Essays in Health Economics

by

Yubraj Acharya

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
(Health Services Organization and Policy)
in the University of Michigan
2017

Doctoral Committee:

Professor Edward C. Norton, Chair
Professor Richard A. Hirth
Assistant Professor Andrew D. Jones
Professor Dean Yang

Yubraj Acharya

yubraj@umich.edu

ORCID iD: 0000-0002-9003-636X

Dedication

To June and Yuna

Acknowledgements

I consider myself incredibly lucky to have had mentors who are not only experts in their field but who also understood my strengths and limitations well. First and foremost, I would like to thank Professor Edward C. Norton for chairing the dissertation committee and providing mentorship throughout my time at the University of Michigan, even during his sabbatical. I have benefited tremendously from the guidance from Professor Dean Yang and other members of my dissertation committee: Professors Richard A. Hirth and Andrew D. Jones.

Thanks are also due to Qing Zheng, Anup Das, Betsy Cliff, Morris Hamilton, Ryoko Sato, Dhiraj Sharma and seminar participants at the University of Michigan, the 89th Health Economics Study Group meeting (Spain, June 2016), the Northeast Universities Development Economists Consortium conference (Boston, November 2016), and the International Health Policy Conference (United Kingdom, February 2017) for comments on earlier analyses.

The experiment, based on which I wrote the first chapter, would not have materialized without the support received from Professor Dirgha Ghimire and in the field from Prem Pandit, Ramesh Ghimire, Beena Mahato, Krishna Ghimire, Bishnu Adhikari, Indra Chaudhary, the nurses and the interviewers, the Female Community Health Volunteers and the Health-Post In-Charges in Chitwan, Nepal. I thank them all. At Michigan, I want to thank Mindy Niehaus-Fukuda for her help with numerous administrative requests these past five years.

I would like to acknowledge the generous funding from the Department of Health Management and Policy's McNerney Award, School of Public Health's International Travel Award, and the Rackham Predoctoral Fellowship.

Finally, I want to thank my non-academic collaborators: my family. My 92 year old grandmother and my parents did not complain even when I could not meet them during my trips to Nepal, nor did my siblings when I met them only in transit. My two stress busters—June and Yuna—kept me in high spirits. During the time it took me to put together this short dissertation, my wife raised two beautiful daughters (and a husband) and still managed to get her own graduate degree. Thank you, Munu, for your inspiration, unconditional love, and support.

Table of Contents

List of Tables

List of Figures

Abstract

My dissertation broadly relates to the low uptake of preventive health services in developing countries despite the services' low cost and potential to avert subsequent catastrophic expenses. Using Nepal as a setting, in the first two chapters, I answer two key questions on preventive health that are of general interest to health researchers and policymakers.

Question 1. Can we improve the uptake of health services by the traditionally marginalized groups through the use of differential financial incentives to outreach workers?

For the last three decades or so, the research community has cataloged the differences in health outcomes and access between different groups—based on race, ethnicity, gender and other characteristics—and on a range of medical conditions. And much of the research by economists has focused on improving service utilization in general. In the first chapter of my dissertation, I focus on the *differential* access between individuals from different ethnic groups, and propose and test the use of differential financial incentives as a way to address it. The differential incentives that I propose are ones that depend on the characteristics of the individual to whom the outreach workers reach.

I answer the question using a field experiment in Nepal. The medical condition of interest in the study is diabetes, the prevalence of which is nine percent in the country in 2016 (World Health Organization, 2016). Anecdotal evidence shows that individuals do not go for the diagnosis and treatment of diabetes until the conditions become severe. The resulting health costs, disability and sometimes death affect not only the patients but also their families and communities. Like in many other countries, health outreach workers, called the Female Community Health Volunteers, are used to encourage the use of health services in Nepal. In general, the health outreach workers target a certain geographic area, provide information about available health services, and encourage individuals to utilize those services. Literature in sociology suggests that interactions, such as those that the outreach workers engage in, are more

difficult when they involve individuals from different identities such as ethnicity. If the advantaged ethnic groups reach out primarily to their own groups, inequality could be worsened by the existing outreach efforts; more individuals from the advantaged groups would access services with only a small increase, if any, in the number of individuals from disadvantaged groups.

In the experiment, I varied the amount of financial incentives provided to the health outreach workers by the ethnicity of the client they recruited for a free sugar-level assessment. I also varied the amount of incentives the clients received for appearing for the assessment. With this set up, I measure the extent of barriers to outreach effort and to healthcare utilization that individuals face because of their ethnicity and investigate the role of differential and non-differential incentives in offsetting those barriers. I also examine the asymmetric nature of the barriers that health outreach workers and individuals from traditionally disadvantaged and advantaged ethnic groups face. I find that the barriers due to ethnicity are high. Even a highly skewed differential incentive (in the ratio of 5:2) favoring cross-ethnic interactions is insufficient to offset the barriers. Encouragingly, differential incentives to the advantaged workers, geared toward encouraging them to refer disadvantaged individuals, have the potential to improve access for the disadvantaged groups.

In addition to answering important research questions, the findings from the experiment have immediate policy implications for how financial incentives should be structured to encourage the diagnosis of non-communicable medical conditions, both in Nepal and other countries. The health outreach workers in Nepal have been praised in the international health community for their contribution in reducing maternal and child mortality in the country. This study generates insights on the extent to which the experience of these workers can be extended to other conditions such as diabetes, obesity and mental health that were traditionally not known to be common in the country (likely due to under-diagnosis). My results imply that the health outreach workers can continue to play an important role in encouraging preventive health behavior. The policy challenge now is to build an incentive structure so that the significant disparities prevalent in the uptake of common, communicable diseases and their outcomes do not extend to the newer, non-communicable conditions, such as diabetes.

<u>Question 2</u>. What are the long-term consequences of preventive health measures undertaken in childhood?

There is now a critical threshold of evidence documenting the relationship between one's exposure to shocks in early life and outcomes in adulthood. In a seminal review article in the *Annual Review of Economics*, Janet M. Currie and Tom S. Vogl summarize the work done so far on the relationship between early-life nutrition, famine, rainfall, pollution, disease and war, and long-term health outcomes, primarily height. Based on their extensive review, they argue that "[F]uture research should focus on identifying pathways and mechanisms; measuring the relative magnitudes of the effects of different health shocks; examining interactions between shocks; and revisiting the question of critical periods" (p. 29). My second chapter contributes to the existing literature responding to that call. The first goal of the chapter is to examine the effect of an early-life nutritional intervention on health outcomes that the intervention is intended to affect directly as evidenced by the medical literature, thereby elucidating on a clear mechanism. A second goal is to assess the effect of age at first exposure to the intervention, again on expected health outcomes, to get at the role of critical periods. In many countries with a history of discrimination and unequal access to resources based on gender and ethnicity, it is natural to expect different effects of the program on these dimensions. Therefore, the final goal is to evaluate heterogeneous effects of the program by gender and by ethnicity. In addition to the effect on health outcomes, I also evaluate the effect on education outcomes to check whether the findings here are consistent with the vast amount of literature showing that healthier children tend to be healthier adults with better educational and labor market outcomes.

I make use of Nepal's vitamin A supplementation program, the primary goal of which was to reduce mortality associated with the nutrient's deficiency. Vitamin A deficiency affects nearly 21 percent children below the age of five years in developing countries and leads to the deaths of over 800,000 women and children each year (West, 2002). The sequential rollout of Nepal's vitamin A supplementation program between 1993 and 2001 and the age eligibility provide an exogenous variation in exposure to the program. I utilize that variation to estimate the causal effect of the program on long-term health and economic outcomes. While such programs have had significant positive short-term benefits in reducing mortality in Nepal and elsewhere, as documented in the medical literature, the study aims to provide additional insights on potential mechanisms through which early-life interventions affect long-term outcomes. I find that the program reduced the probability of having a disability or blindness, kept children in school longer, and enabled them to complete different grades by an expected age. The positive effects on disability and education seem to have improved marriage prospects. The program also had different effects on individuals based on their timing of the exposure to the program, with a longer exposure usually strengthening the positive effects. As expected, effects also differed by the individual's gender and ethnicity. They were more pronounced for men and individuals from traditionally advantaged ethnic groups.

Question 3. What is the causal effect of Nepal's Community-Based Neonatal Care Package intervention?

In the third chapter, I evaluate an existing program broadly aimed at reducing child mortality and improving women's health behavior using a rigorous econometric technique. The goal of this chapter is to contribute to ongoing efforts on evidence-based policymaking in Nepal. I evaluate the impact of Community-Based Neonatal Care Package, which the government piloted in 2009 in 10 of the 75 districts. The causal effect of the program is established using a before-and-after comparison of outcomes in program districts relative to those in non-program districts. I find that the program was successful in encouraging cleaner deliveries for births that took place at home and in increasing prenatal visits to the health center by pregnant women significantly. Despite these positive effects on intermediate outcomes, the program's overall effect on neonatal mortality was limited. There is also no evidence that the program increased institutional or professional-attended deliveries. The lack of an effect on other supply-dependent indicators suggests that supply-side constraints may have dampened the program's overall effect. While this program has been evaluated before, I put it to a more rigorous test, and show that the effects may be less impressive than what previous analyses—many of them qualitative or based on a simple pre-program versus post-program comparisons within the program districts—have shown. More importantly, consistent with the international shift in efforts toward improving the quality of health services—from current efforts focused on access—my findings call for improvements in the supply-side if the health of Nepalese women and children is to improve more rapidly.

References

The World Health Organization, 2016. Diabetes country profiles 2016. URL (Accessed 26 April 2017): http://www.who.int/diabetes/country-profiles/npl_en.pdf?ua=1

West, K.P. Jr., 2002. Extent of vitamin A deficiency among preschool children and women of reproductive age. *Journal of Nutrition,* 132:2857S–2866S.

Chapter 1

Barriers to Inter-Ethnic Interactions in Healthcare: Evidence from a Field Experiment

1.1 Introduction

The uptake of healthcare services in developing countries is low, even for simple cost-effective technologies (Kremer and Glennerster, 2011). Supply side efforts to raise uptake include reducing distance to services (Thornton, 2008), improving the quality of services (Clasen *et al*., 2007), and improving the reliability of supply (Banerjee *et al.*, 2010), among others. On the demand side, the dominant interventions include providing information, financial rewards, or both, for seeking care (Jacobs *et al*., 2011; Dupas, 2011).

This study focuses on the issue of *differential* access and uptake of healthcare services among individuals from different groups, although its findings also help understand barriers to uptake in general (as the average uptake is usually reduced by the low uptake of the minority groups). Unequal access and uptake of healthcare services is a major problem in both developed and developing countries (Braveman and Tarimo, 2002). Minority groups tend to have a lower access to and uptake of healthcare services than majority groups (O'Hara and Caswell, 2010). Outreach workers are often used to solve the problem of low access for minority groups. The expectation from policymakers is that outreach workers would reach individuals from minority groups who would not otherwise access services. However, the majority groups are overrepresented in the health workforce (Snyder *et al*., 2015; AHRQ 2013). Literature in sociology suggests that individuals find it easier to reach out to others like themselves (Barnes-Mauthe *et al.*, 2013). If the majority groups reach out primarily to their own groups, inequality could be worsened by the existing outreach efforts; more individuals from majority groups would access services with only a small increase, if any, in the number of individuals from minority groups.

Several studies have evaluated the effect of ethnic matching on treatment outcomes, particularly in mental health (e.g., Cabral and Smith, 2011). However, such matching is not always possible, especially in a resource-poor setting. Matching becomes particularly difficult for preventive health because the risk profile of an individual is not known beforehand—often, one does not know who to encourage diagnosis, let alone how best to encourage and incentivize such a behavior. In general, political and cultural issues can undermine efforts aimed at making the composition of outreach workers reflective of the target population (Rao and Flores, 2007).

Therefore, in order to improve access for minority groups and thus address inequality in general, it is important to find a mechanism to encourage outreach workers from one group to reach out to those from another. Financial incentives have proven to be effective in nudging individuals toward a socially preferred behavior in many settings (Giles *et al*., 2014), but the extent of the effect on outreach effort vis-à-vis multiethnic interactions is poorly understood. We are unaware of any study that attempts to offset inter-ethnic barriers among health workers with differential financial incentives.

Against this background, this study is designed to answer four key questions. First, what is the extent of the barrier to encouraging and seeking preventive health care (in this case, diagnosis of diabetes) that is caused by the difference in the ethnicity of the individual and that of health outreach workers? Second, can we incentivize health outreach workers from one ethnic group to recruit individuals from another group for the diagnosis, either through higher, non-differential incentives or through differential incentives favoring recruitment from a different ethnicity? Third, does the extent of barriers to outreach effort differ by the outreach worker's ethnicity? In other words, is the extent of the barriers faced by a health worker from a traditionally advantaged ethnic group different from the one faced by a worker from a disadvantaged group? Finally, on the demand side, does the clients' decision to use healthcare services depend on the ethnicity of the outreach worker and if so, can financial incentives to the clients help increase the chances that a client utilizes the services?

In economics, the paper is related the most closely to the literature on discrimination, which is also a form of barrier. Economists continue to debate the dominant form, the measurement and the mitigation of discrimination since the seminal works of Becker (1957) and

Arrow (1973). The two strands that have progressed over the years—one on taste-based discrimination and another on statistical—both recognize that individuals may be willing to pay a positive amount in order to interact with individuals from their own ethnic groups or race. In our study, when looking at the barriers that advantaged health workers face when reaching out to disadvantaged clients, the dominant barrier is discrimination, although we are unable to rule out other factors conclusively. Nonetheless, we are able to estimate how much individuals are willing to forego in order to interact with individuals like themselves (i.e., from their own ethnic category) and extend the literature in three other ways. First, while prior studies have looked at discrimination—a form of barrier—from the dominant group to a dominated group, our setup allows us to compare the extent of barriers a traditionally advantaged outreach worker faces when reaching out to a disadvantaged individual and that of barriers a disadvantaged outreach worker faces when reaching out to an advantaged individual. In the US, this would be analogous to asking: how is the extent of the barrier that a White physician faces when interacting with an minority patient different from the extent of the barrier that a minority physician faces when interacting with a White patient? Second, we are also able to measure the barrier at multiple stages of the healthcare seeking process. Finally, as discussed below, we evaluate the barriers from the perspective of the service providers as well as the seekers.

We recruited all health volunteers within a geographic territory in a semi-urban district in Nepal, randomized them into four arms, provided them a basic training on diabetes and asked them to recruit clients from the community for a free sugar-level assessment at their local health center. We varied the amount of financial incentives they received. In two of the arms, the amount depended on the ethnicity of client the health volunteers recruited. In one of these two arms, we provided a higher amount for recruiting a client from their own ethnicity (which we call an own-type referral) than for recruiting a client from a different ethnicity (an other-type referral), whereas in the other, we provided a higher amount for recruiting a client from a different ethnicity. That variation allows us to compare how much additional effort health volunteers make when they are incentivized to recruit own-type clients and when they are incentivized to recruit other-type clients. The comparison of the number of own-type and other-type referrals in the first arm, in which the amount of incentive does not depend on the type of the referral, allows us estimate the extent of the barrier at baseline. Likewise, comparison between the first and the fourth arm, in which the amount of incentive is higher, allows us to answer additional research questions on the role of non-differential incentives in presence of ethnic heterogeneity.

We included a second level of randomization to evaluate barriers due to ethnicity from the clients' perspective. For an individual to increase his or her uptake of health services based on a health worker's persuasion, the individual should be receptive to the health worker's message. It is possible for the individual not to act on the health worker's suggestion, even if the health worker does not face any barrier to reaching out to that individual. The same level of effort on the part of the health worker can then lead to different outcomes (in terms of whether the individual increases his/her uptake of health services) based on whether the health worker and the individual are from the same ethnicity. By randomizing incentives received by clients for showing up for a checkup, we are able to assess if incentives can help offset the barriers faced by a prospective patient when his/her ethnicity does not match with that of the health volunteer.

To preview the results, we find that the barriers due to ethnicity are high. At baseline, the health volunteers recruited only three other-type clients for every five own-type clients. Even a highly differential incentive in the ratio of 5:2, geared toward encouraging the health volunteers to recruit clients from an ethnic group different than their own, is insufficient to offset the barriers. In sub-group analysis, we find suggestive evidence that differential incentives to the traditionally advantaged health volunteers have the potential to improve access for the disadvantaged groups. We also find that the advantaged and disadvantaged health volunteers face different amounts of barriers to outreach efforts, with the latter facing a "stereotype threat". Financial incentives to the clients had no effect on their decision to appear for the assessment.

1.2 The Study Setting

The subjects in this study are the Female Community Health Volunteers (health volunteers) in a semi-urban area in western Nepal and the clients they recruited for a free sugar-level assessment. Nepal is an appropriate site for this study because of the persistent prevalence of health disparities between ethnic groups and the low uptake of preventive health services. Significant disparities exist in both access to healthcare services and health outcomes between ethnic groups (Pandey *et al*., 2013). In fact, widespread discrimination and inequality, in all spheres of life, catalyzed the Maoist insurgency, which claimed 15,000 lives between 1996 and 2006 (Nepal, Bohra and Gawande, 2011). Following a protracted peace process, in September

2015, the country adopted a new constitution, which has renewed the commitment to addressing inequality.

The prevalence of diabetes, the medical condition of interest in this study, is rapidly rising in Nepal, with current prevalence at 9.1% (World Health Organization, 2016). In general, the burden of disease is shifting quickly from communicable to non-communicable conditions such as cardiovascular diseases, cancer, chronic respiratory diseases and mental order (Ministry of Health and Population, 2015).

The government created the health volunteers in 1989 to help administer vitamin A supplements to children. There are nearly 48,000 health volunteers, all female, in the country (Andersen *et al.*, 2013). Each health volunteer is responsible for her Ward, which is the lowest administrative unit in the country. The health volunteers are primarily tasked to create awareness about available health services and to encourage individuals in their Ward to utilize those services. Over the years, the health volunteers' role has expanded significantly and they have been praised in the international development community for their success in reducing child and maternal mortality (Center for Global Development, 2011). Based on the country's past experience in reducing child and maternal mortality, the health volunteers can potentially play an integral role in the management of the new conditions as well. The extent to which this can happen, however, has not been evaluated. Apart from answering the research questions listed earlier, this paper also helps fill that gap.

The Nepalese government has categorized the country's more than 100 ethnicities into 6 main categories based on religion, caste and ethnicity, and further into advantaged and disadvantaged groups based on historical access to resources. In this study, we use these two broad categories. The differences—both in access and outcomes—are pronounced between these categories (Pandey *et al.*, 2013). The categorization also has a political appeal. Other studies have also used this categorization as a basis for ethnicity (e.g., Mishra, Joshi and Khanal, 2014). The advantaged or disadvantaged status of an individual is known to the health volunteers. The general public can also infer it from the individual's last name.

1.3 The Study Design

We randomly assigned 72 health volunteers into four arms stratified by their ethnic category (advantaged versus disadvantaged), education, and age. We invited the health volunteers for one-day training on diabetes at their local health center. After the training, two days before the checkup, the research team visited the health volunteers at their home and explained to them the incentive structure in private. We explained the incentive structure in private so that one health volunteer's behavior was not influenced by the knowledge of what other volunteers were receiving. We did not reveal the specific objectives of the study and the incentive structure even to the research staff. Additional details on the implementation are in Appendix 1.A.

Each health volunteer was told that she would receive an amount of money based on the number of clients who came for the checkup at their local health center on the pre-specified date and time, and according to the schedule in Table 1.1. To summarize, in arms 1 and 4, the amount of incentive per referral did not depend on the ethnicity of the client the health volunteers recruited. In arm 1, which we refer to as the *Low* arm in the rest of the paper, the health volunteers received Nepalese rupees (Rs) 20 per referral. The exchange rate between the US dollar and the Nepalese rupee was $1: Rs100 at the time of the experiment. Therefore, Rs 20 is approximately $0.2 (or 20 cents). In arm 4, which we refer to as the *High* arm, they received Rs 50 per referral. In arms 2 and 3, the amount depended on the ethnicity of client the health volunteers recruited. In arm 2, which we refer to as the *NudgeOther* arm, the amount was higher for recruiting a client from a different ethnicity (an other-type referral) than for recruiting a client from their own ethnicity (an own-type referral). In arm 3, which we refer to as the *NudgeOwn* arm, the amount was higher for recruiting a client from their own ethnicity. As discussed in Section 1.1, arms *NudgeOwn* and *NudgetOther* allow us to compare how much additional effort health volunteers make when they are incentivized to recruit own-type clients and when they are incentivized to recruit other-type clients, relative to a baseline effort (arm *Low*). The comparison of arm *Low* and arm *High* allows us to examine the effect of higher, non-differential incentives on motivation in presence of ethnic heterogeneity. This examination is important in view of the common use of such incentives as a way to raise uptake of preventive health services in many programs, by governments as well as non-governmental organizations.

To put the incentive amount in context, the health volunteers are generally not paid a salary but receive some incentives (not based on performance) from the government, including transport stipends for training and meeting allowances. In this study, the health volunteers were provided a lump sum of Rs 600 (approximately, US$ 6) on the day of the training to cover the cost of transportation and to offset their opportunity cost of time that day. A semi-skilled worker in the area earns approximately Rs 400 per day, close to the Rs 8,000 per month minimum wage set by the government. If a health volunteer in arm *Low* recruited 50 clients, and if all showed up, she would receive Rs 1,000, which is 2.5 times the daily wage of a semi-skilled worker in the area.

The second level of randomization is at the client level. We randomized incentives received by the clients for showing up for the sugar-level assessment. As discussed in Section 1.1, this additional randomization allows us to evaluate the effect of incentives on the decision to appear for the sugar-level assessment and if incentives can help offset the barriers due to ethnicity from the clients' perspective. We sent each client an invitation letter that specified a randomly-assigned amount between Rs 20 and Rs 50 (in intervals of Rs 10) that the client would receive if she or he came to the health center for the checkup. The health volunteers gave clients the letter along with the referral card.

We collected additional information on the health volunteers and the clients who came to the checkup using a survey. We held checkups in eight health centers. Nurses recruited for this experiment tested the blood sugar levels using a handheld Nova-Stat Glucometer. The Nova-Stat Glucometer has been found to be reliable and accurate for the determination of blood glucose levels (Rabiee et al., 2010). Nonetheless, the nurses advised those with high sugar levels to go to a hospital for further diagnosis.

1.4 Empirical Approach

1.4.1 Supply (Health Volunteer's) Response

In order to evaluate the extent of barriers to outreach effort that a health volunteer faces due to her ethnicity, we estimate two different equations below—one for own-type referrals and another for other-type referrals—and perform a number of tests.

$$(1)\quad Y_{own,\,j} = \beta_{1,\,own} + \beta_{2,\,own}\, NudgeOther_j + \beta_{3,\,own}\, NudgeOwn_j + \beta_{4,\,own}\, High_j + \delta_n \mathbf{X}_j + \varepsilon_j$$

$$(2)\quad Y_{other,\,j} = \beta_{1,\,other} + \beta_{2,\,other}\, NudgeOther_j + \beta_{3,\,other}\, NudgeOwn_j + \beta_{4,\,other}\, High_j + \delta_n \mathbf{X}_j + \varepsilon_j$$

In equations (1) and (2), $Y_{own,\,j}$ and $Y_{other,\,j}$ are the natural log of the number of own-type referrals and other-type referrals, respectively, made by a health volunteer j. To recapitulate, an own-type referral is defined as a referral in which a health volunteer recruits a client from her own ethnic group. The arms differ by *j* and are mutually exclusive. Arm *Low* is the excluded arm in both equations. **X** includes a set of health volunteer characteristics which may influence their ability to recruit clients or their choice of the client. These include the health volunteer's age, education level, experience, distance to the health center, ethnicity, occupation and household income. They also include the amount of money they received for their work as a health volunteer in the previous month, the number of households they usually visit per month and the number of advantaged and disadvantaged households in their ward based on the 2011 census. ε is the usual error term.

To further check the validity of randomization and the stability of coefficients, we estimate equations (1) and (2) first without any covariates, then with variables used for stratification (ethnicity, age and education) and finally with additional covariates (age, occupation, number of households normally visited per month, amount received for work as a health volunteer the previous month, distance to the health center and the proportion of advantaged and disadvantaged households in the ward). For interpretation of the results, we use the coefficients from the fully-specified regressions.

Using these two equations, we predict the number of own-type and other-type referrals for each arm. A formal test of the difference between the predicted number of own-type and other-type referrals in arm *Low*—i.e., at baseline—evaluates if there are barriers to outreach effort due to ethnicity. If there are no barriers, then the number of own-type and other-type referrals should not be different from each other at baseline (i.e., $\beta_{1,\,own} = \beta_{1,\,other}$). Likewise, a formal test of the difference between the predicted number of own-type referrals in arm *Low* (from equation (1)) and the predicted number of other-type referrals in *NudgeOther* (from equation (2)) assesses if it is possible to attain the same number of other-type referrals, through a differential incentive, as the number of own-type referrals at baseline. In other words, a test of whether $\beta_{1,\,other} + \beta_{2,\,other} \geq \beta_{1,\,own}$ evaluates if we can eliminate the barriers that a health volunteer faces due to her ethnicity by providing her a differential incentive geared toward encouraging an other-type referral. A difference between the two numbers would also confirm further that the barriers are large—so large that even a differential incentive in the ratio of 2.5:1 (i.e., Rs 50/ Rs 20) cannot fully eliminate. Finally, a test between the number of other-type referrals in arm *Low* and the number of other-type referrals in arm *High* (i.e., $\beta_{1,\,other} = \beta_{4,\,own}$) can be used to evaluate if a higher, non-differential incentive helps offset the barriers that the health volunteer faces.

We estimate equations (1) and (2) separately using two different samples: first using all the clients who received a referral card from the health volunteers and then using only the clients who showed up to the checkup. We do so because the health volunteers' effort can be understood as a combination of two parts: the effort she puts in reaching out to a client and the effort in convincing the client to visit the health center for the checkup. The first part can be measured by the number of referral cards the health volunteers distributed to the clients. The overall effort—sum of the effort in reaching out and in convincing the client to go to the checkup—can be measured by the number of clients who showed up. The health volunteers were told that the amount of incentive they received would depend on the number of clients who showed up. However, the type of clients to whom they gave the referral cards differed in a number of characteristics (discussed in Section 1.6), making the decision to show up potentially endogenous. Therefore, it is logical to conduct analyses using both samples. One can also think of raising the uptake of healthcare services as a three-step process: reaching out to the clients, getting them to the healthcare center and providing them care. Ethnicity-related barriers can limit access and uptake at any of these points. Evaluating findings using the both outcomes, therefore, enables us to assess the relative strength of the barrier at two of these three steps.

To estimate the asymmetric nature of the barrier—i.e., to compare the extent of barriers faced by an advantaged outreach worker when reaching out to a disadvantaged individual and by a disadvantaged outreach worker when reaching out to an advantaged individual—we include interaction terms between arms and the ethnic category of the health volunteer in equations in (1) and (2). We estimate:

(3) $$Y_{own,\,\mathrm{j}} = \beta_{1,\,own} + \beta_{2,\,own}\, NudgeOther_j + \beta_{3,\,own}\, NudgeOwn_j + \beta_{4,\,own}\, High_j + \beta_{5,\,own}\, \mathrm{Ethnicity}_j + (\beta_{6,\,own}\, NudgeOther_j \times \mathrm{Ethnicity}_j) + (\beta_{7,\,own}\, NudgeOwn_j \times \mathrm{Ethnicity}_j) + (\beta_{8,\,own}\, High_j \times \mathrm{Ethnicity}_j) + \delta_n \mathbf{X}_j + \varepsilon_j$$

(4) $$Y_{other,\,\mathrm{j}} = \beta_{1,\,other} + \beta_{2,\,other}\, NudgeOther_j + \beta_{3,\,other}\, NudgeOwn_j + \beta_{4,\,other}\, High_j + \beta_{5,\,own}\, \mathrm{Ethnicity}_j + (\beta_{6,\,other}\, NudgeOther_j \times \mathrm{Ethnicity}_j) + (\beta_{7,\,other}\, NudgeOwn_j \times \mathrm{Ethnicity}_j) + (\beta_{8,\,other}\, High_j \times \mathrm{Ethnicity}_j) + \delta_n \mathbf{X}_j + \varepsilon_j$$

In equations (3) and (4), Ethnicity=1 if the health volunteer is from an advantaged group. If $\beta_{7,\, own} \neq 0$ in equation (3), an advantaged health volunteer and a disadvantaged health volunteer differ in terms of the amount of effort they put toward recruiting an own-type client. Likewise, if $\beta_{6,\, other} \neq 0$ in equation (4), an advantaged health volunteer and a disadvantaged health volunteer differ in terms of the amount of effort they put toward recruiting an other-type client.

Differential incentives provided to health volunteers to change their behavior have the potential to distort the individuals' behavior in a way that is inefficient. In this study's setting, the health volunteers in *NudgeOther* and *NudgeOwn* arms, driven by financial motivation, can recruit other-type and own-type clients who are less likely to be diabetic. A health volunteer in the *NudgeOther* arm, for example, may recruit a healthy other-type client to receive the additional financial incentive, even though there may be other less healthy own-type clients. In order to test if such behavior occurs, we compare the characteristics of the clients who came to the checkup between the arms (information on these characteristics is not available for clients who received a referral card but did not come to the checkup). In particular, we are interested in the diabetic status of patients on the extensive margin and the sugar level on the intensive margin.

We compare the general characteristics and diabetic status of the clients recruited by the health volunteers in two ways. We compare the characteristics of clients in *NudgeOther* to those in *NudgeOwn* in order to see the difference in the composition and severity of clients based on who the health volunteers were incentivized to recruit with the differential incentives. Then we compare the characteristics of clients in *NudgeOther* and *NudgeOwn* to those in *Low* and *High*, in order to see the difference in the composition and severity of clients based on the nature of the incentives—differential versus non-differential. In these analyses, we cluster the standard errors at the health volunteer level.

1.4.2 Demand (Clients') Response

On the clients' side, the key outcome of interest is whether a client who received a referral card from a health volunteer showed up for the checkup. A client is either from the health volunteer's ethnicity or not. In order to evaluate the general effect of the incentives and if a higher incentive encourages a client whose ethnicity is different than that of the health volunteer to come to the checkup, we estimate the following equation.

$$(5)\quad Y_{ij} = \alpha + \beta_{1,\text{demand}}\, \textit{Unmatched}_{ij} + \beta_{2,\text{demand}}\, \textit{Amount of incentive}_i + \beta_{3,\text{demand}}\, (\textit{Amount of incentive}_i \times \textit{Unmatched}_{ij}) + \delta_n \mathbf{X} + \varepsilon_j.$$

In equation (5), Y_{ij} is a binary variable that equals 1 if an individual *i* referred by health volunteer *j* showed up for the checkup and 0 otherwise. Unmatched=1 if health volunteer and the individual are from different ethnic categories and 0 otherwise. **X** is a vector of health volunteer characteristics. **X** also includes a categorical variable for the arm that the health volunteer belongs to because health volunteers in different arms may put different effort toward convincing the client to come to the checkup, which in turn may affect the client's decision. Here, too, we cluster the standard errors at the health volunteer level.

Clients from the same ethnic category as that of the health volunteer can be expected to be more likely to show up than those from a different ethnic category. Therefore, in equation (5), we expect $\beta_{1,\text{demand}}<0$. Because the clients receiving a higher incentive should be more likely to show up, we expect $\beta_{2,\text{demand}}>0$. We hypothesize that, with higher incentives, clients who are from a different ethnic category than that of the health volunteer will be more likely to show up than at lower incentives, therefore $\beta_{3,\text{demand}}>0$.

1.5 Descriptive Statistics and the Validity of Randomization

Of the 72 health volunteers who had been randomized into four groups, 69 showed up for the training and were recruited for the experiment. The three health volunteers who did not show up were one each from arm *Low*, arm *NudgeOther* and arm *High*. Of the 69 health volunteers, 43 (62 percent) were from the advantaged ethnic category, while the remaining 26 (38 percent) were from the disadvantaged category (Table 1.2). In the analytical sample, on average, a health volunteer is 46 years old and has 19 years of experience. All are women. Less than one-third of health volunteers have education equivalent to the school leaving certificate (equivalent to the sophomore year of high school in the United States) and 10% have only informal education. On average, a health volunteer in the sample visited 50 households in the month preceding the survey and lives half an hour away from the nearest local health center. Seventy-eight percent of health volunteers received honorarium for their work in the month before the survey, 82 percent of health volunteers reported agriculture as their main occupation, and 20 percent said at least one of their nearest five neighbors was from a different ethnic group than their own.

The health volunteers distributed the referral cards to 2,825 clients (average = 40.9 cards per health volunteer). Of these, 2,403 (85.1 percent) showed up for the checkup and 2,365 (98.4 percent of all those who showed up) were interviewed. The remaining 38 include clients who showed up after the interviewers had left. For clients who received the cards from the health volunteers but did not show up, we have data on their ethnicity and the amount of incentive they would have received for showing up. Of the 2,365 clients who were interviewed, information on some of the covariates is missing for a total of 29 clients, leaving a final, complete analytical sample of 2,336 (97.2 percent of all clients who showed up and 98.7 percent of all clients who were interviewed).

Among those who showed up and provided complete information, 60 percent of clients are women, the average age is 52 years, and 56 percent are from advantaged ethnic category (Table 1.2). Sixty-six percent are from the same ethnic category as that of the health volunteer. Eighty-eight percent are married and the average education level is grade 4. On average, clients live 27 minutes away from the nearest health center, 82 percent are engaged in agriculture and 61 percent had heard about diabetes before. Almost all of them heard about the sugar-level checkups from their health volunteer. In 98 percent cases, the health volunteer visited the individual at home to talk about diabetes and to give the referral card and the letter.

Randomization divided the health volunteers into four similar arms (Appendix 1.B, Table B1). For many health volunteer's, their actual experience, age and the level of education—self-reported during the interviews—were different from the information collected from the health centers before randomization (not shown). However, the arms are generally balanced based on self-reported experience, age and the level of education collected from the health volunteers individually at the time of the training. Surprisingly, there is a monotonic decrease in age and experience going from arm *Low* to arm *High*, but that is due to chance, and we control for these—and other characteristics of the health volunteers—in the regression analysis. The first set of p-values is from a joint orthogonality test of all arms. The p-values in the last column are from the test of difference in means between arms *NudgeOther* and *NudgeOwn*, the critical two arms required to draw inference on the health volunteers' differential response to differential financial incentives.

On the demand side, the health volunteers in all arms had similar probabilities of receiving letters offering Rs 20, Rs 30, Rs 40 and Rs 50, which confirms the validity of randomization of incentives to the clients (Appendix 1.B, Figure B1). There is no evidence that the health volunteers opened the envelopes beforehand to give letters mentioning a higher amount to their own-type clients—the proportion of envelopes going to own-type clients were 64.7 percent for Rs 20, 66 percent for Rs 30, 63.2 percent for Rs 40 and 64.4 percent for Rs 50. In fact, the characteristics of clients who received different amounts are also balanced, except in the proportion of clients who reported that they had heard about diabetes even before the health volunteers visited them (Appendix 1.B, Table B2).

1.6 Main Empirical Results

1.6.1 Extent of the Barriers due to Ethnicity

Based on the clients who received a referral card and controlling for the characteristics of the health volunteers, at baseline (i.e., in the *Low* arm), they recruited 20 own-type clients and 12 other-type clients. The own-type and other-type referrals at baseline were, therefore, in the ratio of 5:3 (Table 1.3). The difference is statistically significant at the 5% significance level ($p<0.001$).

With a differential incentive in the ratio of 5:2 (i.e., Rs 50/Rs 20) geared toward encouraging the outreach workers to recruit a client from an ethnicity different than their own, other-type referrals increase by 11.6% (statistically insignificant). Even with this increment, however, the other-type referrals are lower than own-type referrals at baseline (p value from a test of difference, between own-type referrals at baseline and other-type referrals with a differential incentive, is 0.014).

Both of these findings suggest that the barriers due to ethnicity are high in this setting. It is not the case of the health volunteers not responding at all to the incentives. In fact, they are very responsive to incentives in general, as reflected by their response to a differential incentive geared toward encouraging an own-type referral. On the log scale, the mean number of own-type and other-type referrals at baseline translate to 3 and 2.5, respectively. Own-type referrals increased by approximately 48% from *Low* (baseline) to *NudgeOwn* (higher incentives for own-type referral) (Table 1.3, Panel A). We obtain this estimate by taking the exponent of the coefficient and subtracting one from the result. The corresponding change in other-type referrals from *Low* to *NudgeOther* (higher incentives for other-type referrals) is 12%. If these proportional increments reflect the additional amount of effort made by health volunteers in response to the incentives, the effort they made when they were incentivized to recruit other-type clients was about one-fourth (=12/48) of the effort they made when they were incentivized to recruit own-type clients. The price change was 150% (i.e., went up from Rs 20 to Rs 50) in both cases, which means that the elasticity is 0.32 (=48/150) for own-type referrals and 0.08 for other-type referrals.

The extent of barrier is similarly high when it is measured based on the number of clients who showed up for the checkups (Appendix 1.C, Table C1). In this case, the baseline own-type and other-type referrals are in the ratio of 3:2 and are statistically different from each other (p-value = 0.003). The difference persists even with a differential incentive geared toward encouraging an other-type referral; the p-value from the test of difference, between own-type referrals at baseline and other-type referral with a differential incentive, is less than 0.001.

1.6.2 Effect of Non-Differential Incentives

We find that higher, non-differential incentives can be counterproductive in offsetting the barriers due to ethnicity. When the incentive amount is increased from Rs 20/referral (*Low*) to Rs 50/referral (*High*), the number of own-type referrals remains unchanged. However, the number of other-type referrals falls by a statistically significant amount (Table 1.3); the coefficient of -0.791 on *High* in Table 1.3 corresponds to an approximately 55 percent reduction in the number of other-type referrals from baseline. The corresponding decline based on the clients who came to the health center for the sugar-level assessment is 59 percent. In terms of the ethnic composition of the clients, from *Low* to *High*, the share of other-type clients falls from 42 percent to 27.5 percent based on the sample of clients who received a referral card. We return to these striking results in the discussion section.

1.6.3 Asymmetry in the Extent of the Barrier

The study's setup enables us to analyze the asymmetric nature of the barriers that traditionally advantaged individuals face when they interact with those from traditionally disadvantaged individuals, and *vice versa*. The regression results from estimating equations (3) and (4) are in Appendix 1.D, Table D1.

For a visual comparison of the difference in the behavior of advantaged and disadvantaged health volunteers, we plotted the natural log of the predicted number of referrals from equations (3) and (4) against the amount of incentive provided to the health volunteers. As seen in Figures 1.1(a) and 1.1(b), the two types of health volunteers had similar responses when they were provided a higher incentive for own-type referrals. However, they differed in their response to the incentive geared toward encouraging an other-type referral. The coefficients on *NudgeOwn* and *NudgeOther* are not statistically different from one another for the advantaged health volunteers but they are for the disadvantaged health volunteers. For the advantaged health volunteers, the number of own-type referrals at baseline and the number of other-type referrals with a *NudgeOther* incentive are statistically not different from each other. This suggests that, although the barriers due to ethnicity are high in general, it is possible to improve the access of disadvantaged groups to health services by providing differential incentives to the advantaged health volunteers. For the disadvantaged health volunteers, an incentive geared toward encouraging an other-type referral *decreased* the number of other-type referrals (compared to *Low*). According to Table D1, based on the sample of clients who came to the checkup, the disadvantaged health volunteers reduced the number of other-type (advantaged to them) referrals by a statistically significant 65 percent.

1.6.4 Effect of Incentives on the Type of Clients Reached

On the extensive margin, among the clients who showed up, clients recruited by the health volunteers in *NudgeOther* and *NudgeOwn*—in which the health volunteers received differential incentives—were more likely to be diabetic (Table 1.4). The mean probability of being diabetic among those in *Low* and *High* is five percent. The clients recruited by *NudgeOther* and *NudgeOwn* health volunteers were about two percentage points more likely (or about 6.9 percent likely) to be diabetic.

However, we do not find any difference in the diabetic status of clients recruited by health volunteers in *NudgeOwn* and *NudgeOther* arms. We also do not find any effect of the type of incentives on the intensive margin—conditional on the client being diabetic, there was no difference in the sugar level of the client recruited by the health volunteers in *NudgeOwn* and *NudgeOther* arms, or between those recruited by health volunteers receiving differential or non-differential incentives.

In terms of the general characteristics, health volunteers in *NudgeOther* recruited older, less-educated clients and fewer women than did health volunteers in *NudgeOwn* (Appendix 1.E,

Table E1). In this setting, the incentives to recruit other-type clients also seem to have encouraged the health volunteers to reach out to clients who otherwise are usually less likely to go for the checkup, such as women and older, less-educated individuals. Compared to the non-differential incentives (*Low* and *High*), the differential incentives (*NudgeOther* and *NudgeOwn*) encouraged the health volunteers to reach out to younger, slightly more educated clients but who lived further from the health posts.

1.6.5 Effect of Incentives on Demand (Decision to Access Services)

On the demand side, overall, the incentives to the clients—in the range tested, i.e., between Rs 20 and Rs 50—were inconsequential in affecting the clients' behavior (Table 1.5). Even controlling for health volunteer's arms and other characteristics, the standard errors on the key variables—interaction of matching ethnicity and incentive amount—are large, and the R-squared values reveal that a negligible portion of the variation in outcome is explained by the incentives, ethnic match between the health volunteer and the client, and the characteristics of the health volunteer. The estimated coefficients are statistically insignificant at the conventional 5% significance level. There is suggestive evidence that financial incentives to the clients may have *reduced*, not increased, their chances of coming to the checkup. As expected, the mismatch in the ethnicity of the health volunteer and the client seems to reduce the chances of the client appearing for the checkup. Because the coefficient on incentives and ethnic match are not in the opposite direction, the discussion of whether incentives help offset the barriers on the part of the client is not relevant.

1.7 Discussion and Conclusion

Using a unique experimental setup, we showed that the difference in the ethnicity of health outreach workers and the prospective patients constitutes a significant barrier to health services utilization. At baseline, health outreach workers recruited a significantly higher number of clients from their own ethnic category than from a different ethnic category. Even a differential incentive in the ratio of 5:2, geared toward encouraging the health outreach workers to recruit an other-type client, was insufficient to offset the barrier. The health volunteers do respond to financial incentives in general, as suggested by the statistically significant increase in the number of own-type referrals in response to incentives encouraging such referrals; they just do not respond in a similar manner when they are incentivized to reach across ethnic lines. At a higher, non-differential incentive, the health volunteers reduced the number of other-type referrals, suggesting that such incentives can be counterproductive in reducing barriers due to ethnicity.

In many cases, the policy goal is to reach out to the traditionally disadvantaged groups—rather than simultaneously encouraging advantaged workers to reach out to disadvantaged groups and encouraging the disadvantaged workers to reach out to advantaged ones. Our sub-group analysis shows that differential incentives to the advantaged outreach workers have the potential to meet such goals. For the advantaged health volunteers, the differential incentive in the ratio of 5:2, geared toward encouraging them to recruit disadvantaged clients enabled them to offset the baseline differences in own-type versus other-type referrals. Such differential incentives do not have adverse effects on efficiency—in fact, the health volunteers receiving a differential incentive recruited clients who were more likely to be diabetic than those recruited by health volunteers receiving a non-differential incentive.

The study's findings, especially the magnitude of the effects, may have limited external validity given that it was conducted in a specific setting in Nepal, and therefore should be interpreted accordingly. Nepal's health volunteers are anecdotally known for working effectively even across ethnic lines; if that is the case, the estimates of the barrier we present here should be taken as the lower bound of the barriers that prevail in many other settings. Nonetheless, the methodological approach we adopted – differential incentives based on the ethnicity of the individual that a client interacts with – may be applied to several settings, both as a way to evaluate the extent of barriers due to ethnicity and to reduce those barriers. Examples of potential applications include efforts to raise diversity in universities and to raise the uptake of government services by minority groups.

Our study has a number of striking findings that warrant further research. Perhaps the most striking finding is that the disadvantaged health volunteers' recruited fewer of the other-type (i.e., advantaged clients) when they were incentivized to recruit the other-type, even compared to the baseline. This was not expected, but is consistent with the presence of "stereotype threat", a phenomenon in which emphasizing the status of an individual and making it more salient reinforces a behavior associated with that status. Such effects have been found elsewhere. In India, for example, Hoff and Pandey (2005) find that publicly emphasizing students' caste to them created a large and robust gap in performance. Lower caste students performed more poorly when their caste was mentioned to them publicly. Incentives may have played a similar role in this study by making the notions of identity and intra-ethnic bonds more salient for the disadvantaged health volunteers. When we explained to the disadvantaged health volunteers that they would receive a higher amount if they referred an advantaged client, they may have inferred that we expected them to face barriers when making such referrals. In general, the asymmetric nature of the barrier warrants further research, including along other demographic differences, such as gender, race and economic status, which are key determinants of health status.

Second, the high, non-differential incentive encouraged health volunteers to reduce the number of other-type referrals compared to the number of referrals at the low, non-differential incentive. This is consistent with the target income hypothesis, whereby the health volunteers may have recruited the number of clients necessary for them to meet their target income. One can also hypothesize that with higher incentives, the stakes of the client not showing up increase and as a result the health volunteers opted to invest more time to convince their own-type clients to come to the checkup, thereby reducing the number of other-type referrals. From an immediate policy perspective, this finding raises questions about the effectiveness of financial incentives in improving access to care for minority groups. The dominant form of incentives used currently in many part of the world is non-differential. If the outreach workers are predominantly from traditionally advantaged or majority groups, the current incentives may be exacerbating, not ameliorating, the existing health disparities.

Third, the finding that incentives to the clients—which were exogenous—had no or even a negative effect on the decision to come to the checkup also warrants further research. It is possible that the amount of incentive offered to the clients signaled the service's quality, with a higher incentives signaling lower quality. It is also possible that the lowest incentive amount provided to the client – Rs 20 – was already high enough in terms of offsetting the costs they faced when going for the checkup, and therefore the additional amount had no effect on their decision. This second argument is consistent with the high uptake in this study—approximately 85% of the clients approached by the health volunteers came to the sugar-level assessment.

Returning to the policy issue of whether the health volunteers in Nepal can be mobilized in response to the shifting burden of diseases toward non-communicable ones—in a manner they were so successfully mobilized to help reduce child and maternal mortality—the high uptake found in this study is encouraging. It suggests that the health volunteers can continue to play an important role in encouraging preventive health behavior. The policy challenge now is to build an incentive structure so that the significant disparities prevalent in the uptake of common, communicable diseases and their outcomes do not extend to the newer, non-communicable conditions, such as diabetes.

References

The Agency for Healthcare Research and Quality, 2014. National Healthcare Disparities Report 2013. US Department of Health and Human Services. URL (accessed 25 April 2017): http://www.ahrq.gov/sites/default/files/publications/files/2013nhdr.pdf.

Andersen, K., Singh, A., Shrestha, M.K., Shah, M., Pearson, E., Hessini, L., 2013. Early pregnancy detection by female community health volunteers in Nepal facilitated referral for appropriate reproductive health services. *Global Health: Science and Practice* 1, 372–81.

Arrow, K.J., 1973. The theory of discrimination. In O. Ashenfelter & A. Rees (Eds.), *Discrimination in Labor Markets*, pp. 3–33, Princeton, NJ: Princeton University Press.

Banerjee, A., Duflo, E., Glennerster, R., Kothari, D., 2010. Improving immunization coverage in rural India: A clustered randomized controlled evaluation of immunization campaigns with and without incentives. *British Medical Journal*, 340, C2220.

Barnes-Mauthe, M., Arita, S., Allen, S. D., Gray, S. A., Leung, P.S., 2013. The influence of ethnic diversity on social network structure in a common-pool resource system: implications for collaborative management. *Ecology and Society* 18(1): 23.

Becker, G.S., 1957. The economics of discrimination, Chicago, IL: The University of Chicago Press.

Braveman, P., Tarimo, E., 2002. Social inequalities in health within countries: not only an issue for affluent nations. *Social Science and Medicine* 54, 1621-1635.

Brown, D. J., DeCorse-Johnson, A. L., Irving-Ray, M., Wu, W. W., 2005. Performance evaluation for diversity programs. *Policy Politics and Nursing Practice*, 6 (4), 331–334.

Cabral, R. R., Smith, T. B., 2011. Racial/ethnic matching of clients and therapists in mental health services: A meta-analytic review of preferences, perceptions, and outcomes. *Journal of Counseling Psychology*, 58(4), 537-554.

Center for Global Development, 2011. Reducing child mortality with vitamin A in Nepal. URL: http://www.cgdev.org/doc/millions/MS_case_4.pdf.

Clasen, T., Haller, L., Walker, D., Bartram, J., Cairncross, S., 2007. Cost-effectiveness analysis of water quality interventions for preventing diarrhoeal disease in developing countries. *Journal of Water and Health* 5:599–608.

Dupas, P., 2011. Health Behavior in Developing Countries. *Annual Review of Economics* 3, 425-449.

Giles, E.L., Robalino, S., McColl, E., Sniehotta, F.F., Adams. J., 2014. The effectiveness of financial incentives for health behaviour change: systematic review and meta-analysis. *PLoS ONE* 9(3): e90347.

Hoff, K., Pandey, P., 2005. Belief systems and durable inequalities: an experimental investigation of Indian caste. World Bank Policy Research Working Paper No. 3351.

Jacobs, B., Ir, P., Bigdeli, M., Annear, P. L., Van Damme, W., 2011. Addressing access barriers to health services: an analytical framework for selecting appropriate interventions in low-income Asian countries. *Health Policy and Planning,* 1-13.

Kremer, M., Glennerster, R., 2011. Improving Health in Developing Countries. Evidence from Randomized Evaluations. *Handbook of Health Economics*, 2, 201-315.

Ministry of Health and Population, Government of Nepal, 2015. Nepal Health Sector Strategy 2015-2020. URL (accessed 25 April 2017): http://nhsp.org.np/wp-content/uploads/2016/05/NHSS-English-Book-Inside-final.pdf

Mishra, S., Joshi. M.P., Khanal. V., 2014. Family planning knowledge and practice among people living with HIV in Nepal. *PLoS ONE* 9(2): e88663.

Nepal, M., Bohara, A. K., Gawande, K., 2011. More inequality, more killings: the Maoist insurgency in Nepal. *American Journal of Political Science,* 55, 886–906.

O'Hara, B., Caswell, K., 2012. Health status, health insurance, and medical services utilization: 2010. *Current Population Reports,* 70–133.

Pandey, J. P., Dhakal, M.R., Karki, S., Poudel, P., and Pradhan, M.S., 2013. Maternal and child health in Nepal: the effects of caste, ethnicity, and regional identity. URL (accessed 25 April 2017): http://www.dhsprogram.com/pubs/pdf/FA73/FA73.pdf

Rabiee, A., Magruder, J.T., Grant, C., Salas-Carrillo, R., Gillette, A., DuBois, J., Shannon, R.P., Andersen D.K., Elahi, D, 2010. Accuracy and reliability of the Nova StatStrip® glucose meter for real-time blood glucose determinations during glucose clamp studies. Journal of Diabetes Science and Technology, 4, 1195–1201.

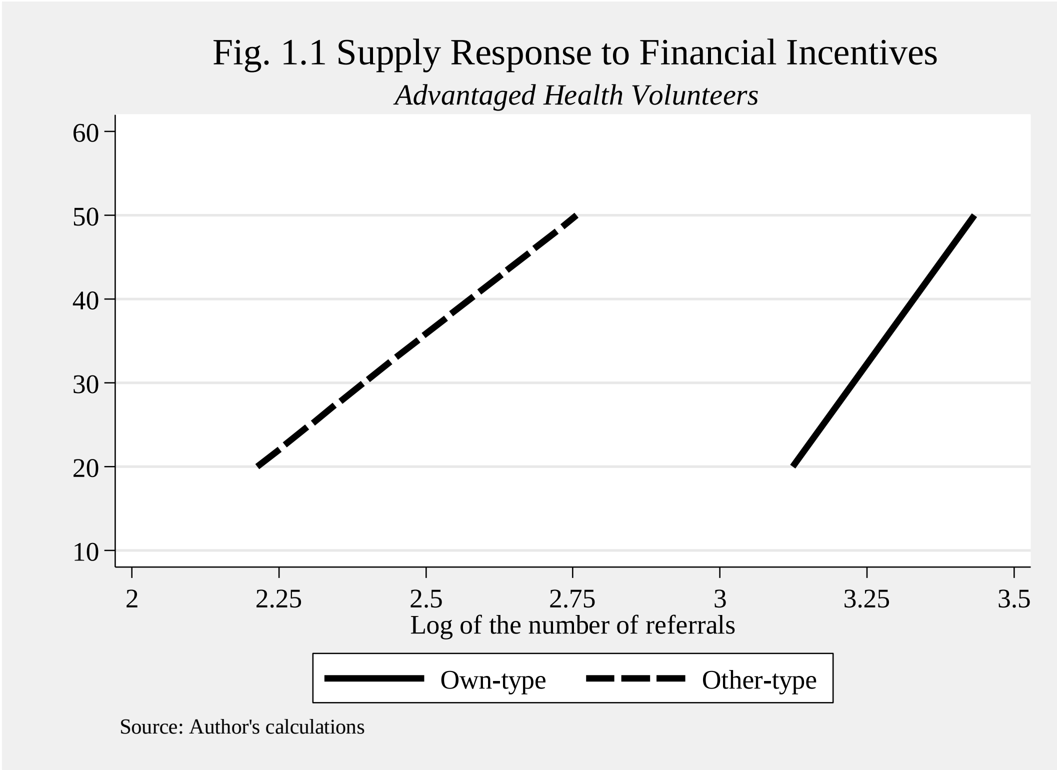
Rao, V., Flores, G., 2007. Why aren't there more African-American physicians? *Journal of National Medical Association,* 99(9), 986–93.

Snyder, C.R., Stover, B., Skillman, S.M., Frogner, B.K., 2015. Facilitating racial and ethnic diversity in the health workforce. Center for Health Workforce Studies, University of Washington. URL (accessed 25 April 2017): http://depts.washington.edu/uwrhrc/uploads/FINALREPORT_Facilitating%20Diversity%20in%20the%20Health%20Workforce_7.8.2015.pdf

Thornton, R. L., 2008. The Demand for, and Impact of, Learning HIV Status. *American Economic Review* 98(5), 1829-1863.

The World Health Organization, 2016. Diabetes country profiles 2016. URL: http://www.who.int/diabetes/country-profiles/npl_en.pdf?ua=1

Tables and Figures



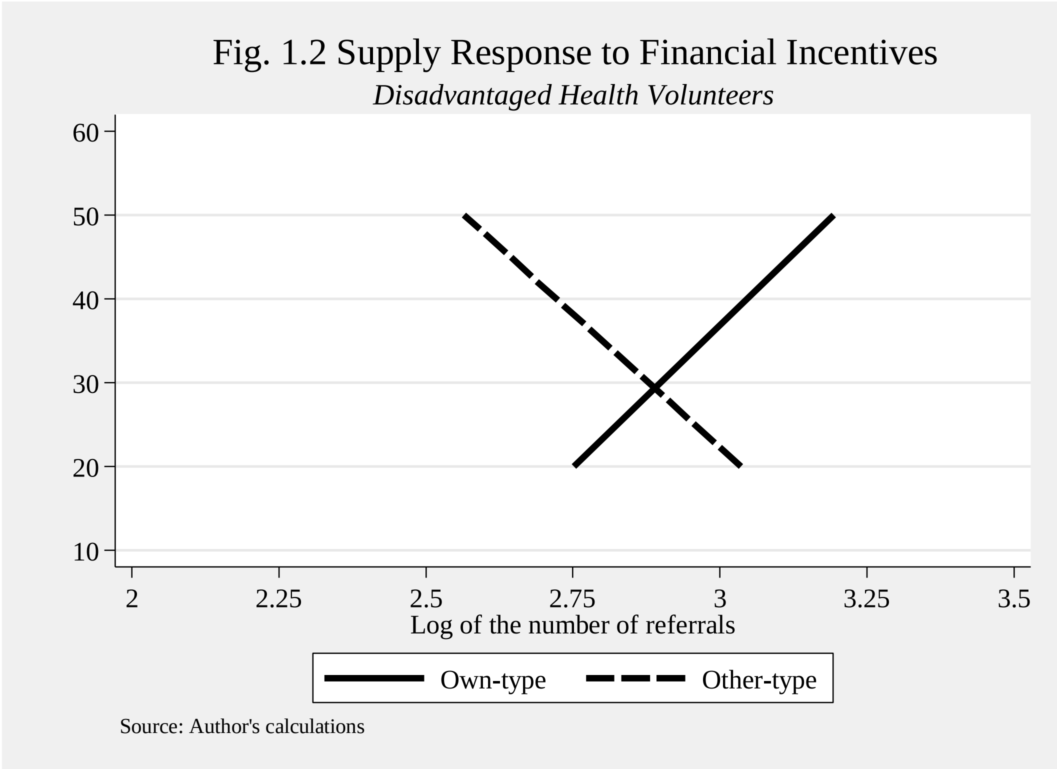


Table 1.1. Incentives Provided to the Health Volunteers

	Arm 1 (*Low*)	Arm 2 (*NudgeOther*)	Arm 3(*NudgeOwn*)	Arm 4 (*High*)
Refer own-type	Low	Low	High	High
Refer other-type	Low	High	Low	High

The exchange rate at the time of the experiment was approximately US$ 1: Nepalese rupees (Rs) 100. Low – Rs 20/referral; High – Rs 50/referral.

Table 1.2. Summary Statistics for the Analytic Sample

	Mean	SD
Health Volunteers (N=69)		
Age, years	46.09	9.28
Experience, years	18.96	7.54
Education higher than grade 10 (yes=1)	0.28	0.45
Had informal schooling (yes=1)	0.10	0.30
Ethnicity (Advantaged=1)	0.62	0.49
Number of household visited per month	50.26	42.65
Received money for work as HV in the previous month	0.78	0.42
Distance to the health center, minutes	29.74	19.89
Primary occupation is agriculture (yes=1)	0.83	0.38
Has one of five neighbors from a different ethnicity	0.20	0.41
Clients (N=2,336)		
Gender (female=1)	0.60	0.49
Age, years	52.07	12.34
Ethnicity (advantaged=1)	0.56	0.50
Same ethnic category as that of the HV	0.66	0.47
Marital status (married=1)	0.89	0.31
Years of schooling	4.11	4.64
Distance to the health center, minutes	26.94	24.30
Primary occupation is agriculture (yes=1)	0.82	0.38
Knew about diabetes before the HV's visit	0.61	0.49
Knew about the checkup from the HV	0.99	0.09
HV informed the client by visiting the client's house	0.98	0.12

Note: Clients include individuals who received a referral card from a HV, showed up for the checkup and answered the questionnaire administered by the research team. As mentioned in the text, of the 2,803 individuals who received a referral card, 2,403 showed up. Of those, 2,336 provided complete information on the various indicators above.

Table 1.3. Regression Results of Log of Referrals on Incentives			
	(1)	(2)	(3)
A. Own-type referrals			
Baseline mean = 2.99 (no. of referrals = 19.9)			
NudgeOther	-0.066	-0.102	-0.034
	(0.206)	(0.203)	(0.185)
NudgeOwn	0.334	0.392*	0.389**
	(0.203)	(0.201)	(0.193)
High	0.255	0.315	0.106
	(0.206)	(0.214)	(0.201)
R-squared	0.08	0.17	0.45
B. Other-type referrals	(4)	(5)	(6)
Baseline mean = 2.50 (no. of referrals = 12.2)			
NudgeOther	0.185	0.189	0.110
	(0.318)	(0.319)	(0.333)
NudgeOwn	-0.519	-0.567*	-0.897**
	(0.313)	(0.316)	(0.348)
High	-0.434	-0.634*	-0.791**
	(0.318)	(0.337)	(0.362)
R-squared	0.10	0.16	0.27
Additional covariates (for both panels)			
Stratification variables	No	Yes	Yes
Other HV characteristics	No	No	Yes

Note. The results in this table are from estimating equations (1) and (2), and the sample is based on all clients who received a referral card from their health volunteer. * p<0.10, ** p<0.05, *** p<0.01. Standard errors are in parenthesis. Stratification variables include ethnicity, experience and education. Other health volunteer characteristics include age, annual household income, number of households the health volunteer visited in the past month, the amount of money she received for working as a health volunteer, distance to the nearest health center, primary occupation and the share of own-type households in the health volunteer's ward.

Table 1.4. Effect of Type of Incentives on the Types of Clients Recruited				
	Diabetic status		Sugar level (for diabetic patients)	
	(1)	(2)	(1)	(2)
NudgeOther	0.004	0.003	-12.5	-15.4
(comparison group: *NudgeOwn*)	(0.012)	(0.012)	(14.2)	(16.7)
Excluded group mean	0.07	0.07	96.3	96.3
R-squared	0.02	0.02	0.19	0.24
N	1155	1155	80	80
NudgeOwn and *NudgeOther*	0.020**	0.022***	3.9	1.6
(Comparison group: *Low* and *High*)	(0.008)	(0.008)	(8.0)	(8.7)
Excluded group mean	0.05	0.05	95.0	95.0
R-squared	0.01	0.01	0.11	0.14
N	2297	2297	136	136
Additional covariates (both panels)				
Health volunteer characteristics	Yes	Yes	Yes	Yes
Individual characteristics	No	Yes	No	Yes

Notes: Each coefficient is from a separate regression. * p<0.10, ** p<0.05, *** p<0.01. Standard errors, clustered at the health volunteer level, are in parenthesis. The health volunteer characteristics include ethnicity, experience, education, age, annual household income, number of households the health volunteer visited in the past month, the amount of money she received for working as a health volunteer, distance to the nearest health center and primary occupation. Individual characteristics include gender, age, marital status, education, distance to the health center, occupation and ethnic category.

Table 1.5. Regression Results of the Decision to Show up for the Checkup by the Clients

	All clients	Advantaged clients	Disadvantaged clients
Mean probability of coming to the checkup	0.85	0.88	0.82
Incentive amount, Rs	-0.0011	-0.0003	-0.0026*
	(0.0008)	(0.0009)	(0.0015)
Client and HV from different ethnic categories	-0.0442	-0.0101	-0.0828
	(0.0466)	(0.0804)	(0.0770)
Incentive amount × Client and HV from different groups	0.0011	-0.0007	0.0034*
	(0.0012)	(0.0019)	(0.0020)
Health volunteer's characteristics	Yes	Yes	Yes
Health center fixed effects	Yes	Yes	Yes
Health volunteer's incentive arm	Yes	Yes	Yes
R-squared	0.06	0.05	0.08
N	2,755	1,507	1,248

Notes: Each column represents a separate regression. * p<0.10, ** p<0.05, *** p<0.01. Standard errors are in parenthesis. They are clustered at the health volunteer level. The health volunteer characteristics include ethnicity, experience, education, age, annual household income, number of households the health volunteer visited in the past month, the amount of money she received for working as a health volunteer, distance to the nearest health center and primary occupation. Individual characteristics include gender, age, marital status, education, distance to the health center, occupation and ethnic category.

Appendices

Appendix 1.A. Implementation Details

As mentioned in Section 1.3, we randomly assigned 72 health volunteers into four arms stratified by their ethnic category (advantaged vs disadvantaged), education, and age. We stratified in order to ensure that each arm had a reasonable number of health volunteers from the traditionally advantaged and disadvantaged ethnic groups. We collected information on ethnicity, education, and age before the experiment from local health centers.

We invited the health volunteers for one-day training at their local health center. A practicing endocrinologist provided information to the health volunteers—in Nepali, the dominant local language—on basic risk factors for diabetes, prevention, symptoms and implications if not treated on time.

Two days before the checkup, the research team visited the health volunteers at their home and explained them the incentive structure in private. The health volunteers were requested not to share their incentive structure with other health volunteers, so that one health volunteer's behavior was not influenced by the knowledge of what other volunteers were receiving. Anecdotal evidence showed that the health volunteers complied with this request, partly because the health volunteers themselves did not want the community to know that they were receiving a monetary reward for their work. We did not reveal the specific objectives of the study and the incentive structure even to the research staff.

Each health volunteer was told that she would receive an amount of money based on the number of clients who came for the checkup at their local health center on the pre-specified date and time, and according to the schedule in Table 1.1.

The second level of randomization is at the client level. We randomized incentives received by the clients for showing up for the sugar-level assessment. We sent each client an invitation letter which specified a randomly-assigned amount between Rs 20 and Rs 50 (in intervals of Rs 10) the client would receive if she or he came to the health center for the checkup. We put letters mentioning these amounts in envelopes, shuffled them and created stacks of 50 envelopes each.

We gave these 50 letters along with 50 referral cards to each health volunteer. We told them not to open the letters to the clients so that they did not selectively give letters with higher amounts to clients who were more (or less) likely to show up. Without opening the envelopes, it was not possible to know the amount mentioned in the letter.

We told the health volunteers that they could call the research team if they needed more cards and letters or if the clients had questions. None of them called. We gave each health volunteer a day to recruit clients. If the checkup was scheduled for Friday morning, for example, the health volunteer received the referral cards and the letters on Wednesday afternoon. We kept this window for recruitment short partly to ensure that the health volunteers in smaller wards did not visit all households in their wards (for, if they did, we would not know if the mix of clients received is because of a differential effort made by the health volunteer or simply because she referred everyone in her ward) and to reduce the chances of interaction between the health volunteers.

To keep track of all clients to whom the health volunteers provided the referral cards, the referral card had the design of a boarding pass (Figure 1.A1). The health volunteers gave one part of the card to clients and kept the other part. In the part that she kept, the volunteer was asked to write the name and contact information of the individual she spoke to and the code on the envelope that she gave the individual. The research team collected the cards from the health centers at the time of the checkup and from the volunteers the same morning.

Figure 1.A1. The Referral Card

ID Number ☐☐ ☐☐ ☐☐☐ For use by the Female Community Health Volunteer Information on the recipient: Full name:.................................... Phone no.:.................................... Envelop no:..................................	Free Diabetes Checkup ID Number ☐☐ ☐☐ ☐☐☐ Dear Mr/Mrs ...., Please bring this card, along with the letter provided to you by your health volunteer, when you come to the free diabetes (sugar) checkup at your health post. Venue:........................... Date:............................. Time: 7 am (please fast overnight and do not eat anything before coming to the checkup) • For use by the Female Community Health Volunteer: Envelop number:............... Thank you!

We collected additional information on the health volunteers and the clients who came to the checkup using a survey. We administered the survey to the health volunteers on the day of their training and to the clients when they appeared for the checkup.

The coding system in referral cards, the envelopes and the survey questionnaire allowed us to match each individual client to the health volunteer, to know how many clients each health volunteer recruited and how many showed up, and to know the financial incentives the clients received (or would have received, for those who did not come).

We held checkups in eight health centers. On a pre-specified date—which we communicated to the health volunteers after finalizing it with the health center administrators—the research team consisting of eight practicing nurses, 20 trained interviewers, two other research staff, and the author went to the centers to conduct the checkup and to administer the survey to the clients who came. We reached each health post by 7 am. Each individual who appeared was first read the consent form, interviewed and then sent to a separate room for the sugar-level assessment and to receive the financial incentive. The nurses tested the blood sugar levels using a handheld Nova-Stat Glucometer and advised those with high sugar levels to go to a hospital for further diagnosis. Interviews stopped around 9:45 am to allow the health centers to open for regular business at 10 am. Individuals who showed up after the interviewers had left the health center were read the consent form, administered the test and provided with the financial incentive, but were not interviewed.

We paid the incentive to the clients at the time of the checkup and to the health volunteer three weeks after the experiment. For ethical reasons, we provided the same amount to all health volunteers. However, throughout the experiment, the health volunteers did not know that they would eventually receive the same amount.

Appendix 1.B. Evidence on the Validity of Randomization

Table 1.B1. Balance in Key Characteristics of the Health Volunteers between the Arms

	Low	*Nudge-Other*	*Nudge-Own*	*High*	p-value (all arms)	p-value (*NudgeOwn vs NudgeOther*)
Age, years	49.35	47.76	45.78	41.47	0.07	0.54
	(1.74)	(2.64)	(1.83)	(2.36)		
Experience, years	21.29	20.47	19.17	14.88	0.06	0.58
	(1.72)	(1.46)	(1.81)	(1.97)		
Education higher than grade 10 (yes=1)	0.24	0.18	0.28	0.41	0.48	0.48
	(0.11)	(0.10)	(0.11)	(0.12)		
Had informal schooling (yes=1)	0.12	0.18	0.06	0.06	0.62	0.26
	(0.08)	(0.10)	(0.06)	(0.06)		
Ethnicity (Advantaged=1)	0.65	0.65	0.56	0.65	0.93	0.58
	(0.12)	(0.12)	(0.12)	(0.12)		
Income category	2.29	2.18	1.94	1.88	0.45	0.31
	(0.19)	(0.15)	(0.21)	(0.26)		
Number of household visited per month	38.29	56.59	58.78	46.88	0.48	0.88
	(9.69)	(10.82)	(9.39)	(11.30)		
Received money in the previous month	0.88	0.88	0.67	0.71	0.27	0.13
	(0.08)	(0.08)	(0.11)	(0.11)		
Distance to the health center, minutes	27.94	37.65	29.56	23.82	0.23	0.28
	(3.79)	(6.16)	(4.10)	(4.63)		
Primary occupation is agriculture	0.76	0.88	0.78	0.88	0.70	0.41
	(0.11)	(0.08)	(0.10)	(0.08)		

Note: The p-values in column (5) are from the joint orthogonality test of the arms. The p-values in column (6) are from the t-test of the difference in means between the *NudgeOwn* and the *NudgeOther* arms. All variables reported here were self-reported by the HVs. Income was categorized into four groups: 1 - less than 50,000 per year; 2 - 50,000-100,000 per year; 3 - 100,000-200,000 per year; 4 - 200,000-500,000 per year; and 5 - more than 500,000 per year. The mean income category reported in this table is based on those categories.

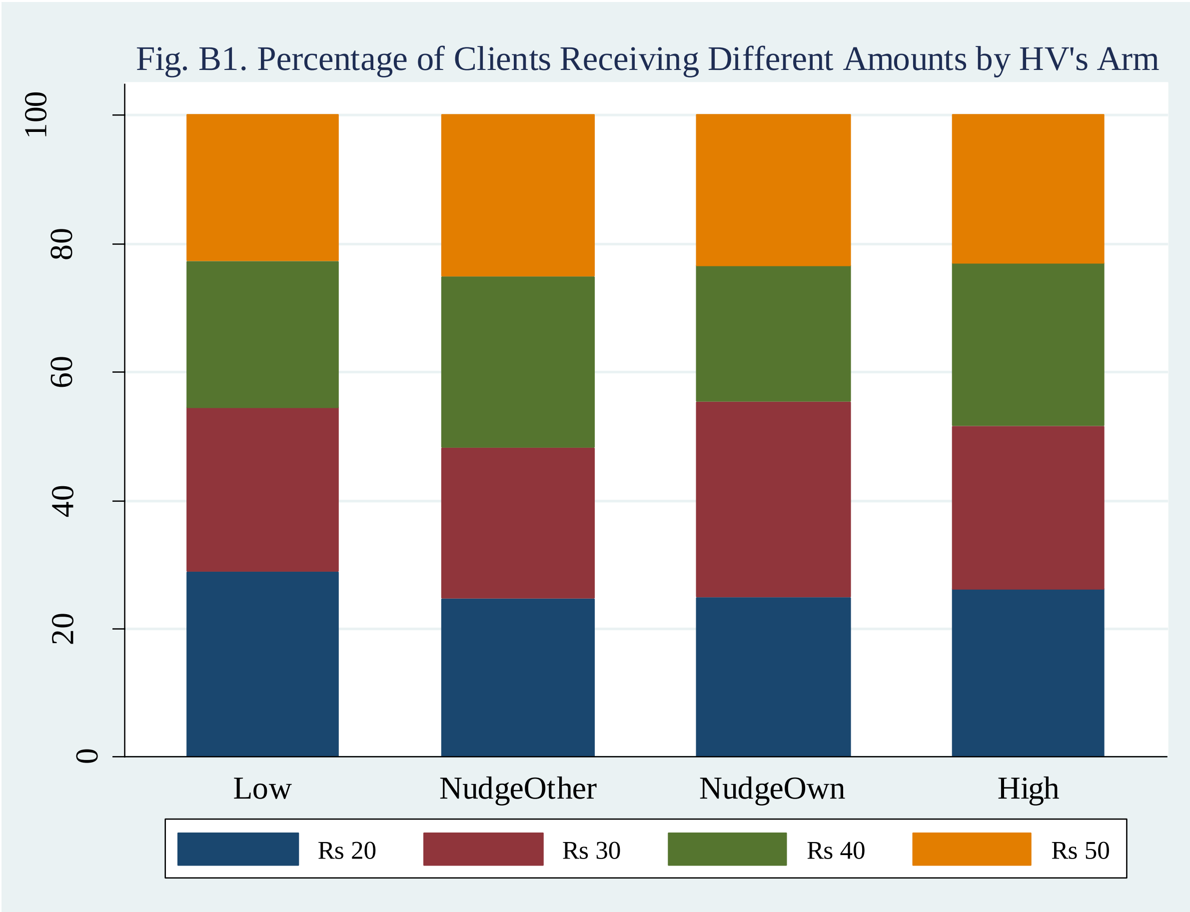


Fig. B1. Percentage of Clients Receiving Different Amounts by HV's Arm
0
20
40
60
80
100
Low
NudgeOther
NudgeOwn
High
Rs 20
Rs 30
Rs 40
Rs 50

Table 1.B2. Balance in Key Covariates between the Clients Receiving Different Amounts

	Rs 20	Rs 30	Rs 40	Rs 50	p-value (all arms)
Age, years	52.54	51.68	51.85	52.41	0.448
	(0.51)	(0.49)	(0.53)	(0.53)	
Women, proportion of total	0.61	0.62	0.59	0.59	0.795
	(0.02)	(0.02)	(0.02)	(0.02)	
Currently married, proportion	0.88	0.89	0.90	0.89	0.875
	(0.01)	(0.01)	(0.01)	(0.01)	
Education, years	3.95	4.26	4.43	3.87	0.149
	(0.19)	(0.19)	(0.20)	(0.20)	
Distance to the health center, minutes	28.54	26.67	25.71	27.11	0.227
	(1.07)	(0.97)	(0.96)	(1.10)	
Farming as main occupation, proportion	0.80	0.84	0.82	0.82	0.164
	(0.02)	(0.02)	(0.02)	(0.02)	
Had heard about diabetes, proportion	0.59	0.62	0.68	0.56	0.002
	(0.02)	(0.02)	(0.02)	(0.02)	
Blood sugar level, mg/dL	94.84	95.79	96.73	95.72	0.780
	(0.91)	(0.93)	(1.38)	(0.98)	
Blood sugar level > 110 mg/dL, proportion	0.15	0.14	0.15	0.14	0.934
	(0.02)	(0.01)	(0.02)	(0.02)	

Note: The p-values in column (5) are from the joint orthogonality test of the arms. The numbers here are for individuals who came to the checkup.

Appendix 1.C. Main Results Based on the Sample of Clients Who Came to the Checkup

Table 1.C1. Regression Results of Log of Referrals on Incentives

	(1)	(2)	(3)
A. Own-type referrals			
Baseline mean = 2.80 (no. of referrals = 16.5)			
NudgeOther	-0.126	-0.166	-0.091
	(0.228)	(0.222)	(0.204)
NudgeOwn	0.372	0.443**	0.448**
	(0.225)	(0.220)	(0.213)
High	0.308	0.368	0.125
	(0.228)	(0.234)	(0.222)
R-squared	0.09	0.20	0.47
B. Other-type referrals	(4)	(5)	(6)
Baseline mean = 2.39 (no. of referrals = 10.9)			
NudgeOther	-0.125	-0.117	-0.074
	(0.327)	(0.331)	(0.347)
NudgeOwn	-0.491	-0.541	-0.717*
	(0.323)	(0.328)	(0.363)
High	-0.529	-0.710**	-0.897**
	(0.327)	(0.350)	(0.377)
R-squared	0.06	0.10	0.22
Additional covariates			
Stratification variables	No	Yes	Yes
Other HV characteristics	No	No	Yes

Note: N=69. The results in this table are from estimating equations (1) and (2), and the sample is based on all clients who came to the health center for the sugar-level assessment. * $p<0.10$, ** $p<0.05$, *** $p<0.01$. Standard errors are in parenthesis. Stratification variables include ethnicity, experience and education. Other health volunteer characteristics include age, annual household income, number of households the health volunteer visited in the past month, the amount of money she received for working as a health volunteer, distance to the nearest health center, primary occupation and the share of own-type households in the health volunteer's ward.

Appendix 1.D. Results on the Asymmetric Effect of the Incentives on Advantaged and Disadvantaged Health Volunteers

Table 1.D1. Response to Incentives by Health Volunteers' Ethnicity

	Based on all clients who received a card		*Based on clients who came to the health center*	
	Own-type (1)	Other-type (2)	Own-type (3)	Other-type (4)
NudgeOther (β2)	0.168	-0.51	-0.018	-1.010*
	(0.330)	(0.581)	(0.365)	(0.591)
NudgeOwn (β3)	0.466	-1.397**	0.644*	-1.316**
	(0.309)	(0.544)	(0.342)	(0.554)
High (β4)	0.155	-1.698**	0.25	-2.093***
	(0.365)	(0.643)	(0.404)	(0.654)
Advantaged (β5)	0.372	-0.823	0.523	-0.904
	(0.310)	(0.546)	(0.343)	(0.556)
Advantaged × NudgeOther (β6)	-0.319	1.014	-0.115	1.526*
	(0.432)	(0.761)	(0.478)	(0.774)
Advantaged × NudgeOwn (β7)	-0.133	0.777	-0.324	0.924
	(0.393)	(0.692)	(0.435)	(0.704)
Advantaged × High (β8)	-0.071	1.294*	-0.19	1.699**
	(0.431)	(0.759)	(0.476)	(0.772)
Additional covariates				
Stratification variables	Yes	Yes	Yes	Yes
Other health volunteer characteristics	Yes	Yes	Yes	Yes
R-squared	0.46	0.31	0.47	0.30
N	69	69	69	69

Note. The Greek letters next to the variable names correspond to those in equations (3) and (4). * p<0.10, ** p<0.05, *** p<0.01. Standard errors are in parenthesis. Stratification variables include ethnicity, experience and education. Other health volunteer characteristics include age, annual household income, number of households the health volunteer visited in the past month, the amount of money she received for working as a health volunteer, distance to the nearest health center, primary occupation and the share of own-type households in the health volunteer's ward.

Appendix 1.E. Effect of Incentives on the General Mix of Clients

Table 1.E1. Effect of Incentives on the Type of Patients Reached by the Health Volunteers

					p-values from the test of difference in means		
	Low *(1)*	*NudgeOther* *(2)*	*NudgeOwn* *(3)*	*High* *(4)*	All arms	*(2) vs* *(3)*	*(1)* and *(4)* vs. *(2)* and *(3)*
Age, years	53.08	52.28	50.11	53.11	<0.01	<0.01	<0.01
	(0.48)	(0.57)	(0.49)	(0.49)			
Women, proportion of total	0.61	0.55	0.65	0.60	0.01	<0.01	0.90
	(0.02)	(0.02)	(0.02)	(0.02)			
Married clients, proportion of total	0.90	0.89	0.89	0.88	0.55	1.00	0.99
	(0.01)	(0.01)	(0.01)	(0.01)			
Education, years	3.84	4.04	4.54	3.98	0.04	0.07	0.03
	(0.19)	(0.20)	(0.18)	(0.20)			
Distance to the health center, minutes	25.43	29.99	27.93	24.21	<0.01	0.17	<0.01
	(0.91)	(1.26)	(0.87)	(0.97)			
Farming as main occupation, proportion	0.82	0.83	0.81	0.83	0.72	0.38	0.86
	(0.02)	(0.02)	(0.02)	(0.02)			
Had heard about diabetes, proportion	0.60	0.62	0.63	0.58	0.38	0.68	0.10
	(0.02)	(0.02)	(0.02)	(0.02)			

Note: All numbers in this table are for individuals who received a referral card from a HV and showed up for the checkup. All variables are self-reported.

Chapter 2

Early Childhood Nutrition to Adult Outcomes: An Exploration of Mechanisms, Duration of Exposure, and Heterogeneous Effects

2.1 Introduction

There is now a critical threshold of evidence documenting the relationship between one's exposure to shocks in early life and outcomes in adulthood. In a seminal review article, Currie and Vogl (2012) summarize the work done so far on the relationship between early-life nutrition, famine, rainfall, pollution, disease and war, and long-term health outcomes, primarily height. Based on their extensive review, they argue that "[F]uture research should focus on identifying pathways and mechanisms; measuring the relative magnitudes of the effects of different health shocks; examining interactions between shocks; and revisiting the question of critical periods" (p. 29). This paper contributes to the existing literature responding to that call. The first goal of this paper is to examine the effect of an early-life nutritional intervention on health outcomes that the intervention is intended to affect directly as evidenced by the medical literature, thereby elucidating on a clear mechanism. A second goal is to assess the effect of age at first exposure to the intervention, again on expected health outcomes, to get at the role of critical periods. In many countries with a history of discrimination and unequal access to resources based on gender and ethnicity, it is natural to expect different effects of the program on these dimensions. Therefore, the final goal is to evaluate heterogeneous effects of the program by gender and by ethnicity. In addition to the effect on health outcomes, I also evaluate the effect on education outcomes to check whether the findings here are consistent with the vast amount of literature showing that healthier children tend to be healthier adults with better educational and labor market outcomes.

I make use of a vitamin A supplementation program in Nepal, which the government of Nepal implemented with funding support from development partners, primarily the United States Agency for International Aid (USAID). The program was rolled out in a sequential manner across districts between 1993 and 2001 and targeted 6-60 months old children. Children who were already five years of age at the time of program's implementation did not benefit from the program. The main empirical strategy in this paper capitalizes on this rollout and its differential effect on children of different ages within the same household based on the geographic location (district) of their birth. I link individuals from the 2011 census to their district and year of birth, which allows me to determine whether they benefited from the program and, if they did, the age which they were exposed to the program. The self-reported measures in 2011 are used as the long-term outcomes.

Immediate health effects are likely the most important channels through which early-life interventions such as this supplementation program can affect educational and labor market outcomes. The primary result of vitamin A deficiency is blindness. With secondary effects such as reduced immune system, temporary low vision can easily translate to permanent disability if the condition is not treated on time. Therefore, I start by evaluating the effect of the program on blindness and disability. I then evaluate the effect on several educational outcomes, as discussed in Section 2.4.

To preview the results, I find that the program reduced the probability of having a disability or blindness, kept children in school longer, and enabled them to complete different grades by an expected age. The positive effects on disability and education seem to have improved marriage prospects, as reflected in the individual's marital status in 2011. The program also had different effects on individuals based on their timing of the exposure to the program, with a longer exposure usually strengthening the positive effects. As expected, effects also differed by the individual's gender and ethnicity. They were more pronounced for men and individuals from traditionally advantaged ethnic groups.

2.2 Nepal's Vitamin A Supplementation Program

The details of Nepal's vitamin A Supplementation Program (hereafter, the program), including how it was conceived, have been discussed extensively elsewhere (e.g., Thapa, Choe and Retherford, 2005; Center for Global Development, 2014). Therefore, I provide only a summary here. The program evolved from an extensive consultation between Nepal's government, development partners and the health community and a recognition of prior evidence—mostly from other countries—illustrating that vitamin A supplementation can help reduce child mortality significantly. The program's primary goal was to reduce child mortality and morbidity related to vitamin A deficiency by providing twice-yearly supplements of vitamin A capsules to children who were 6-60 months old; treating xerophthalmia, severe malnutrition, prolonged diarrhea and measles; and encouraging dietary intake of vitamin A and breast-feeding (United Nations Children's Fund (UNICEF), 2003). The program started in October 1993 in eight districts and was subsequently rolled out to other districts, covering 69 of the 75 districts by April 2001 (Appendix 2.A1). The program was implemented by the government with support primarily from UNICEF and USAID.

The key vehicles of this program were the Female Community Health Volunteers (FCHVs). The FCHVs were trained to identify children in their communities; provide nutritional information to community members; and to mobilize local groups (such as mothers' groups and farmers groups) to encourage participation by communities. Currently, there are nearly 48,000 FCHVs in the countries, performing tasks that range from raising awareness about preventive health to delivering basic healthcare functions (Andersen *et al.*, 2013).

The take-up rate was high. In the first year of implementation, 6,500 FCHVs provided vitamin A capsules to 470,000 children, representing 90 percent of the target population in eight districts. By 1995, 86% of all children below the age of five had received supplementation in 23 districts. Seventy of the 75 districts were covered by 2001.

2.3 Health Effects of Vitamin A Deficiency

As mentioned in Section 2.1, the primary result of vitamin A deficiency is blindness. If the human body lacks vitamin A in sufficient amount, a condition called xerophthamia or a dryness of eyes develops. This condition manifests first as night blindness and progresses into softening of the cornea and total blindness if the vitamin A deficiency continues.

More generally, vitamin A deficiency is associated with the weakening of the tissues and the immune system, both translating to greater risk of respiratory, measles and diarrheal morbidity, and subsequently mortality. In fact, vitamin A deficiency affects about 21% children below the age of 5 years in developing countries and leads to the deaths of over 800,000 women and children each year (West, 2002). Furthermore, vitamin A deficiency is responsible for 20–24 percent of global child mortality from measles, diarrhea, and malaria and for 20 percent of all cause maternal mortality (Rice et al., 2004). It also increases the severity and fatality of measles (Sommer and West, 1996).

The primary food sources of vitamin A are ripe yellow fruits; carrots, spinach and green leafy vegetables; and animal products such as eggs, milk and liver. Recognizing the critical role these food items play in strengthening body functions, the World Health Organization has included a separate category to reflect intake of these items in its Diet Diversity Index. In absence of sufficient vitamin A intake through these sources, children 6-11 months of age are recommended to receive an oral dose of 100,000 International Units (IU), and children 12–59 months of age are recommended to receive a 200,000 IU dose every four to six months (Rose, 2002). There are recommended doses for adults as well, but are not provided here since the focus of this paper are children.

From this discussion, in addition to blindness and disability, ideal measures of health effects would also include respiratory, measles and diarrhea-related mortality specifically caused by vitamin A deficiency. However, such data are not available in the census, the primary source of data for this paper and discussed in the next section.

2.4 Data

The primary data source used in this paper is Nepal’s National Population and Housing Census 2011. Like many national censuses, this census collected information on demographics, education, housing, asset ownership and employment from all individuals living in the country at the time of the survey.

A 15 percent sample of the census was obtained from the Central Bureau of Statistics of Nepal. Of the 4,037,885 individuals included in the sample, information on the district of birth was missing for 127,456 individuals. These individuals were dropped because without information on the district of birth, it was not possible determine whether they were exposed to the program or, if they were, the timing of the exposure. Another 6,597 individuals were dropped because they were foreign citizens. For the remaining 3,903,832 individuals, their birth year was calculated as the difference between 2011 and their current age, and their age at first exposure was calculated as the difference between the year when the program was rolled out in their district and the birth year. Children whose age at first exposure was less than five completed years are considered to be treated through the program.

I limit the sample further in two other ways. First, I include only those individuals who were between 13 and 22 years of age in 2011. Everyone below 13 years in the sample was treated through the program, while everyone above 22 was not treated. In the remaining sample of 736,392 individuals, the age at exposure ranges from -4 to 13 years (i.e., those who were born four years after the program rollout to those who were already 13 years old at the time of the rollout). Second, when estimating equation (2), I limit the analysis to children who were already born at the time of the program rollout in their district. I do so because the program had a nutritional component through which mothers were encouraged to consume healthier food. Some of the long-term effect observed on the health of children who were not already born at the time of the program rollout can be through the health of the mother and not directly through the vitamin A supplementation administered to the child. To parse out the two effects (the effect through the mother's health and the one directly through the vitamin A supplementation received by the child), it is important to limit the analysis to children who were already born and hence were not affected by mother's diet. Given that different samples are used for different outcomes, in Appendix Table 2.A1, I provide a table showing how the final sample of each outcome was derived.

I look at blindness and disability as the primary intermediate outcomes as those are the health conditions that vitamin A directly influences. For each individual, the census asks the form of disability the individual has. Specifically, it asks "What the physical and mental disability of (name)?" and provides nine options, including 'Not disabled'. For this paper, an individual is categorized as having vision disability if he or she indicated having "Blind/low vision". The individual is categorized as having a disability in general if he or she indicated having some form of disability, including "mentally disabled" or "speech problem."

For education outcomes, I look at whether the individual is currently in school. Here, the oldest individual is 22 years of age. Nepal does not have a culture of taking a year off after completing high school (equivalent to sophomore year of high school in the United States). Individuals who do relatively well tend to continue their higher education without an interruption. If an individual starts formal schooling at age six and progresses through the education system without repeating a grade, the individual would still be at a university until age 22. Therefore, being in school would reflect a positive effect of the program.

One downside of the measure above is that individuals may remain in school simply because their performance was weak which made them repeat grades. In absence of a good measure for cognitive ability which the program may have potentially improved, I also examine the effect of the program on whether the individual is in a grade appropriate for his or her age. I do so for children who are in high school (grade 12) or lower. I define the education-for-age (a binary variable) as one if the child has been in grade one by age six and did not repeat a class (that is, the child was at least in grade 2 by age 7, grade 3 by age 8, and so on). Finally, for those who left school, I look at the highest grade they completed before leaving.

2.5 Identification Strategy

In order to evaluate the long-term effects of exposure to the vitamin A supplementation program, I rely on a method similar to the one adapted by the vast number of studies that have evaluated the long-term effects of early-life exposure to different schooling environments—starting with Duflo (2001). Specifically, I compare the long-term effects for children born in the same household around the time of the program rollout. As mentioned before, the program was rolled out across districts between 1993 and 2001, and children under five years were treated. Because of this arrangement, children born in the same household either benefited from the program or did not, depending on their year of birth and the timing of the program rollout in their district. Consider two children born, in 1991 and 1993, to a mother in Arghakhanchi district in the western part of the country. The program started in that district in October 1997. The child born in 1991 would be more than five years old in 1997, so would not benefit from the program. The other child, born in 1993, would be less than five and therefore would benefit from the program. In 2011, the child born in 1991 would be 20 years old while the one born in 1993 would be 18. I compare the outcomes—as measured in 2011—for these two individuals. In econometrics terms, I utilize the within-household variation in exposure to the program and the outcomes reported in 2011 to estimate the causal effect of the early-life vitamin A supplementation on long-term outcomes.

Although one can generally assume that children born within the same household face similar external environments, such as parental care and health risks, there are two threats to identification one needs to address. The first is that, irrespective of the exposure to the program, girls might be treated differently than boys, which could lead to different long-term outcomes for men and women. The discrimination against girls in South-Asian societies and its implications for the girls' health has been widely documented (e.g., Jayachandran and Kuziemko, 2011). To account for such differences, I control for the gender of the child in all specifications. The second threat is that children born to the same household may be exposed to different external environments based on their year of birth. For example, parental employment status and incomes could change over time, possibly altering the time and money investments that they can make on their children. Macro-level changes, such as budget allocation to the health sector, could affect the type of health services the two children receive. To account for such differences, I control for the birth year of the child in all specifications. The regression equation I estimate, then, is:

$$Y_{ijt} = \pi_0 + \pi_1 BelowFive_{ijt} + \pi_2 Male_{ij} + \theta_j + \eta_t + v_{ijt} \dots\dots\dots (1)$$

In this equation, Y_{ijt} is the outcome measured in 2011 for a child *i* born to household *j* in year *t. BelowFive* is a binary variable which equals one if the child was less than five years old in the year the program was rolled out in his or her district. *Male* is a binary indicator for the gender of the child. In this equation, π_1 is the key coefficient of interest and captures the relationship between child's exposure to the program and his or her outcomes later in life (specifically in 2011). The expected sign on π_1 depends on the outcome. When the outcome is disability, for example, the expected sign is negative because we expect the program to reduce the probability of having a disability. θ_j captures the time-invariant characteristics of the household, while η_t captures effects specific to the child's birth year. v_{ijt} is the usual error term, while π_0 reflects the baseline outcome after accounting for gender, birth year and the household-specific factors. Standard errors are clustered at the household level.

Given the inclusion of gender and birth year in the regressions, the key identifying assumption is that, without exposure to the program, two children who are of the same gender and born in the same year within a household would have similar long-term outcomes. The characteristics of the household faced by two children are more likely to vary when the window between births is wider. In robustness check, I conduct additional analysis by limiting the sample to children who are either four or six at the time of the program rollout in their districts, thus comparing long-term outcomes for children who were born very close apart.

In order to evaluate the effect of the age at exposure to long-term outcomes, I estimate an equation similar to equation (1) above but instead of a binary variable *BelowFive* to represent treatment status, I include multiple indicators for the treated children. The equation I estimate is:

$$Outcome_{ijt} = \pi_0 + \delta AgeAtExposure_{ijt} + \pi_2 Male + \theta_j + \eta_t + v_{ijt} \text{……….(2)}$$

The notations in equation (2) are the same as in equation (1). The only difference between the two equations is that in equation (2), *AgeAtExposure* consists of several binary indicators representing the age at which the child was first exposed to the program. The indicators are 'from birth to one year', 'one to two years', 'two to three years', 'three to four years' and 'four to five years'. All untreated children, i.e., those who were already more than five years old at the time of the program rollout in their district, are in the excluded group. Using the example from earlier, the child born in 1991 in Arghakhanchi district would be in the excluded group because he or she would be six years old in 1997, while the child born in 1993 would be in the 'between four to five years' category. The identifying assumption here is that, after controlling for time-invariant characteristics of the household, birth year and the gender of the child, two children within a household differ only in terms of when they were exposed to the program and the difference in outcomes is only due to this difference in the timing of the exposure.

To evaluate the heterogeneous effect of the program by gender and ethnicity, I estimate equation (1) separately for men and women, and for individuals from advantaged and disadvantaged ethnic groups. The Nepalese government has categorized the country's more than 100 ethnicities into six main categories based on religion, caste and ethnicity, and further into advantaged and disadvantaged groups based on historical access to resources. In this study, I use these two broad categories.

2.6 Results

2.6.1 Main Results

The program had large, discernible effects on long-term health and education outcomes. Children who were exposed to the program in childhood were 0.05 percentage points less likely to be suffering from blindness in adulthood (Table 2). While this effect is only marginally significant, it represents a decrease in blindness by approximately 25 percent from the mean (=100*0.0005/0.002). Likewise, the program reduced disability in adulthood by approximately 14.5 percent from the mean (=100*0.0019/0.013). This effect on disability is statistically significant even at the one percent level. Note that these effects, as well as the effects discussed in the rest of the paper, are closer to being treatment-on-the-treated effects, rather than intent-to-treat ones, because of the high uptake of the program.

Given the reduced chances of being blind or disabled, it is not surprising that the children who benefited from the program were also more likely to be in school in 2011 than those who did not benefit. Based on Table 2.2, treated children were 0.41 percentage points (statistically significant at the 10 percent level) more likely to be in school at the time of the survey. In the sample, 64 percent of individuals were in school at the time of the survey. The program increased that rate by 0.6 percent. In the sample, approximately 25 percent children below the age of 19 are in a grade appropriate for their age. The program increased this rate by 1.5 percentage points, or approximately six percent of the mean (=100*1.3/25). The program, however, was inadequate to influence education attainment as measured by the highest level completed for the sample of individuals who had already left school. One possible explanation for this finding is that improved health—as measured by reduced probabilities of having blindness or a disability—also improves labor market prospects. In other words, some children may have left school to go to work. The data do not allow for a formal test of this explanation, however.

Children who were exposed to the program were more likely to be married in 2011 than those who were not. On average, the program increased the probability of being married by their age in 2011 by 2.1 percentage points (statistically significant at one percent level). In the sample, 23 percent individuals are married, which implies that the program increased marriage rate by approximately nine percent relative to the mean. Among those who were married, however, the program had had no effect on the age at which they got married.

2.6.2 Effect of the Duration of Exposure

As mentioned in the introduction, one of the goals of this paper is to examine the effect of the duration of exposure to the program on long-term outcomes. Relative to the untreated children, the timing of the exposure does not seem to affect the health-related outcomes, except marginally the disability status when the child was exposed to the program for either two or three years (Table 2.3).

However, the effects are visible on other outcomes—shown more clearly by the marginal effects' plots in Figure 2.1. Relative to the untreated group, the program improved the probability of staying in school when the child was exposed for three years and marginally when the child was exposed for all four years. The effect on meeting the education-for-age requirement is almost linear; the higher the duration of exposure to the program, the higher the probability of meeting the education-for-age requirement. Marriage prospects seem to change in the opposite direction. Relative to the untreated children, children who were exposed to the program for one year were more likely to be married by 2011 but the magnitude of the effect gradually falls with the duration of the exposure. By the time the children are exposed for four years, they end up having a lower chance of being married than the untreated children. This trend is consistent with the effect seen on the age at marriage for the sample of children who are already married. Relative to the untreated children, the age at first marriage is lower for treated children but the difference falls gradually with the duration of exposure; by the time the child is exposed for four years, he or she is likely to get married later than an untreated child.

2.6.3 Heterogeneous Effects by Gender and Ethnicity

Between men and women, the primary difference in long-term effects is in the probability of having a disability and the probability of staying in school (Table 2.4, panel A). The program did not reduce the probability of being disabled for women while the effect was substantively large and statistically significant for men. Specifically, for men, the program reduced the chances of being disabled in the long term by 0.3 percentage points at a base of about 1.3 percent (thus a reduction of almost 25 percent). It did not help men stay in school until the time of the census, however. In contrast, for women, the program increased the chances of staying in school by 0.7

percentage points at a base of about 64 percent (thus an increase of about one percent). For other outcomes, the program's effects were in the same direction for men and women.

The program also had different effects on individuals from advantaged and disadvantaged ethnic groups (Table 2.4, panel B). It reduced the probability of having a blindness for advantaged groups but not for disadvantaged groups. Specifically, for the advantaged groups, it reduced the probability by about 0.1 percentage point at a base of 0.2 percent (thus a reduction of about 45 percent). This effect is statistically significant at the five percent level. The program helped individuals from disadvantaged groups stay in school at the time of the census but not for the advantaged groups. The effect on the disadvantaged groups was 2.3 percentage points at a base of 64 percent (thus an improvement of about 3.5 percent. This effect is statistically significant at the one percent level. Similarly, the program helped individuals from disadvantaged groups to meet the education-for-age requirement, raising it by 1.2 percentage points or by about 5 percent from the mean of 25 percent.

By improving the probability of staying in school and meeting the education-for-age requirement for individuals in the disadvantaged group, the program also seems to have improved marriage prospects. Treated individuals from disadvantaged groups were more likely to be married at the time of the survey compared to the untreated individuals from the same group. The program had an opposite effect on individuals from the advantaged group. In this group, the treated individuals were less likely to be married—and those who were married had been married a year later—than untreated individuals.

The takeaway from the sub-group analysis by gender and ethnicity is that the program had different long-term effects based on individual's gender and status in the society. Health effects were clearly more pronounced for men and the advantaged ethnic groups. Education benefits were better for women and the disadvantaged ethnic groups. Marriage prospects improved as a result of the program for both men and women, although by different magnitude. There is also suggestive evidence that, for those who were already married, the age at first marriage was lower for treated men and women than those who were untreated. The difference in marital outcome is less clear between advantaged and disadvantaged ethnic groups.

2.7 Validity Check

One can raise a number of concerns about the validity of the main results discussed in the previous sections. The first concern can be that, although I estimate the effects using the district of birth, households might have migrated after the children were born, which can potentially affect the outcomes for younger (thus treated) and older (thus not treated) children differently. To address this concern, I estimated equation (1) using the sample of individuals who are currently living in the same district as the ones in which they were born. The overall effects remain largely unchanged (Table 2.5).

In addition, the chances of two children within the same household having different treatment status solely due to migration of the household are relatively low. Among those who migrated, 69 percent individuals moved to a district where the program was rolled out later than in their district of birth. Take a household with two children, a four year old and a six year old, living in a district where the program has not been rolled out. If the household migrated between the two births, the probability of them moving to a district where the program was already in place and, as a result, exposing the four year old to the program is only 31 percent. The probability of them moving to a district where the program was rolled out later—thus not exposing either of the children to the program—is 69 percent.

A second concern is selection. If the program helped reduce mortality, the marginal children saved might be less healthy. If this happened, in the treatment group, we would end up with a greater proportion of children who have a low vision or are disabled than we would without the program. In that case, our estimated effect would be lower than the actual effect. This is an important concern because the program has been found to reduce mortality substantially in the near-term (Thapa, Choe, and Retherford, 2005).

To assess the extent of this problem when evaluating long-term effects, I estimate the relationship between the size of the cohort and exposure to the program. More specifically, I estimate the following equation:

$$CohortSize_{aj} = \pi_0 + \pi_1 Treatment_{aj} + \theta_j + \eta_t + \ v_{ijt} \text{.........}(3)$$

In equation (3), the outcome is the natural log of the total number of individuals of age *a* born in district *j* who have survived until the time of the survey (i.e., 2011). *Treatment* is a binary variable which equals one if a certain age group in district *j* was exposed to the program in the district (therefore, it varies by age and by district). District fixed effects θ_j control for time-invariant district characteristics that affect mortality, while birth year fixed effects η_t control for district-invariant year-specific factors. The standard errors are clustered at the district level to allow arbitrary correlation between observations within a district.

Cohort size, measured as the total number of a certain age who are currently living, is a measure of survival or cumulative mortality (Jayachandran, 2009; Miller and Urdinola, 2010). To calculate the cohort size here, the population at each age group (at annual intervals) was aggregated to the district level. Individuals born between 1989 and 1997 (the same group of individuals as those in the analytic sample for the main analysis) are included. Data on nine age groups in 75 districts yields a sample of 675 observations, which is the effective sample size for estimating equation (3).

The coefficient of interest (π_1 in equation (3)) is -0.04 and statistically significant at the five percent significance level (not shown in a separate table). The coefficient implies that the population of the treated cohorts was four percent lower than the population of the untreated cohort [=100-(100*exp(0.04))]. As shown by the plots shown in Figure 2.2, which were generated from the regression equation, the cohort size was lower for the treatment group than for the untreated group at each birth year. At least in the long-term, therefore, the effects discussed in the main section are not due to selection on mortality; if anything, long-term mortality in the sample has risen (by a statistically significant amount of approximately four percent).

2.8 Caveats and Conclusion

In this paper, I evaluated the long-term effect of a vitamin A supplementation program in Nepal on primarily health and education outcomes. Although the program's effects varied by the timing of exposure, by gender and ethnicity of the individual, in general it had positive effects on the health and education outcomes.

The findings presented here should be understood in light of a number of caveats, however. First, the two intermediate health outcomes evaluated here – poor vision and disability – and the educational outcomes are only a few of the many components of individuals' wellbeing. Evaluation of the program using a more exhaustive list of outcomes was not possible given the limited data that census collects. It was not possible to replicate the same analysis using other sources from Nepal, including the 2011 Nepal Living Standard Survey (NLSS).

Although the NLSS collects information on a wide range of indicators, including disability, overall health and employment status, the sample size (6,000 households) is much lower than what would be required to detect an effect on these outcomes. It would also be difficult to identify individuals' exposure to the program using the NLSS because the information on the district of birth is not available. A natural next step, therefore, is to conduct a similar analysis with different outcomes, including income and employment (which are not evaluated here because the adults are still young to be employed), once the next round of the census becomes available.

Second, the identification strategy relies on the assumption that a household—and multiple children born into it over time—face similar external environments over time, which may not always be true. One possible threat to this identification originates in the predominantly agrarian nature of the Nepalese economy. Nearly 70 percent of Nepal's population is employed in agriculture which contributes 34 percent to the Gross Domestic Product (International Labor Organization, 2016). Households rely heavily on rainfall for irrigation. Fluctuations in rainfall and the resulting variation in food availability can easily alter a household's investments on children's health depending on when they are born. Maccini and Yang (2009) examine the impact of early-life rainfall on a range of adult outcomes in Indonesia and find that higher early-life rainfall leads to improved health, schooling and socioeconomic status for women. In the current study, if different children within a household are exposed to different rain-fall shocks, their long-term education and health outcomes might differ, irrespective of their exposure to the vitamin A program. Birth year fixed effects should capture some of this effect. Nonetheless, subsequent further analysis can be strengthened by controlling for weather (mainly, rainfall and temperature around the time of birth) and crop output as well.

Finally, although the data are nationally representative, one will need to be cautious when extrapolating the findings to other countries, especially ones with a much lower prevalence of disability, vision problems and higher educational status. It is also important to recognize that the effect of the program is estimated off of the households where there are at least two children—with one exposed to the program and the other not exposed to it. In this study, 25.2 percent of households meet that requirement. The program's effect is essentially driven by the health and education outcomes of children in these households, although there is no strong reason to believe that these households are different than the overall national sample to threaten external validity.

Despite these limitations, this study has generated important insights on the intermediate health outcomes through which an early-life intervention can affect other outcomes. To my knowledge, this is the first study to evaluate the long-term effects of an early-life intervention from Nepal in a way that allows for a causal interpretation of the estimated effects (under certain assumptions discussed in Section 3.5). By evaluating the program's effects of health outcomes that are directly related to the intervention (here, blindness and disability), I have established that one primary channel through which nutritional interventions affect educational outcomes is health. I have also shown that long-term effects can be different for different segments of the population. Strikingly, other studies have found that coverage rates of the program are similar for boys and girls (Thapa, 2010). This implies that long-run effects can be different even when boys and girls are equally likely to be treated. When designing new programs, differential effects the program can have on long-term outcomes to different segments of the population is an important consideration the policymakers should keep in mind. I am unaware of any study that has evaluated differences in long-term outcomes by ethnicity. The findings from this paper suggest that the effects—even on outcomes like blindness which the program directly catered to—varied by ethnic status.

References

Andersen, K., Singh, A., Shrestha, M.K., Shah, M., Pearson, E., Hessini, L., 2013. Early pregnancy detection by female community health volunteers in Nepal facilitated referral for appropriate reproductive health services. *Global Health: Science and Practice* 1, 372–81.

Center for Global Development, 2014. Case 4: Reducing child mortality through vitamin A in Nepal. URL (accessed 26 April 2017): http://www.cgdev.org/page/case-4-reducing-child-mortality-through-vitamin-nepal.

Currie, J., Vogl, T., 2013. Early-Life Health and Adult Circumstance in Developing Countries. *Annual Review of Economics* 5, 1–36.

Duflo, E., 2001. Schooling and Labor Market Consequences of School Construction in Indonesia: Evidence from an Unusual Policy Experiment. *American Economic Review*, 91(4): 795– 813.

Jayachandran, S., 2009. Air Quality and Early-Life Mortality during Indonesia's Massive Wildfires in 1997. Journal of Human Resources, 44(4):916–954.

Jayachandran, S., Kuziemko, I., 2011. Why do mothers breastfeed girls less than boys? Evidence and implications for child health in India. *Quarterly Journal of Economics*, 126:1485–538.

Maccini, S., Yang, D., 2009. Under the Weather: Health, Schooling, and Economic Consequences of Early-Life Rainfall. *American Economic Review,* 99(3):1006-26

Miller, G., Urdinola, P., 2010. Cyclicality, Mortality, and the Value of Time: The Case of Coffee Price Fluctuations and Child Survival in Colombia. Journal of Political Economy, 118: 113-155.

Rice, A.L., West, K.P., Black, R.E., 2004. Vitamin A deficiency. In: Ezzati, M., Lopez, A.D., Rodgers, A., Murray, C., eds. *Comparative Quantification of Health Risks: Global and Regional Burden of Disease Attributable to Selected Major Risk Factors.* Vol. 1. Geneva, Switzerland: World Health Organization; 211–256.

Ross, D.A., 2002. Recommendations for Vitamin A supplementation. *Journal of Nutrition,* 132(9 suppl): 2902s–2906s.

Sommer, A., West, K.P., 1996. *Vitamin A Deficiency: Health, Survival and Vision.* New York, NY: Oxford University Press.

Thapa, S., Choe, M. K., Retherford, R. D., 2005. Effects of vitamin A supplementation on child mortality: evidence from Nepal's 2001 Demographic and Health Survey. *Tropical Medicine & International Health*, 10: 782–789.

Thapa, S., 2010. Nepal's vitamin A supplementation programme 15 years on: Sustained growth in coverage and equity and children still missed. *Global Public Health: An International Journal for Research, Policy and Practice*, 5(4): 325-334.

The United Nations Children's Fund, 2003. Getting to the roots – mobilizing community volunteers to combat Vitamin A Deficiency Disorders in Nepal. URL (accessed 26 April 2017): https://www.unicef.org/rosa/Getting.pdf.

West, K.P. 2002. Extent of vitamin A deficiency among preschool children and women of reproductive age. *Journal of Nutrition,* 132:2857S–2866S.

Tables and Figures

Table 2.1. Summary Statistics for the Overall Sample

	N	Mean	SD
Has a blindness	716,283	0.002	0.045
Has a disability	724,406	0.013	0.114
Currently in school	708,075	0.64	0.48
Meeting education-for-age requirement	443,724	0.25	0.43
Highest grade completed before leaving school	182,486	7.6	3.0
Married	733,586	0.23	0.42
Age at first marriage	170,094	17.2	2.3
Gender (male = 1)	733,586	0.48	0.50
From advantaged ethnic group	730,386	0.39	0.49

Source: Nepal Housing and Population Census 2011

Table 2.2. Linear Probability Model Results for the Effect of Exposure to the Program on Outcomes

	Blindness	Disability	
Treatment	-0.0005*	-0.0019***	
	(0.0003)	(0.0007)	
N	716,283	724,406	
R-squared	0.0001	0.0006	
F statistic	2.8	17.4	
	Currently in school	**Education-for-age**	**Grades completed**
Treatment	0.0041*	0.0150***	0.0483
	(0.0022)	(0.0031)	(0.0340)
N	708,075	443,724	182,486
R-squared	0.20	0.18	0.05
F statistic	5,951	2,796	186
	Married	**Age at first marriage**	
Treatment	0.0213***	0.0131	
	(0.0019)	(0.0212)	
N	733,586	170,094	
R-squared	0.25	0.55	
F statistic	6,422	6,277	

* p<0.10, ** p<0.05, *** p<0.01

Standard errors are clustered at the household level

Table 2.3. Linear Probability Model Results for the Effect of Age at Exposure on Health Outcomes

	Blindness	Disability	Currently in school	Education for age	Married	Age at marriage
Exposed for one year	0.0000	0.0004	0.0013	0.0080**	0.0216***	-0.0841***
	(0.0003)	(0.0008)	(0.0024)	(0.0034)	(0.0020)	(0.0199)
Exposed for two years	-0.0004	-0.0015*	-0.0020	0.0091**	0.0146***	-0.0678***
	(0.0003)	(0.0008)	(0.0026)	(0.0036)	(0.0021)	(0.0214)
Exposed for three years	-0.0005	-0.0015*	0.0080***	0.0092**	0.0095***	-0.0576**
	(0.0004)	(0.0009)	(0.0028)	(0.0040)	(0.0023)	(0.0227)
Exposed for four years	-0.0005	-0.0016	0.0066*	0.0121**	-0.0118***	0.0514*
	(0.0004)	(0.0010)	(0.0035)	(0.0050)	(0.0030)	(0.0276)
Constant	0.0023***	0.0107***	0.8395***	0.3834***	0.5908***	18.1162***
	(0.0005)	(0.0011)	(0.0032)	(0.0043)	(0.0023)	(0.0186)
N	570,611	577,228	563,292	331,457	583,970	162,306
R-squared	0.0001	0.0008	0.1704	0.1952	0.2119	0.5378
F statistic	2.0	12.6	2689.6	1560.9	2993.2	4459.2

* p<0.10, ** p<0.05, *** p<0.01

Standard errors are clustered at the household level

Comparison group are children who were not exposed to the program because of their age

All regressions include gender and birth year fixed effects

Table 2.4. Heterogeneous Effects of the Exposure to the Program, by Gender and Ethnicity

		Blindness	Disability	Currently in school	Education for age	Grade completed	Married	Age at marriage
A. By ethnicity								
Advantaged groups		-0.0009**	-0.0021*	0.0003	0.0009	0.0376	-0.0111***	0.0974**
		(0.0004)	(0.0011)	(0.0034)	(0.0046)	(0.0621)	(0.0028)	(0.0389)
	N	275,951	279,306	274,297	199,622	60,551	282,268	51,962
Disadvantaged groups		-0.0002	-0.0015*	0.0231***	0.0120***	0.0668	0.0259***	0.0177
		(0.0004)	(0.0009)	(0.0030)	(0.0042)	(0.0414)	(0.0025)	(0.0261)
	N	437,226	441,969	430,773	242,568	121,059	448,118	117,335
B. By gender								
Men		-0.0005	-0.0033**	0.0058	0.0282***	0.0901	0.0175***	-0.0117
		(0.0006)	(0.0014)	(0.0043)	(0.0060)	(0.0639)	(0.0033)	(0.1257)
	N	342,630	347,239	340,297	227,837	86,906	351,648	45,791
Women		-0.0001	-0.0010	0.0073*	0.0138**	0.0336	0.0346***	-0.0320
		(0.0005)	(0.0012)	(0.0042)	(0.0060)	(0.0756)	(0.0040)	(0.0653)
	N	373,655	377,172	367,795	215,853	95,596	381,938	124,406
Overall mean		0.002	0.013	0.64	0.25	7.6	0.23	17.2

* p<0.10, ** p<0.05, *** p<0.01
Each coefficient shown above is from a separate regression.
Standard errors are clustered at the household level
All regressions include birth year fixed effects

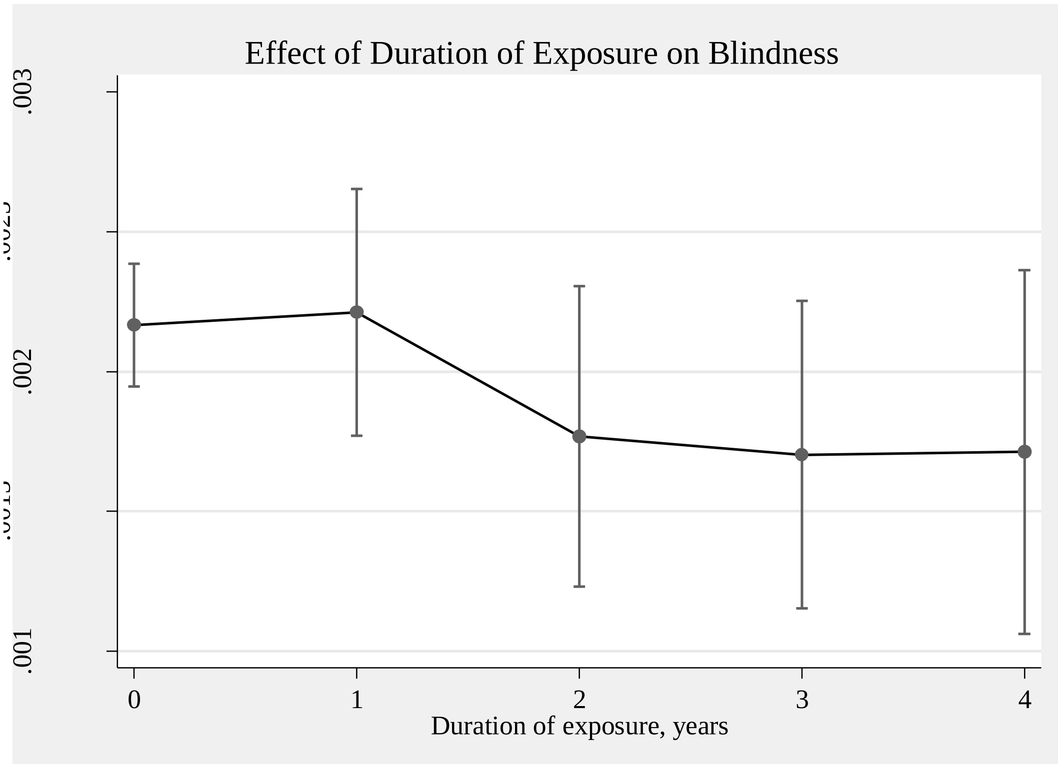
Table 2.5. Linear Probability Model Results for the Effect of Exposure to the Program on Outcomes for the Non-migrating Sample of the Population

	Blindness	Disability	
Treatment	-0.0004	-0.0021***	
	(0.0003)	(0.0008)	
N	633,105	640,684	
R-squared	0.0001	0.0006	
F statistic	2.0	14.5	
	Currently in school	Education-for-age	Grades completed
Treatment	0.0084***	0.0128***	0.0414
	(0.0025)	(0.0034)	(0.0387)
N	626,320	395,877	154,729
R-squared	0.20	0.16	0.05
F statistic	5470	2188	166
	Married	Age at first marriage	
Treatment	0.0225***	0.0190	
	(0.0020)	(0.0247)	
N	649,152	142,521	
R-squared	0.25	0.54	
F statistic	5504	4923	

* p<0.10, ** p<0.05, *** p<0.01

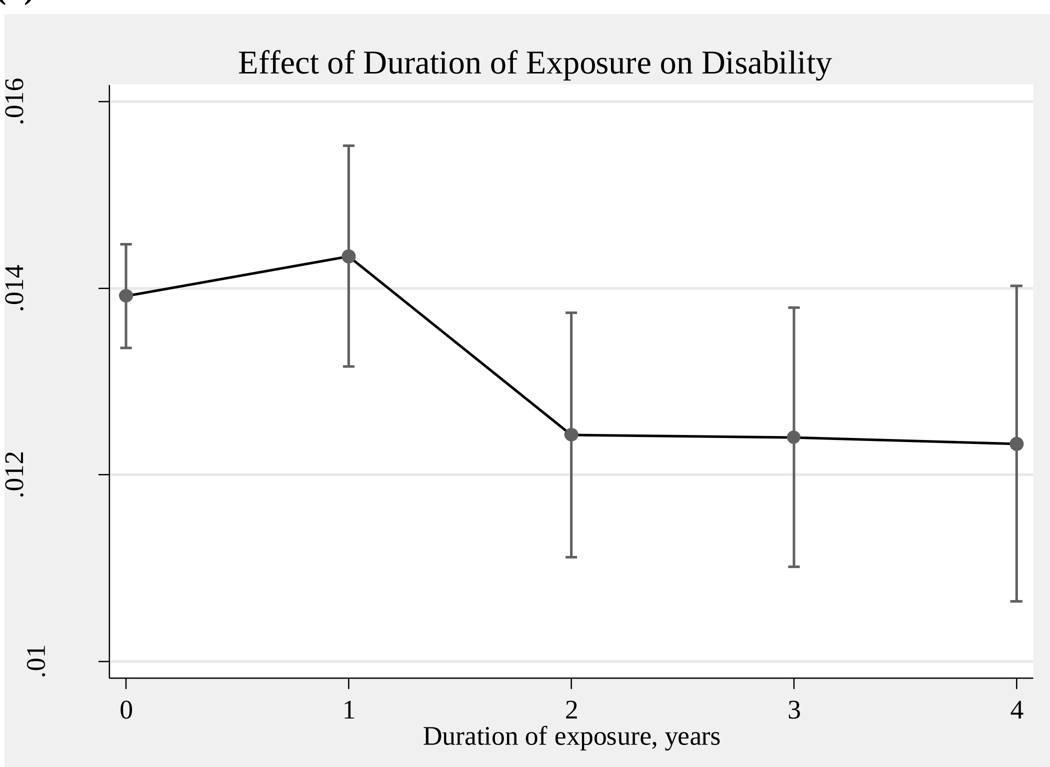
Standard errors are clustered at the household level

Figure 2.1.
(a)



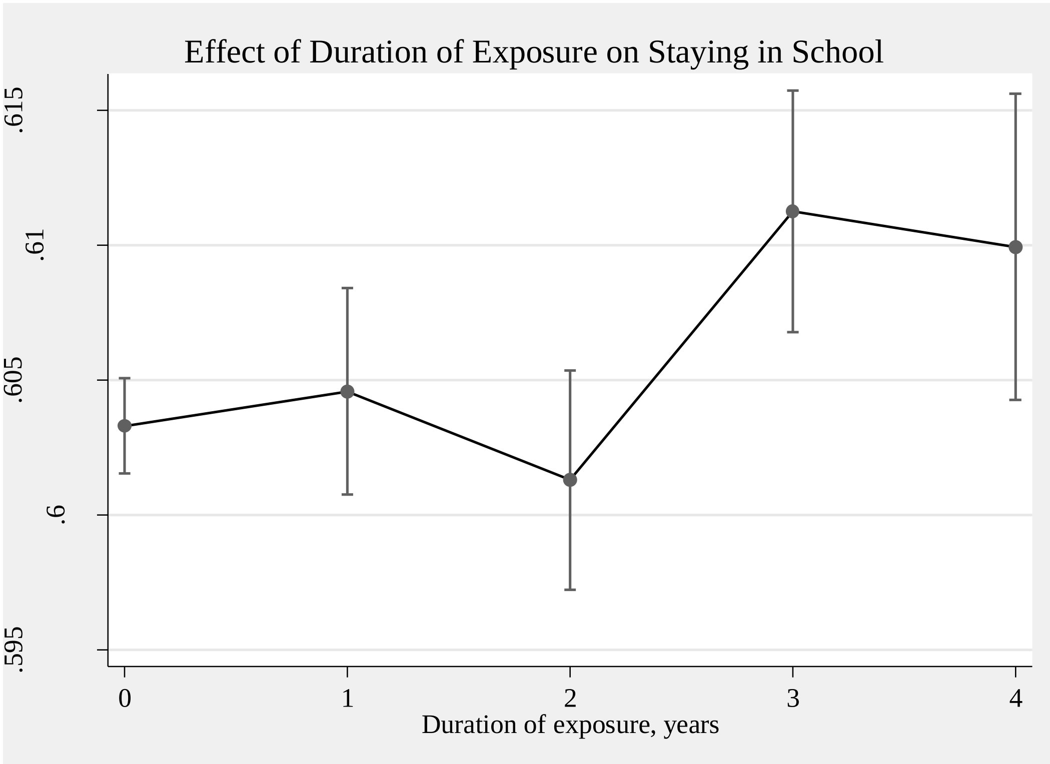
Effect of Duration of Exposure on Blindness
.003
.0025
.002
.0015
.001
0
1
2
3
4
Duration of exposure, years

(b)



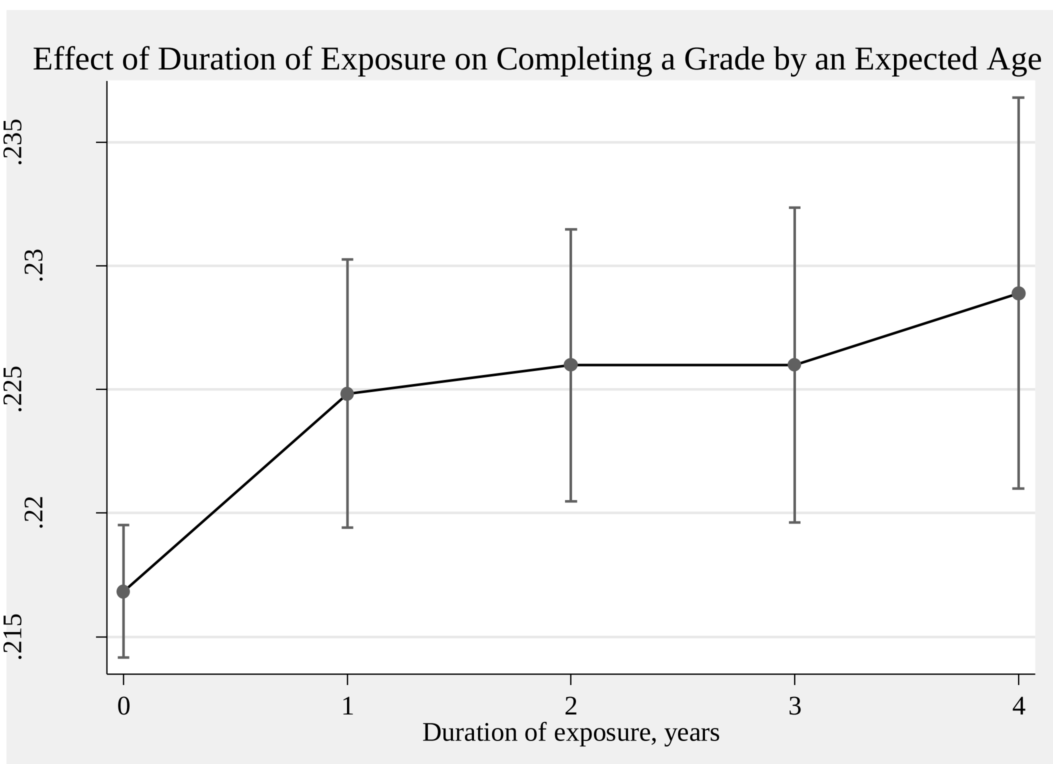
Effect of Duration of Exposure on Disability
.016
.014
.012
.01
0
1
2
3
4
Duration of exposure, years

(c)



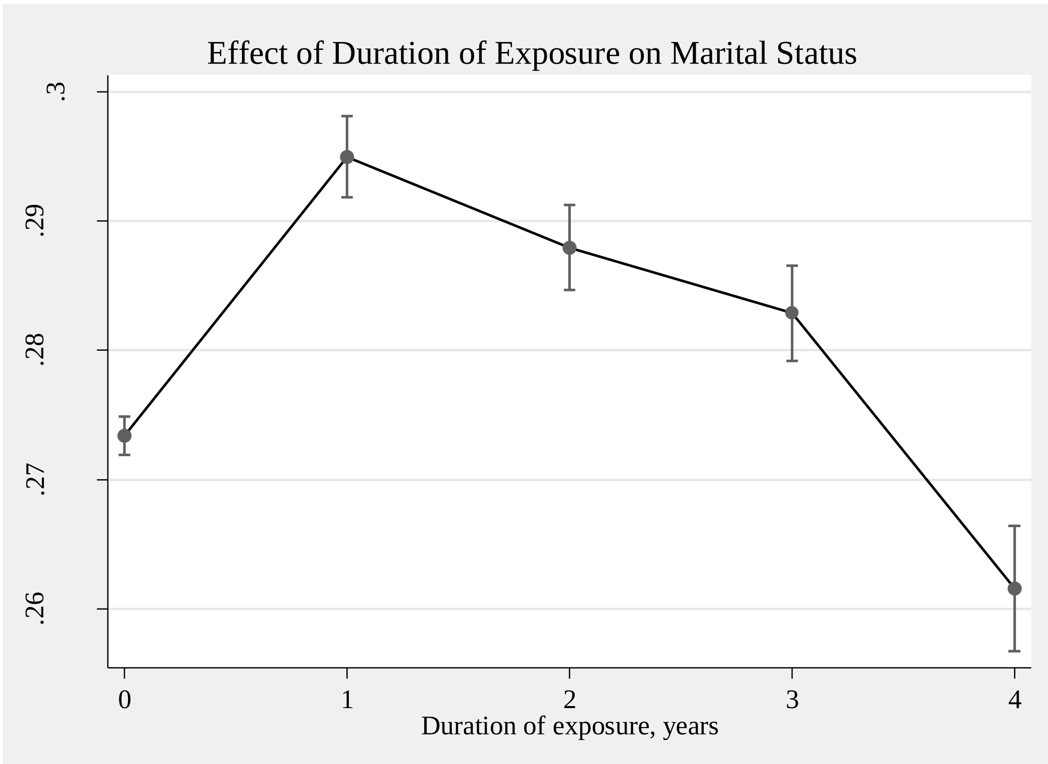
Effect of Duration of Exposure on Staying in School
.595
.6
.605
.61
.615
0
1
2
3
4
Duration of exposure, years

(d)



Effect of Duration of Exposure on Completing a Grade by an Expected Age
.215
.22
.225
.23
.235
0
1
2
3
4
Duration of exposure, years

(e)


Effect of Duration of Exposure on Marital Status
.3
.29
.28
.27
.26
0
1
2
3
4
Duration of exposure, years

(f)

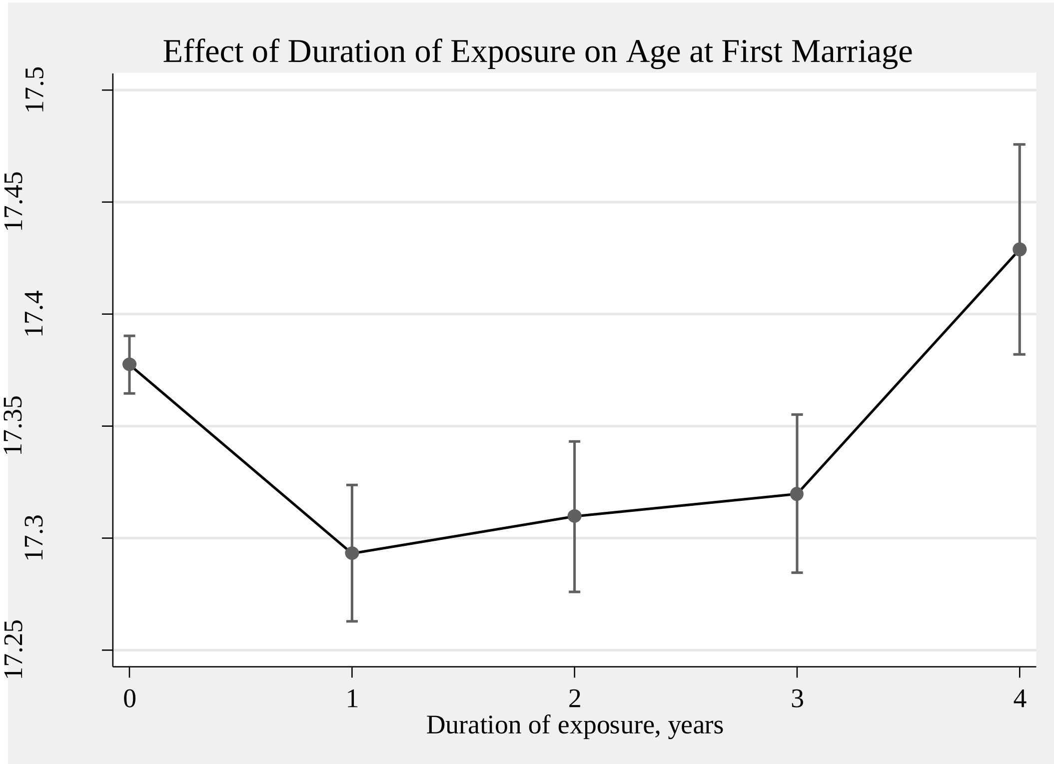
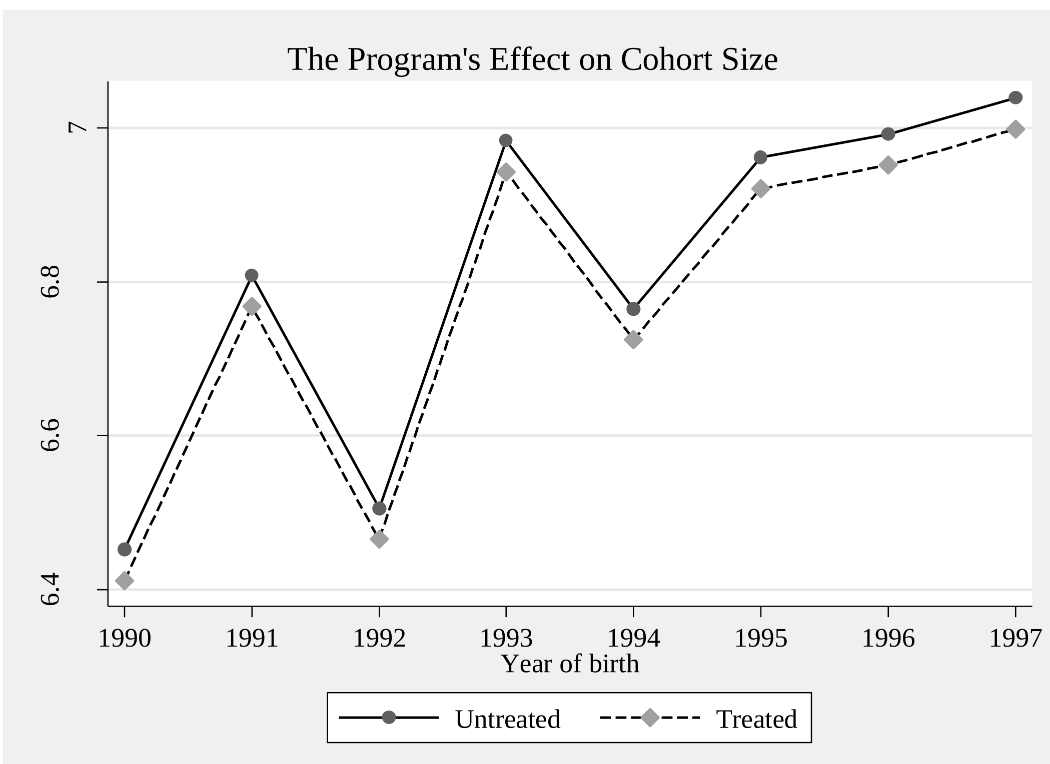

Effect of Duration of Exposure on Age at First Marriage
17.5
17.45
17.4
17.35
17.3
17.25
0
1
2
3
4
Duration of exposure, years

Figure 2.2


The Program's Effect on Cohort Size
7
6.8
6.6
6.4
1990
1991
1992
1993
1994
1995
1996
1997
Year of birth
Untreated
Treated

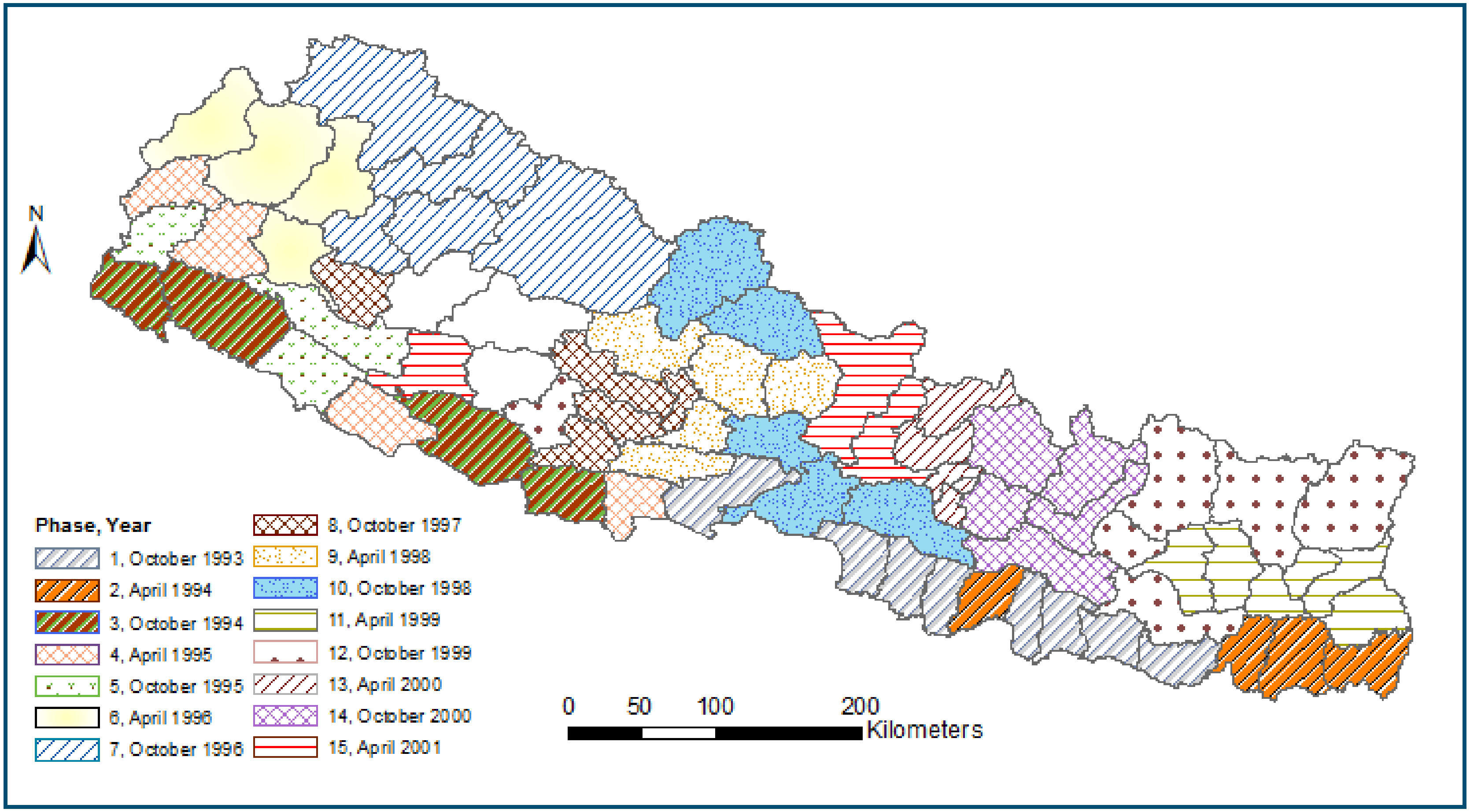
Appendices

Table 2.A1. Analytic Sample by Outcome

	N
Population of Nepal according to the 2011 census	26,494,504
15 percent sample	4,037,885
Non-missing information on district of birth	3,910,420
Nepalese citizens	3,903,832
Between 14 and 22 years of age in 2011	733,586
Non-missing information on blindness	716,283
Non-missing information on disability	724,406
Non-missing information on 'currently in school'	708,075
Under 19 & non-missing information on education-for-age	443,724
Has left school & non-missing information on grades completed	182,486
Non-missing information on marital status	733,586
If married, non-missing information on age at marriage	170,094

Source: Nepal Housing and Population Census 2011

Figure 2.A1. Map Showing the Rollout of the Vitamin A Supplementation Program



Source: Constructed based on information available in MoHP (2002)

Chapter 3

Effects of Nepal's Community-Based Neonatal Care Intervention

3.1 Introduction

For every 1,000 children born, 48 die before their fifth birthday, the majority of them in developing countries (World Bank, 2013). Neonatal deaths—deaths which take place before a child reaches one month of age—make up about 40 percent of these deaths (You et al., 2010). In absolute terms, nearly four million neonates die every year worldwide (Lawn, Cousens and Zupan, 2005; Lawn et al., 2009). Almost all neonatal deaths take place in low- and middle-income countries, with Africa and Southeast Asia accounting for two-thirds of these deaths (Lawn, Cousens and Zupan, 2005). The primary causes of neonatal death are preterm birth, severe infections and asphyxia, which together account for more than 85% of all neonatal deaths (Lawn, Cousens and Zupan, 2005). In Nepal, which is the focus of this paper, nearly 35,000 children die before their fifth birthday each year, with two-thirds of these deaths occurring in the first month of their life (Pradhan et al., 2012). The national under-5 mortality rate stands at 54 per 1,000, which makes Nepal one of the least safe places to be born.

Globally, access to skilled delivery care is important for improving newborn survival (Ngoc et al., 2006). Maternal risk factors such as anemia and hypertension, and delivery complications such as prolonged or obstructed labor, are associated with a higher risk of neonatal mortality (Chalumeau et al., 2000), as they also increase the likelihood of preterm birth, infections and asphyxia. Not surprisingly, current global health efforts to reduce mortality are focused on increasing access to, and utilization of, maternal health care services during pregnancy and delivery (Lawn et al., 2009).

This paper evaluates one such effort: the Community-based Neonatal Care Package (CBNCP) in Nepal. The CBNCP was aimed at reducing child mortality through a range of interventions, such as the provision of a clean kit to be used at the time of delivery at home and the management of newborns' health. Nepal is one of the first countries in South Asia to pilot such a comprehensive strategy to reduce neonatal death at the national level (KC et al., 2011), although Similar interventions have been piloted in India (e.g., Tripathy et al., 2010) and Bangladesh (e.g., Baqui et al., 2008) at local levels.

National-level evaluation of programs such as the CBNCP is often difficult. The programs may be piloted in the poorest districts for which it is difficult to find a reasonable counterfactual. The amount of resources required to collect program-specific nationally representative data may also be high. Not surprisingly, many of the evaluations so far are done on a small scale and test the effectiveness of specific component (such as an educational program or a financial incentive) within a small local area rather than that of the entire intervention, thus limiting the external validity of the findings.

Against this background, this paper contributes to the literature on the effectiveness of community-based health interventions in multiple ways. First, I evaluate Nepal's CBNCP using a sample that is similar to a nationally representative sample, thus obviating the need to justify the external validity of the findings. Second, I evaluate the program's effect on intermediate outcomes in addition to neonatal mortality, thus shedding light on the pathways through which a program such as the CBNCP affects mortality. Finally, the method employed in this paper allows us to make a causal interpretation of the estimated effect.

To preview the results, I find that the CBNCP had limited or no effect on neonatal mortality. This finding contradicts earlier claims of significant positive effects of the program on health behavior and outcomes (e.g., Pradhan et al. (2011)). In terms of the intermediate outcomes, there is no evidence that the program increased institutional deliveries or the skilled-professional attended births. However, the program had significant effects on delivery practices at home, illustrated by the notable increase in the use of a clean kit, which was provided through the program. The program was also influential in changing health behavior of pregnant women such as encouraging them to visit health facilities for prenatal checkups. Similar effects are not seen in the uptake of iron pills, folic acid and tetanus shots—all of which depend on the quality of the providers, including availability of these drugs in local health facilities—suggesting that supply-side constraints prevalent in the Nepali health system may have limited the effects of the program.

3.2 The Community-Based Neonatal Care Package

The details of the CBNCP, including how it was conceived and brought into the national health agenda, have been discussed extensively elsewhere (Poudel et al., 2012; KC et al., 2011; Pradhan et al., 2012). Therefore, I only provide a summary here. The CBNCP evolved from an extensive consultation between Nepal's government, development partners and the health community. The program's primary goal was to reduce neonatal mortality through community-based interventions. The government and three non-government organizations piloted the program in 10 districts across the country in 2009. It is not clear how the districts were chosen. As shown later, on average, the program districts are similar to the districts surrounding them.

The program has seven components, ranging from broad, cross-cutting approaches such as communication for changing behavior to specific interventions such as the management of sepsis, which is the presence of bacteria and their toxins in the body due to infections of a wound. The seven components are: (i) behavior change and communication for newborns' health, (ii) promotion of institutional delivery and clean delivery practices, (iii) postnatal follow up of neonates, (iv) community case management of neonatal infections, (v) management of low birth weight, (vi) prevention and management of hypothermia, and (vii) recognition and resuscitation of an asphyxiated (lacking sufficient oxygen) baby.

The key vehicles of this program are the Female Community Health Volunteers (FCHVs). The government created the FCHVs in 1988 primarily to distribute vitamin A supplements and help reduce childhood pneumonia and diarrhea. They were subsequently instrumental in implementing the community-based Integrated Management of Childhood Illness program (Pradhan et al., 2012). With their pronounced success in delivering health services at the local level, they are often the first network on which the implementers tap.

Under the CBNCP's first three components, the FCHVs were trained to provide face-to-face guidance to pregnant women about healthier delivery practices, to accompany them to a health facility for delivery and, if the delivery took place at home, attend to it along with a skilled birth attendant. The FCHVs were also trained to provide home-based postnatal care and to encourage women to visit health centers if necessary. Information on institutional delivery and clean delivery practices (if delivered a home) were also shared through local radios and social mobilizers. The fourth component was included based on a pilot done in 2007 in one of the districts, Bardia. Under this component, the FCHVs were trained to identify infections, administer oral cotrimoxazole, which helps prevent infections, and refer the sick newborns to the health center for gentamicin injections, which reduce the spread of bacteria. Under the fifth component, the FCHVs were trained to identify cases of low birth weight among newborns using color-coded weighing scales and refer extreme cases to health centers. The key aspect of the sixth component was to encourage women to prevent hypothermia (low body temperature) through skin-to-skin contact between the mother and her baby. This approach has been used in other low-resource settings as an alternative to conventional neonatal care (McCall et al., 2010). The FCHVs were also trained to encourage immediate initiation of breastfeeding. Finally, under the seventh component, the FCHVs were trained to recognize asphyxia, perform step-by-step approach of initial stimulation suctioning and resuscitation using a bag-and-mask. Taken together, the CBNCP was expected to reduce mortality by identifying health problems early and by encouraging women to adopt safer delivery practices. By estimating the effect on mortality and on other intermediate outcomes, I evaluate the program on both types of results.

3.3 Data

I use data from the 2011 Nepal Demographic and Health Survey (NDHS) for the main analysis and also the 2006 NDHS in the robustness check. The NDHS is a nationally representative survey conducted approximately every five years. The NDHS collects detailed information from women between the ages of 15 and 49 years about their pregnancies and births within the five years preceding the survey date. In addition to detailed birth information, the NDHS also collects information on women's characteristics including age, religion, highest level of schooling completed, and household attributes including access to electricity, source of drinking water, type of toilet facilities, and type of roofing and flooring materials. The NDHS provides a wealth index, calculated based on asset ownership using principal component analysis, and associated wealth quintiles.

I evaluate the effect of the CBNCP on several outcome variables (Appendix Table 3.A1). Whether the child survived the first month of birth (neonatal mortality) is the primary outcome of interest for this paper. In the survey, for all births within the preceding five years (including still births), women were asked where the birth took place and if the child is alive. For children who die, the NDHS provides the age of the child at death. Neonatal mortality—and not under-5 mortality—is the primary outcome because neonatal death more accurately reflects the quality of care received by the mother and the child during childbirth compared to under-5 mortality (Ngoc et al., 2006).

Cases of abortion are dropped. They were distributed evenly between treatment and control districts before and after the program. Before the program started (average between 2006 and 2008), 6.04% pregnancies in treatment districts and 6.82% pregnancies in control districts ended in abortion. After the program (average between 2009 and 2011), 2.52% and 3.46% pregnancies in treatment and control districts, respectively, ended in abortion.

In exploring the intermediate outcomes, I look at the probability of institutional birth and, for births that take place at home, whether the birth was attended by a skilled professional. Both of these have been coded as binary variables. A birth that took place in a health center, hospital or a NGO facility has been counted as an institutional birth. Likewise, a birth is assumed to be professional-attended if a health professional (doctor, nurse or another person trained on birth) was present at the time of birth. In addition, I look at the use of a clean kit during delivery (if the delivery took place at home), prenatal and postnatal visits, and intake of iron and folic acid pills and tetanus shots. The program provided the kit to all pregnant women. For prenatal and postnatal visits, I construct binary variables equal to one if, respectively, the woman made four or more prenatal visits over the course of the pregnancy and went for postnatal checkup within two months of delivery. While two months is a long time after delivery to go for postnatal checkup by developed country standards—in developed countries, such as the United States, women usually go for a well-child visit within a week of birth—this is the only information available on postnatal checkup in the survey. For tetanus shots, the variable equals one if the woman took at least two tetanus shots during pregnancy. Two shots of tetanus during delivery, one month apart, are recommended by the World Health Organization for women who have had no prior tetanus shot (World Health Organization, 1999).

The choice of the intermediate outcome variables is driven by the program goals and evidence in the literature on the variables' association with child mortality. Access to skilled attendance at delivery is critical in reducing deaths that occur during pregnancy, delivery and the post-partum period (World Health Organization, 1999). On institutional delivery, a series of articles in the *Lancet* have argued in favor of adopting health center-based intrapartum care for reducing mortality (Filippi et al., 2006). Institutional deliveries, the argument goes, may give women access to skilled service providers who are better able to diagnose and treat complications, thus reducing child mortality. In line with this argument, Maitra (2004) and Panis and Lillard (1994) find a strong effect of institutional delivery on child mortality in India and Malaysia, respectively. Beneficial effects of prenatal care on infant health outcomes have also been shown by Jewell (2007). I look at the use of a clean kit during delivery for births that take place at home because the kit was provided through the program to prevent infections during and immediately after birth. Finally, prenatal visits, postnatal visits and the intake of iron and folic acid pills and tetanus shots are standard prescriptions that international health community, including the WHO, has provided for better health of mothers and newborns.

In the sample used for analysis, the neonatal mortality rate is 34 per 1,000 (Table 3.1). Note that the sample used in the analysis is not the entire NDHS sample. The choice of which observations to use was determined by the identification strategy discussed in Section 3.4. Only 38% of births in the sample take place in health centers and 55% are attended by skilled professionals. Of the deliveries that take place at home, only 19% use a clean kit. Approximately 55% of women make at least four prenatal visits to the health center and 50% make a postnatal visit within two months of delivery. Roughly 80% of women take folic acid/iron tablets during pregnancy (the survey does not ask for information on iron tablet and folic acid tablet intakes separately).

In the sample, 47% of the children are boys. The average birth order is 2.6, close to Nepal's fertility rate. Mother's average education level is 3.5 years, which attests to the necessity of programs such as the CBNCP for communication and health behavior change in Nepal. Mother's age at first birth is about 20 years. About 20% of the children are from households in urban areas, 46% have access to piped water, 67% to electricity and 46% to a toilet. Approximately 51% of children are from the poorest two quintile households based on the wealth index. Mothers of about 55% of children identify getting to the nearest health center, which on average is 55 minutes away from home, as a problem for them.

3.4 Identification Strategy

In evaluating the CBNCP, I capitalize on the fact that it was piloted in 10 districts in 2009. The districts surrounding these 10 districts provide the counterfactual as the program was not implemented in those surrounding districts. Nepal has 75 districts. I do not include all 65 districts in which the CBNCP was not implemented in the control group so as to keep the control districts as similar as possible to the treatment districts. Instead, I only use the 44 districts surrounding the treatment districts as the control districts, as shown in Figure 3.1. Using the surrounding districts as control districts reduces the chance that treatment and control districts may be differentially exposed to another program or a different policy environment. The program districts are also more similar in terms of education, health and wealth to the surrounding districts than they are to all non-program districts (not shown here).

In order to estimate the program's effect, I employ a difference-in-difference strategy where the effect of the program is identified based on the pre-CBNCP and post-CBNCP differences in outcomes between treatment and control districts.

For each of the outcomes discussed in Section 3.3, I estimate the following equation:

$$Outcome_{ijt} = \pi_0 + \pi_1 Treat_j + \pi_2 Post_t + \pi_3 [Treat_j * Post_t\ ] + \ \pi_4 X'_{ijt} + v_{ijt} \text{……….(1)}$$

In this equation, $Outcome_{ijt}$ is the outcome for a child *i* born in district *j* in year *t*, and *Treat*=1 if the CBNCP was piloted in district *j* in 2009 and 0 if it is a district surrounding one of the CBNCP pilot district. *Post=1* for 2009 and after (2009, 2010 and 2011) and 0 for periods before 2009 (2006, 2007 and 2008). X_{ijt} includes child's, mother's, household's and community's characteristics that are different between the treatment and control districts at the time of the survey, and those that may have influenced the outcome. *v* is the disturbance term. The coefficient of interest is π_3, which reflects the effect of the program, i.e., the difference in the outcome between treatment and control districts after the program relative to the difference in the outcome before the program. The expected sign of π_3 depends on the outcome; for neonatal mortality, we expect a negative sign because the program should reduce such mortality.

The key identifying assumption is that, without the CBNCP, the treatment and the control districts would have experienced similar changes in the outcomes. This holds if the treatment and control districts are similar in terms of the observable factors at baseline and if there are no differences in pre-program trends in the outcomes. These assumptions are discussed next.

Before the CBNCP went into effect in 2009, the treatment and control districts are similar in the majority of the aspects (see Tables 3.2 for the covariates and Table 3.3 for the outcomes). However, there are statistically significant differences between the two categories of districts in terms of birth order of the child, urbanicity, access to water, access to electricity and the distribution of wealth. A greater share of households in the treatment districts are from urban areas and have access to electricity. Urban households have fewer children than rural households which the lower birth order of a child in the treatment districts in the survey reflects. The treatment districts also have disproportionately more households in the richer wealth quintiles (quintiles 4 and 5). In contrast, a greater share of households in the control districts have access to water, which may be reflective of water shortages in urban areas.

The treatment and control districts are also different in terms of the proportions of birth that take place in hospitals (institutional delivery) and the proportion of skilled professional-attended births (Table 3.3). I assume that, after controlling for differences in the covariates, the treatment and the control districts are similar in terms of the outcomes. This assumption generally seems to hold as shown later by the statistically insignificant coefficient on *Treat* (π_1).

In order to check if there are different trends in neonatal mortality in treatment and control districts before the program went into effect, I conduct a formal test of the differential time trends in the analytical sample. More specifically, following Antwi, Moriya and Simon (2013), I estimate a regression of the key outcomes of interest on an interaction term between treatment and birth year for the years before 2009 and control for the set of covariates used in the subsequent analysis. A statistically significant coefficient on the interaction term would indicate that, conditional on the covariates, the trends in outcomes between treatment and control districts are different before the program. Appendix Table 3.A2 shows the coefficients from this analysis. The coefficients are all statistically insignificant at 5% level, implying that once I control for the covariates, the treatment and control districts can be assumed to have similar trends in outcomes before the program went into effect.

The estimated effects should be interpreted as intent-to-treat estimates. The program was implemented throughout each pilot district, but there is no information on if, and the extent to which, the program reached all pregnant women in those districts.

A brief note in the choice of covariates is in order. The covariates have been added mostly to account for differences between treatment and control districts prior to the program. However, a few covariates have been added in view of the effect they would have on the outcome of interest. For example, institutional deliveries have been found to be positively correlated with maternal schooling in other settings (Houweling et al., 2007). Likewise, in India, women have been found more likely to give birth in a hospital when they are carrying a son than when they are carrying a daughter (Bharadwaj and Nelson, 2010), although this finding assumes that women know the gender of the child before the child is born (probing into the child's gender before delivery is illegal in Nepal). Ethnicity is also a strong determinant of health-seeking behavior in Nepal, with disparities across ethnicities in health outcomes widening in the recent decade (RTI International, 2008). Therefore, covariates such as mother's education level, sex of the child and ethnicity are included even though there is no statistically significant difference between treatment and control districts in these variables before the program went into effect in 2009. Inclusion of these covariates helps reduce the error term and improve the precision of the estimate of the program effect.

The number of districts included in the analysis is lower than 50. Clustering the standard errors at the district level can accounts for the lack of independence between the observations in a given district. However, the statistical significance on the coefficient needs to be interpreted cautiously when the number of districts is lower than roughly 50 (Duflo, Glennerster and Kramer, 2006, p. 61). I address the small number of clusters (districts in this case) by reporting bootstrapped standard errors. In all estimation results, I report bootstrapped standard errors with 1,000 iterations and seed 1. The standard errors are bootstrapped by district.

3.5 Results

3.5.1 Impact of CBNCP

In the main results tables (Tables 3.4 to 3.12), I show coefficients on the main effects and the interaction term (π_1, π_2 and π_3) from a Linear Probability Model (LPM) estimated on equation (1). The first column reports the coefficients from the regression without any covariates. I then include child-related, mother-related, household-related and community-related covariates in a step-wise manner. Although the coefficient on the interaction term (π_3) is identified in the first column itself, the step-wise addition of variables allows us to check the stability of the coefficient and, if the coefficient is stable, interpret the magnitude of the effect more confidently. Reassuringly, the coefficients are fairly stable in columns 1-5 in the majority of tables. For interpretation, I focus on column (5) since it fully controls for initial differences in the covariates between treatment and control districts. The coefficients should be read as the percentage point change from baseline given in Table 3.3.

In all tables, the coefficients on the covariates (not shown) have the expected signs and magnitudes. I find, for example, that twins are more likely to die than singletons. Likewise, children born to mothers who were older at their first birth and those born in larger households are less likely to die than their respective counterparts. Finally, children from richer households are less likely to die than those from poorer households. All of these findings agree with the NDHS final report (MOHP, 2012) which looks at simple correlations between several factors and under-5 mortality.

Relative to the control districts, neonatal mortality in treatment districts decreased by about 1.4 percentage points due to the program, but the coefficient is not statistically significant at 5% (π_3 in Table 3.4). The value of π_1 is -0.01, meaning that controlling for the covariates, mortality in treatment districts was about 1 percentage point lower than in the control districts before the program started. The coefficient is not statistically significant at 5%. On average, neonatal mortality fell by about 0.5 percentage point after the program relative to before (coefficient on *Post* is -0.005). However, this improvement could be due to several other factors that changed during the period. The R-squared is low, at about 3%, even with all the covariates in the specification, meaning that only about 3% of the variation in mortality is captured by the covariates. Unfortunately, the dataset does not provide information on a few other factors critical to reducing mortality, such as food availability.

Although the program's effect on neonatal mortality was limited, it is possible that the program affected intermediate factors, such as institutional birth or prenatal behavior, which may be beneficial to mother and the child. I now proceed with an exploration of these intermediate outcomes.

Looking at institutional birth, the first row of the first four columns of Table 3.5 reflects the pre-program difference in this variable between treatment and control districts. The difference disappears when we control for urbanicity—an aspect in which the treatment and control districts differ significantly based on Table 3.2. This provides us more confidence in the fully controlled specification (column 5). For all districts, on average, the institutional deliveries increased by as much as 14 percentage points during the period from the baseline of about 30%. The change is statistically significant. However, the effect of the program is small and statistically insignificant, as shown by the coefficient on the interaction term. The lack of the program's effect on institutional birth may reflect the relative difficulty in taking women to hospital in Nepal's difficult geographic terrain. In the analysis sample, about 55% women report that getting to the health facility is difficult for them. This figure is for all women in the sample irrespective of their current pregnancy status and would likely be higher if the question on difficulty in getting to the health center was asked at the time of delivery.

The effect on professional-attended deliveries was also low and statistically insignificant (Table 3.6). The R-squared values in Tables 3.5 and 3.6 are much larger than those in Table 3.4, potentially reflecting the direct effect that a program such as the CBNCP can have on promoting institutional delivery and professional-attended births—unlike in the case of mortality where additional events after birth may be influential (and thus a smaller share of variation in mortality is explained by the CBNCP compared to the variation in institutional or skilled professional-attended delivery).

One area in which the program had a significant effect (although only at 10% level) is the use of a clean kit during delivery for births that took place at home (Table 3.7). Evaluating the effect on the use of a clean kit is important because many newborns die due to infections around the time of birth and because a kit was provided to pregnant women through the CBNCP. Controlling for the covariates, there was no significant pre-program difference in the use of a clean kit between treatment and control districts. The use of a clean kit declined during the period. The decline is statistically significant at 5%, contradicts the improvement in general health behavior in Nepal during the period, and is an area for further research. The program's net effect is evident in the coefficients in the third row, which shows that the program increased the use of a clean kit by about 9 percentage points relative to the overall decline (at the baseline of 24%). However, the effect is statistically significant only at 10%.

Encouraging prenatal visits by pregnant women was an integral part of the program. The program increased the percentage of pregnancies with at least four prenatal visits by about 9 percentage points (Table 3.8) (mean at baseline = 49.5%). The effect is statistically significant at 5%. There was a general rise in women's visit to health care centers for prenatal visits during the period even without the program, as shown by the positive, statistically significant coefficient on *post* (row 2). On the other hand, the program did not increase post-delivery checkups within two months of birth (Table 3.9, row 3), although there was a significant rise in such check-ups generally during the period (Table 3.9, row 2). It is likely that the FCHVs were too focused on promoting safe delivery only up to the time of birth while post-delivery behavior was overlooked.

Although the program increased prenatal visits, there was no increase in the probability of taking folic acid or iron tablets during pregnancy (Tables 3.10 and 3.11). This lack of effect likely reflects supply-side constraints in Nepal's health system, primarily the lack of medicines and regular operation of health centers in rural areas. Table 3.12, which shows that there was no increment in receiving tetanus vaccines—a WHO recommendation for all pregnant women who have not had tetanus shots before—also points to the supply side constraints as a possible barrier because of which the program's overall effect was limited.

3.5.2 Robustness check

It is possible that the positive effect of the CBNCP observed above on the use of a clean kit and prenatal visits, and the lack of effect on other outcomes, is not due to the CBNCP but some other event occurring in the treatment districts. To determine this, I first perform the same analysis as above, but using the 2006 NDHS data under a hypothetical scenario that the CBNCP was implemented in 2004 in the same districts in which it was actually implemented in 2009. I chose the year 2004 because it is the midpoint of the birth years in the 2006 NDHS (the survey covers children born between 2001 and 2006), just as 2009 is the midpoint of the birth years in the 2011 NDHS.

My paper's main findings contradict earlier studies which have found a significant effect of the program on mortality and many of the intermediate outcomes. Although these studies have looked at specific geographic areas and are mostly descriptive, it is important to address the contradictory findings. Therefore, I perform an additional test by evaluating the program's effect on immediate nutritional outcomes. These are outcomes which the program was intended to influence only indirectly. If the nutritional status of mothers and children are similar between treatment and control districts before the program (including the trend), but changed afterwards, it raises concerns that the observed effect—both the significant effects on the use of a clean kit and prenatal checkups and the lack of effects on other outcomes— may also be due to other factors that may not have been fully captured in the estimation.

Finally, as a robustness check on the choice of the LPM over a logistic or probit model, I run a logistic regression for key outcome variables using the same set of covariates as in column 5 of Tables 3.4 to 3.12 and compare the results with the LPM results.

3.5.2.1 Effect assuming the program started in 2004

The effect of the CBNCP observed in the use of a clean kit and prenatal visits in the 2006-2011 sample is unlikely to occur by random chance or because of some underlying characteristic of the treatment districts. When using 2004 as the program year in the sample of children born between 2001 and 2006 (which the 2006 NDHS captures), the coefficient on the interaction term is statistically insignificant in all cases including in the specifications with use of a clean kit and prenatal visits as the outcomes (Table 3.13). If the estimated effect earlier was due to some underlying characteristics of the program districts and not the program itself, the interaction term in Table 3.13 could have been statistically significant.

3.5.2.2 Effect on nutritional outcomes

If, contrary to this paper's findings, the program did reduce child mortality, then one would expect children in the treatment districts to have poorer nutritional status than those in control districts. This is because the saved babies are most likely marginal babies who would have died in absence of the program. These babies are likely to be of poorer nutritional status. If more of these babies are saved in treatment districts than in control districts, then the overall nutritional status of children in treatment districts should fall relative those in control districts.

For this argument to work, the nutritional status before the program should be similar as should the pre-program trends. The nutritional outcomes evaluated here include whether mother is anemic, whether child had low weight at birth, whether child is anemic, whether the child had diarrhea within two weeks prior to the survey, whether the child is underweight, and the z-score for the child's weight-for-age. These are all short-term outcomes. Table 3.14 confirms that the nutritional outcomes are balanced before the program went into effect; the p-values of the difference between treatment and control districts' means are all bigger than 0.05. Likewise, Appendix Table 3.A3 confirms that the trends in these outcomes are similar for treatment and control districts before the program.

Information on nutritional outcomes is missing for a large portion of the sample. If individuals in treatment districts are more likely to have their nutrition information missing, or vice versa, our estimates can be biased. This may happen if, for example, children with low birth weight were consistently less likely to be weighed in control districts because of absence of the CBNCP than in treatment districts. In such a scenario, the estimated effect would be an overestimate of the true effect. Therefore, I check if nutrition data are missing for different proportions of the sample in treatment and control districts. I do so by regressing *missingness* (=1 if information is missing) of each of the nutritional outcome variables, separately, on the treatment status. A statistically significant coefficient on the treatment status for a given outcome would indicate that the missing data are differential across treatment and control districts for that outcome. Table 3.15 shows that this is not the case. The reported coefficient on the CBNCP (treatment) on all of the nutritional outcomes is small and statistically insignificant.

The coefficients on the interaction term (post*treatment) when nutritional indicators are used as outcomes are all small and statistically insignificant (Table 3.16). While this is not a conclusive falsification test (for the reasons provided in Section 3.6), it does provide additional evidence that the program's effect on mortality may have been minimal.

3.5.2.3 Results on the choice of specification

In order to assess if the results discussed above are driven by the choice of the specification, I estimated equation (1) for all outcomes using logistic regression, instead of a linear probability model. I estimate the interaction effects using Ai and Norton (2003). The results (not shown here) confirm findings from the LPM: the program had a positive, significant effect on the use of a clean kit at the time of birth and a strong, positive effect on prenatal visits, and no effect on other outcomes.

3.6 Conclusion, Caveats and Areas for Further Research

In this paper, I evaluated a community-based neonatal intervention aimed at reducing neonatal mortality in Nepal. In contradiction to earlier studies that evaluated parts of this program in select geographic locations, I find that the program had limited effect on neonatal mortality, the primary outcome the program aimed to influence. In terms of the intermediate outcomes, there is no evidence that the program increased institutional deliveries or the skilled-professional attended births. However, by providing a clean kit to be used at the time of delivery, the program encouraged cleaner delivery practices for births that took place at home. The program also encouraged pregnant women to go to the health facilities for prenatal checkups. Such effects, however, are not seen in the uptake of iron pills, folic acid, and tetanus shots.

The NDHS data do not allow us to explore why the use of a clean kit or an increase in prenatal visits did not translate to reduced mortality and why the program did not affect other intermediate outcomes significantly. Nonetheless, one can conjecture that the lack of effect on many of the outcomes reflect supply-side constraints, such as shortage of health workers and medicines, prevalent in the Nepali health system. It is possible, for example, that women visited health centers more often than before in response to the FCHVs' persuasion, but when they went to the health centers, there may not have been anyone to administer tetanus shots. Likewise, pregnant women may be fully aware that it would be safer to deliver a baby in the hospital, but the time and monetary costs of going to the hospital at the time of delivery may be too high. The key weakness of the program was then the lack of sufficient strengthening of the health system commensurate with the rise in demand for services.

The findings in this paper should be understood in light of a number of caveats. First, although I have used institutional delivery and skilled birth attendance as some of the intermediate outcomes through which the CBNCP may influence neonatal mortality, the causal link between institutional delivery and skilled birth attendance and mortality is still being debated in the literature. Walraven and Weeks (1999) argue that the skilled attendant in the local health facilities may be no more skilled than the traditional community midwife. Likewise, Harvey et al. (2007) find significant skill gaps in a study of skilled birth attendants in Benin, Ecuador, Jamaica, Rwanda and Nicaragua. They find that knowledge of a procedure by health workers did not necessarily lead to the correct application of the procedure. If the FCHVs or local health workers were not trained adequately in the CBNCP, simply having them present at the time of delivery, or even taking women to health centers for delivery, may not help reduce mortality.

Second, the baseline differences between treatment and control districts shown in Tables 3.2, 3.3 and 3.17 are the differences for households and for women of children born before 2009 reported in 2011 (the time of the survey). My identification strategy assumes that the differences in the covariates are time-invariant, and thus that the difference observed in 2011 for children born before 2009 is in fact the actual difference in 2009. If individuals in treatment and control districts have different recall bias, the baseline characteristics may be less similar than those reported here. For example, it is possible that women in treatment districts are more likely to keep their health records and records of their child's health since they are in urban areas, and as such they might provide more accurate information than women in control districts. In that case, the identifying assumption is partly violated.

Third, another program, called the Safe Delivery Incentive Program (SDIP), which was implemented throughout Nepal in 2009, may have interacted differently with treatment and control districts owing to initial differences in the distribution of wealth among households in the two categories of districts. On the one hand, the delivery and transportation incentives provided through the SDIP may affect poorer households (and hence control districts) more since those incentives constitute a higher share of their income. On the other hand, the incentives may be small enough to affect only the households with some level of existing resources. Such households may be more evenly distributed among control and treatment districts (Table 3.2 shows that the treatment and control districts have a similar share of households in the third wealth quintile), thus there would be limited, if any, differential effect of the SDIP on treatment and control districts.

The SDIP may have also altered incentive structures differently in treatment and control districts, mainly with respect to the decisions between institutional delivery and professional-attended home delivery. The SDIP provided free delivery in hospitals and reimbursed a fixed amount (based on the region) to women to offset transportation costs if they delivered in hospitals. These incentives would encourage institutional deliveries. Incentives were given to the health workers through the CBNCP for each delivery they attended at home. These incentives would encourage home deliveries. While these two different and offsetting effects would be taking place in treatment districts, only the former effect would be taking place in control districts—thus affecting the two categories of districts differently. The net direction of the effect due to the interaction in treatment districts is an empirical question.

Understanding of the interaction between the SDIP and the CBNCP in the CBNCP and the non-CBNCP districts can provide insights into the net effect of different types of incentives on health workers and households on promoting institutional delivery. From the side of the households alone, the CBNCP primarily provided information, while the SDIP is catered to providing subsidies. Further research could parse out these two effects, thus contributing to current literature on the role of information and subsidies on health products intake (see e.g. Ashraf et al., 2013).

Finally, this paper assumes that there is no spillover of the program's influence from treatment to control districts. There is no certainty that this assumption holds. In fact, in one intermediate outcome with significant effects—the use of a clean kit at the time of delivery—there is limited scope for spillover (it is difficult to imagine a clean kit given to a household in a treatment district being passed to a household in a control district). Conversely, spillover is very possible in institutional delivery, since women are likely to go to the nearest hospital for delivery rather than a hospital in their own district if the latter is further.

In order to ascertain findings from this paper, future research could also look at the interaction of the CBNCP with the SDIP and attempt to parse out spillover effects. Another potential research area is the degree to which the CBNCP may have crowded out other existing programs. An earlier evaluation of the program (Pradhan et al., 2012) has suggested that the quality of the FCHV's work may have been compromised in some tasks due to their overstretched workload; a recent news report suggests that the FCHVs may have been involved in as many as 81 different activities (Setopati, 2014). It is also possible that the FCHVs may have prioritized the CBNCP because of the incentive they received through the program and neglected other important national programs without similar incentives (Pradhan et al., 2011). If this is true, by introducing the CBNCP without due consideration of its implications for other programs, the government may have spent resources on a program that did not produce discernible results and adversely affected other programs. In general, supply constraints—in terms of human resources—are a major problem in Nepal's health system. As such, any program that does not take into account the existing resource constraints during design and implementation is likely to crowd out other initiatives or simply fail. Such crowding out in this case would be particularly worrying given that the program does not seem to have significant positive effects even on reducing mortality, its primary goal. Therefore, a better understanding of the extent to which the CBNCP may crowd out other programs is crucial for future policy design.

Further research can also be conducted on the potential heterogeneity of the treatment effect across districts. One of the reasons why the identification strategy in this paper works is that the districts selected for the program ranged from the poorest to some of the most well-off ones—thus making the district selection for the CBNCP fairly random (as the pre-program balance Tables 3.2 and 3.3 show). However, there are significant differences across the 10 program districts in terms of wealth distribution, literacy and health indicators. It is possible that the effects reported in this paper are driven by effects in a few districts. Apart from initial differences in social and economic conditions across the treatment districts, heterogeneous effects could also originate from variation in implementation. The program was implemented by different agencies (Save the Children, CARE Nepal, Plan International and the Government of Nepal) in different districts, likely creating variation in reach and intensity of the program. The key local implementing agency was the District Public Health Office, whose capacity also varies widely across districts.

References

Ai, C., Norton, E.C., 2003. Interaction Terms in Logit and Probit Models. *Economics Letters*, 80(1):123–9.

Antwi, Y. A., Moriya, A.S., Simon, K., 2013. Effects of Federal Policy to Insure Young Adults: Evidence from the 2010 Affordable Care Act's Dependent-Coverage Mandate. *American Economic Journal: Economic Policy*, 5(4): 1-28.

Ashraf, N., Jack, B.K., Kamenica, E., 2013. Information and subsidies: Complements or substitutes? *Journal of Economic Behavior & Organization*, 88: 133–139

Baqui, A.H., El-Arifeen, S., Darmstadt, G.L., Ahmed, S., Williams, E.K., Seraji, H.R., Mannan, I., Rahman, S.M., Shah, R., Saha, S.K., Syed, U., Winch, P.J., Lefevre, A., Santosham, M., Black, R.E., for the Projahnmo Study Group, 2008 (June). Effect of community-based newborn-care intervention package implemented through two service-delivery strategies in Sylhet district, Bangladesh: a cluster-randomised controlled trial. *The Lancet*, 371(9628):1936-44.

Bharadwaj, P., Nelson, L.K., 2010. Discrimination Begins In The Womb: Evidence Of Sex-Selective Prenatal Investments. Mimeo, University of California, San Diego. URL (accessed 15 December 2013): http://dss.ucsd.edu/~prbharadwaj/index/Papers_files/Bharadwaj_Nelson_Oct24_2010.pdf.

Chalumeau, M., Salanave, B., Bouvier-Colle, M.H., Bernis, L. de., Prua, A., Breart, G., 2000. Risk factors for perinatal mortality in West Africa: a population-based study of 20326 pregnancies. *Acta Pædiatrica*, 89(9): 1115–1121.

Duflo, E., Glennerster, R., Kremer, M., 2006. Using Randomization in Development Economics Research: A toolkit", NBER Technical Working Paper Series, Paper # 333.

Filippi, V., Ronsmans, C., Campbell, O.M.R., Graham, W.J., Mills, A., Borghi, J., Koblinsky, M., Osrin, D., 2006. Maternal health in poor countries: the broader context and a call for action. *The Lancet*, 368(9546): 1535–1541.

Harvey, S.A., Blandon, Y.C.W., McCaw-Binns, A., Sandino, I., Urbina, L., Rodrıguez, C., Gomez, I., Ayabaca, P., Djibrina, S., and the Nicaraguan maternal and neonatal health quality improvement group, 2007. Are skilled birth attendants really skilled? A measurement method, some disturbing results and a potential way forward. *Bulletin of the World Health Organization*, 85(10): 783–790.

Houweling, T. A. J., Ronsmans, C., Campbell, O. M.R., Kunst, A.E., 2007. Huge poor-rich inequalities in maternity care: an international comparative study of maternity and child care in developing countries. *Bulletin of the World Health Organization*, 85: 745–754.

Jewell, R.T., 2007. Prenatal care and birthweight production: evidence from South America. *Applied Economics*, 39(4): 415–426.

KC, A., Thapa, K., Pradhan, Y.V., KC, N.P., Upreti, S.R., Adhikari, R.K., Khadka, N., Acharya, B., Dhakwa, J.R., Aryal, D.R., Aryal, S., Starbuck, E., Paudel, D., Khanal, S., Devkota, M.D., 2011 (October). Developing Community-Based Intervention Strategies and Package to Save Newborns in Nepal. *Journal of Nepal Health Research Council*, 9(19): 107-18.

Lawn, J. E., Lee, A.C.C., Kinney, M., Sibley, L., Carlo, W.A., Paul, V.K., Pattinson, R., Darmstadt, G.L., 2009. Two million intrapartum-related stillbirths and neonatal deaths: Where, why, and what can be done? *International Journal of Gynecology & Obstetrics*, 107, Supplement (0): S5–S19.

Lawn, J. E, Cousens, S., Zupan, J., 2005. 4 million neonatal deaths: When? Where? Why? *The Lancet*, 365(9462): 891–900.

McCall, E.M, Alderdice, F., Halliday, H.L., Jenkins, J.G., Vohra, S., 2010. Interventions to prevent hypothermia at birth in preterm and/ or low birthweight infants (Review). Cochrane Database of Systematic Reviews, 3.

Maitra, P., 2004. Parental bargaining, health inputs and child mortality in India. *Journal of Health Economics*, 23(2): 259–291.

Ministry of Health and Population (MOHP) [Nepal], New ERA, and Macro International Inc., 2007. *Nepal Demographic and Health Survey 2006.* Kathmandu, Nepal: Ministry of Health and Population, New ERA, and Macro International Inc.

Ministry of Health and Population (MOHP) [Nepal], New ERA, and ICF International Inc. 2012. *Nepal Demographic and Health Survey 2011.* Kathmandu, Nepal: Ministry of Health and Population, New ERA, and ICF International, Calverton, Maryland.

Ngoc, N. T. Nguyen, Merialdi, M., Abdel-Aleem, H., Carroli, G., Purwar, M., Zavaleta, N., Campodonico, L., Ali, M.M., Hofmeyr, J.G., Mathai, M., Lincetto, O., Villar, J., 2006. Causes of stillbirths and early neonatal deaths: data from 7993 pregnancies in six developing countries. *Bulletin of the World Health Organization,* 84(9): 699–705.

Panis, C.W. A., Lee A. L., 1994. Health inputs and child mortality: Malaysia. *Journal of Health Economics*, 13(4): 455–489.

Poudel, D.C., Acharya, B., Pant, S., Paudel, D., Pradhan, Y.V., 2012 (May). Developing, Piloting and Scaling-up of Nepal's Neonatal Care Program. *Journal of Nepal Health Research Council*. 10(21): 95-100.

Pradhan, Y.V., Upreti, S.R., KC, N.P., KC, A., Khadka, N., Syed, U., Kinney, M.V., Adhikari, R.K., Shrestha, P.R., Thapa, K., Bhandari, A., Grear, K., Guenther, T., Wall, S.N., 2012. Newborn survival in Nepal: a decade of change and future implications. *Health Policy and Planning*, 27:iii57-iii71.

Pradhan Y.V., Upreti, S.R., KC, N.P., Thapa, K., Shrestha, P.R., Shedain, P.R., Dhakwa, J.R., Aryal, D.R., Aryal, S., Paudel, D.C., Paudel, D., Khanal, S., Bhandari, A., KC, A., 2011

(October). Fitting Community Based Newborn Care Package into the health systems of Nepal. *Journal of Nepal Health Research Council*, 9(19):119-28.

RTI International, 2008. *Equity Analysis of Health Care Utilization and Outcomes*. Research Triangle Park, NC: RTI International.

Setopati, 2014 (July). Female Community Health Volunteers Declare Protest Programs (unofficial translation). URL (accessed 30 July 2014): http://setopati.com/samaj/14661/.

Tripathy, P., Nair, N., Barnett, S., Mahapatra, R., Borghi, J., Rath, S., Rath, S., Gope, R., Mahto, D., Sinha, R., Lakshminarayana, R., Patel, V., Pagel, C., Prost, A., Costello, A., 2010 (April). Effect of a participatory intervention with women's groups on birth outcomes and maternal depression in Jharkhand and Orissa, India: a cluster-randomised controlled trial. *The Lancet*, 375(9721):1182-92.

Walraven, G., Weeks, A., 1999. The role of (traditional) birth attendants with midwifery skills in the reduction of maternal mortality. *Tropical Medicine and International Health*, 4(8): 527–529.

The World Bank, 2013. URL (accessed 15 December 2013): http://data.worldbank.org/indicator/SH.DYN.MORT/countries?display=default

World Health Organization, 1999. Reduction of maternal mortality: A joint WHO/UNFPA/UNICEF/World Bank statement. URL (accessed 15 December 2013): unfpa.org/upload/lib_pub_file/236_filename_e_rmm.pdf

You, D., Wardlaw, T., Salama, P., Jones, G., 2010. Levels and trends in under-5 mortality, 1990–2008. *The Lancet*, 375(9709):100–103.

Tables and Figures

Table 3.1. Summary statistics for the overall sample

	N	Mean	SD
Neonatal mortality	3305	0.034	0.182
Institutional delivery (yes=1)	3305	0.384	0.486
Skilled birth (for at-home births) (yes=1)	3305	0.546	0.498
Clean kit used at delivery (yes=1)	1517	0.190	0.392
Four antenatal visits (yes=1)	2557	0.544	0.498
Postnatal within 2 months (yes=1)	2554	0.500	0.500
Iron/folic acid in pregnancy (yes=1)	2556	0.805	0.396
Iron/folic acid after delivery (yes=1)	2556	0.444	0.497
Child's gender (male=1)	3305	0.473	0.499
Month of birth	3305	6.489	3.316
Twin (yes=1)	3305	0.013	0.115
Birth order	3305	2.578	1.805
Mother's education in years	3305	3.549	3.977
Mother's age at first birth	3305	19.656	3.122
Household size	3305	6.044	2.696
Urban (yes=1)	3305	0.207	0.405
Access to water (yes=1)	3305	0.458	0.498
Access to electricity (yes=1)	3305	0.670	0.470
Access to latrine (yes=1)	3305	0.544	0.498
Wealth quintile 1 (poorest)	3305	0.313	0.464
Wealth quintile 2	3305	0.208	0.406
Wealth quintile 3	3305	0.178	0.382
Wealth quintile 4	3305	0.156	0.363
Wealth quintile 5 (wealthiest)	3305	0.145	0.352
Problem getting to hospital (yes=1)	3305	0.552	0.497

Source: NDHS 2011

Figure 3.1. Map of Nepal showing the districts in which CBNCP was piloted (red) and those used as the control districts (blue).

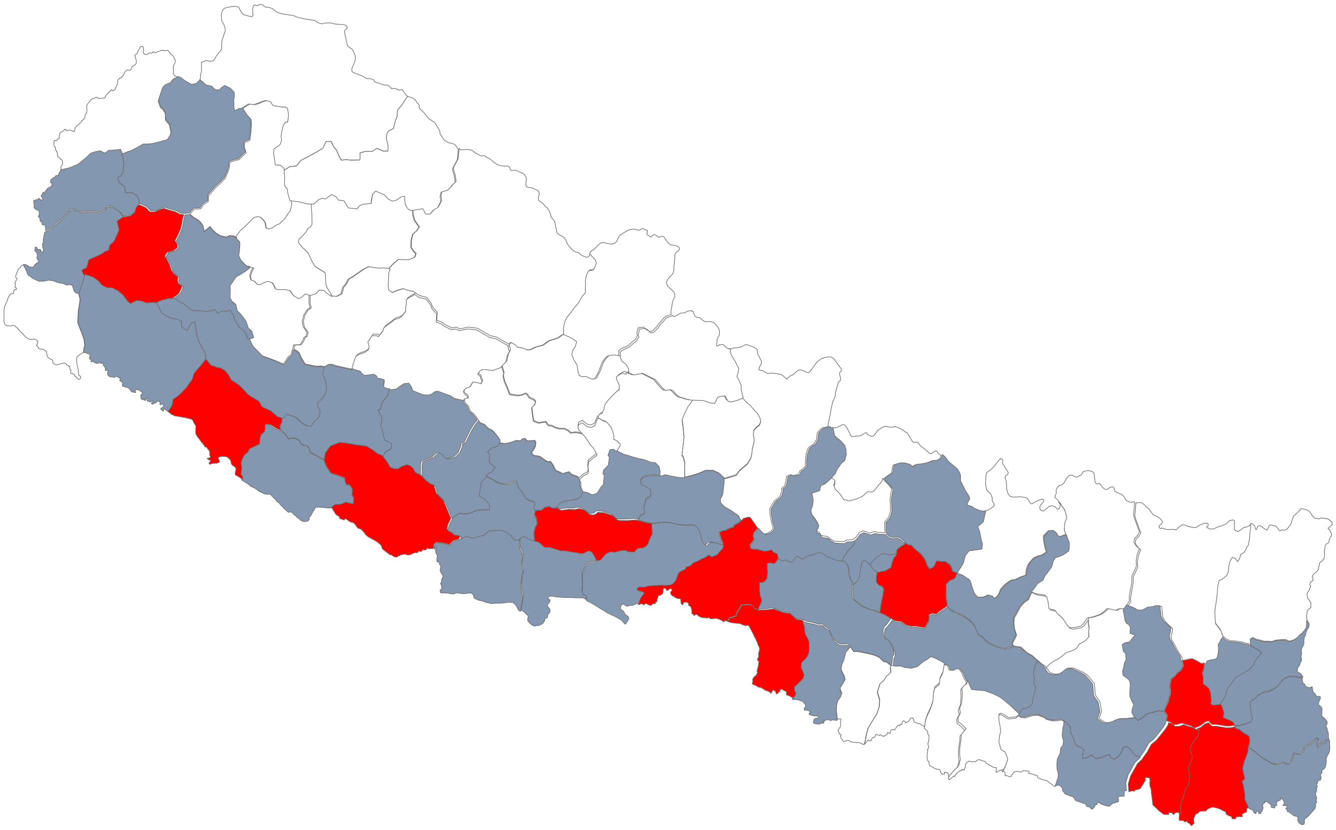


Table 3.2. Comparison of mean between treatment and control districts before CB-NCP	N	Overall	Control	Treatment	p-value
Child's gender	1872	0.471	0.473	0.464	0.723
(male=1)		(0.50)	(0.50)	(0.50)	
Month of birth	1872	6.155	6.148	6.178	0.860
		(3.19)	(3.18)	(3.21)	
Twin	1872	0.015	0.014	0.017	0.650
		(0.12)	(0.12)	(0.13)	
Birth order	1872	2.651	2.712	2.466	**0.013**
		(1.85)	(1.87)	(1.78)	
Mother's education	1872	3.192	3.152	3.313	0.435
		(3.86)	(3.85)	(3.89)	
Mother's age at first birth	1872	19.522	19.588	19.324	0.112
		(3.11)	(3.05)	(3.27)	
Household size	1872	6.007	5.999	6.032	0.817
		(2.71)	(2.57)	(3.11)	
Urban	1872	0.207	0.166	0.330	**0.000**
		(0.41)	(0.37)	(0.47)	
Access to water	1872	0.459	0.511	0.303	**0.000**
		(0.50)	(0.50)	(0.46)	
Access to electricity	1872	0.670	0.652	0.723	**0.005**
		(0.47)	(0.48)	(0.45)	
Access to latrine	1872	0.549	0.543	0.569	0.329
		(0.50)	(0.50)	(0.50)	
Wealth quintile 1	1872	0.322	0.344	0.255	**0.000**
		(0.47)	(0.48)	(0.44)	
Wealth quintile 2	1872	0.207	0.222	0.163	**0.007**
		(0.41)	(0.42)	(0.37)	
Wealth quintile 3	1872	0.173	0.164	0.197	0.101
		(0.38)	(0.37)	(0.40)	
Wealth quintile 4	1872	0.148	0.132	0.197	**0.001**
		(0.36)	(0.34)	(0.40)	
Wealth quintile 5	1872	0.150	0.138	0.187	**0.011**
		(0.36)	(0.35)	(0.39)	
Problem getting to hospital	1872	0.555	0.565	0.524	0.115
		(0.50)	(0.50)	(0.50)	

Notes: Standard deviations in parentheses. Ho: the means are not different. P < 0.05: reject null at 5%.

Source: NDHS 2011

Table 3.3. Comparison of mean between treatment and control districts before CB-NCP

	N	Overall	Control	Treatment	p-value
Neonatal mortality	1872	0.050	0.051	0.047	0.732
		(0.22)	(0.22)	(0.21)	
Institutional delivery	1872	0.316	0.287	0.401	**0.000**
		(0.46)	(0.45)	(0.49)	
Skilled birth (for at-home births)	1872	0.504	0.459	0.639	**0.000**
		(0.50)	(0.50)	(0.48)	
A clean kit used at delivery	784	0.217	0.211	0.240	0.423
		(0.41)	(0.41)	(0.43)	
Four antenatal visits	1194	0.511	0.516	0.495	0.527
		(0.50)	(0.50)	(0.50)	
Postnatal within 2 months	1193	0.491	0.485	0.508	0.494
		(0.50)	(0.50)	(0.50)	
Iron/folic acid in pregnancy	1194	0.763	0.758	0.778	0.489
		(0.43)	(0.43)	(0.42)	
Iron/folic acid after delivery	1193	0.397	0.391	0.418	0.412
		(0.49)	(0.49)	(0.49)	

Notes: Standard deviations in parentheses. Ho: the means are not different. $P < 0.05$: reject null at 5%.

Source: NDHS 2011

Table 3.4. Linear Probability Model Results for the Effect of Treatment on Neonatal Mortality					
	(1)	(2)	(3)	(4)	(5)
CBNCP	-0.008	-0.009	-0.010	-0.009	-0.011
	(0.010)	(0.010)	(0.010)	(0.010)	(0.012)
Post	-0.005	-0.005	-0.003	-0.004	-0.005
	(0.008)	(0.008)	(0.008)	(0.008)	(0.008)
Post*CBNCP	-0.016	-0.016	-0.016	-0.015	-0.014
	(0.013)	(0.013)	(0.013)	(0.013)	(0.013)
Child's characteristics		x	x	x	x
Mother's characteristics			x	x	x
Household characteristics				x	x
Fixed effects					
Birth year		x	x	x	x
Districts		x	x	x	x
N	3305	3305	3305	3305	3305
R-squared	0.002	0.013	0.017	0.028	0.033

* p<0.10, ** p<0.05, *** p<0.01

Bootstrapped standard errors (with 1000 reps and seed 1, by district) are reported.
Child-related covariates include gender, month of birth, twin status and birth order. Mother's characteristics include education and age at first birth. Household characteristics include wealth index, access to water, access to electricity, access to latrine, whether the household reported having trouble getting to the hospital, ethnicity and whether the household is urban.

Table 3.5. Linear Probability Model Results for the Effect of Treatment on Institutional Birth

	(1)	(2)	(3)	(4)	(5)
CBNCP	0.114***	0.095***	0.102***	0.051**	0.013
	(0.026)	(0.026)	(0.024)	(0.024)	(0.030)
Post	0.167***	0.152***	0.117***	0.141***	0.142***
	(0.020)	(0.020)	(0.018)	(0.017)	(0.017)
Post*CBNCP	-0.037	-0.016	-0.006	-0.018	-0.022
	(0.038)	(0.037)	(0.035)	(0.034)	(0.034)
Child's characteristics		x	x	x	x
Mother's characteristics			x	x	x
Household characteristics				x	x
Fixed effects					
Birth year		x	x	x	x
Districts		x	x	x	x
N	3305	3305	3305	3305	3305
R-squared	0.034	0.110	0.222	0.299	0.310

* p<0.10, ** p<0.05, *** p<0.01

Bootstrapped standard errors (with 1000 reps and seed 1, by district) are reported.
Child-related covariates include gender, month of birth, twin status and birth order. Mother's characteristics include education and age at first birth. Household characteristics include wealth index, access to water, access to electricity, access to latrine, whether the household reported having trouble getting to the hospital, ethnicity and whether the household is urban.

Table 3.6. Linear Probability Model Results for the Effect of Treatment on Professional-attended Birth (for deliveries that take place at home)

	(1)	(2)	(3)	(4)	(5)
CBNCP	0.150***	0.141***	0.141***	0.025	0.055
	(0.032)	(0.032)	(0.032)	(0.030)	(0.037)
Post	-0.011	-0.014	-0.020	-0.007	-0.003
	(0.022)	(0.022)	(0.022)	(0.020)	(0.020)
Post*CBNCP	-0.051	-0.041	-0.039	-0.034	-0.044
	(0.054)	(0.053)	(0.053)	(0.048)	(0.048)
Child's characteristics		x	x	x	x
Mother's characteristics			x	x	x
Household characteristics				x	x
Fixed effects					
Birth year		x	x	x	x
Districts		x	x	x	x
N	2036	2036	2036	2036	2036
R-squared	0.015	0.038	0.046	0.227	0.258

* p<0.10, ** p<0.05, *** p<0.01

Bootstrapped standard errors (with 1000 reps and seed 1, by district) are reported.
Child-related covariates include gender, month of birth, twin status and birth order. Mother's characteristics include education and age at first birth. Household characteristics include wealth index, access to water, access to electricity, access to latrine, whether the household reported having trouble getting to the hospital, ethnicity and whether the household is urban.

Table 3.7. Linear Probability Model Results for the Effect of Treatment on the Use of a Clean Kit during Delivery					
	(1)	(2)	(3)	(4)	(5)
CBNCP	0.025	0.014	0.016	-0.008	0.038
	(0.036)	(0.036)	(0.035)	(0.036)	(0.041)
Post	-0.074***	-0.085***	-0.090***	-0.066***	-0.062***
	(0.022)	(0.022)	(0.021)	(0.021)	(0.021)
Post*CBNCP	0.088*	0.103**	0.106**	0.108**	0.091*
	(0.052)	(0.051)	(0.050)	(0.050)	(0.049)
Child's characteristics		x	x	x	x
Mother's characteristics			x	x	x
Household characteristics				x	x
Fixed effects					
Birth year		x	x	x	x
Districts		x	x	x	x
N	1481	1481	1481	1481	1481
R-squared	0.012	0.045	0.083	0.129	0.186

* p<0.10, ** p<0.05, *** p<0.01

Bootstrapped standard errors (with 1000 reps and seed 1, by district) are reported.
Child-related covariates include gender, month of birth, twin status and birth order. Mother's characteristics include education and age at first birth. Household characteristics include wealth index, access to water, access to electricity, access to latrine, whether the household reported having trouble getting to the hospital, ethnicity and whether the household is urban.

Table 3.8. Linear Probability Model Results for the Effect of Treatment on 'At Least Four Antenatal Visits'					
	(1)	(2)	(3)	(4)	(5)
CBNCP	-0.014	-0.028	-0.023	-0.035	-0.019
	(0.025)	(0.025)	(0.023)	(0.024)	(0.029)
Post	0.202***	0.189***	0.155***	0.178***	0.175***
	(0.020)	(0.019)	(0.019)	(0.018)	(0.018)
Post*CBNCP	0.072*	0.091**	0.103***	0.091**	0.094***
	(0.039)	(0.038)	(0.036)	(0.036)	(0.036)
Child's characteristics		x	x	x	x
Mother's characteristics			x	x	x
Household characteristics				x	x
Fixed effects					
Birth year		x	x	x	x
Districts		x	x	x	x
N	3305	3305	3305	3305	3305
R-squared	0.050	0.096	0.192	0.239	0.251

* p<0.10, ** p<0.05, *** p<0.01

Bootstrapped standard errors (with 1000 reps and seed 1, by district) are reported.
Child-related covariates include gender, month of birth, twin status and birth order. Mother's characteristics include education and age at first birth. Household characteristics include wealth index, access to water, access to electricity, access to latrine, whether the household reported having trouble getting to the hospital, ethnicity and whether the household is urban.

Table 3.9. Linear Probability Model Results for the Effect of Treatment on 'Postnatal Visit within Two Weeks of Birth'

	(1)	(2)	(3)	(4)	(5)
CBNCP	0.015	0.004	0.008	-0.002	0.050*
	(0.025)	(0.024)	(0.024)	(0.024)	(0.029)
Post	0.160***	0.147***	0.122***	0.146***	0.141***
	(0.020)	(0.020)	(0.019)	(0.019)	(0.019)
Post*CBNCP	0.039	0.053	0.062*	0.050	0.053
	(0.039)	(0.038)	(0.037)	(0.037)	(0.037)
Child's characteristics		x	x	x	x
Mother's characteristics			x	x	x
Household characteristics				x	x
Fixed effects					
Birth year		x	x	x	x
Districts		x	x	x	x
N	3305	3305	3305	3305	3305
R-squared	0.031	0.057	0.115	0.166	0.182

* p<0.10, ** p<0.05, *** p<0.01

Bootstrapped standard errors (with 1000 reps and seed 1, by district) are reported.
Child-related covariates include gender, month of birth, twin status and birth order. Mother's characteristics include education and age at first birth. Household characteristics include wealth index, access to water, access to electricity, access to latrine, whether the household reported having trouble getting to the hospital, ethnicity and whether the household is urban.

Table 3.10. Linear Probability Model Results for the Effect of Treatment on the Use of Folic Acid before Delivery

	(1)	(2)	(3)	(4)	(5)
CBNCP	0.012	0.002	0.007	-0.017	-0.017
	(0.027)	(0.027)	(0.026)	(0.026)	(0.031)
Post	0.307***	0.297***	0.274***	0.293***	0.292***
	(0.018)	(0.018)	(0.018)	(0.017)	(0.017)
Post*CBNCP	0.025	0.039	0.047	0.037	0.039
	(0.036)	(0.035)	(0.034)	(0.034)	(0.034)
Child's characteristics		x	x	x	x
Mother's characteristics			x	x	x
Household characteristics				x	x
Fixed effects					
Birth year		x	x	x	x
Districts		x	x	x	x
N	3305	3305	3305	3305	3305
R-squared	0.103	0.132	0.180	0.230	0.239

* p<0.10, ** p<0.05, *** p<0.01

Bootstrapped standard errors (with 1000 reps and seed 1, by district) are reported.
Child-related covariates include gender, month of birth, twin status and birth order. Mother's characteristics include education and age at first birth. Household characteristics include wealth index, access to water, access to electricity, access to latrine, whether the household reported having trouble getting to the hospital, ethnicity and whether the household is urban.

Table 3.11. Linear Probability Model Results for the Effect of Treatment on the Use of Folic Acid after Delivery					
	(1)	(2)	(3)	(4)	(5)
CBNCP	0.017	0.009	0.013	0.011	0.038
	(0.023)	(0.023)	(0.023)	(0.024)	(0.027)
Post	0.213***	0.202***	0.175***	0.194***	0.190***
	(0.019)	(0.018)	(0.018)	(0.017)	(0.017)
Post*CBNCP	-0.022	-0.010	-0.001	-0.008	-0.004
	(0.037)	(0.037)	(0.036)	(0.035)	(0.034)
Child's characteristics		x	x	x	x
Mother's characteristics			x	x	x
Household characteristics				x	x
Fixed effects					
Birth year		x	x	x	x
Districts		x	x	x	x
N	3305	3305	3305	3305	3305
R-squared	0.047	0.066	0.133	0.168	0.188

* p<0.10, ** p<0.05, *** p<0.01

Bootstrapped standard errors (with 1000 reps and seed 1, by district) are reported.
Child-related covariates include gender, month of birth, twin status and birth order. Mother's characteristics include education and age at first birth. Household characteristics include wealth index, access to water, access to electricity, access to latrine, whether the household reported having trouble getting to the hospital, ethnicity and whether the household is urban.

Table 3.12. Linear Probability Model Results for the Effect of Treatment on Taking Tetanus Vaccine					
	(1)	(2)	(3)	(4)	(5)
CBNCP	0.034	0.024	0.027	-0.004	-0.016
	(0.027)	(0.027)	(0.026)	(0.027)	(0.032)
Post	0.201***	0.191***	0.168***	0.187***	0.190***
	(0.019)	(0.019)	(0.019)	(0.018)	(0.018)
Post*CBNCP	-0.015	0.000	0.008	-0.003	-0.005
	(0.039)	(0.038)	(0.037)	(0.037)	(0.037)
Child's characteristics		x	x	x	x
Mother's characteristics			x	x	x
Household characteristics				x	x
Fixed effects					
Birth year		x	x	x	x
Districts		x	x	x	x
N	3305	3305	3305	3305	3305
R-squared	0.039	0.066	0.109	0.166	0.174

* p<0.10, ** p<0.05, *** p<0.01

Bootstrapped standard errors (with 1000 reps and seed 1, by district) are reported.
Child-related covariates include gender, month of birth, twin status and birth order. Mother's characteristics include education and age at first birth. Household characteristics include wealth index, access to water, access to electricity, access to latrine, whether the household reported having trouble getting to the hospital, ethnicity and whether the household is urban.

Table 3.13. CBNCP's Effects on Outcomes Assuming the Program Started in 2004				
	Neonatal mortality	Institutional birth	Professional-attended birth	Use of clean kit during delivery
CBNCP	-0.003	0.013	0.048***	-0.027
	(0.010)	(0.022)	(0.017)	(0.031)
Post	-0.010	0.037***	0.013	-0.031*
	(0.007)	(0.013)	(0.013)	(0.018)
Post*CBNCP	-0.010	-0.002	-0.003	0.045
	(0.013)	(0.028)	(0.022)	(0.036)
N	3705	3705	2982	2105
R-squared	0.059	0.269	0.137	0.114
	At least four antenatal visits	Post-natal visit within two weeks	Took tetanus vaccines	Took folic acid/iron during pregnancy
CBNCP	-0.022	-0.055**	-0.088***	-0.064*
	(0.029)	(0.024)	(0.033)	(0.034)
Post	0.02	-0.005	0.007	0.079***
	(0.018)	(0.015)	(0.020)	(0.019)
Post*CBNCP	0.017	0.023	0.041	0.033
	(0.035)	(0.029)	(0.041)	(0.040)
N	2671	2152	2670	2671
R-squared	0.291	0.084	0.222	0.266

* p<0.10, ** p<0.05, *** p<0.01

Bootstrapped standard errors (with 1000 reps and seed 1, by district) are reported.
All individual, parental, household and community controls are included in all specifications. Child-related covariates include gender, month of birth, twin status and birth order. Mother's characteristics include education and age at first birth. Household characteristics include wealth index, access to water, access to electricity, access to latrine, whether the household reported having trouble getting to the hospital, ethnicity and whether the household is urban.

Table 3.14. Comparison of Mean between Treatment and Control Districts before CB-NCP, for Nutritional Indicators

	N	Overall	Control	Treatment	p-value
Mother is anemic	897	0.319	0.308	0.351	0.230
		(0.466)	(0.462)	(0.478)	
Low birth weight (child)	588	0.143	0.133	0.165	0.296
		(0.350)	(0.339)	(0.372)	
Child is anemic	832	0.349	0.358	0.322	0.353
		(0.477)	(0.480)	(0.469)	
Child had diarrhea in past two weeks	1765	0.077	0.077	0.077	0.965
		(0.267)	(0.267)	(0.266)	
Child underweight	836	0.318	0.311	0.339	0.434
		(0.466)	(0.463)	(0.475)	
Child weight-for-age	836	-1.583	-1.588	-1.570	0.821
		(1.000)	(0.976)	(1.067)	

Notes: Standard deviations are in parentheses. Ho: the means are not different. $P < 0.05$: reject null at 5%.

Table 3.15. Results from Regressing Whether Data are Missing on Treatment (for Nutritional Outcomes)

	Mother anemic	Low birth weight	Child anemic	Diarrhea prevalence	Child underweight	Child weight-for-age
CBNCP	0.058	0.015	-0.023	-0.005	0.005	0.054
	(0.039)	(0.024)	(0.041)	(0.016)	(0.042)	(0.138)
N	1558	1247	1327	3146	1467	1467
R-squared	0.003	0.000	0.000	0.000	0.000	0.000

* p<0.10, ** p<0.05, *** p<0.01

Table 3.16. Linear Probability Model Results for the Effect of Treatment on Short-term Nutritional Outcomes

	Mother is anemic	Child had low birth weight	Child is anemic	Child had diarrhea	Child is underweight	Child's weight-for-age
CBNCP	-0.072	-0.013	-0.107**	0.000	0.040	0.000
	(0.047)	(0.037)	(0.046)	(0.020)	(0.048)	(0.107)
Post	0.049*	-0.014	0.301***	0.131***	-0.017	0.196***
	(0.028)	(0.023)	(0.031)	(0.015)	(0.027)	(0.062)
Post*CBNCP	0.050	-0.022	0.037	-0.007	-0.085	0.146
	(0.058)	(0.042)	(0.063)	(0.030)	(0.052)	(0.121)
N	1558	1247	1327	3146	1467	1467
R-squared	0.084	0.086	0.161	0.060	0.109	0.192

* p<0.10, ** p<0.05, *** p<0.01

Bootstrapped standard errors (with 1000 reps and seed 1) are reported.
All individual, parental, household and community controls are included in all specifications. Child-related covariates include gender, month of birth, twin status and birth order. Mother's characteristics include education and age at first birth. Household characteristics include wealth index, access to water, access to electricity, access to latrine, whether the household reported having trouble getting to the hospital, ethnicity and whether the household is urban.

Appendices

Table 3.A1. Construction of Key Outcome Variables

Name	Nature	Description
Under-5 mortality	Binary	=1 if the child is dead at the time of the survey, 0 otherwise. In the overall sample, there are 3,691 children of which 165 died before the age of 5. 3691 matches with the number of observations in Table 4.
Neonatal mortality	Binary	=1 if the child is dead at the time of the survey and was less than or equal to one month old at the time of death, 0 otherwise. In the overall sample, there are 3,691 children of which 123 died within the first month of birth.
Institutional delivery	Binary	=1 if the delivery took place in place other than home. In the survey, respondents were asked where a child was born and given 13 options. Two referred to "respondent's home" and "other home". These are coded as 0. Other options have been coded as 1. Out of the 3,691 children in the overall sample, 1,469 were born in the hospital while the remaining 2,222 were not. 2,222 matches with the number of observations in Table 6 in which the analysis is done on the sample of children who were delivered at home.
Skilled birth attendance	Binary	=1 if the respondents received help from a trained professional (doctor, nurse, health attendant, FCHV) other than her relative or friend, 0 otherwise. In the survey, respondents were asked who, if anyone, provided assistance during birth of her child. Out of the 2,222 children who were born at home, 638 were attended by a skilled professional.
Clean kit used during delivery	Binary	=1 if the birth took place at home and a clean kit was used at the time of delivery, 0 otherwise. In the survey, the respondents were asked if a clean kit was used during delivery. Of the 2,222 births that took place at home, 321 were reported to have used the clean kit, 1,305 were reported to have not used it, while 596 have missing information.
At least four antenatal visits	Binary	=1 if the mother of the child went for at least four antenatal visits. Of 3,691 children in the overall sample, mothers of 1,599 went for at least four antenatal visits, 1,279 did not and 813 had missing information.
Postnatal visit within two months	Binary	=1 if the mother of the child went for a postnatal checkup within two months of birth. Of 3,691 children in the overall sample, mothers of 1,469 went for postnatal checkup within two months of birth, while mothers of 1,406 did not. Information is missing for mothers of 816 children.
At least two tetanus shots	Binary	=1 if the mother of the child took at least two tetanus shots during pregnancy. The survey asks the respondents if, and how many, tetanus shots they took during pregnancy. In the sample, mothers of 1,989 children took two or more shots, mothers of 882 did not, and information is missing for mothers of 820.

Took iron/folic acid during pregnancy	Binary	=1 if the mother of the child took iron/folic acid during pregnancy. In the sample, mothers of 2,337 children reported taking iron/folic acid pills during pregnancy, 540 reported not taking them and information is missing for 814.
Took iron/folic acid after delivery	Binary	=1 if the mother of the child took iron/folic acid after the child was born. In the sample, mothers of 1,294 children reported taking iron/folic acid pills after delivery, 1,583 reported not taking them and information is missing for 814.
Low birth weight (all children)	Binary	=1 if the birth weight was less than 2,490 gram. 2,246 children in the sample had missing data. This is about 60% of the total sample. Of those whose information on birth weight was not missing, 1,252 were underweight, while 193 (13% of those who were measured) were not.
Mother is anemic	Binary	=if mother's anemia status has been categorized as 'severe', 'moderate' or 'mild' and 0 otherwise. In the overall NDHS sample, of those whose anemia status was measured, about 65% women are not anemic, rest are.
Child is anemic	Binary	=if the child's anemia status has been categorized as 'severe', 'moderate' or 'mild' and 0 otherwise. In the overall NDHS sample, of those whose anemia status was measured, about 52% children are not anemic, rest are.
Child had diarrhea in the past two weeks	Binary	=1 if the child was reported to have diarrhea during the two weeks preceding the survey. Approximately 14% children in the overall sample have had diarrhea during the two weeks preceding the survey.
Child is underweight	Binary	=1 if the child's weight is two or more standard deviation lower than the weight of a child of the same age in the reference population (provided by WHO). This is calculated based on the weight-for-age z-scores.
Child's weight-for-age	Continuous	Weight-for-age z-score available in the survey.

Table 3.A2. Results from regressing outcomes on an interaction of treatment and birth year to check the parallel trend assumption

	Neonatal mortality	Institutional birth	Professional-attended birth	Use of clean kit during delivery
Treatment*birth year	0.00000	0.00000	0.00003	0.00003
	(0.00001)	(0.00001)	(0.00002)	(0.00002)
N	2496	2496	1638	1086
R-squared	0.039	0.296	0.260	0.193
	At least four antenatal visits	**Post-natal visit within two weeks**	**Took tetanus vaccines**	**Took folic acid/iron during pregnancy**
Treatment*birth year	-0.00002	0.00002	-0.00003*	-0.00002
	(0.00002)	(0.00002)	(0.00002)	(0.00002)
N	1755	1753	1751	1755
R-squared	0.256	0.178	0.200	0.220

* p<0.10, ** p<0.05, *** p<0.01

Bootstrapped standard errors (with 1000 reps and seed 1) are reported.
All individual, parental, household and community characteristics are included in all specifications. Child-related covariates include gender, month of birth, twin status and birth order. Mother's characteristics include education and age at first birth. Household characteristics include wealth index, access to water, access to electricity, access to latrine, whether the household reported having trouble getting to the hospital, ethnicity and whether the household is urban.

Table 3.A3. Results from Regressing Nutritional Outcomes on an Interaction of Treatment and Birth Year to Check the Parallel Trend Assumption

	Mother anemic	Low birth weight	Child anemic	Child had diarrhea	Child underweight	Child's weight-for-age
Treatment*birthyear	-0.00003	-0.00004	-0.00008***	0.00000	0.00003	-0.00004
	(0.00003)	(0.00002)	(0.00003)	(0.00001)	(0.00003)	(0.00005)
N	897	588	832	1765	836	836
R-squared	0.076	0.137	0.110	0.032	0.106	0.162

* p<0.10, ** p<0.05, *** p<0.01

Bootstrapped standard errors (with 1000 reps and seed 1) are reported.
All individual, parental, household and community controls are included in all specifications. Child-related covariates include gender, month of birth, twin status and birth order. Mother's characteristics include education and age at first birth. Household characteristics include wealth index, access to water, access to electricity, access to latrine, whether the household reported having trouble getting to the hospital, ethnicity and whether the household is urban.