are as follows: *10%, **5%, ***1%.

Table 2.7: Financial Vulnerability

	(1)	(2)	(3)	(4)
Panel A: Internal SHG Financial Records -- Internal and External Finances				
	input into fund	output from fund	loan disburse-ment	outstanding amount
Treatment Village	-124.3	-4880.1	28.01	-3790.7***
	(135.3)	(4017.2)	(84.76)	(1103.3)
N (SHG Members Month)	20988	20988	20970	4352
Panel B: MFI Financial Records -- External Finances				
	outstanding loan amount	overdue principle	overdue interest	amount paid
Treatment Village	363.6	136.3	2.815**	27.42
	(770.8)	(137.1)	(1.137)	(66.28)
N (SHG Members Month)	36792	36792	36792	34164

All regressions include health status variables from the February 2011 SHG Pilot Survey as independent control variables. Regressions include block area fixed effects. Standard errors are clustered at the village level. Source: Panel A: SHG Financial Records, Panel B: MFI Financial Records. Statistical significance levels are as follows: *10%, **5%, ***1%.

Table 2.8: Psuedo-Baseline Health Status

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	February, 2011				July, 2011			
	Ill	Prolonged Bedrest	High Expenditure	Collected Data	Ill	Prolonged Bedrest	High Expenditure	Collected Data
	Panel A							
Treatment Village	-0.0836	-0.00239	-0.00127	-0.0349	-0.129	0.0128	0.0455	0.0337
	(0.0665)	(0.0368)	(0.0385)	(0.102)	(0.0639)	(0.0481)	(0.0420)	(0.0788)
Obs (SHGs)	99	98	98	160	64	64	61	85
	Panel B: Block 1 (Most Urban)							
Treatment Village	0.374	0.151	0.159	0.867				
Obs (SHGs)	37	37	37	45				
Control Village	0.446	0.152	0.185	0.867				
Obs (SHGs)	26	26	26	30				
	Panel C: Block 2 (Semi-Urban)							
Treatment Village	0.430	0.088	0.227	0.316	0.407	0.108	0.110	0.789
Obs (SHGs)	6	6	6	19	14	14	14	19
Control Village	0.598	0.218	0.224	0.350	0.504	0.126	0.066	0.800
Obs (SHGs)	7	6	6	20	16	16	13	20
	Panel D: Block 3 (Rural/Tribal)							
Treatment Village	0.438	0.154	0.167	0.438	0.273	0.108	0.108	0.875
Obs (SHGs)	7	7	7	16	14	14	14	16
Control Village	0.505	0.083	0.094	0.533	0.432	0.067	0.061	0.800
Obs (SHGs)	16	16	16	30	20	20	20	30

Observations include all SHGs for which at least one member was included in the baseline. Illness is a recall period of one month. Prolonged bedrest and high expenditure has a recall period of 3 months. July 2011 excludes Block 1 where a significant amount of enrollments had already occurred to date. Panel A: Regressions include block area fixed effects. Standard errors are clustered at the village level.

Table 2.9: Difference-in-Difference using Pseudo-Baseline

	(1)	(2)	(3)	(4)
	Proportion of households in the SHG who experienced:			
Dependent Variable:	illness (1 month recall)	prolonged bedrest (3 month recall)	illness (1 month recall)	prolonged bedrest (3 month recall)
Treatment Village * Post Intervention	0.00664	-0.0533	0.156	-0.0540
	(0.0660)	(0.0454)	(0.0942)	(0.0430)
Treatment Village	-0.117*	-0.00745	-0.107*	0.0204
	(0.0633)	(0.0328)	(0.0592)	(0.0401)
Post Intervention	-0.0891*	-0.0168	-0.133**	-0.0269
	(0.0461)	(0.0280)	(0.0565)	(0.0305)
N(SHG Month)	197	210	119	149
Month for Pre-Intervention	Feb-11	Feb-11	Jul-11	Jul-11
Includes Block 1	Yes	Yes	No	No

July 2011 excludes Block 1 where a significant amount of enrollments had already occurred to date. Regressions include block area fixed effects. Standard errors are clustered at the village level. Sample limited to SHGs for which Pilot SHG Surveys were conducted. Post Intervention refers to the SHG Monthly Survey conducted in the same month as the Pilot Survey the following year. Statistical significance levels are as follows: *10%, **5%, ***1%.

Table 2.10: Endline Survey Non-response

	(1)
Dependent Variable : Surveyed	
Treatment Village	0.0302
	(0.0838)
Obs (SHG Member)	2068

Observations are limited to households selected and surveyed in the Endline Health Survey and are weighted to be representative of the target population. Regressions include block area fixed effects. Standard errors clustered at the village level. Statistical significance levels are as follows: *10%, **5%, ***1%.

Table 2.11: Reasons for 2012 October Endline Attrition (by fraction)

	(1)	(2)	(3)
		Not Surveyed	
	Control	Treatment	Treatment - Control
Did not give consent	0.019	0.008	-0.00949
Could not find house	0.007	0.030	0.0234*
Passed Away	0.041	0.032	-0.00785
Pilgrimage/Trip	0.010	0.019	0.00986
Out of town for health	0.010	0.000	-0.0108**
Door locked/not home	0.043	0.016	-0.0281
Moved	0.372	0.502	0.154
Unknown	0.497	0.394	-0.131
No. SHG Women	219	169	388

Observations are limited to households selected and surveyed in the Endline Health Survey and are weighted to be representative of the target population. Treatment - Control are based on regressions which include block area fixed effects. Standard errors clustered at the village level. Statistical significance levels are as follows: *10%, **5%, ***1%.

Table 2.12: Attrition in 2012 Endline Survey by February 2011 SHG Health Status

	(1)	(2)	(3)	(4)
Dependent Variable: HH in	Endline Survey		SHG Monthly Survey	
Treatment Village	0.109	0.0376	0.331**	0.108
	(0.103)	(0.0627)	(0.139)	(0.115)
Ill	0.0606		0.388**	
	(0.144)		(0.156)	
Ill * Treatment	-0.230		-0.491**	
	(0.190)		(0.233)	
Health Shock		-0.0303		0.273
		(0.297)		(0.323)
Health Shock * Treatment		-0.142		0.0348
		(0.325)		(0.432)
Obs (SHG Member, SHG Member Month)	2068	2068	26250	26250

Observations are limited to households selected and surveyed in the Endline Health Survey and are weighted to be representative of the target population. Recall period for being ill is one month. Recall period for health shock (proxied by an expenditure over Rs. 1000 or 3 days of consecutive bed rest) is 3 months. Independent variables are proportion of SHG Members who had such an incident in their household. Regressions includes block fixed effects. Standard errors are clustered at the village level. Statistical significance levels are as follows: *10%, **5%, ***1%.

Table 2.13: SHG Monthly Survey Health Responses by Endline Response Status

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Dependent Variable:	surveyed in SHG Monthly Surveys	illness	visiting health facility/ dan doctor	visiting health facility	doctor (not village visit)	doctor (including village visit)	admitted	prolonged bed rest	health expenses
Endline Respondents:									
Treatment Village	0.0535	-0.0757***	-0.0715***	-0.0962***	-0.0852***	-0.0586**	-0.0207***	-0.0162**	-869.6***
	(0.0423)	(0.0245)	(0.0217)	(0.0209)	(0.0209)	(0.0217)	(0.00420)	(0.00670)	(295.4)
Obs (SHG Member Month)	16800	8611	8636	8612	8621	8636	8606	8597	7243
Endline Non-Respondents									
Treatment Village	0.368***	-0.0261	-0.00496	-0.0261	-0.0405	-0.0154	-0.0194*	-0.0273**	-910.7**
	(0.124)	(0.0410)	(0.0398)	(0.0387)	(0.0354)	(0.0367)	(0.0102)	(0.0135)	(419.4)
Obs (SHG Member Month)	3880	1178	1186	1180	1182	1185	1179	1177	1080

Observations are limited to households selected and surveyed in the Endline Health Survey and are weighted to be representative of the target population. Regressions includes block fixed effects. Standard errors are clustered at the village level. Statistical significance levels are as follows: *10%, **5%, ***1%.

Table 2.14: SHG Survey Response Rates and Demographics

	(1)	(3)	(4)	(5)	(7)	(8)	(9)	(10)	(11)
	Panel A: Household Health Characteristics								
	In the past week,		In the past year,						
	no of sick hh members	health expenses	Rs. 1000 expenditure	admitted	prolonged bed rest	no. of health shocks	health expenses	health loan	
SHG HH per Month Surveyed									
Treatment Village	0.0224	31.59	-0.0827**	-0.0781**	-0.0686	-0.0790	-4075.9	-2290.8***	
	(0.0853)	(371.5)	(0.0375)	(0.0352)	(0.0428)	(0.0524)	(2478.7)	(724.0)	
N (SHG Members Month)	8629	8357	8843	8816	8809	7948	8808	8777	
SHG HH Per Month Not Surveyed									
Treatment Village	0.217**	141.0	-0.114***	-0.128***	-0.103***	-0.137**	-5586.8***	-1933.1***	
	(0.0852)	(171.3)	(0.0372)	(0.0374)	(0.0376)	(0.0570)	(1791.3)	(464.6)	
N (SHG Members Month)	6191	6123	6407	6384	6381	5822	6352	6323	
Test of Equality (P-value)	**0.0084**	**0.7589**	**0.454**	**0.1763**	**0.4973**	**0.4026**	**0.5366**	**0.458**	
	Panel B: Household Demographics								
	Forward Caste	Poverty Status: Above	Poverty Status Subsidized	Poverty Status Below	Agricultura l Laborer	House Type	Mean Female	Mean Age	Mean Education
SHG HH per Month Surveyed									
Treatment Village	0.156	0.0114	-0.0265	0.0107	-0.0219	0.0258	0.00949	-0.137	-0.404
	(0.101)	(0.00921)	(0.0556)	(0.0507)	(0.0649)	(0.0646)	(0.0104)	(0.889)	(0.322)
N (SHG Members Month)	8470	8469	8469	8469	8821	8852	9823	9847	9833
SHG HH Per Month Not Surveyed									
Treatment Village	0.262***	0.0224***	-0.0515	0.0237	-0.0346	0.120*	0.00850	-0.428	-0.262
	(0.0887)	(0.00692)	(0.0465)	(0.0450)	(0.0686)	(0.0671)	(0.0136)	(1.027)	(0.270)
N (SHG Members Month)	6080	6041	6041	6041	6389	6408	6887	6893	(0.270)
Test of Equality (P-value)	**0.2668**	**0.2487**	**0.6529**	**0.8286**	**0.7346**	**0.0297**	**0.9893**	**0.6098**	(0.270)

Observations are limited to households selected and surveyed in the Endline Health Survey and are weighted to be representative of the target population. Regressions includes block fixed effects. Standard errors are clustered at the village level. Tests of equality are not weighted. Source: 2012 October Endline Survey. Statistical significance levels are as follows: *10%, **5%, ***1%.

Table 2.15: Length of Exposure to Insurance Offer

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	illness	visiting facility/ dan doctor	visiting health facility	doctor (not village visit)	doctor (including village visit)	admitted	bed rest	health expenses
Treatment Village	3.127	0.646	1.164	0.531	0.259	-0.548	-1.814	-30210.1
	(8.432)	(8.770)	(8.578)	(8.353)	(8.579)	(2.036)	(1.783)	(76944.6)
Month Number	-0.0336***	-0.0325***	-0.0323***	-0.0306***	-0.0307***	-0.00587**	-0.0121***	-53.00
	(0.00979)	(0.0101)	(0.0101)	(0.00975)	(0.00981)	(0.00272)	(0.00188)	(118.7)
Month * Treatment	-0.00510	-0.00110	-0.00199	-0.000975	-0.000475	0.000855	0.00289	47.18
	(0.0135)	(0.0140)	(0.0137)	(0.0133)	(0.0137)	(0.00325)	(0.00284)	(122.9)
Obs (SHG Member Month)	12765	12797	12766	12777	12796	12755	12742	10852

Regressions include block area fixed effects and health controls. Standard errors are clustered at the village level. Source: SHG Monthly Surveys. Statistical significance levels are as follows: *10%, **5%, ***1%.

Table 2.16: Variation on Time of Village Enrollment on Health Variables

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	In the past month,							
	illness	visiting facility/ dan doctor	visiting health facility	doctor (not village visit)	doctor (including village visit)	admitted	prolonged bed rest	health expenses
				Panel A: Village Enrollment				
Indicator for Village Enrollment	-0.0615	-0.0461	-0.102**	-0.0938*	-0.0346	-0.0340	-0.0626**	-248.7
	(0.0519)	(0.0528)	(0.0516)	(0.0516)	(0.0529)	(0.0243)	(0.0307)	(264.6)
N (SHG Member per month)	12765	12797	12766	12777	12796	12755	12742	10852
				Panel B: Household Enrollment				
Enrolled	-0.0779*	-0.105*	-0.125**	-0.121**	-0.100*	-0.0421**	-0.0573**	-246.0
	(0.0403)	(0.0575)	(0.0509)	(0.0483)	(0.0548)	(0.0195)	(0.0235)	(230.7)
Obs (SHG Member Month)	12765	12797	12766	12777	12796	12755	12742	10852

Regressions include member fixed effects and a time trend. Standard errors are clustered at the village level. Indicator for Village Enrollment is an indicator as to whether at least one household in the village has been enrolled. Enrolled is an indicator for whether the household was enrolled during the given month.Source: SHG Monthly Surveys. Statistical significance levels are as follows: *10%, **5%, ***1%.

Table 2.17: Enrolled Members in Control versus Treatment Villages

	Source: Endline Survey on Demographics						Source: February 2011 SHG Survey on Proportion of Illnesses	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Forward Caste	Poverty Status: Above	Poverty Status: Subsidized	Poverty Status: Below	Agricultural Laborer	House Type	Proportion Ill	Proportion with Health Shock
Treatment Village	0.268***	0.00618	-0.0677	0.0480	0.0466	-0.00803	-0.100	-0.00652
	(0.0981)	(0.00816)	(0.0625)	(0.0548)	(0.0694)	(0.0932)	(0.0758)	(0.0353)
Enrolled Household	-0.0417	-0.000932	-0.0128	-0.00129	0.0233	-0.0621	0.102**	0.0438
	(0.116)	(0.0112)	(0.0544)	(0.0475)	(0.0517)	(0.0674)	(0.0491)	(0.0393)
Treatment * Enrolled	-0.0509	0.0156	0.0743	-0.0804	-0.108	0.139	-0.0369	-0.0312
	(0.143)	(0.0142)	(0.0718)	(0.0618)	(0.0743)	(0.101)	(0.0824)	(0.0471)
Obs (SHG Member)	1594	1615	1615	1615	1687	1675	1243	1243

Observations are limited to households selected and surveyed in the Endline Health Survey and are weighted to be representative of the target population. Data from February 2011 Pilot SHG Survey limited to those SHGs for which data was collected. Health shock is defined as having been on prolonged bed rest or high health expenditure. Regressions include block area fixed effects. Standard errors are clustered at the village level. Statistical significance levels are as follows: *10%, **5%, ***1%.

Table 2.18: Endline Survey -- Treatment Effect on the Treated

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	In the past week,		*In the past year,*					
	no. of sick hh members	health expenses	Rs. 1000 health shock	admitted	prolonged bed rest	no of health shocks	health expenses	health loan
Treatment Village	0.345*	1928.5	-0.0987**	-0.107**	-0.0958*	-0.0812	-5875.6**	-2532.4***
	(0.201)	(1539.4)	(0.0462)	(0.0482)	(0.0499)	(0.0807)	(2644.3)	(670.3)
Enrolled Household	0.112	173.7	0.0204	0.0254	0.00430	0.0566	-935.3	1068.8
	(0.0740)	(304.4)	(0.0352)	(0.0392)	(0.0446)	(0.0500)	(2247.1)	(1080.2)
Treatment * Enrolled	-0.270	-2479.0	0.0312	0.0332	0.0667	-0.0366	-15.13	-50.14
	(0.200)	(1735.2)	(0.0498)	(0.0538)	(0.0593)	(0.0883)	(3695.6)	(1352.4)
N (SHG Member)	1643	1591	1690	1685	1684	1526	1677	1670

Observations are limited to households selected and surveyed in the Endline Health Survey and are weighted to be representative of the target population. Enrolled is an indicator for whether the household was enrolled during the given month. Regressions include block fixed effects. Standard errors are clustered at the village level. Source: 2012 October Endline Survey. Statistical significance levels are as follows: *10%, **5%, ***1%.

Table 2.19: SHG Monthly Survey -- Treatment Effect on the Treated

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	In the past month,							
	illness	visiting facility/ dan doctor	visiting health facility	doctor (not village visit)	doctor (including village visit)	admitted	prolonged bedrest	health expenses
Treatment Village	-0.0788**	-0.0857***	-0.0985***	-0.0920***	-0.0782***	-0.0229***	-0.0237**	-716.0***
	(0.0369)	(0.0296)	(0.0298)	(0.0274)	(0.0274)	(0.00527)	(0.0104)	(196.9)
Enrolled Household	0.0643**	0.0597**	0.0582**	0.0529**	0.0538**	0.00977	0.0121	436.5
	(0.0286)	(0.0265)	(0.0267)	(0.0248)	(0.0250)	(0.00617)	(0.00951)	(390.5)
Treatment * Enrolled	-0.00154	0.0195	0.00253	0.00662	0.0255	0.00797	0.0121	-194.2
	(0.0397)	(0.0343)	(0.0345)	(0.0317)	(0.0325)	(0.00807)	(0.0129)	(453.3)
Obs (SHG Month Member)	12765	12797	12766	12777	12796	12755	12742	10852

Enrolled is an indicator for whether the household was enrolled during the given month. Regressions include block fixed effects. Standard errors are clustered at the village level. Source: SHG Monthly Surveys. Statistical significance levels are as follows: *10%, **5%, ***1%.

Table 2.20: Enrolled Members Versus Unenrolled Members

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Panel A: Household Illnesses							
	In the past week,		*In the past year,*					
	no. of sick hh members	health expenses	Rs. 1000 health shock	admitted	prolonged bed rest	no. of health shocks	health expenses	health loan
Enrolled HH Members	0.0202	-535.8	0.0100	0.0131	0.0160	0.00551	-461.3	-141.4
	(0.0367)	(392.3)	(0.00903)	(0.00906)	(0.00990)	(0.0196)	(547.4)	(118.5)
Obs (SHG Members)	736	702	746	742	741	684	743	740
	Panel B: Individual Illness, Week Recall							
	poor health (today)	was sick	days ill	days missed	health facility	doctor	was admited	total health expenses
Enrolled	0.0513***	0.0629***	0.333***	0.196**	0.0630***	0.0587***	0.0283	-148.7
	(0.0170)	(0.0195)	(0.0940)	(0.0744)	(0.0173)	(0.0157)	(0.0172)	(135.9)
Obs (Individuals)	4146	4285	4274	4254	4279	4276	4285	4285
	Panel C: Individual Illness, Year Recall							
	experienced a health shock	days missed	health facility	doctor	total days admitted	health expense	received health loan	amount borrowed
Enrolled	0.0457**	0.863**	0.0437**	0.0431**	0.285***	684.1	0.0142***	148.9
	(0.0177)	(0.411)	(0.0196)	(0.0194)	(0.0769)	(524.8)	(0.00489)	(93.39)
Obs (Individuals)	4289	4278	4287	4287	4290	4282	4237	4290

Observations are limited to households selected and surveyed in the Endline Health Survey and are weighted to be representative of the target population. Health Shock is defined as high expenditure, being admitted, or prolonged bed rest. Enrolled HH Members is the maximum number of household members on the policy during the time of the study. Enrollment is an indicator for whether or not the individual was ever enrolled. Sample is limited to only treatment villages. Regressions include block fixed effects and health controls from February 2011 Pilot SHG Surveys. Standard errors are clustered at the village level. Source: 2012 October Endline Survey. Statistical significance levels are as follows: *10%, **5%, ***1%.

Table 2.21: Indirect Effects of Insurance

	(1)	(2)	(3)	(4)	(5)
	Week Recall		*Year Recall*		
	days waited before seeking care	forewent treatment	days waited before seeking care	forewent treatment	recovered
	Panel A: Intent to Treat				
Treatment Village	-0.898	0.0414*	1.219	0.0369	0.00667
	(0.736)	(0.0215)	(3.325)	(0.0236)	(0.0288)
Obs (Individuals)	1145	1131	846	892	892
	Panel B: Intensity of Doctor Visits to Villages				
Treatment Village	-0.490	0.0443*	1.947	0.0161	0.0340
	(0.536)	(0.0254)	(3.431)	(0.0380)	(0.0482)
Number of months village visited by OPD	-0.193	-0.00137	-0.331	0.00939	-0.0124
	(0.169)	(0.00953)	(1.060)	(0.0128)	(0.0183)
Obs (Individuals)	1145	1131	846	892	892

Observations are limited to households selected and surveyed in the Endline Health Survey and are weighted to be representative of the target population. All regressions include health status variables from the February 2011 SHG Pilot Survey as independent control variables. Regressions include block area fixed effects. Standard errors are clustered at the village level. Source: 2012 October Endline Survey. Statistical significance levels are as follows: *10%, **5%, ***1%.

Table 2.22: Endline Survey - Village Doctor Visits

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	In the past week,		*In the past year,*					
	no. of sick hh members	health expenses	Rs. 1000 ($US 20) health shock	admitted	prolonged bed rest	no of health shocks	health expenses	health loan
Treatment Village	0.0950	312.8	-0.138***	-0.140***	-0.134***	-0.148**	-7680.2**	-2363.5***
	(0.111)	(327.1)	(0.0396)	(0.0440)	(0.0483)	(0.0675)	(2877.2)	(695.2)
Number of months	0.0365	6.727	0.0218	0.0215	0.0279	0.0196	448.4	77.66
village visited by OPD	(0.0494)	(150.9)	(0.0156)	(0.0162)	(0.0191)	(0.0280)	(968.0)	(252.5)
Obs (SHG Members)	1643	1591	1690	1685	1684	1526	1677	1670

Observations are limited to households selected and surveyed in the Endline Health Survey and are weighted to be representative of the target population. Regressions includes block fixed effects and health controls. Standard errors are clustered at the village level. Source: 2012 October Endline Survey. Statistical significance levels are as follows: *10%, **5%, ***1%.

Table 2.23: SHG Monthly Survey - Village Doctor Visits

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	In the past month,						
	illness	dan network doctor	visiting health facility	doctor (not village visit)	doctor (including village visit)	health shock	health expenses
Village Visited by OPD	0.0245	0.0202*	0.0126	0.0173	0.0391	0.0104	181.8*
	(0.0250)	(0.0103)	(0.0206)	(0.0210)	(0.0242)	(0.00697)	(96.25)
Number of villages visited by OPD	0.00486***	0.00164**	0.00465***	0.00429***	0.00421**	0.00177***	33.02**
	(0.00145)	(0.000679)	(0.00149)	(0.00144)	(0.00161)	(0.000502)	(13.22)
Obs (SHG Member Month)	12765	11502	12766	12777	12796	12757	10852

Each row is a separate regression. Health shock is defined as having either been admitted or on prolonged bed rest. Regressions include treatment indicator, time trend, health controls and block fixed effects. Standard errors are clustered at the village level. Source: SHG Monthly Survey. Statistical significance levels are as follows: *10%, **5%, ***1%.

Table A.2.1: Health Illness Summary Statistics (HH)

	Control Villages				Treatment Villages			
	Obs (SHG Members)	Mean	SD	Max	Obs (SHG Members)	Mean	SD	Max
In the past week,								
number of sick hh members	907	0.74	0.83	6	736	0.88	1.06	12
health expenses	889	618	1792	23000	702	1085	9734	152270
In the past year,								
Rs. 1000 ($US 20) health shock	944	0.56	0.50	1	746	0.47	0.50	1
admitted	943	0.53	0.50	1	742	0.44	0.50	1
bed rest	943	0.50	0.50	1	741	0.43	0.50	1
health expenses	934	19189	34859	500500	743	15669	39520	550000
health loan	930	3683	15537	300000	740	1330	10097	200000

Observations are limited to households selected and surveyed in the Household Health Survey and are weighted to be representative of the target population. Expenses given in Ind Rupees (1USD = 50 Rs). Source: 2012 October Endline Survey.

Table A.2.2: SHG Monthly Summary Statistics

	Control Villages					Treatment Villages				
	Obs	Mean	SD	Min	Max	Obs	Mean	SD	Min	Max
present at meeting	6425	0.881	0.324	0	1	8089	0.821	0.384	0	1
In the past month, the household has experienced										
illness	5819	0.383	0.486	0	1	6958	0.317	0.465	0	1
visiting health facility/dan docto	5834	0.376	0.484	0	1	6975	0.322	0.467	0	1
visiting health facility	5829	0.374	0.484	0	1	6949	0.299	0.458	0	1
doctor (including village visit)	5833	0.354	0.478	0	1	6975	0.310	0.462	0	1
doctor (not village visit)	5832	0.353	0.478	0	1	6957	0.284	0.451	0	1
admitted	5825	0.041	0.199	0	1	6942	0.024	0.155	0	1
bed rest	5817	0.052	0.221	0	1	6937	0.033	0.180	0	1
health expenditure	4861	1067	11527	0	700000	5999	433	3240	0	150000

Source: SHG Monthly Surveys.

Table A.2.3: Summary Statistics on Assets

	Control Villages					Treatment Villages				
	Obs	Mean	SD	Min	Max	Obs	Mean	SD	Min	Max
In the past year,										
Mortgaged Assets	919	0.209	0.406	0	1	738	0.198	0.398	0	1
Sold Assets	915	0.142	0.349	0	1	735	0.131	0.338	0	1
Purchased Assets	918	0.298	0.458	0	1	736	0.273	0.446	0	1
Conditional upon selling or mortgaging assets, reasons given for doing so:										
Health	315	0.182	0.378	0	1	217	0.129	0.332	0	1
Business Requirements	315	0.203	0.389	0	1	217	0.284	0.440	0	1
Current Outstanding Loans										
Number of Loans	835	0.903	0.855	0	7	694	0.831	0.821	0	6
Has health loan	826	0.123	0.329	0	1	690	0.091	0.288	0	1
Has home/business improvement loan	826	0.390	0.488	0	1	690	0.418	0.494	0	1
Total Loan Outstanding	831	36135	107212	0	1200000	689	53076	453380	0	11500000

Observations are limited to households selected and surveyed in the Household Health Survey and are weighted to be representative of the target population. Source: 2012 Endline Survey.

Table A.2.4: Summary Statistics of SHG Finances (Internal and External)

	Control Villages				Treatment Villages			
	Obs (SHG Members)	Mean	SD	Max	Obs (SHG Members)	Mean	SD	Max
SHG Financial Records								
input into funds	10139	882	2023	37840	10849	791	2154	109740
output into funds	10139	5390	490469	49400000	10849	569	3605	72000
loan disbursement	10132	477	2814	50000	10838	501	3020	50000
outstanding amount	2195	12614	19604	544000	2157	10994	11903	80000
MFI Financial Records								
outstanding loan amount	18438	2501	7759	50000	18354	2305	7432	50600
overdue principle	18438	137	1037	28533	18354	170	1058	26533
overdue interest	18438	3	53	3047	18354	4	65	2773
amount paid	17121	276	1253	31000	17043	239	1148	50000

Source: Panel A - SHG Financial Records, Panel B - MFI Financial Records.

Table A.2.5: Distribution of Visits by Insurance Doctor to Villages

Number of months village visited by doctor	Number of villages	Percentage of villages	Number of households	Percentage of households
0	5	23.81	119	9.5
1	3	14.29	254	20.29
2	3	14.29	118	9.42
3	5	23.81	497	39.7
4	3	14.29	196	15.65
5	1	4.76	13	1.04
6	1	4.76	55	4.39
Number of months, household visited by doctor	**Number of villages**	**Percentage of villages**	**Number of households**	**Percentage of households**
0			1,100	87.86
1			107	8.55
2			23	1.84
3			20	1.6
4			2	0.16

Observations include only treatment villages.

References

Arnott, Richard J., and Joseph E. Stiglitz. 1988. The basic analytics of moral hazard. *The Scandinavian Journal of Economics* 90, (3, Information and Incentives. Vol. 1: Organizations and Markets) (Sep.): 383-413.

Banerjee, Abhijit, Esther Duflo, Rachel Glennerster, and Cynthia Kinnan. 2009. The Miracle of Microfinance? Evidence from a Randomized Evaluation. http://econ-www.mit.edu/files/4161

Chankova, S., Sulzbach, S., and Diop F. 2008. Impact of mutual health organizations: evidence from West Africa. *Health Policy and Planning.* 23(4): 264 – 276.

Das, Jishnu, Jeffrey Hammer and Carolina Sanchez – Paramo. 2011. The Impact of Recall Periods on Reported Morbidity and Health Seeking Behavior. Policy Research Working Paper. 5778. World Bank.

Dupas, Pascaline. 2011. Health Behaviors in Developing Countries. Prepared for the Annual Review of Economics, Vol 3.

Dupas, Pascaline, and Jonathan Robinson. 2009. Savings Constraints and Microenterprise Development: Evidence from a Field Experiment in Kenya. National Bureau of Economic Research Working Paper No 14693.

Ekman, Bjorn. 2004. Community-based health insurance in low-income countries: a systematic review of the evidence. Health Policy and Planning. 19(5): 249-270.

Jakab, Melitta and Chitra Krishnan. 2003. Review of the Strengths and Weaknesses of Community Financing. In Health Financing for Poor People: Resource Mobilization and Risk Sharing, edited by A. S. Preker and G. Carrin. Washington D.C.: The World Bank.

Jalan, Jyotsna and Martin Ravallion. 1999. Are the poor less well insured? Evidence on vulnerability to income risk in rural China? *Journal of Development Economics,* 58: 61 – 81.

Jutting, J. 2004. Do community-based health insurance schemes improve poor people's access to health care? Evidence from rural Senegal. *World Development* 32: 273 – 88.

Morduch, Jonathan. 1999. Between the State and the Market: Can Informal Insurance Patch the Safety Net? *World Bank Research Observer*, 14(2): 187 – 207.

Morduch, Jonathan. 2006. "Micro-insurance: The Next Revolution?" in Understanding Poverty, edited by Abhijit Banerjee, Roland Benabou, and Dilip Mookherjee. Oxford University Press.

Townsend, Robert M. 1994. Risk and insurance in village India. *Econometrica* 62, (3) (May): 539-91.

Wagstaff, Adam and Magnus Lindelow. 2008. Can insurance increase financial risk? The curious case of health insurance in China. Journal of Health Economics, 27: 990 – 1005.

Chapter 3: The Distributional Consequences of Micro Health Insurance: Can a Pro-Poor Program Prove to be Regressive?

Abstract: Despite the rapid growth of micro health insurance (*MHI*) programs, there is little empirical evidence of their distributional consequences. If health care consumption increases with income and households pay identical premiums, then MHI may result in unintended regressive premium redistribution from poorer to wealthier households. This paper assesses the effect of household responses to MHI on the targeting and progressivity of MHI income transfer and health care benefits. Partnering with a microfinance MHI distributor in rural Maharashtra, India, I exploit an imbalanced randomized controlled trial to separately identify the following: 1) the initial health care distribution by income among eligible households, 2) differential demand for MHI, and 3) heterogeneous effects of MHI on health care by income. I estimate that among those eligible for MHI, households Below the Poverty Line (BPL) spend 34% less on health care than non-BPL households prior to MHI. This relationship persists even among the sub-sample of households that choose to enroll in the MHI program, suggesting that MHI will lead to poorer households subsidizing wealthier households unless there are offsetting heterogeneous effects of MHI on health care by income. Using household fixed effects and a difference-in-difference analysis, I estimate that MHI improves health care consumption more for BPL households. These results suggest that gains in health care are concentrated among poorer households, thereby narrowing the gap in health care among the insured. Strikingly, almost two years after the introduction of MHI, there is no significant relationship between health care and income among the insured. Thus, even though *ex-ante* health care consumption suggests MHI will result in regressive premium redistribution, *ex-post* behavior suggests the poor will not subsidize wealthier households.

1. Introduction

Poor households in low income countries suffer large health burdens and are particularly vulnerable to health shocks (Banerjee et al. 2013, Dupas and Robinson 2009). Yet health insurance has traditionally been missing in most developing countries, despite informal risk pooling offering insufficient protection against health shocks (Townsend 1994, Morduch 1999, Jalan and Ravallion 1999). In the past two decades, insurance specifically targeted at low income households, known as micro health insurance (MHI) or community based health insurance (CBHI), has attempted to fill this gap. In the same way that microfinance innovated to make credit available where traditional banking had previously been absent, microinsurance now aims to use similar innovations to bring insurance markets to low income households. Already, MHI covers over 40 million lives in Africa, the Americas, and Asia, and continues to expand rapidly (McCord et al. 2013, McCord et al. 2012, Oza et al. 2013 , Morduch 2006).[1]

Though there is a general consensus that MHI improves access to health care and increases financial protection (Leatherman et al. 2012, Chankova et al. 2008), there is little empirical evidence on the distribution of benefits and whether they are concentrated among poorer households, the target demographic. Specifically, how do health care and income transfer benefits distribute among insured households? As MHI continues to expand, understanding how household responses to MHI influence the distribution of such benefits is critical in improving MHI contracts to maximize benefits and limit unintended adverse effects for poorer households.

The relationship between health care and income informs the expected redistribution of premiums across income groups. Health care consumption generally rises with income in developing countries, despite poorer households having a higher burden of illness (Wagstaff 2002).[2] This health care

---

[1] Micro health insurance has become so significant in India, the context of this study, that the government of India's 11th Five Year Plan specifically advocated community based health insurance to improve access to health for poor rural households (GoI 2008). India is considered to be a powerhouse in microinsurance and currently is the leading Asian microinsurance market (ILO 2012).

[2] This is not true in most OECD countries, where the poor have the highest burden of illness, but also utilize the most health care (Wagstaff 2002).

inequality suggests that MHI could result in premiums redistributing regressively such that poorer households cross-subsidize the health care consumption of relatively wealthier households. However, along any segment of the income distribution there is a tension between increased health care demand from a higher income and reduced health care demand from a lower health burden. Thus, it is unclear which effect dominates within the left tail of the income distribution, the population eligible for MHI.

Additionally, if enrollment patterns and MHI effects on health care differ by income, then MHI will alter the relationship between health care and income among the insured. If these household responses increase (*decrease*) health care inequality among insured MHI members, it suggests an increase (*decrease*) in the regressivity of premium redistributions. This paper first identifies the preexisting distribution of health care consumption by income among the eligible population as targeted by the administrating organization. It then separately assesses the effect of the following household behaviors (*self-targeting*) on the final distribution of health care consumption by income among insured MHI members: differential demand for MHI, and heterogeneous effects of MHI on health care consumption by income. The identification of this latter component additionally informs whether MHI is progressive in health care benefits – that is, whether the gains in health care are highest for poorer households and thus reduces health care inequality (and *horizontal inequity*[3]) among insured members. These effects are identified using primary data and exploiting the staggered expansion and enrollment of a typical MHI program in rural India.

This paper confirms a negative relationship between health care consumption and income at baseline even among the eligible target MHI population. Below the Poverty Line (BPL) households consume only 66% of non-BPL households' annual health care expenditure, suggesting a health care

[3] Horizontal inequity is commonly defined as when persons with the same health need receive different levels of care. It is based on the belief that the demand for health care should be primarily a function of health need and preferences rather than socioeconomic factors such as income and power distributions. See Fleubaey and Schokkaert (2012) and Wagstaff and van Doorslaer (2000).

expenditure concentration index[4] of .15. This underlying distribution of health care suggests that poorer households could systematically pay disproportionately more and cross-subsidize wealthier households under MHI through regressive premium redistributions. Differential enrollment into MHI and heterogeneous effects on health care by income will either exacerbate or reduce this concern.

Similar to other microinsurance products in developing countries (Dercon et al. 2008), I find that MHI demand increases with income, but contrary to expectations, is not influenced by past health care consumption. I find that health care inequality prior to MHI is the same among the entire eligible population and among the sub-sample of households that choose to enroll in MHI. This suggests that the choice of enrollment neither increases nor decreases the regressivity of MHI transfer benefits.

Using panel data with household fixed effects, I find that among the insured, BPL households experience greater gains in health care from MHI. I estimate that MHI differentially increases BPL households' health care expenditure by almost USD 3 per month of coverage. This is further supported using a difference-in-difference estimation which compares health care inequality across MHI enrollers and MHI coverage. This suggests that MHI narrows the gap in health care consumption among the insured, and the poor are more sensitive to a reduction in the cost of health care. A striking result is that twenty months after the initial MHI offer, there is no significant relationship between health care consumption and income among the insured. These results suggest that heterogeneous effects of the MHI, rather than differential enrollment choices into the insurance program, is the primary factor in reducing health care inequality among the insured. Thus, even though *ex-ante* health care behavior raised the possibility of redistribution of premiums away from poorer households, *ex-post* health care behavior suggests a single priced premium will not lead to regressive transfers.

[4] The concentration index is twice the area between a concentration curve (i.e., Lorenz Curve) and the line of equality (i.e., 45 degree line). In this case, the concentration curve plots the cumulative percentage of total health care expenditure against the cumulative percentage of the population, ranked by income (from poorest to richest) (O'Donnell et al. 2008).

This paper adds to a sparse, but evolving, literature on the distribution of benefits of MHI. Previous literature has highlighted concerns of wealthier households capturing a greater share of net benefits, but health care inequality being a potential mechanism has not yet been explored. Additionally, the results on differential health care gains have been mixed. Focusing on claims data from VimoSEWA, a large micro health insurance program in India, Ranson et al. (2006) finds that the rural poor are less likely to submit and receive claims. This suggests that insurance results in an income transfer away from poorer households, but the authors are unable to comment on how these transfers compare with the underlying distribution of health care by income. Sneider and Hanson (2006) find a large reduction in inequality in health care utilization when comparing the insured with the uninsured in assessing MHI in Rwanda. However, Wang et al. (2005) study community based health insurance in urban China and find that the gains in health care utilization and outpatient expenditure increase with income, suggesting an increase in health care inequality.

The main contribution of this paper is improved identification of self-targeting household responses on the distribution of benefits. This paper is the first to disentangle the effects of administrative targeting, enrollment, and the MHI program itself on the distribution of health care. A primary concern in estimating the effects of MHI is the difficulty in isolating the program's effect from the endogenous choice of enrollment into the program. The above studies are limited to comparisons between insured and uninsured households, or are descriptions of correlations among the insured, raising concerns about whether results reflect the type of households which enroll into the program or the effect of the program itself on the household. The potentially confounding effect of endogenous enrollment is minimized in this paper in three ways: 1) estimating the effect of MHI using *within* household comparisons, 2) identifying enrollment patterns among households that received MHI after the study period, and 3) the program's eligibility requirements. Though the villages offered MHI were randomly selected, the identification of the heterogeneous health care effects does not rely on comparisons across control and treatment arms due to potential imbalance. Rather, the heterogeneous effects documented in this paper are identified using within household comparisons. Using household fixed effects and variation in the timing of enrollment, I assess changes in health care consumption by observing within household differences after a change in MHI status. This methodology minimizes the concern of confounding estimates of the MHI effect with the type of household that chooses to enroll. Secondly, by observing enrollment patterns among control households (after MHI was made available to all households), I can discern how the enrollment would bias estimates of MHI's effect on the distribution of health care by income. Finally, an eligibility requirement for a household to purchase the MHI was that 80 percent of the household's preexisting group within the distributing organization (15 to 20 individuals) also had to enroll. Because enrollment is determined at the group level, selection into MHI by household characteristics (such as health status) is reduced due to heterogeneity among members within a group.

The remainder of this paper is organized as follows: Section 2 describes the specifics of the microinsurance contract, data sources, and health care in the context of India; Section 3 provides a theoretical framework of the effect of insurance on the distribution of health care consumption by income; Section 4 outlines the estimations and identifying assumptions; Section 5 reviews the results; and Section 6 concludes.

2. Context of Micro Health Insurance

Increased concerns of low-income households' vulnerability to health shocks and the success of the microfinance movement have given rise to the micro insurance movement. Though MHI contracts differ across the globe, they tend to share three key features. The first is that they are distributed through pre-existing local community organizations, the most common of which are micro finance institutions. By building upon the infrastructure of preexisting local community organizations, MHI programs successfully overcome barriers such as low trust, high transaction costs, adverse selection, and moral hazard – challenges that have traditionally prevented formal insurance markets from serving this demographic (Morduch 2006). Second, MHI aims to reduce, but does not eliminate, the cost of health care.[5] And finally, MHI tends to charge a single premium price for all clients (Jakab and Krishnan 2003, Morduch 2006, Ekman 2004).

2.1 Overview of the Micro Health Insurance Contract

In January 2011, Chaitanya, a non-profit microfinance institution (MFI) working on women's empowerment and microfinance in Junnar sub-district of rural Maharashtra, expanded its community based MHI program, Dipthi Arrogya Nidhi (DAN). Though micro health insurance contracts differ in design, DAN shares many of the characteristics common to MHI. These include distributing through an existing MFI infrastructure (the most common provider of micro health insurance), reducing the cost of health care, implementing a coverage cap and co-pay, and charging a single premium price. The reduction in the cost of health care includes both price reductions and mechanisms such as improved signals of health care quality (e.g., empanelling facilities), easier access to health care, and increased saliency of health.

The cost of membership to DAN is INR 200 (USD 4) per person per year if the household insures 1 or 2 persons, or INR 150 (USD 3) per person per year if the household insures 3 or more persons. The main provisions of the health insurance contract are discounted prices (5 to 20%) negotiated at private network medical facilities, which include hospitals, medical laboratories, and pharmacies. Additionally, for in-patient treatment, the member receives 60 percent reimbursement of their medical fees at network private hospitals, and 100 percent reimbursement at government medical facilities, up to a limit of INR 15,000 (USD 300) per event.[6] The product also includes a 24-7 medical help-line, health camps, and monthly village visits by a doctor to offer referrals and basic medicines. However, village

---

[6] Specific illnesses may have lower coverage caps based on predefined categories of illness type. Relative to other micro health insurance plans, this limit is relatively generous. For example, VimoSEWA, a large micro insurer in India, has a limit of INR 2,000 – 6,000 (USD 40 – 120) and RSBY (government insurance for BPL households) has a limit of INR 30,000 (USD 600) for the entire household (SEWA 2013, RSBY 2013a).

visits by a doctor were intermittent and only one health camp was implemented during the timeframe of the research study.

DAN capitalizes on Chaitanya's preexisting microfinance Self Help Groups[7] (SHGs) structure. The option to purchase the contract is limited to SHGs in which at least 80 percent of members purchase the MHI, though women can decide the number of family members to enroll. This eligibility requirement reduces concerns of adverse selection by reducing the likelihood of household characteristics being correlated with enrollment into the program. In addition to improving financial sustainability, this feature reduces concerns of endogeneity from the enrollment decision when estimating effects of MHI. If all population heterogeneity was within SHGs, then the eligibility requirement would be the most effective in ensuring enrollment is uncorrelated with household characteristics. The eligibility requirement falls short of preventing household characteristics from being correlated with enrollment by three factors: heterogeneity across SHGs, members being free to choose additional household members to enroll, and requiring 80 percent compliance (as opposed to 100).

DAN does not involve a third party insurer, and health claims and operational costs are financed by the premiums collected. A team of medical doctors, who are able to judge the technical validity of the claim, reviews the reimbursement claims. Afterward, the claims are sent to a committee composed of local women from the Self Help Groups to determine the final disbursement amount.

2.2 Timeline and Data Sources

In January 2011, Chaitanya offered the MHI in 21 randomly selected treatment villages (1,314 households[8]). Almost two years later in November 2012, the program was made available to the 22

---

[7] SHGs are groups of 15 – 20 women who voluntarily come together to save and access micro credit from Chaitanya.

[8] Each member is assumed to be a separate household for the analysis in this paper. However, it may be the case that more than one member belongs to the same household.

remaining villages (1,311 households), referred to as control villages.[9] The households evaluated are limited to members of the MFI as of January 2011, ensuring that results are not a function of changes in the eligible population's composition.

In October 2012, a large comprehensive Household Health Survey was conducted on a subset of 1,702 households.[10] This survey collected basic demographic indicators and household health incidents, health care utilization and health care expenditure in the past year. Additionally, from October 2011 to July 2012, SHG Monthly Health Surveys were conducted during the member's monthly SHG meeting.[11] This panel collected information on household illness, health care utilization and health care expenditure with a one month recall period.

Both surveys collect household level information, but are conducted in different settings by different enumerators. The Household Health Survey was conducted at the member's home by a team of fifteen hired surveyors (only one local to Junnar sub-district), and the SHG Monthly Health Surveys were conducted during standard SHG meetings by the regular field staff of the MFI.

Unfortunately, baseline data was not collected prior to the initial expansion of the MHI program. However, the surveys collect *Midline* and *Endline* data for treatment villages, but can be considered to be *Baseline* surveys for control villages where the MHI had not yet been offered. Furthermore, the demographic data collected in the Household Health Survey can be considered descriptive of baseline characteristics for *both* control and treatment households. The variables collected include such descriptors

[9] The initial randomization was done on 61 villages (30 treatment villages and 31 control villages). However, after the start of the program, it was determined that 18 villages were non-functional (9 treatment villages and 9 control villages).
[10] In larger villages, randomly selected households were chosen to be surveyed. In total, 2,068 members were selected to be surveyed.
[11] The SHG Monthly Surveys are limited to households in attendance at the meetings. On average, 57 percent of households were in attendance in a given month. 85 percent of households were represented at least once during the panel period.

as the type of ration card issued by the government (indicators of income levels)[12], whether the household belongs to a disadvantaged caste, whether anyone in the household participates as an agricultural laborer, and the condition of the house infrastructure. These measures would be difficult to change in the time period of the study, particularly the issuance of a new ration card (the primary variable for poverty status used in this paper). Thus, even if the MHI affected poverty status and income, these measures are still likely to reflect baseline socioeconomic characteristics in treatment villages.

Institutional data is used for initial MFI membership[13] and MHI enrollment and claims from January 2011 to June 2013.

2.3 Health Expenditure as a Proxy for Health Care

This paper uses household health care expenditure as a proxy for health care consumption. This is possible in the unique setting of India because user fees are the primary financing mechanism for health care (La Forgia and Nagpal 2011). As long as household health expenditure is mapped into health care in the same way for all income groups, differences in health care expenditure are a valid measure of differences in health care consumption.

One concern in using health care expenditure might be differential access to discounted care by poorer households. However, access to public facilities in India is not limited to any sub-population (i.e., they are available to all income levels). Though the government has stated its commitment to increasing health, public health expenditure as a percentage of GDP hovers around 1%, one of the lowest in the world (GoI 2013). The capacity of the public health system is weak and, as a result, services are far from free (GoI 2008). Though public facilities are less expensive due to government support, the failing state of the public health infrastructure suggests that differences in expenditure, even between public versus

[12] The four types of ration cards are as follows (increasing in income): Below the Poverty Line with an AAY stamp, Below the Poverty Line without an AAY stamp, Subsidized Ration Card, Above the Poverty Line.

[13] Initial MFI membership was constructed by reviewing SHG meeting records in the three months prior to the introduction of MHI. It could be the case that some members were included in the baseline because their name was still on the official meeting roster even though they had already left the MFI.

private facilities, still accurately reflect differences in the quality of health care consumed.[14] The lack of restrictions on the population served by the public health infrastructure, combined with the high level of user fees to finance health care, suggests that differences in expenditure are valid proxies for differences in health care.

Existing health insurance is another potential source that differentially changes the mapping between household health expenditure and health care consumption by income. If poorer households receive different prices for health care through these programs, then a comparison of health care expenditure by income would no longer be a viable proxy for comparisons in health care by income. In 2006, the Government of India began providing public sponsored insurance schemes for households Below the Poverty Line; As of 2010, approximately 20% of the population has been insured under such schemes (La Forgia and Nagpal 2011). In Pune District of Maharashtra, the setting of this study, the relevant public program is Rashtriya Swasthya Bima Yojna, RSBY.[15] However, RSBY seems to have limited reach in the area during the time of the study (RSBY 2013b). In the Household Health Survey, households were asked whether they received any financial assistance or reimbursements for their health care expenditures in the past year recall. Only 5% of households claimed to have received any assistance (none mentioning RSBY), and the receipt of financial assistance did not differ by poverty status, a primary income variable used in this paper.

2.4 Summary Statistics

Table 3.1 provides the demographic details of MFI household members (i.e., households targeted by the MHI program). Columns 1 – 2 provide an overview of the entire sample, and Columns 3 – 6 separates the sample into households from treatment villages and control villages. Reflecting the rural

[14] Some evidence suggests that though public facilities are often thought to serve poorer households, it is wealthier households that receive the greatest benefit. Balarajan et al. (2011) reports that richer households use a greater share of public services, particularly for tertiary care and hospital based services.
[15] The state government of Maharashtra launched a health insurance program called RGLHS (Rajiv Gandhi Lifesaving Health Scheme) in August 2011. However, this program has not yet been implemented in Pune District.

setting, the majority of households are involved in agriculture. As expected in a population of non-profit MFI clients, households have low socioeconomic characteristics with a significant proportion being designated Below the Poverty Line (BPL). Income Rank is created using principal component analysis and is a composite of the following variables, all of which increase with poverty: ration card, disadvantaged caste, house status, and household participation in agricultural labor. Both treatment and control villages are typical of the population targeted by MHI, but as Column (7) indicates, treatment households have slightly higher socioeconomic characteristics.

Table 3.2 provides an overview of enrollment into the MHI program. Households are considered *Enrolled* if at least one household member for at least one year enrolled in the MHI program. The program take up is high, with over 60% enrollment among households introduced to MHI for almost two years. In comparison, most micro insurance programs struggle to reach 30% enrollment (Matul et al. 2013). One likely explanation for the relatively high take-up rate is the requirement that 80 percent of members in an SHG become enrolled for any single member to be eligible to purchase the MHI. As expected, the insurance is more prevalent in treatment villages even after control villages become eligible for the MHI. However, 7 months after the initial offer of MHI, treatment and control villages had very similar enrollment rates of approximately 30 percent. This suggests that the difference in enrollment is primarily a result of the operational time and cost involved in expanding to new villages as opposed to differences in the underlying factors of program demand.

Table 3.3 summarizes the enrollment and claims data from January 2011 to October 2012. On average, each member enrolls almost two household members. This reflects members typically enrolling either only themselves or at least 3 household members to receive the discounted premium. Among these enrolled households, 9.8 percent submitted a claim for reimbursement and INR 253 (USD 5) per enrolled household was disbursed on average. Table 3.4 describes the claims in greater detail. For those claims under the coverage cap,[16] the average claimable expenditure was INR 5,533 (USD 111) and the average amount disbursed was INR 2,911 (USD 58). Over 92% of claims provided a reimbursement greater than 50 percent, and 75% of claims provided a reimbursement greater than 60 percent. The most common types of incidents were illnesses related to enteric fever (typhoid) and malaria.

Table 3.5 summarizes the health variables used in this paper.[17] In the SHG Monthly Health Surveys, 34 percent of households sought health care and health expenditures averaged INR 604 (USD 12) per month. In the year recall, almost half the sample experienced a large health shock resulting in either an expenditure of at least INR 1,000 (USD 20) or being admitted to a health facility. Annual health expenditure averaged over INR 15,000 (USD 300). Health expenditure refers to the amount households report paying for health care in the previous year or month. As a result, any financial assistant received after the care was paid for, including claims reimbursements, will not be deducted from the reported amount. This ensures that the reported health expenditure is a valid proxy for the quality of health care consumed. Both surveys suggest health care to be a common event with significant expenditure.

The self-reported data suggests that claims should be larger than the observed 10 percent. It is possible that the discrepancy is due to measurement error in self-reported data or not all health incidents being claimable. However, it is also likely that members are still learning the claims process and that claims will continue to rise to reach the rates reported in the surveys.

3. Theoretical Framework

The following summarizes the results of a simple theoretical framework on the role of insurance and the relationship between health care consumption and income. The details of the model are provided in an attached Appendix.

3.1 Health Care Consumption and Income

[16] Over 95% of claims submitted were below the coverage cap.
[17] Health expenditure is winsorized at the 99.9% for monthly recall and the 99% for annual recall.

I assume the optimal level of health care is determined by households maximizing a general utility function with arguments for health care (m) and non-health care consumption (N), subject to income (W):

$$(1)\ \max_{N,m} U(N, m, \theta, s)\,, \qquad s.t.\ N + P_m m \leq W$$

where P_m is the price of health care, θ is a permanent health status, and s is a random health shock. Higher θ and s signify worse health. I assume the following: 1) utility is increasing and concave in N and m: $\frac{\partial U}{\partial N} > 0$, $\frac{\partial U}{\partial m} > 0$, $\frac{\partial^2 U}{\partial N^2} < 0$ and, $\frac{\partial^2 U}{\partial m^2} < 0$; 2) increasingly poor health reduces utility: $\frac{\partial U}{\partial \theta} < 0$ and $\frac{\partial U}{\partial s} < 0$, but increases the marginal utility of health care: $\frac{\partial U}{\partial m \partial \theta} \geq 0$ and $\frac{\partial U}{\partial m \partial s} \geq 0$; 3) utility derived from health care consumption and non-health care consumption is additively separable: $\frac{\partial \mathrm{U}}{\partial \mathrm{m}\, \partial \mathrm{N}} = 0$; and 4) the marginal utility of non-health care consumption is unaffected by health status: $\frac{\partial \mathrm{U}}{\partial \mathrm{N}\, \partial \theta} = 0$ and $\frac{\partial \mathrm{U}}{\partial \mathrm{N}\, \partial \mathrm{s}} = 0$. The last two assumptions are simplifying assumptions.

If the only difference between poorer and wealthier households is their endowment W, then poorer households will have lower health care consumption (i.e., $\frac{dm^*}{dW} > 0$). However, if poorer households are also more likely to have greater health needs ($\frac{d\theta}{dW} \leq 0$), the relationship between observed health care consumption and income becomes ambiguous: wealthier households consume *more* health care because of their higher income, but consume *less* health care because of their reduced marginal value of health care from having a higher health status.

It is important to note that *both* cases exhibit horizontal inequity (i.e., for a given health status, determined by θ and s, poorer households demand less health care). Horizontal equity suggests differences in health care due directly from income will be eliminated. The additional assumption that there exists an $\bar{M}(\theta, s)$ such that $\frac{dU(N,m,\theta,s)}{dm} = 0|_{m \geq \bar{M}(\theta,s)}$ is one mechanism that could lead to such a result.

The following cases assume $\frac{d\theta}{dW} = 0$ and instead separately identifies the heterogeneous effects of W and θ.

3.2 Insurance with a Predetermined Indemnity Schedule

It follows from the concavity of the utility function that all households will prefer an insurance contract in which the household chooses an indemnity schedule *ex-ante* and pays the corresponding actuarially fair premium. If households can commit to a pre-specified health care amount for a given health shock s, households will generally reduce $m^*(W, P, \theta, s)$ in healthier states and increase $m^*(W, P, \theta, s)$ in sicker states relative to when uninsured (i.e.,$\frac{\partial(m^*_{insured} - m^*_{uninsured})}{\partial s} \geq 0$).[18] Though such insurance increases welfare, health care demand will still increase with income, $\frac{\partial m^*}{\partial W} > 0$. However, it is no longer necessary that health care increases with poor health. Whether health care demand would become more or less sensitive to income, W, or health status, θ, for a given health state is ambiguous, depending on the relative change in $m^*(W, P, \theta, s)$ and $N^*(W, P, \theta, s)$, the probability of the shock s, and the curvature of the utility function in each parameter. If the additional assumption of $\bar{M}(\theta, s)$ is made, then the likelihood of decreasing health care inequality by income is increased.

3.3 Insurance with Price Reductions

Most insurers, including micro health insurers, cannot observe the correct form of treatment for every s and instead provide an indemnity schedule based on health care consumption. The insurer continues to charge an actuarially fair premium based on the household's expected payout, but the household can no longer credibly commit to an amount of health care prior to the realization of s. Unlike the previous case, the household does not choose the level of health care that maximized its expected utility prior to the realization of s, and instead determines its optimal level of health care by maximizing its utility conditional on s. Because the price faced by the household is less than P_m, the marginal benefit

[18] This holds true when $0 < m^*_s < W$. If for a given s, households optimally choose $m^*_s = 0$ or $m^*_s = W$, then it is not necessary that m^*_s will change with insurance.

of the household's demand for health care is less than its marginal cost (across all states s).[19] Relative to the previous case, welfare will decrease due to this "overconsumption" of health care.

Because households now face a lower price for health care, they will substitute away from non-health care, N, towards health care, m. However, households will also have to pay a higher premium because of this increased demand in health care. Thus, relative to an insurance contract with a predetermined indemnity schedule, this contract increases health care demand by reducing prices but decreases health care demand by charging higher premiums. The demand for health care will increase most in the states which have the lowest probability of being realized, because the consequential increase in the premium will be relatively low. Recall from the previous case that even if choosing an *ex-ante* indemnity schedule, households would prefer to transfer resources into sicker states and potentially consume greater health care overall.

Health care inequality among the insured would now be based on these combined price and income effects and remains ambiguous. The additional assumption of $\bar{M}(\theta, s)$ increases the likelihood of decreasing health care inequality by income. There will be no expected redistribution of premiums across households because each household pays a premium to reflect their own expected health care expenditure.

It is no longer the case that all households will prefer to purchase such an insurance contract. As Wolfe and Goddeeris (1991) show, under such a contract it is ambiguous whether wealthier households will be more or less likely to demand insurance because of the trade-off between the benefit of greater savings from higher health care expenditure versus the cost of an increased premium and lower risk aversion. It is also ambiguous whether sicker households have greater demand for the insurance because of the trade-off between increased savings and the higher premium.

[19] The Appendix shows the case in which the household pays nothing for health care but internalizes the consequential increase in the premium price. The results are similar if households are required to pay a co-pay – the marginal cost experienced by the household increases but remains less than the true marginal cost (P_m).

3.4 Charging a Single Premium

If the insurer is unable to identify households' wealth (W) and non-state-dependent health status (θ), then a single premium will be charged for the insurance contract. While the insurance contract is still priced to be actuarially fair across all households that enroll, it is no longer actuarially fair for each household.

When the insurance is pooled across health status, insurance will be unambiguously demanded more by those with poor health, assuming the marginal expected disutility from one's health status is less when uninsured than insured. It still remains the case that the demand for insurance by income is ambiguous because of the tension between increased savings and lower risk aversion. However, insurance demand for higher income (*sicker*) households is higher relative to the demand when insurance was priced based on household characteristics (because the premium has been lowered for higher income (*sicker*) households). The reduced premium for wealthy (*sick*) households is similar to an increase in income, increasing health care demand. The opposite will be true for poorer (*healthier*) households – they will demand insurance less when the premium is pooled across income (*health*) groups, and their premiums will increase resulting in reduced health care demand.

This added feature results in poorer (*healthier*) households subsidizing the health care of wealthier (*sicker*) households, thereby redistributing income from poorer to wealthier households and from healthier to sicker households. Relative to premiums priced according to household characteristics, the reduced income effect from a single pooled premium increases health care inequality among the insured. This suggests that a singled priced premium adversely affects the income and health care consumption of households that have lower health care expenditure. By a similar argument, a singled priced premium positively affects the income and health care consumption of households that have higher health care expenditure.

4. Estimation

4.1 Measuring Health Care Inequality

Two methods are used to measure the relationship between health care consumption and income: 1) a simple comparison of mean health care expenditures by households above and below the poverty line using ordinary least squares, and 2) the concentration curve and index for health care expenditure. In estimating the concentration curve and index, rather than using a binary indicator of poverty, households are ranked by income using the Income Rank variable. The health care consumption concentration index is defined as follows:

$$(2)\ CI = 1 - 2\int_0^1 L_m(x)dx$$

where $L_m(x)$ represents the corresponding concentration curve for health expenditure. The concentration curve plots the cumulative proportion of health care expenditure ($L_m(x)$) against the cumulative proportion of the population ranked by income (x); formally, $L_m(x) = \frac{\sum_{i=i}^{x} \lambda_i m_i}{\sum_{i=1}^{N} \lambda_i m_i}$, where λ_i is the number of households in income group i, m_i is the amount of health care expenditure in income group i, N is the total number of income groups, and i is ranked by income (determined by Income Rank). The standard errors of the concentration index are approximated using robust standard errors (O'Donnell et al. 2008).

4.2 Measuring Health Care Inequality: Sensitivity to Differential Enrollment and Price Changes

Depending on the observed relationship between income, W, and MHI demand, the relationship between health care consumption and income may differ between the insured versus the general population, even if the MHI itself has no effect on health care.

If insurance demand was only a function of W (and not θ), then a comparison of means would not be affected by the enrollment choice. This is because the means are not weighted by the number of observations in each income group. However, this is not the case for the concentration curve and index which is a function of the proportion of households in a given income group. Thus even though differential enrollment will not change the mean health care consumption of a given income group, it will change the proportional health expenditure of the income group. If θ influences MHI demand differentially by W, then a comparison of means will also be affected by enrollment. This is because the mean health care consumed in each income group now differs among the insured relative to the general population.

A change in the price of health care will reduce health care expenditures for all households, even if there is no actual change in the amount of health care consumed. This change in price will not affect concentration curve and index estimations because the measure is formulated using *proportional* health expenditure. However, a comparison of mean health expenditure will be affected. If ϑ is the original difference in means of health care expenditure by poverty status, then a price reduction of $(1-c)$ will reduce this difference to $c\vartheta$.

4.3 Initial Health Care Inequality and Insurance Demand

I observe control household enrollment decisions after the MHI is made available to all villages (from December 2012 to June 2013). Therefore, health care inequality in the MHI target population prior to the introduction of the insurance program can be documented using survey data from control villages (collected prior to December 2012). The comparison of means is estimated using the following equation:

$$(3)\ m_h = \alpha + \delta BPL_h + \varepsilon_h$$

where m is health care consumption, BPL is an indicator for whether the household is Below the Poverty Line, and h indicates the household. Both (3) and the corresponding concentration curve and index are estimated using health variables from the Household Health Survey and are limited to baseline households in the control villages.

The control households also provide identification for measuring the effect of poverty and health behaviors on the demand for health insurance. This is tested using a logit to predict the household's decision to enroll in the micro health insurance.

$$(4)\ Enrollment_{ht} = F\left(\gamma_1 BPL_{h,t-1} + \gamma_2 HealthCare_{h,t-1} + \gamma_3 BPL * HealthCare_{h,t-1} + \varepsilon_{ht}\right)$$

where $F(\cdot)$ is the logistic function, and ε is an error term; h indicates household and t indicates a time period of one year. Because 80 percent of SHG members were required to join the MHI for a household to be eligible to enroll in the insurance, estimates of $\hat{\boldsymbol{\gamma}}$ can be considered a lower bound (in magnitude) for completely voluntary MHI.

The estimates $\hat{\delta}$ and $\hat{\boldsymbol{\gamma}}$ are unbiased if households did not alter their health care or poverty status in anticipation of becoming insured. As mentioned earlier, it is unlikely that households were able (or would have an incentive) to change poverty variables in anticipation of the health insurance. It is similarly unlikely that households altered their health behavior. The timing of the MHI offer for a particular village was unknown, and on average, households did not begin coverage until 4 months after the Household Health Survey. This makes it unlikely that households were able to avoid illness or delay health care utilization until insured. This is particularly true given that the variables have a year recall period. If it were the case that households enrolled in anticipation of increased health care, one should expect to see greater claim submissions directly after enrollment. However, as Table 3.6 documents, the average length of time between a household's enrollment and first claim submission in treatment villages was 7 months. Furthermore, this duration does not differ by income. Nonetheless, if the type of household that enrolled had successfully delayed consuming health care, this would bias $\widehat{\gamma_2}$ downward. And if it were more likely for BPL households relative to non-BPL households to delay health care in anticipation of enrollment, this would bias $\widehat{\gamma_3}$ downward. Similarly, $\hat{\delta}$ would be a lower (*upper*) bound if the poor (*rich*) are less likely to enroll in the insurance, assuming the anticipation effect was similar across income groups.

To understand how health care inequality among the enrolled is affected by the choice to become insured, Equation (3) and the corresponding concentration index can be estimated separately for the sub-sample of control households who chose to enroll and those who did not:

$$(5)\ H_o: CI_{Enrolled} - CI_{NonEnrolled} = 0$$

CI is the health care expenditure concentration index, and only control households are used to estimate test Eq (5).

4.4 Heterogeneous Effects of Health Care

The panel data from the SHG Monthly Surveys provides monthly health care expenditure and MHI status. Using household and month fixed effects, I estimate heterogeneous effects of the MHI on health care consumption by poverty status:

$$(6)\ HealthCare_{ht} = \beta_1 insured_{ht} + \beta_2 BPL * insured_{ht} + \alpha_h + \tau_t + \eta_{ht}$$

where *insured* is either 1) an indicator of whether the household is covered under the MHI in the given month, or 2) the number of months the household has been insured to date.[20] BPL is an indicator for whether the household is Below the Poverty Line; h indicates the household and t indicates the month; η represent the error term.

Household fixed effects capture time-invariant differences between households and month fixed effects capture average health care consumption in the given month. Therefore, β_1 is estimated using the change in health care consumption after changing insurance status *within* a given household relative to the monthly average. β_2 is how this change differs for BPL households relative to non-BPL households.

Using comparisons within households helps protect against the concern that households that choose to enroll in the MHI are different from those that choose to not enroll. The underlying assumption for identification of β_2, the parameter of interest, is that among households that enrolled in MHI, BPL households do not have different health care time trends from non-BPL households (in the absence of MHI). However, identification does not require similar health care time trends among enrolled and non-enrolled households. One concern may be that households enroll at a time when they are most ill or foresee being ill. This will only bias $\widehat{\beta_2}$ if the anticipation effect on enrollment differs by poverty status. This seems unlikely as Table 3.6, Panel B, illustrates there is no difference in the average duration between enrollment and the first claim submission by poverty status. Furthermore, in this context, flexibility in the timing of enrollment is limited by the scarce staff resources of the MFI and the

[20] It is likely that in a low-income context where MHI programs are new, it takes time for households to learn and use the product. For this reason, one may expect that health care consumption increases do not necessarily occur immediately in the first month.

requirement of SHGs to enroll together. It is therefore unlikely that households were able to time their enrollment based on changing health needs.

In addition to household fixed effects and panel data, I use a difference-in-difference technique to control for the endogenous choice of enrollment. Measures of health care inequality can be estimated in both treatment and control villages among households that chose to enroll and those that did not enroll:

$$(7)\ H_0: \left(CI_{Treatment_{Enrolled}} - CI_{Treatment_{Non-enrolled}}\right) - \left(CI_{Control_{Enrolled}} - CI_{Control_{Non-enrolled}}\right) = 0$$

where CI is the health care expenditure concentration index, the preferred measure of health care inequality because it is unaffected by changes in price. The first term is the difference in health care inequality between the enrolled and non-enrolled in treatment villages and reflects both the effect of selection into MHI and the effect of MHI. In contrast, the second term is the difference in health care inequality between the enrolled and non-enrolled in control villages and reflects only the effect of selection. This is because the health care expenditure is a post-MHI measure for treatment households and a pre-MHI measure for control households. Therefore, subtracting the second term from the first isolates the effect of MHI on health care inequality.

The key identification assumption for Eq (6) is the enrollment decision in treatment and control villages is comparable and that non-enrolled households' health care in treatment villages was unaffected MHI. The similar enrollment rates documented in Table 3.2 is suggestive that the enrollments between the two villages are similar. Section V also provides suggestive evidence of Eq (7) estimates being relatively robust to the small differences observed in enrollment patterns.

5. Results

5.1 Health Care Inequality of Target Population

Table 3.7 estimates the differences in health care utilization by poverty status for control households, the households not offered the MHI at the time of the Household Health Survey. Panel A reports mean differences estimated by OLS and Panel B estimates the log likelihood ratio. The estimates indicate no statistically significant difference in health shock incidents as proxied by having a household member admitted or experiencing a health event that resulted in expenditure of over INR 1,000 (USD 20) in the past year. These results suggest that households Below the Poverty Line (BPL) are not seeking health care at lower rates than non-BPL households.[21]

Unlike Table 3.7, Table 3.8 suggests large differences in health care consumption by poverty status prior to MHI.[22] BPL households spend approximately INR 7,500 (USD 150) less on health care annually, a difference of over 30%. Based on the estimates from Table 3.7, these differences are unlikely to be from the extensive margin of seeking health care, but rather come from the intensive margin of the quality of health care sought. Table 3.7 additionally suggests the inequality observed in Table 3.8 is not likely to be due to non-BPL households having greater health needs. Figure 3.1 depicts this relationship using the finer measure of Income Rank. Because health care expenditure is able to capture these differences more than the health care utilization indicators, the remainder of this paper focuses on health care expenditures as the variable of interest and analogous estimates for the utilization variables are presented as Appendix Tables.

Table 3.9 quantifies these differences by estimating the concentration index and yields a statistically significant estimate of .15, suggesting that 11% of health care consumption should be redistributed (linearly) from wealthier to poorer households for health care equality (O'Donnell et al. 2008). Figure 3.2 depicts the underlying concentration curve and confirms that at almost every income rank the proportional health care expenditure falls below the line of equality. The health care expenditure concentration curve in Figure 3.2 is statistically dominated by the 45˚ line of equality.[23] Recall, if poorer

[21] However, it may be the case that poorer households are more ill. If so, then a lack of differences in health care utilization would suggest the poor are foregoing treatment at higher rates than wealthier households, conditional on health need.
[22] Similar health care expenditure inequality is found among enrolled treatment households in the first months of the SHG Monthly Surveys (Table A.3.1) and among non-enrolled treatment households.
[23] All concentration curve dominance tests used in this paper's analysis uses the multiple comparison approach decision rule at 19 quantile points and a 5 percent statistical level (O'Donnell et al. 2008). I am unable to reject the null hypothesis of non-dominance for all concentration curves (against the 45˚line of equality) using the alternative intersection union principle (O'Donnell et al. 2008).

households have a higher burdens of illness then these estimates are lower bounds of health care inequity (i.e., horizontal inequity).

5.2 Demand for MHI

Table 3.10 estimates the likelihood of becoming enrolled in the control villages within 7 months of the MHI becoming available. BPL households are 13 percentage points less likely to enroll with an estimated .58 odds ratio of enrollment. As Column 2 indicates, this result is robust even when controlling for the previous year's health care expenditure. Though this result is somewhat surprising given that poorer households are commonly believed to be more risk averse, it is similar to the uptake of other micro insurance products in low income countries without a long history of insurance (Dercon et al. 2008).

Column 2 highlights that health expenditure in the past year is not a determinant of enrollment. This surprising result suggests that MHI is not demanded by households who have the greatest demand for health care. Alternatively, it may be the case that the recall period does not accurately account for current health status, either because incidents are too long ago or more representative of larger idiosyncratic health shocks than general health status. Column 3 suggests that poorer households may be more responsive to poor health in their demand for MHI, but the differences are small in magnitude and statistically insignificant.[24]

As expected from the enrollment patterns estimated in Table 3.10, Table 3.11 confirms that health care inequality by poverty status prior to MHI is similar across the type of households that choose to enroll and the type of households that choose not enroll. Table 3.12 also finds no difference in the concentration index of enrolled versus not enrolled households. Figure 3.3 confirms that the concentration curves for the enrolled and not enrolled are similar; Both curves are dominated by the 45˚ line of equality,

[24] Using health care expenditure in the past week, as opposed to the past year, does suggest that households are sensitive to health status when enrolling in MHI. Additionally, Table A.3.2 suggests that BPL households are more likely to enroll conditional upon having sought significant health care in the past year than non-BPL households. However, this result is not robust to health care consumption.

but neither curve dominates the other. This suggests that the endogenous choice of enrolling in MHI may not affect measures of health care inequality when comparing insured versus uninsured households.

5.3 Heterogeneous Effects on Health Care Consumption

Table 3.13 estimates the heterogeneous effect of the MHI on health care consumption by poverty status. Controlling for seasonal trends by including month fixed effects, and differences between households by including household fixed effects, the estimates indicate that MHI increases health care consumption more for BPL households. This suggests that conditional upon enrollment, MHI reduces health care inequality as measured by differences in health care consumption.

Panel A estimates the difference in health care expenditure as a function of months insured by MHI by poverty status. Column 1 first limits the identifying sample to households that change their MHI status during the panel period (October 2011 – July 2012). I find that BPL households have INR 143 (approximately USD 3) higher health expenditure per month of MHI coverage relative to non-BPL households. Accordingly, the health expenditure gap would decrease by approximately INR 3,003 (USD 60) after half a year of coverage. It is likely that this differential in health care expenditure by poverty status would decrease over time, but the limited time period of the panel makes it difficult to identify the long term curvature of the change in health expenditure. Column (2) expands the identifying sample to include those households that enrolled within 5 months prior to the start of the panel period, and Column (3) expands it to include all households that enrolled prior to July 2012. The estimated $\hat{\beta}_2$ slightly decreases in magnitude across each additional column, consistent with differential marginal gains in health care for an additional month of MHI coverage in the earlier, but not later, months of enrollment. These results should be interpreted as the short term effects of MHI on the change in health care consumption by poverty status.

Panel B estimates a similar regression using the independent variable of whether or not the household is insured in the given month. Though the estimates lack precision, the point estimates suggest a similar pattern to Panel A. Because identification is based only using the subset of households which changed MHI enrollment status during the panel period, Column (2) and Column (3) are intentionally not estimated.

Due to the majority of households in the identifying sample enrolling within the first three months of the panel, it is not possible to test the identifying assumption of similar time trends in health care prior to MHI coverage. If the underlying trend in enrollment is driving the estimates of β_2 in Table 3.13, then the pattern of enrollment in control households should lead to similar results. In Table A.3.3, I estimate the same equation as Table 3.13, Panel A, on the subsample of control households using the order of the enrollment in control households as a measure of months enrolled. Estimates of β_2 in Table A.3.3 reflect only health care expenditure trends by poverty status (among the type of households which demand MHI) because the health expenditure measures are collected prior to actual enrollment of control households. Unfortunately, the estimates of β_2 from Table A.3.3 are so imprecise that the standard errors are close to the estimated parameters in Table 3.13 and the test is not very informative. Though the point estimates suggest a higher time trend among BPL households, the lack of change from Column (1) to Column (2) suggest no differences in early versus later enrollers, unlike that observed among treatment households in Table 3.13.

Because health care expenditure may be a function of discounted prices negotiated through the insurance, it is not possible to interpret the difference in health expenditure between insured and uninsured households as differences in health care consumption. Therefore, estimates of β_1 are potentially confounded by the discounted price of health care. Because wealthier households are estimated to have higher health expenditure initially, discounted health prices through MHI will provide greater savings in expenditure to higher income households. As a result, part of the estimate of β_2 may be due to BPL households having fewer savings from their lower initial health care consumption. If $\hat{\beta}_2$ was entirely due to this difference in savings, its magnitude should be proportionate to the gap in health care consumption by poverty status: $\hat{\beta}_2 = .34\hat{\beta}_1$, as suggested by Table 3.8. This is tested in Table 3.13, Panel

A, and the null hypothesis is rejected at the 5 percent level for Column (1). Though the p-value is relatively low, the estimates of Column (2) and Column (3) are unable to reject the null hypothesis.

There are a variety of reasons that poorer households may be increasing their overall health care expenditure upon enrolling in MHI. First, the health variables are out-of-pocket (OOP) health expenditures and thus do not deduct reimbursements awarded through the claims process. However, there are additional reasons that households may increase their health care consumption by more than the expected claim reimbursements. For example, MHI may reduce the cost of seeking higher quality health care by creating a network of empanelled facilities or induce health care demand by doctor visits to the village or increased saliency of health. Though the results potentially suggest a high price elasticity, similar results of increased OOP expenditure were found in China's health insurance program (Wagstaff and Lindelow 2008).

Table 3.14[25], Column (1), quantifies the composite effect of *both* selection and the effect of MHI on health care inequality. Among insured households in treatment villages, there is no statistically significant difference in health care consumption by poverty status, with the proportion of health expenditure by BPL households being only 5 percent less than non-BPL households. This is starkly contrasted to Column (2) which quantifies the difference in health care expenditure among the uninsured in treatment villages. Among households which chose not to enroll in the MHI, BPL households report spending less than half of non-BPL households.

Table 3.15 quantifies the reduced health care inequality in treatment villages using the concentration index. Among the insured, the estimated concentration index is .02 (statistically insignificant), again suggesting no relationship between income and health care consumption. However, among the uninsured the concentration index is .254, even higher that the inequality estimated among those not enrolled in control villages. The difference between these concentration indices is .232 and is statistically significant at the 1 percent level. Figure 3.4 highlights this stark difference by graphing the

[25] Table A.3.4 reports analogous estimations for health care utilization.

health care expenditure concentration curve for the insured and uninsured in treatment villages. Consistent with Table 3.15, the concentration curve for the insured in treatment villages follows the line of equality, with some points even above the line. However, the uninsured fall well below the line of equality at every income ranking. The concentration curve of the uninsured is dominated by both the 45° line of equality and the concentration curve of the insured (neither of which dominates the other).

Table 3.11 and Table 3.12 suggested that endogenous selection of enrollment was unlikely to result in different estimates of health care inequality among the insured and uninsured. Indeed, the difference-in-difference outlined in Eq (7) yields a reduction in health inequality by .236, statistically significant at the 5 percent level, when differencing out the enrollment effect ($.0220 - .254$ – $(.159 - .155)$).

Figure 3.4 highlights this by overlaying the concentration curves for both uninsured and insured from control households. Recall, for control households these are health care expenditures prior to actually being enrolled in the MHI, whereas for treatment households the expenditures reflect health care consumption after being covered by MHI. This suggests that the reduction in health care inequality among the insured in treatment villages is more likely a result of heterogeneous effects of the MHI on health care by income rather than differential enrollment decisions.

Given that treatment village enrollment is double the enrollment in control villages, it may be that later enrollment into MHI is driving the reduction in health care inequality among the insured. As a robustness check, Table 3.16 estimates the relationship between health care expenditure and poverty among treatment households who became enrolled in the MHI within the first 7 months of the offer, the same length of time for which I observe control households' enrollment choices. After 7 months, treatment villages had an enrollment rate of only 30%, nearly identical to that in control villages. Table 3.16, Panel A, documents that among these early enrollers in treatment villages, BPL households have *higher* health care expenditures than non-BPL households. Panel B estimates a concentration index of -.01. Figure 3.5 corroborates the estimates in Table 3.16 by graphically depicting the concentration curve of early enrollers in treatment villages. As before, this concentration curve dominates that of non-enrollers in treatment villages and is not dominated by the 45° line of equality. This suggests that the reduced health care consumption inequality documented among insured treatment households is not due to the type of household which enrolled a few months after the initial offer of MHI.

Though enrollment rates are similar across treatment and control villages after the same length of exposure to MHI, it may still be the case that the underlying factors of MHI demand differ between the two. As Table 3.17 depicts, among the first 7 months of exposure, the proportion of BPL households choosing to enroll was 9.5 percentage points less in treatment villages relative to control villages. I estimate upper and lower bounds of the estimates in Table 3.16, Panel A, assuming that 9.5 percentage points more BPL households had enrolled early in treatment villages. The upper (*lower)* bound assumes early enrollment of BPL households who reported the highest (*lowest*) health care expenditures (and were not actually enrolled in the first 7 months). Even assuming that those BPL households with the lowest health expenditures had enrolled earlier, the gap increases to only 13% and remains statistically insignificant. The estimated bounds of the concentration index of Panel B are larger in range, but both bounds remain statistically insignificant and depict lower health care inequality than estimated among the non-enrolled. This suggests that the elimination of the gap in health care inequality conditional upon enrollment in MHI is relatively robust to differential proportional enrollment by poverty status between treatment and control villages.

6. Conclusion

This paper suggests that the concern of MHI leading to regressive transfers from poor to relatively wealthier households is lower than expected from the health care distribution by income prior to MHI. This is most likely due to heterogeneous effects of MHI on health care consumption by income.

Prior to MHI, this paper documents significantly lower health care consumption among households Below the Poverty Line. This suggests significant health care inequalities even among households targeted by MHI, and that MHI may lead to poorer households cross-subsidizing the health expenditure of relatively wealthier households. However, after twenty months of exposure to MHI, there is no difference in household health care consumption by poverty status. The results in this paper provide suggestive evidence that this reduction is primarily due to heterogeneous effects on health care by poverty status (as opposed to sorting into MHI). I find no evidence of differential enrollment reducing health care inequality among the insured, but I do find support for greater impact of MHI on increased health care consumption among poorer households, conditional upon enrollment.

The lack of health care inequality among the insured suggests that, all else equal, the redistribution of premiums to households through MHI's claim process will not result in a transfer away from poorer households. Though *ex-ante* health care consumption raised concerns that MHI could increase income inequality, the *ex-post* relationship between health care and income indicates no expected redistribution of premiums by poverty status. In general, these results suggest that MHI reduces health care inequality by improving access to health care most for the poor, which in turn diminishes the concern that the underlying health care distribution will lead to poorer households subsidizing wealthier households.

To the extent that MHI may not always eliminate health care consumption inequality, consideration should be given to reducing the price of premiums for poorer households to be more actuarially fair in such cases. Not only will this protect against cross-subsidization of wealthier households by poorer households, it will likely also increase enrollment rates among poorer households. However, greater understanding of households' willingness to pay for MHI is required to ensure there such changes in cost-benefit ratio would not lead to reduced enrollments of wealthier households and threaten the financial viability of the risk pool. As this paper illustrates, using *ex-ante* health care behavior will underestimate premium amounts unless heterogeneous MHI effects on health care consumption are taken into consideration.

Though this analysis is unable to comment on health care inequality in the entire target population, the results suggest that an added benefit of greater inclusion of the poor is the reduction of health care inequalities in the larger population. Without greater inclusion of poorer households, even if MHI increases health care consumption more among poorer households conditional upon enrollment, it could still be the case that it increases health care inequality among the entire population. Additionally, further research using measures of health outcomes would further inform whether the heterogeneous effects of health care consumption translate into greater improvements in health. For example, a reduction in health care consumption inequality is also consistent with a reduction in the health burden of wealthier households. Thus, to fully understand whether reduced health care consumption inequality also reduced horizontal inequity, greater research is required on outcomes of health.

The increased consumption of health care also brings into question the financial sustainability of MHI programs. Whether these programs are able to sustain such increases is still unknown. If the increases in health care are highest for poorer households, this reduces the expected benefit for wealthier households and may lead to lower enrollment rates among wealthier households. The MHI studied in this paper has been able to cover claim reimbursements from collected premiums for over three years. However, further research into how changing health care and health status affects premium requirements and enrollment patterns would help inform long run sustainability of MHI programs.

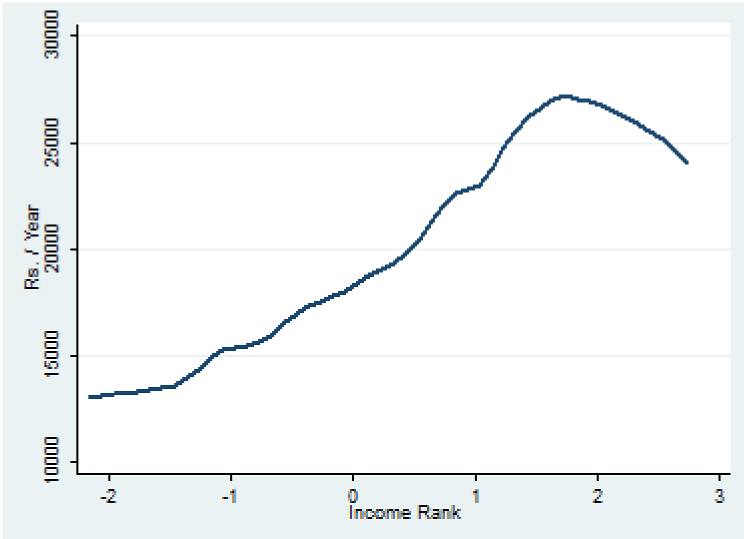

Figure 3.1: Yearly Health Care Expenditure and Income (Control Households)

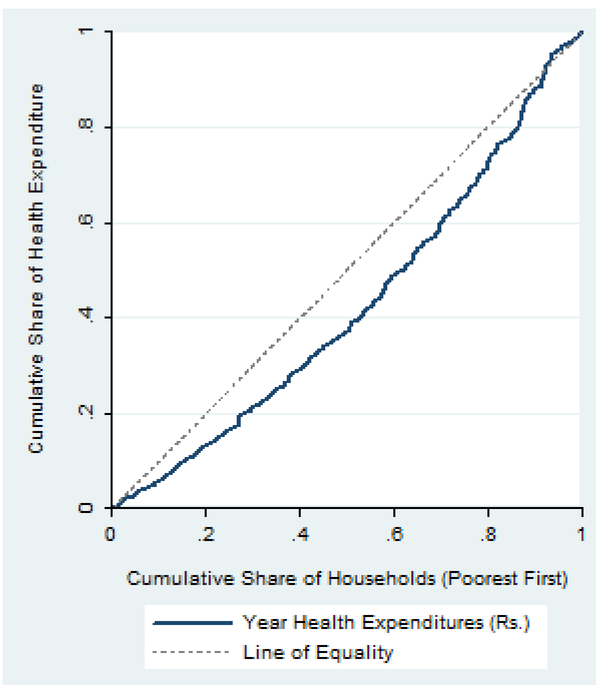

Figure 3.2: Yearly Health Care Expenditure Concentration Index (Control Households)

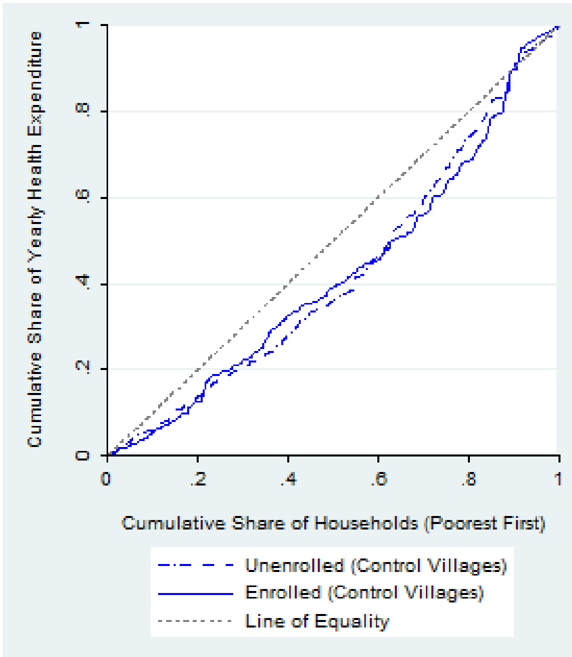

Figure 3.3: Yearly Health Care Expenditure Concentration Index by Enrollment Status (Control Households)

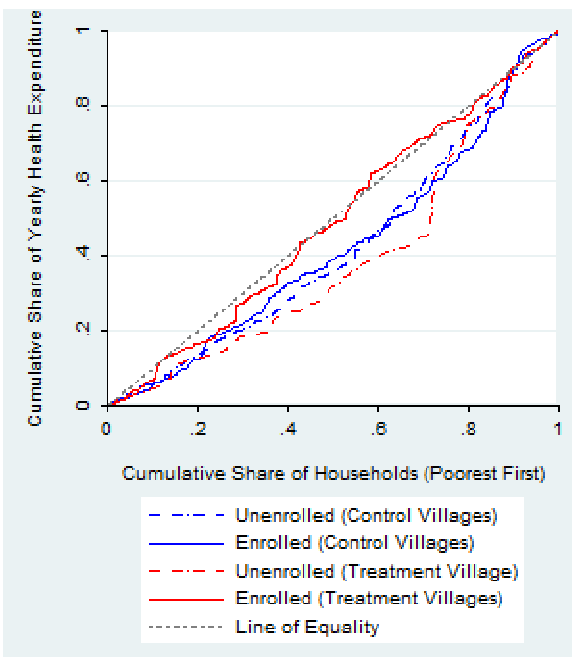

Figure 3.4: Yearly Health Care Expenditure Concentration Index by Enrollment Status

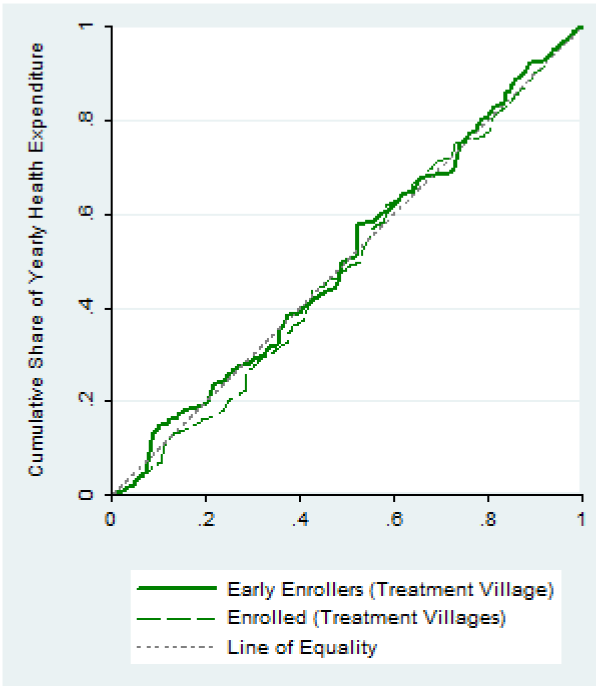

Figure 3.5: Yearly Health Care Expenditure Concentration Index of Early Insured (Treatment Households)

Table 3.1: Demographic Summary Statistics

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	All Households		Treatment Households		Control Households		Treatment-Control
	Obs (HH)	Mean	Obs (HH)	Mean	Obs (HH)	Mean	
Below the Poverty Line (BPL)	*1,615*	*0.411*	*710*	*0.358*	*905*	*0.462*	*-0.104*
Poverty Status (Ration Card)	1,615	2.419	710	2.360	905	2.477	-0.117
House Type	1,675	1.635	744	1.555	931	1.716	-0.162*
Agricultural Laborer (in past year)	1,687	0.670	749	0.639	938	0.701	-0.0622
Disadvantaged Caste	1,594	0.533	707	0.365	887	0.703	-0.338***
Agricultural Cultivator	1,682	0.768	748	0.768	934	0.767	0.00118
Laborer (Primary Occupation)	1,682	0.147	748	0.147	934	0.148	-0.00140
Income Rank	*1,509*	*0.298*	*665*	*0.628*	*844*	*-0.0298*	*0.658**

Observations are limited to households selected and surveyed in the Household Health Survey and are weighted to be representative of the target population.

All variables are indicators, except for Poverty Status (1 to 4, increasing with poverty), and House Type (1 to 3, increasing with worse infrastructure); Income Rank is a variable created using the bolded variables (with an indicator for each category) through principal component analysis and ranges from -2.170245 to 2.72455, increasing with income.

Column (7) reports standard errors clustered at the village level.

Statistical significance levels are as follows: *10%, **5%, ***1%.

Table 3.2: Enrollment Summary Statistics

	(1)	(2)	(3)	(4)	(5)	(6)
	All Households		Treatment Households		Control Households	
Households Enrolled as of:	Obs	Percent	Obs	Percent	Obs	Percent
October 2012	2,625	30%	1,311	61%	1,314	0%
(Household Health Survey)						
June 2013	2,625	47%	1,311	62%	1,314	33%
(7 months after Initial MHI Offer to Control HH)						
7 months after Initial MHI Offer	2,625	31%	1,311	30%	1,314	33%

Enrolled: Indicator for whether the household had at least one member insured for at least one year.

Table 3.3: Enrolled Households

	(1) Mean	(2) SD	(3) Min	(4) Max
HH Members Enrolled	1.932	1.40	1	8
Submitted Claim	0.0977	0.30	0	1
Claim Disbursement (*INR per household*)	253.3	1,087	0	9,985
	(USD 5)	(USD 22)	(USD 0)	(USD 200)

Summary statistics are given for the subset of households which had at least one household member insured for at least one year (798 households) as of October 2012.

Table 3.4: Claim Summary Statistics

	(1) Claimable Expense	(2) Disbursed Amount
Mean Amount (INR)	5,533 (USD 111)	2,911 (USD 58)
25% (15 cases) Malaria Claims (INR)	5,161 (USD 103)	2,371 (USD 47)
42 % (25 cases) Enteric Fever/Typhoid Claims (INR)	4,812 (USD 96)	2,863 (USD 57)
Amount Disbursed/Claimable Expense > .60: 75%		
Amount Disbursed/Claimable Expense > .50: 92%		

Summary statistics are based on 62 claim cases; Claims for which the claimable expense exceeding INR 25,000 are excluded (3 claims: mean claimable expendable INR 45,135; mean disbursed amount INR 8,662).

Observations limited to accepted and settled claims of treatment households enrolled as October 2012.

Table 3.5: Health Summary Statistics

	(1)	(2)	(3)	(4)	(5)
	Obs	Mean	SD	Min	Max
	Panel A: Month Recall				
Visited Health Facility	12,797	0.347	0.476	0	1
Admitted	12,755	0.032	0.177	0	1
Health Expenditure (INR)	10,852	604	3,015	0	52,775
		(USD 12)	(USD 60)	(USD 0)	USD (1,055)
	Panel B: Year Recall				
Admitted	1,685	0.486	0.500	0	1
Health Shock Over INR 1,000 (USD 20)	1,690	0.515	0.500	0	1
Health Expenditure (INR)	1,677	16,055	25,079	300	170,450
		(USD 321)	(USD 502)	(USD 6)	(USD 1,136)

Month Recall variables are sourced from SHG Monthly Health Surveys; Year Recall are sourced from Household Health Survey and weighted to be representative of the the target population.

Visited Health Facility, Admitted, and Health Shock Over INR 1,000 are indicators for whether any household member experienced such an event; Health expenditures are total expenses of the household and are windsored at the 99% for year recall and 99.9% for month recall.

Unit of observation is household for year recall, and household-month for month recall.

Table 3.6: Duration Between Enrollment and First Claim (Conditional upon Claim Submission)

	Panel A: Summary Statistics				
	Obs (HH)	Mean	SD	Min	Max
Duration Between Enrollment and First Claim (Months)	64	7.39	4.27	0	18

	Panel B: Duration by Poverty Status
Dependent Variable:	Duration Between Enrollment and First Claim (Months)
Below the Poverty Line (BPL)	0.459
	(2.047)
Constant	8.472***
	(0.961)
Obs (HH)	31

Observations are limited to treatment households which submitted a claim as of October 2012.

Observations in Panel B are limited to treatment households selected and surveyed in the Household Health Survey and are weighted to be representative of the target population.

Enrollment date refers to the initial enrollment of the household.

BPL is an indicator for whether the household is Below the Poverty Line.

Standard errors are in parentheses and are robust standard errors.

Statistical significance levels are as follows: *10%, **5%, ***1%.

Table 3.7: Poverty and Health Care Utilization, Year Recall (Control Households)

	(1)	(2)
Dependent Variable:	Admitted	Health Shock Over INR 1,000 (USD 20)
	Panel A: OLS	
Below the Poverty Line (BPL)	0.0178	0.0255
	(0.0311)	(0.0330)
Constant	0.531***	0.554***
	(0.0255)	(0.0250)
Obs (HH)	904	905
	Panel B: Logit (Odds Ratio)	
Below the Poverty Line (BPL)	1.074	1.109
	(0.135)	(0.149)
Obs (HH)	904	905

Observations are limited to control households selected and surveyed in the Household Health Survey and are weighted to be representative of the target population.

BPL is an indicator for whether the household is Below the Poverty Line.

Standard errors are in parentheses and are clustered at the household's Self Help Group (SHG).

Statistical significance levels are as follows: *10%, **5%, ***1%.

Table 3.8: Poverty and Health Care Expenditure (Control Households)

Dependent Variable: Yearly Health Expenditure (INR)	
Below the Poverty Line (BPL)	-7559.6***
	(2059.6)
Constant	21965.8***
	(1835.1)
Obs (HH)	897
BPL Fractional Expenditure	0.66

Observations are limited to control households selected and surveyed in the Household Health Survey and are weighted to be representative of the target population.

BPL is an indicator for whether the household is Below the Poverty Line.

Standard errors are in parentheses and are clustered at the household's Self Help Group (SHG).

Statistical significance levels are as follows: *10%, **5%, ***1%.

Table 3.9: Poverty and Health Care Consumption (Control Households)

Concentration Index	0.154***
(Yearly Health Expenditure)	(0.0337)
Obs (HH)	839

Observations are limited to control households selected and surveyed in the Household Health Survey and are weighted to be representative of the target population.

Standard errors are in parentheses and are approximated using robust errors (O'Donnell et al. 2008).

Statistical significance levels are as follows: *10%, **5%, ***1%.

Table 3.10: Demand for MHI (Control Households)

	(1)	(2)	(3)
Dependent Variable:		Household Enrolled in MHI	
Below the Poverty Line (BPL)	0.584**	0.586**	0.581*
	(0.158)	(0.159)	(0.179)
Yearly Health Expenditure		0.999	0.999
		(0.00313)	(0.00359)
BPL * Yearly Health Expenditure			1.001
			(0.00716)
Obs (HH)	905	897	897

Table reports odds ratio from logistic regression.

Observations are limited to control households selected and surveyed in the Household Health Survey and are weighted to be representative of the target population.

BPL is an indicator for whether the household is Below the Poverty Line; Yearly Health Expenditure is reported in INR 1,000.

Standard errors are in parentheses and are clustered at the household's Self Help Group (SHG).

Statistical significance levels are as follows: *10%, **5%, ***1%.

Table 3.11: Poverty and Health Care Expenditure by Enrollment Status

	(1)	(2)
Dependent Variable:	Yearly Health Expenditure (INR)	
	Enrolled Households	Non-Enrolled Households
Below the Poverty Line (BPL)	-7203.5**	-7870.7***
	(3233.3)	(2942.0)
Constant	21579.7***	22291.9***
	(2503.0)	(2781.1)
Obs (HH)	336	561
BPL Fractional Expenditure	0.666	0.647

Observations are limited to control households selected and surveyed in the Household Health Survey and are weighted to be representative of the target population.

BPL is an indicator for whether the household is Below the Poverty Line.

Standard errors are in parentheses and are clustered at the household's Self Help Group (SHG).

Statistical significance levels are as follows: *10%, **5%, ***1%.

Table 3.12: Poverty and Health Care Expenditure by Enrollment

	(1) Enrolled Households	(2) Non-Enrolled Households
Concentration Index	0.159**	0.155***
(Yearly Health Expenditure)	(0.0627)	(0.0367)
Obs (HH)	308	531

Observations are limited to control households selected and surveyed in the Household Health Survey and are weighted to be representative of the target population.

Households which enrolled at least one member for at least one year are considered to be Enrolled Households.

Standard errors are in parentheses and are approximated using robust errors (O'Donnell et al. 2008).

Statistical significance levels are as follows: *10%, **5%, ***1%.

Table 3.13: Heterogeneous Effect of Insurance by Poverty Status

	(1)	(2)	(3)
Dependent Variable:	Monthly Health Expenditure (INR)		
	Panel A: Months of Coverage		
Months Insured (β_1)	-73.01	-48.48	-22.38
	(45.36)	(39.92)	(36.93)
Months Insured * Below the Poverty Line (BPL) (β_2)	**142.8*** **	**96.75***	**81.22**
	(53.48)	(52.44)	(57.60)
Obs (HH Month)	5945	6691	7160
$\beta_2 + .34*\beta_1$	117.98	81.63	73.61
p-value (H_0: $\beta_2 + .34*\beta_1 = 0$)	0.0129	0.1025	0.1868
	Panel B: Change in MHI Status		
Insured (β_1)	-174.2		
	(238.6)		
Insured * Below the Poverty Line (BPL) (β_2)	**300.6**		
	(288.9)		
Obs (HH Month)	5945		
Treatment HH: Changed Enrollment Status (Oct 2011 - Jul 2012)	Yes	Yes	Yes
Treatment HH: Enrolled Prior to Panel Jun - Oct 2011)	No	Yes	Yes
Treatment HH: Enrolled Prior to Panel Feb - Oct 2011)	No	No	Yes

Regressions include household and month fixed effects.

Observations are weighted to be representative of the target population.

Months Insured are the number of months the household has had coverage from January 2011 to the current month; Insured is an indicator for whether any household member is insured in the current month.

BPL is an indicator for whether the household is Below the Poverty Line.

Standard errors are in parentheses and are clustered at the household level.

Statistical significance levels are as follows: *10%, **5%, ***1%.

Table 3.14: Poverty and Health Care Expenditure by Insurance Status (Treatment Households)

	(1)	(2)
Dependent Variable:	Yearly Health Expenditure (INR)	
	Enrolled Households	Non-Enrolled Households
Below the Poverty Line (BPL)	-754.3	-11153.5**
	(2527.0)	(4359.6)
Constant	14097.5***	19453.3***
	(1268.1)	(4332.8)
Obs (HH)	433	270
BPL Fractional Expenditure	0.95	0.43

Observations are limited to treatment households selected and surveyed in the Household Health Survey and are weighted to be representative of the target population.

BPL is an indicator for whether the household is Below the Poverty Line.

Standard errors are in parentheses and are clustered at the household's Self Help Group (SHG).

Statistical significance levels are as follows: *10%, **5%, ***1%.

Table 3.15: Poverty and Health Care Expenditure by Insurance Status (Treatment Households)

	(1) Enrolled Households	(2) Non-Enrolled Households
Concentration Index	0.0220	0.254***
(Yearly Health Expenditure)	(0.0469)	(0.0918)
Obs (HH)	403	255

Observations are limited to treatment households selected and surveyed in the Household Health Survey and are weighted to be representative of the target population.

Households which enrolled at least one member for at least one year are considered to be Enrolled Households.

Standard errors are in parentheses and are approximated using robust errors (O'Donnell et al. 2008).

Statistical significance levels are as follows: *10%, **5%, ***1%.

Table 3.16: Poverty and Health Care Expenditure of Early Enrollers (Treatment Households)

Dependent Variable:	Yearly Health Expenditure (INR)
	Panel A: OLS
Below the Poverty Line (BPL)	2819.3
	(4743.0)
Constant	12897.1***
	(1651.0)
Obs (HH)	169
Upper Bound	*11719.8****
	(4336.5)
Lower Bound	*-1740.7*
	(3588.3)
	Panel B: Concentration Index
Concentration Index	-0.0148
(Yearly Health Expenditure)	(0.0720)
Obs (HH)	166
Upper Bound	*-0.0731*
	(0.0659)
Lower Bound	*0.0763*
	(0.0863)

Upper and lower bounds are estimated by assuming an additional 9.5 percentage points of BPL households are enrolled.

Observations are limited to treatment households selected and surveyed in the Household Health Survey and are weighted to be representative of the target population.

Households which enrolled at least one member for at least one year in the first 7 months of the MHI offer are considered Early Enrollers.

BPL is an indicator for whether the household is Below the Poverty Line.

Standard errors are in parentheses and are clustered at the household's Self Help Group (SHG) for Panel A and are robust standard errors for Panel B (O'Donnell et al. 2008).

Statistical significance levels are as follows: *10%, **5%, ***1%.

Table 3.17: Enrollment Patterns (Early Enrollers)

	(1)	(2)	(3)
	Percent Enrolled:		
	Below the Poverty Line	Non-Below the Poverty Line	Total
Control Villages	33%	46%	40%
Treatment Villages	21%	43%	35%

These enrollment rates are limited to the sub-sample from which the Household Health Survey data is collected and are weighted to be representative of the target population; This accounts for the difference among these rates versus those for the entire population documented in Table 2a.

Chapter 3 Appendix: Theoretical Framework

Health Care Consumption and Income

The solution to the maximization problem, $\max_{N,m} U(N, m, \theta, s)\ s.t. N + P_m m \le W$, is characterized by the following:

$$(1)\ P_m U_1(W - P_m m^*(W, P_m, \theta, s), m^*(W, P_m, \theta, s), \theta, s) - U_2(W - P_m m^*(W, P_m, \theta, s), m^*(W, P_m, \theta, s), \theta, s) \begin{cases} < 0\ if m^* = W \\ = 0\ if\ 0 < m^* < W \\ > 0\ if\ m^* = 0 \end{cases}$$

Assuming an interior solution, this implies that health care consumption increases with wealth:

$$(2)\ \frac{dm^*(W, P_m, \theta, s)}{dW} = \frac{P_m U_{11}}{P_m^2 U_{11} + U_{22}} > 0$$

It can be similarly shown that health care demand increases with worse health:

$$(3)\ \frac{dm^*}{d\theta} = -\frac{U_{23}}{P_m^2 U_{11} + U_{22}} > 0$$

$$(4)\ \frac{dm^*}{ds} = -\frac{U_{24}}{P_m^2 U_{11} + U_{22}} > 0$$

If poorer households also have greater health needs ($\theta_1 = \frac{d\theta}{dW} < 0$), then the tension between less income and higher health burden leads to an ambiguous relationship between health care consumption and income:

$$(5)\ \frac{dm^*(W, P_m, \theta, s)}{dW} = \frac{\partial m^*}{\partial W} + \frac{\partial m^*}{\partial \theta}\frac{\partial \theta}{\partial W} = \frac{P_m U_{11}}{U_{11} P_m^2 + U_{22}} - \frac{U_{23}\theta_1}{U_{11} P_m^2 + U_{22}}$$

It may be the case that there exists an $\overline{M}(\theta, s)$ such that $\frac{dU(N,m,\theta,s)}{dm} = 0|_{m \ge \overline{M}(\theta,s)}$. If $\frac{W}{P_m} \ge \overline{M}(\theta, s)$, it may be the case that in certain states of health all households are consuming $\overline{M}(\theta, s)$ and there is no difference in health care consumption by income.

Insurance with Predetermined Indemnity Schedule

If households can commit to pre-specified health care consumption for a given health shock, the optimal level of health care for each state s, m_s, is characterized by the following utility maximization problem:

$$(6)\ \max_{N,m} EU(N, m, \theta, s) = \max_{m_s} \sum_s p_s U(W - P_m \sum_k p_k m_k, m_s, \theta, s)$$

where p_s is the probability of state s,and households pay an actuarially fair premium (i.e., a premium equal to their expected health care expenditure). The general condition that characterizes the optimal m_s^*, is given by:[1,2]

$$(7)\ P_m U_1\left(W - P_m \sum_k p_k m_k\,, m_s, \theta, s\right) = \ P_m U_{1_{insured}} = U_{2_{insured}}$$

$$= U_2\left(W - P_m \sum_k p_k m_k\,, m_s, \theta, s\right)\ \forall\, s \in S$$

Note that this is an analogous optimization condition to Eq (1), but the relationship between m and N now differs. Rather than equate the relative marginal utility of N and m in a given state s, they are equated across all states of s. Condition (7) suggests that under such insurance, households will generally reduce m_s^* in healthier states and increase m_s^* in sicker states relative to when uninsured.[3] In other words, $\frac{\partial(m^*_{insured} - m^*_{uninsured})}{\partial s} \geq 0$.

All households prefer to be insured. Expected utility (*ex-ante* to the realization of s) is unambiguously higher when insured due to the utility function's concavity. If the utility function is rewritten to depict the additive separability, $U(N, m, \theta, s) = U_N(N) + U_m(m, \theta, s)$, then for a given level of health care consumption, $\bar{m}$, concavity yields the following to be true:

$$(8)\ EU_{uninsured}(\bar{N}, \bar{m}, \theta, s) = \sum_s p_s U_N(W - P_m \bar{m}_s) + \sum_s p_s U_m(\bar{m}_s, \theta, s)$$

$$< U_N\left(W - P_m \sum_s p_s\, \bar{m}_s\right) + \sum_s p_s U_m(\bar{m}_s, \theta, s) = EU_{insured}(\bar{N}, \bar{m}, \theta, s)$$

It must then be the case that $EU_{insured}(\bar{N}, \bar{m}, \theta, s)$ is not greater than the expected utility when households optimally choose the level of health care under insurance $(EU_{insured}(\bar{N}, \bar{m}, \theta, s) < EU_{insured}(N^*, m^*, \theta, s)$.

Health care consumption continues to rise with income assuming health care is not an inferior good in any state:

---

1 Eq (7) characterizes the solution for $0 < m_s^* < W$. Alternatively, $m_s^* = 0$ if $P_m U_1(W - P_m \sum_k p_k m_k^*\,, m_s^*, \theta, s) > U_2(W - P_m \sum_k p_k m_k^*\,, m_s^*, \theta, s)$, and $m_s^* = W$ if $P_m U_1(W -$ *Pmkpkmk∗,ms∗,θ, s<U2W−Pmkpkmk∗,ms∗,θ, s.*

2 For ease of composition, the subscript insured indicates the insurance relationship between N and m. Specifically, $U(W - P_m \sum_k p_k m(k)\,, m, \theta, s) = U_{insured}$. Similarly, the subscript uninsured indicates the original relationship between N and m. For example, $U(W - P_m m, m, \theta, s) = U_{uninsured}$.

3 This holds true when Eq (7) characterizes the optimal level of N_s^* and m_s^*. If for a given s, households optimally choose $m_s^* = 0$ or $m_s^* = W$, then it is not necessary that m_s^* will change with insurance.

$$(9)\,\frac{dm^*(W,P_m,\theta,s)}{dW}=\frac{P_m U_{11_{insured}}\left(1-P_m\sum_{\mathrm{k}\neq\mathrm{s}}p_k\frac{dm_k^*}{dW}\right)}{U_{22_{insured}}+\ p_s P_m^2 U_{11_{insured}}}>0$$

Health care no longer necessarily increases with poor health status.

$$(10)\,\frac{dm^*(W,P_m,\theta,s)}{d\theta}=-\frac{P_m^2 U_{11_{insured}}\sum_{\mathrm{k}\neq\mathrm{s}}p_k\frac{dm_k^*}{dW}+U_{23_{insured}}}{U_{22_{insured}}+\ p_s P_m^2 U_{11_{insured}}}$$

Though welfare has increased with insurance, health care consumption continues to be a function of income and health status. Whether either (9) or (10) increases or decreases for a given s under health insurance is ambiguous, depending on the relative change in m_s^* and N_s^*, p_s, and the concavity for each parameter.

For example, a utility function of the form $U=\ln(N)+\ \theta s\ln m$ would result in a decrease (increase) in health care inequality by income and health status in states of worse (better) health shocks (i.e., $\frac{d(m_{insured}^*-m_{uninsured}^*)}{dWds}<0$ and $\frac{d(m_{insured}^*-m_{uninsured}^*)}{d\theta ds}<0$).

If there exists an $\bar{M}(\theta,s)$ and households are relatively close to this level of health care consumption, then insurance has a greater likelihood of decreasing health care inequality by income, even in states of positive health shocks.

Insurance without Commitment (Incentive Compatible)

Suppose that households cannot credibly commit to a predetermined indemnity schedule. Upon realization of the state s, households re-optimize their utility to determine the level of health care to consume:

$$(11)\max_{m_s}U\left(W-P_m\sum_k p_k m_k\,,m_s,\theta,s\right)$$

The level of health care is now characterized by the following equation:

$$(12)\ P_m p_s U_{1_{insured}}-U_{2_{uninsured}}\begin{cases}<0\ if\, m^*=W\\=0\ if\ 0<m^*<W\\>0\ if\ m^*=0\end{cases}$$

Rather than equating the relative marginal benefit of N and m, the relative marginal benefit of health care now appears artificially high by the reduced cost of health care ($P_m p_s$). This reduced cost of health care causes households to shift away from N and towards m (i.e., health care demand increases). Conditional upon realizing s, the only cost to household's health care consumption is the consequential increase in the insurance premium. The demand for health care will increase most in the states which have the lowest probability of occurring, since the consequential increase in the premium will be relatively low. Thus, if the probability of good health is relatively low, it could even be the case that health care demand is now greater in states of good health relative to states of bad health. This is a feature of the household internalizing the increase in the premium price. Not making this assumption would result in an impossible insurance contract in which households consume an infinite amount of health care, unless an alternative assumption of the maximum $\bar{M}(\theta, s)$ is made. Thus, an alternative model would not have households internalize the increase in the premium price, but allow for an $\bar{M}(\theta, s)$ to exist. In this case, the insurance will only exist if $W > \sum_s p_S P_m \bar{M}(\theta, s)$. With either assumption, the insurance premium will increase in price to reflect this "overconsumption" in health care, thereby reducing income in all states *s*. Therefore, the reduced cost of health care increases health care demand, but the reduced income decreases health care demand. Whether or not overall health care consumption increases is ambiguous due to these two competing factors.

Health care continues to increase with income, assuming that health care is not an inferior good in any state:

$$(13)\ \frac{dm^*(W, P_m, \theta, s)}{dW} = \frac{P_m p_s U_{11_{insured}}}{U_{22_{insured}} + {p_s}^2 P_m^2 U_{11_{insured}}} \left(1 - P_m \sum_{\mathrm{k} \neq \mathrm{s}} p_{\mathrm{k}} \left(\frac{dm_k^*}{dW}\right)\right) > 0$$

Whether or not health care increases with health status continues to be ambiguous:

$$(14)\ \frac{dm^*(W, P_m, \theta, s)}{d\theta} = -\frac{P_m^2 p_s U_{11} \sum_{\mathrm{k} \neq \mathrm{s}} p_{\mathrm{k}} \frac{dm_k^*}{d\theta} + U_{23}}{U_{22} + P_m^2 p_S^2 U_{11}}$$

Whether income sensitivity increases continues to be ambiguous. In the example using the log utility function, health care inequality by income and health status will increase or decrease depending on the probability of the health shock. For example, the condition for health care inequality by income decreasing in sicker status is that the probability of the health shock must be lower than some critical point. Similar to the previous case, the existence of $\bar{M}(\theta, s)$ increases the likelihood of decreasing health care inequality by income.

As Wolfe and Goddeeris (1991) show, under such a contract it is ambiguous whether wealthier households will be more or less likely to demand insurance because of the trade-off between their higher health care expenditure, increased premium, and lower risk aversion.

Define maximum willingness to pay, P^*, for an insurance contract as the following:

$$(15)\ E_u V(W - P^*, 0, \theta, s) = E_u V(W, P_m, \theta, s)$$

where V represents the maximized utility function in terms of exogenous variables. Then it follows that $\frac{d(P^*-P)}{dW} = 1 - \frac{EU_1(W - P_m m(W, P_m, \theta, s))}{EU_1(W - P^*)} - \frac{dE(m(W-P, 0, \theta, s))}{dW}$, where P is the actual premium charged.

$\frac{EU_1(W-P_m m(W,P_m,\theta,s))}{EU_1(W-P^*)}$ cannot be signed unambiguously because it is unknown whether the marginal utility of non-health consumption is greater when insured or uninsured.

It is also ambiguous whether sicker households have greater demand for the insurance: $\frac{d(P^*-P)}{d\theta} = \frac{EU_3(W-P^*,m(W-P^*,0,\theta,s),\theta,s)}{EU_1(W-P^*,m(W-P^*,0,\theta,s),\theta,s)} - \frac{EU_3(W-P_m m(W,P_m,\theta,s),m(W,P_m,\theta,s),\theta,s)}{EU_1(W-P^*,m(W-P^*,0,\theta,s),\theta,s)} - \frac{dE(m(W-P,0,\theta,s))}{d\theta}$.

Though the first two terms are positive, it is unclear if the gains from insurance outweigh the higher premium charged due to increased demand in health care.

Single Premium Pooled Risk

If the insurer is unable to charge households a premium based on their wealth and health status, all households will face the same premium for the given contract. For households that have higher health care expenditure, health care consumption will increase from the reduced price of the premium.

Wealthier (*sicker*) households will receive a relative increase in income due to lower premiums and thus increase their health care demand. The opposite will be true for poorer (*healthier)* households. This suggests that health care inequality by income will be increased, and that differences in health care by health status will increase.

It still remains the case that the demand for insurance by income is ambiguous for the same reasons mentioned above[4]: $\frac{dP^*}{dW} = 1 - \frac{EU_1(W-P_m m(W,P_m,\theta,s))}{EU_1(W-P^*)}$.

When the insurance is pooled across health status, insurance will be unambiguously demanded more by those with worse health if we assume that health care demanded is higher when insured for all health states: $m^*(W-P,0,\theta,s) \geq m^*(W,P_m,\theta,s)$:

$$\frac{dP^*}{d\theta} = \frac{EU_3(W-P^*,m(W-P^*,0,\theta,s),\theta,s)}{EU_1(W-P^*,m(W-P^*,0,\theta,s),\theta,s)} - \frac{EU_3(W-P_m m(W,P_m,\theta,s),m(W,P_m,\theta,s),\theta,s)}{EU_1(W-P^*,m(W-P^*,0,\theta,s),\theta,s)} > 0.$$

[4] If $\frac{\partial\theta}{\partial W} < 0, \frac{\partial P^*}{\partial W}$ will decrease, but still cannot be signed.

Table A.3.1: Poverty and Health Care Consumption (Panel Data, Enrolled Treatment Households)

	(1)	(2)
Dependent Variable:	Monthly Health Expenditure (INR), Oct - Dec 2011	
Below the Poverty Line (BPL)	-189.4	-522.0***
	(209.5)	(181.7)
Constant	568.4***	677.2***
	(107.0)	(171.5)
Obs (HH Month)	509	420
BPL Fractional Expenditure	0.67	0.23
Enrollment Date:	Prior to Oct 2011	Oct 2011 - June 2012

Observations are limited to treatment households selected and surveyed in the Household Health Survey and are weighted to be representative of the target population.

BPL is an indicator for whether the household is Below the Poverty Line.

Standard errors are in parentheses and are clustered at the household level.

Statistical significance levels are as follows: *10%, **5%, ***1%.

Table A.3.2: Poverty and Health Care Utilization by Enrollment Status (Control Households)

	(1)	(2)	(3)	(4)
	Enrolled Households		Not Enrolled Households	
Dependent Variable:	Admitted	Health Shock Over INR 1,000 (USD 20)	Admitted	Health Shock Over INR 1,000 (USD 20)
Below the Poverty Line (BPL)	0.0958*	0.0903*	-0.0248	-0.00971
	(0.0490)	(0.0532)	(0.0437)	(0.0452)
Constant	0.519***	0.544***	0.541***	0.561***
	(0.0341)	(0.0337)	(0.0378)	(0.0376)
Obs (HH)	338	338	566	567

Observations are limited to control households selected and surveyed in the Household Health Survey and are weighted to be representative of the target population.

BPL is an indicator for whether the household is Below the Poverty Line.

Standard errors are in parentheses and are clustered at the household's Self Help Group (SHG).

Statistical significance levels are as follows: *10%, **5%, ***1%.

Table A.3.3: Heterogeneous Effect of Insurance by Poverty Status (Control Households)

	(1)	(2)
Dependent Variable:	Monthly Health Expenditure (INR)	
Months Enrolled (β_1)	-63.47	-77.32
	(87.55)	(140.1)
Months Enrolled * Below the Poverty Line (BPL) (β_2)	**76.56**	**75.83**
	(92.78)	(159.6)
Obs (HH Month)	3812	2941
Households Enrolled Dec 2012 - Feb 2013	Yes	No
Households Enrolled March 2013 - June 2013	Yes	Yes

Regressions include household and month fixed effects.

Observations are limited to all control households (unless noted otherwise) selected and surveyed in the Household Health Survey and are weighted to be representative of the target population.

Months Enrolled are the months the household has been enrolled since being offered the MHI (December 2012 - June 2013).

BPL is an indicator for whether the household is Below the Poverty Line.

Standard errors are in parentheses and are clustered at the household level.

Statistical significance levels are as follows: *10%, **5%, ***1%.

Table A.3.4: Poverty and Health Care Utilization by Insurance Status (Treatment Households)

	(1)	(2)	(3)	(4)
	Enrolled Households		Not Enrolled Households	
Dependent Variable:	Admitted	Health Shock Over INR 1,000 (USD 20)	Admitted	Health Shock Over INR 1,000 (USD 20)
Below the Poverty Line (BPL)	-0.0858	-0.117*	-0.0113	-0.0492
	(0.0637)	(0.0626)	(0.0654)	(0.0676)
Constant	0.469***	0.514***	0.448***	0.500***
	(0.0353)	(0.0383)	(0.0485)	(0.0521)
Obs (HH)	431	433	271	273

Observations are limited to treatment households selected and surveyed in the Household Health Survey and are weighted to be representative of the target population.

BPL is an indicator for whether the household is Below the Poverty Line.

Standard errors are in parentheses and are clustered at the household's Self Help Group (SHG).

Statistical significance levels are as follows: *10%, **5%, ***1%.

References

Balarajan, Y. S. Selvaraj, S.V. Subramanian. 2011. Health care and equity in India. *The Lancet*, 377(9764): 505 – 515.

Banerjee, Abhijit, Esther Duflo, Rachel Glennerster, and Cynthia Kinnan. 2013. The Miracle of Microfinance? Evidence from a Randomized Evaluation. NBER Working Paper 18950.

Chankova, Slavea, Sara Sulzbach, and François Diop. 2008. Impact of mutual health organizations: evidence from West Africa. *Health Policy and Planning.* 23 (4): 264 – 276.

Dercon, Stefan, Martina Kirchberger, Jan Willem Gunning, and Jean-Philippe Platteau. 2008. Literature Review on Microinsurance. *International Labour Organization*, Geneva, Switzerland.

Dupas, Pascaline, and Jonathan Robinson. 2009. Savings Constraints and Microenterprise Development: Evidence from a Field Experiment in Kenya. National Bureau of Economic Research Working Paper No 14693.

Ekman, Bjorn. 2004. Community-based health insurance in low-income countries: a systematic review of the evidence. *Health Policy and Planning*. 19(5): 249-270.

Fleubaey, Marc and Erik Schokkaert. 2012. Equity in Health and Health Care. In *Handbook of Health Economics* (Volume 2, pp. 1003-1092). Amsterdam: Elsevier (North-Holland).

GoI. 2008. Eleventh Five Year Plan 2007 – 2012. *Government of India*, New Delhi: Oxford University Press.

GoI. 2013. Twelfth Five Year Plan 2012 – 2017, Volume III. *Government of India*, New Delhi: Sage Publications.

ILO. 2012. Micro Insurance Innovation Facility Annual Report 2011. *International Labour Organization*, Geneva, Switzerland.

Jalan, Jyotsna and Martin Ravallion. 1999. Are the poor less well insured? Evidence on vulnerability to income risk in rural China? *Journal of Development Economics,* 58: 61 – 81.

Jakab, Melitta and Chitra Krishnan. 2003. Review of the Strengths and Weaknesses of Community Financing. In *Health Financing for Poor People: Resource Mobilization and Risk Sharing*, edited by A. S. Preker and G. Carrin. Washington D.C.: The World Bank.

La Forgia, Gerard and Somil Nagpal. 2011. Government-Sponsored Health Insurance in India: Are You Covered? *The World Bank*, Washington D.C.

Leatherman, Sheila, Lisa Jones Christensen, Jeanna Holtz. 2012. Innovations and barriers in health microinsurance. In C. Churchill and M. Matul (Eds.), *Protecting the Poor: A microinsurance compendium* (Vol 2, Chapter 5). Geneva: ILO.

Matul, Michal, Aparna Dalal, Ombeline De Bock, and Wouter Gelade. 2013. Why people do not buy microinsurance and what can we do about it. *International Labour Organization*, Geneva, Switzerland.

McCord, Michael J., Clemence Tatin-Jaleran, Molly Ingram. 2012. The Landscape of Microinsurance in Latin America and the Caribbean Briefing Note. *Munich Re Foundation*.

McCord, Michael J., Roland Steinmann, Molly Ingram. 2013. Briefing Note: The Landscape of Microinsurance in Africa 2012. *Munich Re Foundation* and *GIZ-Program Promoting Financial Sector Dialogue in Africa*.

Morduch, Jonathan. 1999. Between the State and the Market: Can Informal Insurance Patch the Safety Net? *World Bank Research Observer*, 14(2): 187 – 207.

Morduch, Jonathan. 2006. Micro-insurance: The Next Revolution?” in *Understanding Poverty*, edited by Abhijit Banerjee, Roland Benabou, and Dilip Mookherjee. Oxford University Press.

O’Donnell, Owen, Eddy van Doorslaer, Adam Wagstaff, and Magnus Lindelow. 2008. Analyzing Health Equity Using Household Survey Data. A Guide to Techniques and Their Implementation. *The World Bank*, Washington D.C.

Oza, Arman, Prernasis Mukherjee, Rupalee Ruchismita. 2013. The Landscape of Microinsurance in Asia and Oceania 2013 Briefing Note. *Munich Re Foundation* and *GIZ-RFPI Asia*.

Ranson, Kent, Tara Sinha, Mirai Chatterjee, Akash Acharya, Ami Bhavsar, Saul S. Morris, Anne J. Mills. 2006. Making health insurance work for the poor: Learning from the Self-Employed Women's Association's (SEWA) community-based health insurance scheme in India. *Social Science & Medicine*, 62 (3): 770 – 720.

RSBY. 2013a. “About the Scheme.” n.d. Web. 30 Oct. 2013. http://www.rsby.gov.in/about_rsby.aspx

RSBY. 2013b. “Scheme Status/State Wise.” n.d. Web 30 Oct. 2013. <http://www.rsby.gov.in/Statewise.aspx?state=35>

SEWA. 2013. “SEWA Services: VimoSEWA.” n.d. Web. 30 Oct. 2013. http://www.sewa.org/Services_Work_Security_Insurance.asp

Sneider, Pia and Kara Hanson. 2006. Horizontal equity in utilization of care and fairness of health financing: a comparison of micro-health insurance and user fees in Rwanda. *Health Economics*, 15: 19-31.

Townsend, Robert M. 1994. Risk and insurance in village India. *Econometrica* 62, (3) (May): 539-91.

Wagstaff, Adam. 2002. Poverty and health care sector inequalities. Policy and Practice. *Bulletin of the World Health Organization*, 80:97-105.

Wagstaff, Adam and Eddy van Doorslaer. 2000. Equity in health care finance and delivery. In A. Culyer & J. Newhouse (Eds.), *Handbook of health economics* (Vol. 1B, pp. 1803-1862). Amsterdam: Elsevier (North-Holland).

Wagstaff, Adam and Magnus Lindelow. 2008. Can insurance increase financial risk? The curious case of health insurance in China. *Journal of Health Economics*, 27: 990-1005.

Wang, Hong, Winnie Yip, Licheng Zhang, Lusheng Wang, William Hsiao. 2005. Community-based health insurance in poor rural China: the distribution of net benefits. *Health Policy Plan*, 20(6): 366-374.

Wolfe, John R. and John H. Goddeeris. 1991. Adverse selection, moral hazard, and wealth effects in the Medigap insurance market. *Journal of Health Economics*, 10: 433-459.