Skip all front matter: Jump to Page 1

Dissertation

# Essays on Family Welfare and Indian Development Policy

Sinduja V. Srinivasan

RAND CORPORATION

Dissertation

# Essays on Family Welfare and Indian Development Policy

Sinduja V. Srinivasan

This document was submitted as a dissertation in August 2014 in partial fulfillment of the requirements of the doctoral degree in public policy analysis at the Pardee RAND Graduate School. The faculty committee that supervised and approved the dissertation consisted of Krishna Kumar (Chair), Shanthi Nataraj, and Peter Glick.

RAND PARDEE RAND GRADUATE SCHOOL

*To Amma and Appa: You have supported me in every step of this process; without you, this would not be possible.*



Published 2014 by the RAND Corporation
1776 Main Street, P.O. Box 2138, Santa Monica, CA 90407-2138
1200 South Hayes Street, Arlington, VA 22202-5050
4570 Fifth Avenue, Suite 600, Pittsburgh, PA 15213-2665
RAND URL: http://www.rand.org/
To order RAND documents or to obtain additional information, contact
Distribution Services: Telephone: (310) 451-7002;
Fax: (310) 451-6915; Email: order@rand.org

# Table of Contents

## Abstract

My dissertation is concerned with family welfare and economic development in India. In my first two papers, I consider the potential for India's public works program to contribute to long-term growth and poverty reduction by examining the impact on household entrepreneurship and investments in child health. The third paper analyzes the impact of increasing male earnings inequality on female marital and education outcomes.

Essays I and II: Long-term impact of NREGS

First, I analyze the impact of India's National Rural Employment Guarantee Scheme (NREGS) on selection into entrepreneurship using a model of household occupational choice. My hypothesis is that NREGS allows liquidity constrained individuals to accumulate some savings, enabling subsequent investment in a risky, but more profitable, entrepreneurial venture. Using linear regression methods, I find that rates of non-farm entrepreneurship increase by 3 percentage points in program areas (increasing rates from 15 percent to 18 percent), compared to areas that did not receive NREGS. The effects are concentrated amongst rural non-farm entrepreneurs, who have been often excluded from the formal credit market. The results suggest that by acting as a source of credit, NREGS impacts household occupational choice, contributing to increased income, and ultimately promoting current and future family welfare.

Next, adapting the neoclassical model of labor supply, I examine the impact of NREGS on maternal investments in children's health, and the subsequent impact on child health outcomes. The net effect of a public works program on child health is unclear given the trade-off women face between working and care-giving. NREGS raises the value of employment, reducing the incentive for women to spend time at home on care-giving activities, potentially negatively impacting children's health. But income from NREGS may allow women to purchase high-quality healthcare, resulting in improved health outcomes for children. Using linear, non-linear, and non-parametric methods, I find that women substitute towards employment and reduce investment in time-intensive health activities: they delay their first prenatal visit, reduce the total number of visits, and decrease the duration of breastfeeding. Further, the results show that women are using their additional earned income to purchase money-intensive health inputs: women in NREGS areas are more likely to deliver in a healthcare facility.

Understanding how the program can reduce long-term poverty rates through its impact on the skills, productivity, and health of the labor force is important for economic growth in developing countries with similar policies, in addition to analyzing the sustainability of NREGS itself.

Essay III: Income inequality and female marriage

Early marriage has been associated with negative health, educational, and economic outcomes, especially for women. This paper utilizes the logic of the job search model to explore the role of increased male income inequality in delaying female marriage and the subsequent impact on female educational attainment. The results from linear regression methods indicate that greater male income inequality decreases female marriage rates by about 2 percentage points and delays marriage by one third of a year. In response to their longer search times, women attend school, and are more likely to finish high school or attend college. The acquisition of more education due to increased search duration on the marriage market might increase the future income stream of women, contributing to their economic well-being. In addition, greater educational attainment may improve their bargaining power in the household, allowing women to direct resources towards children's, especially girls', education and health. These findings will inform policy decisions aimed at increasing women's empowerment in India.

# Acknowledgments

I am grateful for the tremendous support of my stellar committee: Krishna Kumar, Peter Glick, and Shanthi Nataraj. They generously gave their time reading drafts, listening to practice presentations, and supporting my job search process. Their guidance and patience have helped me to become a better researcher and I will strive to provide such mentorship to others.

A.V. Chari has been a wonderful mentor; researching and writing with him has been instrumental in improving my own work. Will Carrington kindly agreed to be my outside reader. His meticulous reading and incredibly helpful comments have helped shape the policy discussion in these papers.

My dissertation was funded by the Palevsky Dissertation Award, the Rosenfeld Dissertation Award, and the JL Foundation. This generous support allowed me to procure necessary data and spend more than a year concentrating solely on my thesis work.

Finally, I thank my family and friends for their support and love through the past few years. My parents, remembering their own PhD days, supported me through the lows, celebrated the highs, and encouraged me always.

# Impact of Public Works on Household Occupational Choice: Evidence from NREGS in India

Sinduja V. Srinivasan*

August 2014

**Abstract**

I analyze the impact of India's public employment generation program (NREGS) on entrepreneurship. One of the main barriers to entrepreneurship in India is a lack of access to capital. My hypothesis is that NREGS allows liquidity constrained individuals to accumulate savings, enabling subsequent investment in a risky, but more profitable, venture, and ideally, permanent graduation from poverty. Taking advantage of the quasi-experimental nature of the program, I use a nationally representative data set to estimate the impact of NREGS on selection into entrepreneurship. I find that rates of non-agricultural entrepreneurship increase by 3 percentage points in NREGS districts (increasing rates from 15% to 18%), compared to areas that did not receive the program. This result is robust to various specifications, including two falsification tests. The results suggest that by acting as a source of credit, NREGS impacts household occupational choice, contributing to increased income, and ultimately promoting current and future family welfare.

*Pardee RAND Graduate School. Email: sinduja@rand.org
I thank Krishna Kumar, Peter Glick, Shanthi Nataraj, and participants in the PRGS and Labor & Population Brown Bag Seminars at the RAND Corporation for valuable comments and suggestions. This work was supported by a Rosenfeld Dissertation Award. Any remaining errors are my own.

## 1 Introduction

Social safety net programs are viewed by governments as a powerful policy tool to support economically vulnerable groups. Increasingly, developing countries are relying on one type of social safety net, the employment generation program, as a way to enhance livelihood security for the poor, while also developing local infrastructure. Such programs have been successfully implemented in many middle and emerging economies, including Argentina, Ethiopia, India, and South Africa.[1] The goal is to provide households with increased income and food security in times when finding employment is difficult or to supplement insufficient incomes.

Whether employment generation programs can contribute to permanent reductions in poverty remains unclear. It is difficult to estimate the impact of these programs, as they are not randomized; without a comparable control group, a causal analysis cannot be conducted. In developed countries, there is evidence that public works programs are less successful in improving employment outcomes and fostering growth than other types of active labor market policies (Card et al., 2010; Kluve, 2006). The limited empirical evidence from programs in developing countries has focused on short-term outcomes, such as income, consumption, and participation in public employment (Ravi and Engler, 2012; Azam, 2012; Imbert and Papp, 2013), leaving unanswered the question of how employment generation programs affect growth in the long-run.

India's National Rural Employment Guarantee Scheme (NREGS) is the largest direct employment program in the world and has increased work opportunities for millions of the rural poor in India. Since its inception in 2006, the program has provided employment to nearly 50 million Indian households (about 36 percent of the rural labor force). The program continues to expand: NREGS participation rates more than doubled in the four years after implementation, and current program expenditures have risen to \$7.5 billion, approximately 1 percent of GDP. These costs cannot be maintained over an extended period. NREGS will be a sustainable and cost-effective policy only if it supports a graduation mechanism, enabling program participants to make a lasting change in their lives.

In this paper, I research how NREGS can result in sustained economic development and poverty alleviation, and in doing so, make three important contributions. First, I consider how NREGS

[1] Devereux and Solomon (2006) provide a review of recent employment generation programs. See also Subbarao et al. (2013).

supports the transition to a new occupation by examining the impact of the program on a novel outcome: household entrepreneurship status. I propose that NREGS serves as a conduit for agents to overcome financial constraints and subsequently engage in entrepreneurial activities, fostering economic development. Banerjee and Newman (1993) show that with capital market imperfections, the initial wealth distribution and persistence of economic institutions determine whether an economy achieves prosperity (entrepreneurship) or stagnation (subsistence self-employment). The implication is that a one-time transfer, which changes the wealth distribution, can have permanent growth effects. NREGS may be viewed as such a finite transfer that can break the pattern of stagnation. Little is known about how active labor market policies can support entrepreneurship, apart from limited evidence on direct grants for self-employment in Europe (Carling and Gustafson, 1999).[2] Examining the impact on household occupation provides insight into one channel through which public works programs can provide long-term economic benefits to participants, thus informing policy in India and countries with similar programs.

Second, I adapt a model of household occupational choice (Evans and Jovanovic, 1989) to the context of a public works program, generating two testable predictions to support my empirical analysis. The first prediction is that a public works program will increase the extensive margin of entrepreneurship: the program allows households previously excluded from entrepreneurship to acquire the minimum level of capital necessary for a venture, thus increasing the share of households engaged in entrepreneurship. The second prediction is that a public works program will increase the intensive margin of entrepreneurship: after program implementation, households initially employing sub-optimal levels of capital are able to more intensively engage in the entrepreneurial venture, using income earned from the program to acquire the optimal level of capital. I then estimate the effect of NREGS on different measures of entrepreneurship, comparing Indian districts that received the program to those that did not. I exploit the timing of the program, implementing a differences-in-differences methodology, which allows me to control for time-invariant characteristics within districts that might also be correlated with entrepreneurship.

Third, I examine how NREGS differentially affects the rural non-farm and rural farm sectors. To date, there has been limited research on how government policies can promote rural non-farm

[2]Gilligan et al. (2008) find that participation in Ethiopia's social safety net programs is associated with increased household business activities. However, this study analyzes the impact of the employment generation program PSNP and other food security programs jointly. My research considers the effect of a public works program singularly.

entrepreneurship and current evaluations of NREGS do not analyze the rural non-farm sector separately. The Indian rural non-farm sector has grown steadily over the last thirty years (Coppard, 2001; Himanshu et al., 2011), while the agricultural sector has been shrinking.[3] Within the rural non-farm sector, entrepreneurship has been the most dynamic source of income growth, driven by the expansion of productive household activities, rather than agrarian distress (Binswanger-Mikhize, 2012). However, there is evidence that a significant faction of the rural population is precluded from engaging in such entrepreneurial opportunities. The main barrier to entry is a lack of access to credit (Coppard, 2001). Households are forced to use their own land as capital to finance entrepreneurial ventures (Lanjouw and Murgai, 2008), and households without sufficient assets are effectively barred from engaging in entrepreneurship altogether. Thus, for researchers and policymakers, understanding how government policy can support the expansion of the productive entrepreneurial rural non-farm sector is crucial, as it has played an important role in rural development and poverty reduction. My work analyzing the impact of NREGS on entrepreneurship attempts to shed light on this issue.

My results are consistent with the theoretical predictions. In my analysis of the impact of India's rural public works policy, NREGS, on entrepreneurship, I find that the program differentially affects rural non-farm and rural farm entrepreneurs. NREGS positively impacts rates of extensive rural non-farm entrepreneurship, resulting in a three percentage point increase (from 15 percent to 18 percent). Further, these results are robust to two separate falsification tests, indicating that the program does increase the extensive margin. The impact of NREGS on the intensive margin (measured by the share of household members engaged in entrepreneurship and the time spent on the main entrepreneurial activity) is less stark; I do not find a significant effect. The results provide some evidence that liquidity constrained households previously excluded from entering rural non-farm entrepreneurship are able to use NREGS as a source of credit to acquire the necessary capital for their entrepreneurial venture. In contrast, NREGS has little or no impact on rural farm entrepreneurship. Workers in this sector are often subsistence entrepreneurs who may be using NREGS to leave self-employment, substituting towards the better employment opportunities offered by the program.

[3] See Visaria and Basant (1994), Fisher et al. (1997), Dev and Ravi (2007), Himanshu (2005), Lanjouw and Murgai (2008), Abraham (2011), and Binswanger-Mikhize (2012) for detailed explanations on the historical and economic reasons behind this trend.

The paper proceeds as follows. Section 2 provides a description of the National Rural Employment Guarantee Scheme. In Section 3, I discuss the theoretical framework underlying the econometric approach of Section 4. Section 5 describes the data. In Section 6, I discuss the main findings, and extensions to and robustness of the baseline results. Section 7 concludes.

## 2 National Rural Employment Guarantee Scheme

The National Rural Employment Guarantee Scheme (NREGS) was implemented in 2006, after the passage of the National Rural Employment Guarantee Act (NREGA) in 2005.[4] As the name suggests, NREGS is implemented only in rural areas. The program was rolled out in phases across rural India (see Figure 1). In 2006 (Phase 1), NREGS was implemented in the 200 least developed districts.[5] In 2007 (Phase 2), the program was implemented in another 130 districts. The remaining 285 districts received the program in 2008-2009 (Phase 3).

Any rural household can opt in to the program. Each enrolled household is guaranteed a maximum of 100 days of labor, which may be allocated across the adult household members in any fashion, at any time of the year. Wages are paid according to Minimum Wage Act of 1948, and are not less than 60 rupees ($1.12) per day, with equal wages for men and women.[6] The program focuses on projects that improve agricultural infrastructure and productivity, such as water conservation and water harvesting, drought proofing, land development, flood control and protection, and rural connectivity (Ministry of Rural Development, 2005). As with other employment generation schemes, NREGS is a demand-driven program: work is provided to those who ask (enroll).

The high enrollment rates and vast geographic scale of NREGS make it the largest employment generation program in the world (Ravi and Engler, 2012). In the first year of NREGS (2006-2007), 21 million households (about 15 percent of the rural labor force) received employment, totaling 905 million person days of work, equivalent to 1.0% of government expenditure, about $2.5 billion (NREGS website).[7] In the most recent year of the program (2011-2012), nearly 50 million

[4]The Ministry of Rural Development (MRD) in India renamed NREGA/NREGS to the Mahatma Gandhi National Rural Employment Act/Scheme in 2009. I continue to refer to the act and scheme as NREGA and NREGS, respectively.

[5]A district is an administrative unit in India, similar to a U.S. county.

[6]Recently, NREGS increased the employment ceiling to 150 days per year and raised the daily minimum wage to 100 rupees ($1.50). However, these changes took effect after the period I am studying, 2004-2007.

[7]Author's calculation based on current dollar-rupee exchange rates, using NREGS outlay data obtained from

households (about 36 percent of the rural labor force) received employment, totaling more than 2.1 billion person days of work, equivalent to 3.1% of government expenditure, about $7.5 billion (NREGS website). This rapid growth rate raises questions about the sustainability and cost-effectiveness of NREGS, which I address in this paper by examining the impact on one possible path of program graduation: household entrepreneurship.

As NREGS was not randomized, but was implemented in high need areas first, this will impact my empirical approach. I exploit the quasi-experimental nature of the program: the timing of the NREGS rollout allows me to consider early (Phase 1/2) districts as the treated group, and late (Phase 3) districts as the untreated. To account for time invariant characteristics across early and late districts that may affect the outcome, I include district fixed effects in my econometric specifications. However, I cannot control for household level selection into the program. To address this issue, I use the survey household weights to generate aggregated district rates of entrepreneurship, and estimate an intent-to-treat effect. Table 1 shows there are pre-program differences across treatment and untreated districts (the groups are not balanced). This is to be expected given that NREGS districts were selected due to their poorer economic outcomes in 2004. I detail my econometric approach to address the differences in treatment and untreated groups in Section 4.

## 3 A Model of Household Occupational Choice

In this section, I discuss the analytical framework underlying my empirical approach. I adapt the Evans and Jovanovic (1989) model of household occupational choice under liquidity constraints to the context of a public works program. Consider a household prior to the implementation of a workfare scheme. The household has the option of engaging in a riskless employment option, with a known wage rate, resulting in a fixed income. Alternatively, the household could engage in a (risky) entrepreneurial activity, using an exogenous endowment to procure the required capital. There is no borrowing or lending in this economy, so a household's initial wealth is the only way to obtain capital for the entrepreneurial activity. This assumption reflects the fact that rural households in India, the population covered by NREGS, have very limited access to formal credit institutions. Informal institutions often charge exorbitant interest rates, which further

NREGS website (accessed August 27, 2013), 2001 Indian Census data, and Indian GDP data obtained from the World Bank website (accessed August 27, 2013).

excludes these households from the credit market (Bhattacharjee and Rajeev, 2010; Mahajan and Ramola, 1996; Hoff and Stiglitz, 1990). The profits from such an activity are directly related to the household's ability level (which is known and exogenous). Thus there will be households that become entrepreneurs and those that do not. Among those that do not become entrepreneurs, some will choose not to do so because of low ability; it is not profitable for these households to engage in the entrepreneurial activity. However, other households will be precluded from engaging in entrepreneurship though it is profitable, because they cannot afford the required capital; their initial wealth level is too low and so the liquidity constraint is binding.

With the onset of the workfare program, it is the households restricted from entrepreneurship that are of most interest. Anyone who wants a job in public works will obtain one (I assume there is no competition for work under the program), but there is a cap on the total income a household can earn through the workfare program each year. Thus previously constrained households now have the ability to earn additional income. They will use the program to earn enough to buy the capital required for the entrepreneurial activity, and overcome their financial constraint. Here the workfare program acts as a credit institution where instead of requiring interest for the loan, it requires work.

### 3.1 Household Decision

This is a static model, and in every period, households undertake the occupational decision. Households are endowed with wealth $M$ and ability level $\theta$. There are two employment options. The first is to become a wage household, earning a fixed wage $w$. The second is to become an entrepreneurial household, with profit from the venture given by the (convex) function:

$$y = \theta k^{\alpha}, \alpha \in [0, 1) \tag{1}$$

Thus households with greater ability yield a higher total (and marginal) product, for all levels of capital. The minimum level of capital required to enter into entrepreneurship is $\underline{k}$. The household must decide whether to engage in the entrepreneurial activity or the wage earning option. All households are risk neutral.

In order to choose an occupation, the household must first determine its income under the entrepreneurial activity. It will solve for the profit maximizing level of capital, $k^*$. The price of capital is normalized to one. Households purchase capital with their initial endowment. Households do not have access to credit markets, and there is no borrowing or lending between households. Thus a household can purchase a maximum level of capital equal to $M$. The household solves:

$$\max_k \theta k^\alpha - k$$

$$FOC \Rightarrow k^* = (\theta\alpha)^{\frac{1}{1-\alpha}} \tag{2}$$

Where $k^*$ is a function of household ability.

### 3.2 Implications of Liquidity Constraints

Case 0: Wage households

$$w > \theta k^\alpha - k, \forall k \tag{3}$$

Households facing Condition 3 will never choose entrepreneurship because the outside employment option is more attractive, regardless of the level of investment in entrepreneurship (these households have very low ability).

Case 1: Excluded households

$$w \leq \theta \underline{k}^\alpha - \underline{k}, M < \underline{k} \tag{4}$$

While these households would be better off by engaging in entrepreneurship, their wealth level is too low to procure even the minimum capital requirement, $\underline{k}$. As such, they are excluded from entrepreneurship and are forced to resort to wage employment.

Case 2: Constrained households

$$w \leq \theta \hat{k}^\alpha - \hat{k}, \underline{k} \leq M < k^* \tag{5}$$

These households have enough wealth to overcome the minimum capital requirement, but are constrained from acquiring the optimal level of capital $k^*$. In employing the sub-optimal level of capital $\hat{k}$, constrained households are unable to achieve the profit level possible if they were unconstrained.

Case 3: Unconstrained households

$$w \leq \theta k^{*}\alpha - k^{*}, k^{*} \geq M \tag{6}$$

Households facing Condition 6 are unconstrained: they are better off by engaging in entrepreneurship and they have enough wealth to acquire the optimal level of capital for their venture.

### 3.3 Implications of the Workfare Program

After implementation of a workfare program, households have an additional source of income. They can earn a maximum of $N > w$ through the program, and may invest $\lambda N$ in the entrepreneurial activity, where $\lambda \in [0, 1]$. This yields three predictions of the model.

Prediction 1: Wage work and unconstrained entrepreneurship unaffected

Wage households (Case 0) and unconstrained entrepreneurial households (Case 3) will not be affected by the implementation of a workfare program. For the former group, the program only serves to increase the value of the outside option compared to entrepreneurship, thus reaffirming the decision of a low ability household to choose wage employment. Households in the latter group were not bound by the liquidity constraint prior to program implementation, so the option to earn additional income does not affect their investment and occupational decisions.

Prediction 2: Extensive entrepreneurship increases

Households previously excluded from entrepreneurship (Case 1) now, under the workfare program, may face:

$$(1 - \lambda)N \leq \theta \underline{k}^{\alpha} - \underline{k}, \underline{k} \leq M + \lambda N \tag{7}$$

Excluded households are now able to purchase the minimum level of capital and engage in entrepreneurship. Thus the program will increase rates of entrepreneurship, or the extensive margin.

Prediction 3: Intensive entrepreneurship increases

Households constrained from purchasing $k^*$ (Case 2) will now be able to invest at the optimal level, subject to total wealth (the initial endowment coupled with the workfare income) being greater than the cost of $k^*$.

$$(1-\lambda)N \leq \theta\hat{k}^{\alpha} - \hat{k}, \hat{k} \in [\underline{k}, \min\{k^*, M+\lambda N\}] \tag{8}$$

Thus entrepreneurial households previously employing sub-optimal levels of capital will now be able to more actively engage in their venture; the workfare program extends the intensive margin.

### 3.4 Transformative versus Subsistence Entrepreneurship

The discussion so far assumes that entrepreneurship, for able households, is always a better option than wage work, implicitly focusing on the impact of a workfare program on transformational entrepreneurs, who drive economic growth (Schoar, 2009). In this sector, households can use the workfare program to overcome liquidity constraints. However, the program effect may differ for households forced in to subsistence self-employment because permanent employment opportunities are not available. Households in this sector may substitute away from self-employment upon onset of the workfare program due to improved job prospects.

Thus I hypothesize that the workfare program will promote transformational entrepreneurship and reduce and will reduce subsistence entrepreneurship. This translates directly into my econometric specification where I estimate the impact of NREGS separately for the rural non-farm and rural farm households. The former represent transformational entrepreneurship: the rural non-farm sector in India has been expanding steadily over the last 5-10 years, with a range of high-paying jobs. Rural farm households represent the subsistence sector: proper, salaried employment in agriculture has been shrinking. Further, casual and self-employment opportunities in agriculture are not well-paying nor are they consistently available (Lanjouw and Murgai, 2008).

## 4 Empirical Strategy

I take advantage of the timing of the NREGS phases and survey administration, employing differences-in-differences (DID) to estimate the impact of NREGS on rates of entrepreneurship.

Treated districts include NREGS Phases 1 and 2, where NREGS was in operation by 2007. NREGS Phase 3 districts comprise the untreated group, which received the program only after administration of the national survey was completed in 2008.

The baseline specification is given in Equation 9:

$$ln(y_{qdt}) = \beta treatment_d * post_t + \delta post_t + \gamma X_{dt} + \theta Z_d * post_t + \eta_d + \tau_q + \epsilon_{qdt} \tag{9}$$

Outcome $y_{qdt}$ is a measure of entrepreneurship in quarter $q$, district $d$, period $t$ (households are randomly assigned to each quarter, by the survey). I first consider extensive entrepreneurship, which is defined as rates of total rural entrepreneurship, rural non-farm entrepreneurship, and rural farm entrepreneurship. The rates are calculated, using the survey weights, as the share of households engaged in (a particular category of) entrepreneurship, relative to the total labor force. I then consider two measures of intensive entrepreneurship. Measure 1 is the average share of working age (15-60) household members participating in the entrepreneurial activity, in each category (total, non-farm, and farm). Measure 2 is similarly defined, but considers the share of time working age household members spend in the main entrepreneurial activity. I analyze non-farm and farm entrepreneurs separately since, as noted above, they may engage in entrepreneurship for different reasons, and so are likely to differ in their response to the program.

The $treatment_d$ indicator is equal to 1 if district $d$ is in the treatment group (early districts, NREGS Phases 1 or 2); the $post_t$ indicator is equal to 1 in the post-2007 period. The coefficient of interest is $\beta$, which measures the intent-to-treat effect of NREGS on entrepreneurship.

With the inclusion of district fixed effects $\eta_d$, Equation 9 compares rates of entrepreneurship within districts before and after program implementation to eliminate any time invariant characteristics that might also affect the outcome. This is crucial since NREGS was not randomly implemented; the district fixed effects address selection issues in treated versus untreated districts. The specification also controls for any national level trends in a particular year (with the $post_t$ indicator) that could drive entrepreneurship by comparing rates of entrepreneurship across districts within each year. To account for seasonality, I include quarter fixed effects ($\tau_q$). I account for the correlation of $\epsilon_{qdt}$ over time within each district by clustering the standard errors at the district level.

One criticism of this approach is that differential pre-program trends across NREGS and non-NREGS districts might drive the results. Treated areas were less economically developed than untreated areas. In Table 2, I compare the pre-program trends across the treated and untreated groups for various demographic characteristics; there are significant differences in these trends. If these trends are also related to entrepreneurship, then my estimates of the treatment effect will be biased. I address this issue in multiple ways. First, I include time-varying district characteristics, given by $X_{dt}$ in Equation 9. Second, I interact a vector of pre-program district level characteristics ($Z_d$ in Equation 9) with the post period indicator. In addition, one might be also concerned with regression to the mean, implying that any specifications detecting a relationship between program status and entrepreneurship are actually spurious. I address this by including the pre-program trend in entrepreneurship in my regression specification as a robustness test. Finally, I run two falsification tests to ensure the results are not driven by factors unrelated to the implementation of NREGS.

## 5 Data

I incorporate data from a variety of sources. The outcome data, measures of entrepreneurship, come from the Employment and Unemployment Schedule of the National Sample Survey (NSS), which is administered by the National Sample Survey Office (NSSO) of India. The survey is administered throughout India and is nationally representative. It is conducted annually, with a different topic every year. The quinquennial surveys (conducted every five years) are the largest, collecting data on employment/unemployment, consumer expenditure, and demographics from approximately 125,000 households and 600,000 individuals (usually split 65% rural, 35% urban). However, in 2007-2008 the NSSO conducted a larger Employment and Migration survey ($64^{th}$ round), which was similar in size and scope to the quinquennial surveys.

I exploit the timing of the NSS survey rounds and the phased rollout of NREGS to analyze the impact of the program on entrepreneurship. Round 61 was administered in 2004-2005, prior to the start of NREGS, and so serves as the pre-program baseline period. Phases 1 and 2 of NREGS were rolled out in February 2006 and January 2007, respectively. NSS Round 64 was administered in 2007, prior to NREGS Phase 3, which began in April 2008. Thus Round 64 demarcates the post-program period. Figure 2 depicts the timing of the NSS rounds, relative to the phases of NREGS.

I combine NSS Rounds 61 and 64, aggregating household data to create a district-level panel. I matched districts using borders from the 2001 Indian Census borders, to track consistent geographic units over time. The panel was then matched with the NREGS rollout data, to identify districts that received the program by 2007, and those that did not. Phase 1 and 2 (early) districts combined represent the treated group; Phase 3 (late) districts comprise the untreated group. Details of the data management and variable construction are provided in the Appendix.

District and state covariates come from the 2001 Census, India's statistical agency (Ministry of Statistics and Programme Implementation, MOSPI), and the Indian Planning Commission. The Census provides demographic data, poverty data, labor force characteristics, and rural development measures. From MOSPI, I gather state-level GDP data. In 2003, the Planning Commission developed a Development Index, identifying the 200 least developed districts in India according to various criteria. The Ministry of Rural Development incorporated this ranking when deciding which districts would receive NREGS first. To control for the non-random implementation of NREGS, I include the components of the Development Index in my regressions.

# 6 Results & Discussion

## 6.1 Impact of NREGS on Extensive Entrepreneurship

The results of the baseline regression (Equation 9) estimating the impact of NREGS on extensive entrepreneurship are provided in Table 3, for rates of total rural (column 1), rural non-farm (column 2), and rural farm (column 3) entrepreneurship. This specification includes only time and district fixed effects. Tables 4, 5, and 6 extend the baseline regressions, presenting results for different specifications, in which groups of controls are additively included. In Tables 3-6, the outcome is defined as the share of households engaged in entrepreneurship relative to the total labor force, in a given quarter, district, and period. In all specifications, the standard errors are clustered at the district level.

The baseline results in Table 3 provide evidence consistent with the hypotheses presented in Section 3: NREGS increases rates of extensive entrepreneurship (prediction 2), but in a specific sector (likely promoting transformational ventures over subsistence activities). To see this, first consider column 1, which provides results for the impact of NREGS on rates of total entrepreneurship. The difference-in-difference estimate (-0.02) is small in magnitude and statistically insignificant. However, columns 2 and 3 show the importance of disaggregating into the non-farm and farm sectors. Farm entrepreneurs outnumber non-farm entrepreneurs nearly 3 to 1, and the results in column 1 (total entrepreneurship) reflect this composition. However, it is clear that the impact of the program on non-farm entrepreneurship differs greatly from the impact on farm entrepreneurship. The coefficient for the impact on non-farm entrepreneurs (column 2) is 0.13; NREGS is promoting ventures (likely transformational) in this sector. Further the results indicate that NREGS may pull farm entrepreneurs away from their subsistence activities into the workfare program, as the coefficient in column 3 is negative (-0.02), though it is not statistically significant.

The results in Tables 4 (total entrepreneurship), 5 (non-farm entrepreneurship), and 6 (farm entrepreneurship) are consistent with the baseline results of Table 3. There is no impact of NREGS on total entrepreneurship, consistent with the results in Table 3. Column 1 of Table 5 shows that for the rural non-farm entrepreneurial population, the baseline treatment effect (0.13) is likely an underestimate, when compared to the results in columns 2-5 of Table 5, which control for a range of time-varying and time-invariant district characteristics. The DID estimates increase to 16-18% in columns 2-5, consistent across these specifications. Prior to the program implementation (in 2004), baseline levels of non-farm entrepreneurship were 15%; the program has a net effect of 3 percentage points, increasing the rate to 18%. For farm entrepreneurs (Table 6), the estimated effect is also consistent across the various specifications, at around -3 to -3.5%, but is not significant.

Thus there is clear evidence that NREGS increases rates of rural non-farm entrepreneurship by approximately 3 percentage points. The estimated increase in extensive rural non-farm entrepreneurship is consistent with the prediction that NREGS may relax credit constraints for excluded households, allowing them to invest income earned from the program into their venture. Further, this result supports the hypothesis that NREGS promotes transformational entrepreneurship, the more productive, growing sector. Correspondingly, we would expect to see a negative impact of NREGS on rural farm self-employment, as these households substitute away from subsistence work in favor of higher paying employment opportunities offered by the program. While the estimated impact of NREGS on farm entrepreneurship is negative, it is not statistically significant.

### 6.2 Impact of NREGS on Intensive Entrepreneurship

While there is evidence that NREGS increases the extensive margin of entrepreneurship, the results in Tables 7 and 8 indicate that there is little impact on the intensive margin of entrepreneurship. I first consider if the program affects the share of working age household members engaged in entrepreneurship (Table 7). I also estimate the impact on the share of time working age household members spend on the main entrepreneurial activity (Table 8). In considering measures of intensive entrepreneurship, I analyze whether NREGS can also promote entrepreneurship for households engaged in a venture, but not at an optimal level, prior to program implementation (prediction 3).

From the results in Tables 7 and 8, we see that NREGS does not increase the intensity of the entrepreneurial activity within the household when considering the total entrepreneurial population (Panel A). The coefficient is small and not statistically significant. This result holds under disaggregation into the non-farm and farm sectors (Panels B and C). Although these results differ from the model prediction, they do provide useful insight into the behavior of entrepreneurial households and their response to a workfare program. In particular, the empirical evidence points to the possibility that households involved in transformational entrepreneurship in the non-farm sector prior to program implementation may already be operating at or near optimal levels of capital. As a result, these households do not increase the intensity of entrepreneurship post-NREGS.

It is somewhat surprising that farm entrepreneurs do not exhibit a decrease in the intensity of the entrepreneurial activity, as evidenced by Panel C of Tables 7 and 8. The results from Table 6 indicate that households in farm entrepreneurship do not use NREGS as a complete substitute for their subsistence activities. However, it seems likely that these households could use NREGS as a partial substitute, which would affect the intra-household allocation towards entrepreneurship. This lack of evidence on partial substitution may indicate that even with access to NREGS, households in this sector must still engage in a subsistence activity to achieve a minimum level of income.

Combined, the results in Tables 4-8 provide evidence that NREGS may relax credit constraints, allowing previously excluded households to now afford the minimum capital requirement, thus facilitating a discrete jump into entrepreneurship in the non-farm sector. There is little evidence that the program supports existing entrepreneurial households in acquiring more capital, as there is no increase in the intensity with which households engage in entrepreneurship. This may reflect the fact that the minimum level of capital is relatively high compared to average wealth in this sector and that, in many ventures, this minimum level is very close to the optimal level. Thus we see increases along the extensive margin and not the intensive margin.

### 6.3 Robustness Tests

Given that the evidence is strongest in Tables 4-6, which reflect the impact of NREGS on extensive entrepreneurship, I subject these results to a variety of robustness tests. To determine if the effect estimated in Equation 9 is truly due to NREGS and does not result from differing trends in entrepreneurship across treatment and control groups prior to the program, I conduct a falsification test. The main assumption of the differences-in-differences approach is that, in the absence of treatment, rates of entrepreneurship in the treatment and control groups would have continued along their same relative trajectories. While I cannot observe this counterfactual, I can test for a program effect during years in which NREGS was not implemented. Thus I estimate Equation 9 using data from the two NSS rounds prior to NREGS rollout: Round 55 in 1999-2000 and Round 61 in 2004-2005. I compare the same treatment and control groups as in Equation 9, but there is no program in effect. If the false treatment has no effect on the outcomes of interest, this provides additional evidence that the original estimates are indeed causal effects of NREGS.

The nature of NREGS itself permits another falsification test. Since the program was offered only in rural areas, there should not be any treatment effect in urban parts of the country unless a national secular trend was driving rates of entrepreneurship at the same time NREGS was implemented. I re-estimate Equation 9, comparing the original treated and untreated districts, but restricting the sample to urban areas. Once again, if the false treatment has no effect, this is further proof that NREGS is responsible for the increased rates of entrepreneurship.[8]

The results from the falsification tests (Tables 9 and 10) indicate that the effect of NREGS on rural non-farm rates of entrepreneurship can be attributed to the program and not to other trends. In Table 9, the difference-in-difference coefficient ($\beta$) is not significant, indicating that the results

[8]In this falsification test, I consider only the non-farm entrepreneurial sector, as there is no farm sector in urban areas.

in Table 5 are due to the implementation of NREGS, and not differential trends in the treated and untreated districts.[9] The urban falsification test in Table 10 corroborates this result: there is no evidence of an impact of NREGS in urban areas, where the program was not implemented.

The magnitude of the DID estimates (5-8%) in Table 9 indicates there may be upward bias in my estimate of the effect of NREGS on rural non-farm entrepreneurship. I include the pre-program trend in entrepreneurship (interacted with the post-period indicator) in my specification as a robustness check; the results are presented in Table 11. Qualitatively, the results remain the same as in Tables 4, 5, and 6; the effect of NREGS on rural non-farm entrepreneurship now increases to more than 18%.

It is worth noting that the coefficient on the entrepreneurship trend in Table 11 is negative and significant. This indicates possible reversion to the mean: districts that had particularly high positive trends in entrepreneurship are now falling back to the national average. Historically, however, rates of rural self-employment in India have been increasing, especially since the mid-1990s (Lanjouw and Murgai, 2008; NSSO, 2010). In 2004, the rate was quite high, above the value the trend from earlier years would have predicted. It is likely that the abnormally high rate of entrepreneurship in the pre-period (2004) may be manifesting as mean reversion within the current specification. To investigate this possibility, I estimate Equation 9, on extensive entrepreneurship, but now using 1999 as the pre-program period (instead of 2004). The results in Table 12 are consistent with the results in Tables 4-6. In fact, the estimated impact on rates of rural non-farm entrepreneurship (column 2) more than doubles (to 0.36), although the coefficient on the post-period indicator is negative and significant.

### 6.4 Extension: Dosage Response

The original specification given in Equation 9 groups NREGS Phase 1 and 2 districts together, comparing them to Phase 3 districts, which had not received the program in 2007, when NSS Round 64 was administered. However, by 2007, Phase 1 districts had been exposed to NREGS for more than a year longer than Phase 2 districts. To allow for a dosage response, that the impact of

[9] In the regression presented in Table 9, I exclude measures drawn from the 2001 Indian Census, as I consider the period 1999-2004. However, I also conduct the falsification test with the entire set of controls (including those from 2001) and find that the results are the same (available upon request): there is no effect of NREGS on entrepreneurship outside of the program era.

NREGS on entrepreneurship may be related to the length of time the program had been in effect, I modify Equation 9 as follows:

$$ln(y_{qdt}) = \beta_1 Phase_1 * post_t + \beta_2 Phase_2 * post_t + \delta post_t + \gamma X_{dt} + \theta Z_d * post_t + \eta_d + \tau_q + \epsilon_{qdt} \quad (10)$$

Differences in $\beta_1$ and $\beta_2$ will indicate the extent to which the impact of NREGS on rates of entrepreneurship is related to the length of exposure to the program.

Table 13 presents the results for the differential effect, or dosage response, of NREGS on Phase 1 and Phase 2 districts. The results are similar to the cumulative effect of NREGS presented in Tables 4, 5, and 6, but now the contribution of each phase to increases in entrepreneurship is apparent. For the rural non-farm population (column 2), Phase 1 increases the rate of extensive entrepreneurship by 10% (1.5 percentage points) and Phase 2 results in about a 20% increase in the rate of entrepreneurship, although insufficient power of the regression prevents detecting statistical significance for the first phase. The lack of power is evident in the F-test, which indicates that the Phase 1 and Phase 2 coefficients are not statistically different. Similar to the baseline results, there is no effect of the program on the total rate of entrepreneurship or rates of rural farm entrepreneurship (columns 1 and 3).

## 7 Conclusion

I utilize a nationally representative dataset to analyze the impact of India's rural public works policy, NREGS, on entrepreneurship. I find that the program differentially affects rural non-farm and rural farm entrepreneurs. NREGS positively impacts rates of extensive rural non-farm entrepreneurship, resulting in a three percentage point increase. The rural non-farm sector, which includes transformational entrepreneurial activities, has been expanding rapidly in India. The results presented provide some evidence that liquidity constrained households that had been excluded from entering rural non-farm entrepreneurship are able to use NREGS as a source of credit to acquire the necessary capital for their entrepreneurial venture. In contrast, NREGS has little or no impact on rural farm entrepreneurship. Workers in this sector are often subsistence entrepreneurs who may be using NREGS to leave self-employment, substituting towards the better employment opportunities offered by the program. In ongoing work, I am researching sector-specific responses to NREGS, to clarify if the program is truly promoting transformational entrepreneurship over subsistence activities, and other (non-parametric) estimation methods.

Over the period 1983-2004, prior to NREGS, the annualized growth rate of employment in the rural non-farm sector was 4.0% each year (Himanshu et al., 2011). From my results, I estimate that in NREGS districts, the size of the rural non-farm self-employed workforce grew by 6.2% every year, more than a fifty percent increase over the previous two decades. Effects of this magnitude could translate into considerable wealth gains, pulling individuals out of poverty, even if the baseline share of rural non-farm self-employment in total GDP is low. Further, there may be intergenerational effects: children in households that have graduated from poverty are less likely to be poor themselves. In addition, the implications for the sustainability of NREGS are significant. The estimates indicate that the net number of rural non-farm entrepreneurial households has increased. To the extent that this population reduces its dependency on NREGS and contributes to the Indian economy, the per capita cost of NREGS will decrease, making the program more economically viable. Thus analyzing this issue has important consequences for the program itself and economic development in India.

## References


Abraham, Vinoj (2011). Agrarian distress and rural non-farm sector employment in India. Paper No. 35275, Munich Personal RePEc Archive.

Afridi, Farzana, Mukhopadhyay, Abhiroop, Shaoo, Soham (2012). Female Labour Force Participation and Child Education in India: The Effect of the National Rural Employment Guarantee Scheme. Discussion Paper No. 6593, Institute for the Study of Labor (IZA).

Alatas, Vivi, Cameron, Lisa A. (2008). The Impact of Minimum Wages on Employment in a Low-Income Country: A Quasi-Natural Experiment in Indonesia. *Industrial & Labor Relations Review*, 61(2): 201-223.

Athey, Susan, Imbens, Guido W. (2006). Identification and Inference in Nonlinear Difference-in-Differences Models. *Econometrica*, 74(2): 431-497.

Azam, Mehtabul (2012). The Impact of Indian Job Guarantee Scheme on Labor Market Outcomes: Evidence from a Natural Experiment. Discussion Paper No. 6548, Institute for the Study of Labor (IZA).

Banerjee, Abhijit, Duflo, Esther (2007). The Economic Lives of the Poor. *Journal of Economic Perspectives*, 21(1): 141-168.

Banerjee, Abhijit, Newman, Andrew F. (1993). Occupational Choice and the Process of Development. *Journal of Political Economy*, 101(2): 274-298.

Barnett, Adrian G., van der Pols, Jolieke C., Dobson, Annette J. (2005). Regression to the mean: what it is and how to deal with it. *International Journal of Epidemiology*, 34(1): 215-220.

Basu, Arnab, Chau, Nancy, Kanbur, Ravi (2007). The National Rural Employment Guarantee Act of India, 2005. In K. Basu (Ed.), *The Oxford Companion to Economics in India*, 365-367. USA: Oxford University Press.

Berg, Erland, Bhattacharyya Sambit, Durgam, Rajasekhar, Ramachandra, Manjula (2012). Can Rural Public Works Affect Agricultural Wages? Evidence from India. CSAE Working Paper WPS/2012-05, Centre for the Study of African Economies.

Bhalla, Surjit S. (2011). Inclusion and Growth in India: Some facts, some conclusions. Working Paper No. 39, Asia Research Centre, London School of Economics and Political Science.

Bharadwaj, Prashant (2010). Quasi experimental methods: Difference in differences. Presentation at the University of California, San Diego.

Bhattacharjee, Manojit, Rajeev, Meenakshi (2010). Interest rate formation in informal credit markets in India: does level of development matter? Working Paper 126, Brooks World Poverty Institute, University of Manchester.

Binswanger-Mikhize, Hans P. (2012). India 1960-2010: Structural Change, the Rural Nonfarm Sector, and the Prospects for Agriculture. Department of Agricultural and Resource Economics, University of California, Berkeley.

Cabral, Luís M.B., Mata, José (2003). On the Evolution of the Firm Size Distribution: Facts and Theory. *American Economic Review*, 93(4): 1075-1090.

Card, David, Kluve, Jochen, Weber, Andrea (2010). Active Labor Market Policy Evaluations: A Meta-Analysis. Working Paper 16173, National Bureau of Economic Research.

Carling, Kenneth, Gustafson, Lena (1999). Self-employment grants vs. subsidized employment: Is there a difference in the re-unemployment risk? IFAU Working Paper 1999:6, Institute for Labour Market Policy Evaluation.

Census of India (2001). Table S00-12: Distribution of households by size and number of dwelling rooms. Ministry of Home Affairs, Government of India. `http://censusindia.gov.in/Tables\_Published/H-Series/H-Series\_link/S00-012.pdf`. Accessed August 27, 2013.

Chaudhuri, Sarbajit (2004). International Migration of Skilled and Unskilled Labour, Welfare and Skilled-Unskilled Wage Inequality: A Simple Model. *Journal of Economic Integration*, 19: 726-741.

Chowdhury, Subhanil (2011). Employment in India: What Does the Latest Data Show? *Economic and Political Weekly*, 46(32): 23-26.

Coppard, Daniel (2001). The Rural Non-farm Economy in India: A Review of the Literature. Natural Resources Institute, NRI Report No. 2662.

Dasgupta, Sukti, Sudarshan, Ratna M. (2011). Issues in labour market inequality and women's participation in India's National Rural Employment Guarantee Programme. Working Paper No. 98, Policy Integration Department, International Labor Office, Geneva.

Deaton, Angus (1991). Saving and Liquidity Constraints. *Econometrica*, 59(5): 1221-1248.

Dev, S. Mahendra, Ravi, C. (2007). Poverty and Inequality: All-India and States, 1983-2005. *Economic and Political Weekly*, 42(6): 509-521.

Devereux, Stephen, Solomon, Colette (2006). Employment creation programmes: The international experience. Issues in Employment and Poverty Discussion Paper No. 24, International Labour Office, Geneva.

Dutta, Puja, Murgai, Rinku, Ravallion, Martin, van de Walle, Dominique (2012). Does India's Employment Guarantee Scheme Guarantee Employment? Policy Research Working Paper No. 6003, The World Bank.

Evans, David S., Jovanovic, Boyan (1989). An Estimated Model of Entrepreneurial Choice under Liquidity Constraints. *Journal of Political Economy*, 97(4): 808-826.

Fan, Shenggen, Hazell, Peter, Thorat, Sukhadeo (1999). Linkages between Government Spending, Growth, and Poverty in Rural India. Research Report 110, International Food Policy Research Institute.

Fisher, Thomas, Mahajan, Vijay, Singha, Ashok (1997). *The Forgotten Sector: Non-Farm Employment and Enterprises in India*. London: Intermediate Technology Publications.

Gilligan, Daniel O., Hoddinott, John, Taffesse, Alemayhe Seyoum (2008). The Impact of Ethiopia's Productive Safety Net Programme and its Linkages. Discussion Paper 00839, International Food Policy Research Institute.

Himanshu (2005). Wages in Rural India: Sources, Trends and Comparability. *Indian Journal of Labour Economics*, 48(2): 375-406.

Himanshu, Lanjouw, Peter, Mukhopadhyay, Abhiroop, Murgai, Rinku (2011). Non-Farm Diversification and Rural Poverty Decline: A Perspective from Indian Sample Survey and Village Study Data. Asia Research Centre, London School of Economics & Political Science, Working Paper 44.

Hoff, Karla, Stiglitz, Joseph E. (1990). Imperfect Information and Rural Credit Markets - Puzzles and Policy Perspectives. *World Bank Economic Review* 4(3), 235-250.

Holtz-Eakin, Douglas, Joulfaian, David, Rosen, Harvey S. (1994). Entrepreneurial decisions and liquidity constraints. *RAND Journal of Economics* 25(2), 334-347.

Imbert, Clément, Papp, John (2013). Labor Market Effects of Social Programs: Evidence from India's Employment Guarantee. CSAE Working Paper WPS/2013-03, Centre for the Study of African Economies.

Jovanovic, Boyan (1982). Selection and Evolution of Industry. *Econometrica* 50, 649-670.

Khera, Reetika, Nayak, Nandini (2011). Women workers and perceptions of the National Rural Employment Guarantee Act in India. In Khera, Reetika (ed.), The Battle for Employment Guarantee. New Delhi: Oxford University Press, 81-104.

Kluve, Jochen (2006). The Effectiveness of European Active Labor Market Policy. Discussion Paper No. 2018, Institute for the Study of Labor (IZA).

Kumar, Anjani, Kumar, Sant, Singh, Dhiraj K., Shivjee (2012). Rural Employment Diversification in India: Trends, Determinants, and Implications on Poverty.

Lanjouw, Peter, Murgai, Rinku (2008). Poverty Decline, Agricultural Wages, and Non-Farm Employment in Rural India: 1983-2004. World Bank, Policy Research Working Paper 4858.

Liu, Yanyan, Deininger, Klaus W. (2010). Poverty Impacts of India's National Rural Employment Guarantee Scheme: Evidence from Andhra Pradesh. Unpublished manuscript, Presented at the Agricultural and Applied Economics Association, Annual Meeting.

Lucas, Robert E., Jr. (1978). On the Size Distribution of Business Firms. *Bell Journal of Economics* 9(2), 508-523.

Mahajan, Vijay, Ramola, Bharti Gupta (1996). Financial services for the rural poor and women in India: Access and sustainability. *Journal of International Development* 8(2), 211-224.

Ministry of Law and Justice (2005). National Rural Employment Guarantee Act, 2005. *The Gazette of India*.

National Sample Survey Office (2001). Concepts and Definitions Used in NSS. Ministry of Statistics & Programme Implementation, Government of India.

National Sample Survey Office (2006). Employment and Unemployment Situation in India, 2004-05. Report No. 515 (61/10/1), Ministry of Statistics & Programme Implementation, Government of India.

National Sample Survey Office (2010). Employment and Unemployment Situation in India, 2007-2008. Report No. 531 (64/10.2/1), Ministry of Statistics & Programme Implementation, Government of India.

National Sample Survey Office (2010). Migration in India, 2007-2008. Report No. 533 (64/10.2/2), Ministry of Statistics & Programme Implementation, Government of India.

Ministry of Rural Development (2005). The Mahatma Gandhi National Rural Employment Guarantee Act 2005: `nrega.nic.in`. Accessed August 30, 2013.

Pankaj, Ashok, Tankha, Rukmini (2010). Empowerment Effects of the NREGS on Women Workers: A Study in Four States. *Economic and Political Weekly* 45(30), 45-55.

Ravi, Shamika, Engler, Monika (2012). Workfare in Low Income Countries: An Effective Way to Fight Poverty? The Case of NREGS in India. Under review, Social Science Research Network.

Ravi, Shamika, Kapoor, Mudit, Ahluwalia, Rahul (2012). The Impact of NREGS on Urbanization in India. Presented at the Northeast Universities Development Consortium Conference, 2012.

Schoar, Antoinette (2009). The Divide between Subsistence and Transformational Entrepreneurship. In In Lerner, Josh and Stern, Scott (Eds.), *Innovation Policy and the Economy* Volume 10, 57-81. University of Chicago Press, 2010.

StataCorp. 2009. *Stata Statistical Software: Release 11*. College Station, TX: StataCorp LP.

Vaidyanathan, A. (1986). Labour Use in Rural India: A Study of Spatial and Temporal Variations. *Economic and Political Weekly*, 21(52): A130-A146.

Visaria, Pravin M., Basant, Rakesh (1994). *Non-agricultural employment in India: Trends and prospects*. New Delhi: Sage.

Walker, Thomas S., Ryan, James G. (1990). *Village and Household Economies in India's Semi-arid Tropics*. Baltimore: Johns Hopkins University Press.

World Development Indicators. World Bank: `data.worldbank.org/country/india`. Accessed August 27, 2013.

Zeldes, Stephen P. (1989). Consumption and Liquidity Constraints: An Empirical Investigation. *Journal of Political Economy* 97(2), 305-346.

## Data Appendix

To create a district-level panel using NSS Rounds 61 and 64 that tracked geographically consistent units over time and also contained NREGS phase data, I converted all district borders to match that of the 2001 Census. In 2001, India had 593 districts. NSS Rounds 61 and 64 used 2001 Census district borders, sampling from 585 and 588 districts, respectively (inaccessible areas in Jammu and Kashmir, Nagaland, and Andaman and Nicobar Islands were not surveyed). I started with the 585 districts common to NSS Rounds 61 and 64. When rolling out the program, NREGS used current district borders, which differed from that of the 2001 Census because of newly created districts. Between the 2001 and 2011 Censuses, 47 districts were created, by splitting existing districts or combining areas across districts. I converted the NREGS districts to 2001 Census borders to match the NSS districts. I dropped 31 districts where the new district and the original district received NREGS in different phases. A further 11 districts where NREGS was not administered were dropped from the sample, leaving a core sample of 543 districts with consistent borders from 2004 to 2007.

I employed the same approach when preparing the data for the first falsification test, comparing NREGS areas and non-NREGS areas from 1999 (NSS Round 55) to 2004 (NSS Round 61), when the program was not implemented. In order to track consistent units, I converted Rounds 55 and 61 to 1991 Census borders. In 1991, India had 466 districts. Again dropping cases where a parent and child district received NREGS in different phases and dropping urban districts, I was left with a sample of 395 districts from 1999 to 2004.

The controls used in the regressions came from three sources: the 2001 Census, India's statistical agency (Ministry of Statistics and Programme Implementation, MOSPI), and the Indian Planning Commission. The District Profiles of the Census provided demographic data (average household size, proportion Scheduled Caste and Scheduled Tribe, literacy rate), poverty data (average household monthly consumption and expenditure), labor force characteristics (share in main employment, share in marginal employment, workforce participation rate), and rural development measures (share of households occupying permanent structures, share of villages with safe drinking water, share of villages with a primary school, share of villages with a medical facility) at the district level. I obtain state GDP and NDP data for 2000 from MOSPI. I also control for the development ranking used by the Ministry of Development to decide which districts would receive NREGS first. This ranking came from an Indian Planning Commission study conducted in 2003, to identify the 200 least developed districts in India. The commission used three criteria in its ranking: share of the population that was Scheduled Caste/Tribe (from the 1991 Census), daily agricultural wages (from 1996-1997), and average agricultural productivity per worker (over 1990-1993).

**Figure 1: Map of NREGS Phased rollout**

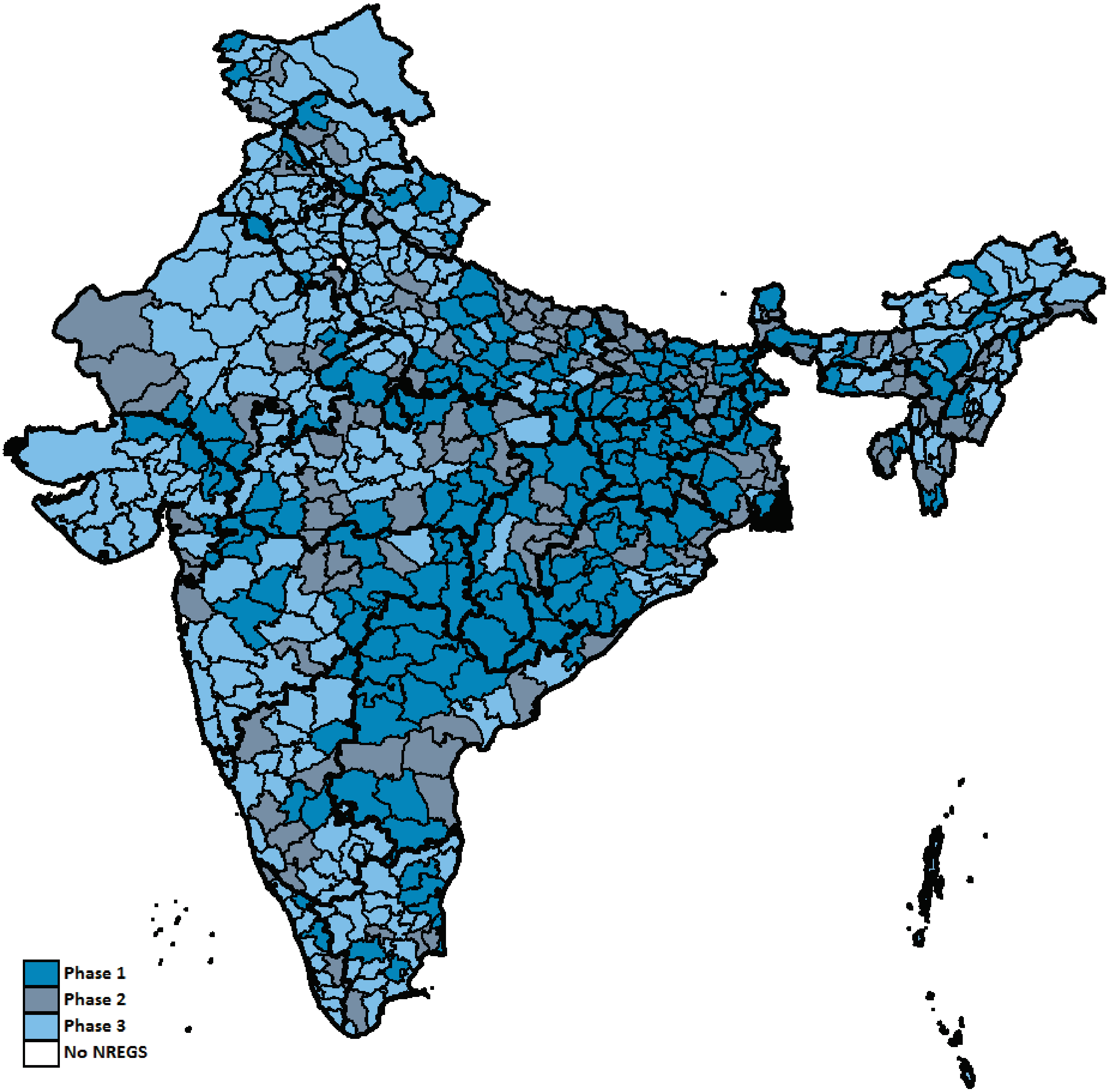



Figure 1 presents a map of the phased implementation of the National Rural Employment Guarantee Scheme (NREGS) in India. In 2006 (Phase 1, dark blue), NREGS was implemented in the 200 least developed districts. In 2007 (Phase 2, gray), the program was implemented in another 130 districts. The remaining 285 districts received the program in 2008 (Phase 3, light blue). Urban areas (white) did not receive the program.

**Figure 2: Timeline of NSS Rounds and NREGS Rollout**

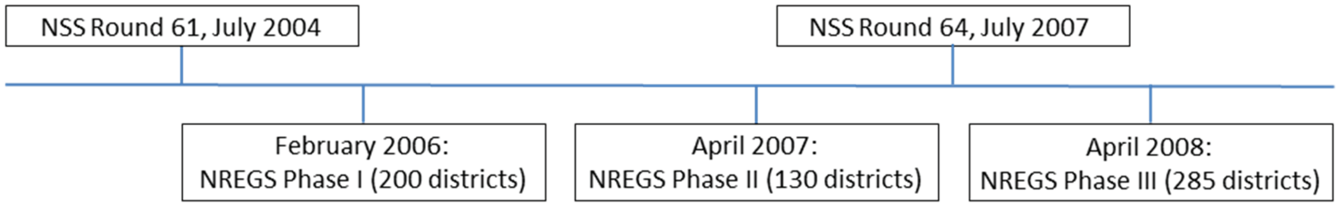



Figure 2 depicts the timing of the National Sample Survey (NSS) rounds, relative to the phases of NREGS. I exploit the timing of the NSS survey rounds and the phased rollout of NREGS to analyze the impact of the program on entrepreneurship.

**Table 1: Treated vs. Untreated Districts in 2004, Pre-NREGS**

| | Treated | Untreated | p-value |
|---|---|---|---|
| **Poverty and Demographics** | | | |
| Avg HH size | 5.43 | 5.42 | 0.937 |
| Scheduled caste share | 15.92 | 14.77 | 0.119 |
| Scheduled tribe share | 19.72 | 12.77 | 0.003 |
| Literacy rate | 59.43 | 69.10 | 0.000 |
| Avg monthly household consumer exp | 2,884 | 3,600 | 0.000 |
| Land owned (hectares) | 1.08 | 1.1 | 0.831 |
| **Labor Force Characteristics** | | | |
| Main worker share | 30.48 | 31.84 | 0.014 |
| Marginal worker share | 10.97 | 8.37 | 0.000 |
| Workforce participation rate | 41.39 | 40.22 | 0.053 |
| **Rural Development Measures** | | | |
| Share of HH in permanent housing | 38.79 | 57.55 | 0.000 |
| Share of villages with safe water | 96.57 | 96.33 | 0.794 |
| Share of villages with primary school | 80.97 | 85.42 | 0.001 |
| Share of villages with medical facility | 33.94 | 46.51 | 0.000 |
| Share of villages with communication facility | 41.29 | 60.71 | 0.000 |
| Share of villages with bus services | 35.70 | 57.24 | 0.000 |
| **State Expenditures** | | | |
| Real GDP trend, 2000-2005 | 4.23 | 4.76 | 0.014 |
| Real per capita NDP trend, 2000-2005 | 3.14 | 3.03 | 0.567 |
| **Development Index** | | | |
| Agricultural wages, 1996-1997 | 33.89 | 47.72 | 0.000 |
| Agricultural output per worker, 1990-1993 | 5,962 | 12,126 | 0.000 |
| N (districts) | 248 | 294 | |

Table 1 compares the treated and untreated districts in 2004, prior to NREGS implementation, for a variety of characteristics. There are pre-program differences, as Phase 1 and 2 districts were selected due to their poorer economic outcomes in 2004, e.g. treated districts were less developed and faced lower rates of workforce participation in 2004 (compared to untreated districts).

Notes: 1 hectare = 10,000 square meters.

*Source*: Author calculations using data from the National Sample Survey, 2001 Indian Census, Ministry of Statistics and Programme Implementation, and the Indian Planning Commission.

**Table 2: Pre-NREGS Trends in Treated and Untreated Districts, 1999-2004**

| **1999-2004 Change** | **Treatment** | **Untreated** | **p-value** |
|---|---|---|---|
| Average HH size | -0.21 | -0.22 | 0.799 |
| Scheduled caste share | -.01 | 0.02 | 0.024 |
| Scheduled tribe share | -0.01 | 0.01 | 0.062 |
| Literacy rate | 0.07 | 0.04 | 0.000 |
| Average monthly HH expenditure | -73.84 | -146.76 | 0.295 |
| Average land holdings (hectares) | 0.06 | -0.08 | 0.182 |
| N (districts) | 205 | 184 | |

Table 2 compares the trends in various demographic characteristics across the treated and untreated districts over the 1999 to 2004 period, prior to NREGS implementation. There are significant differences in these trends; to the extent they are related to entrepreneurship, my estimates of the program effect will be biased. I address this issue by including time-varying district characteristics and interacting a vector of pre-program district characteristics with a post period indicator, in addition to implementing robustness tests (see Section 6.3).
Notes: 1 hectare = 10,000 square meters. Household expenditure adjusted to 2004 rupees.
*Source*: Author calculations using data from the National Sample Survey, 2001 Indian Census, Ministry of Statistics and Programme Implementation, and the Indian Planning Commission.

**Table 3: Baseline: Effect of NREGS on Rates of Entrepreneurship**

| | **Total** (1) | **Non-farm** (2) | **Farm** (3) |
|---|---|---|---|
| NREGS ($\beta$) | -0.020 | **0.126**$^{*}$ | -0.024 |
| | (0.031) | (0.066) | (0.046) |
| Post indicator ($\delta$) | -0.050$^{**}$ | -0.240$^{***}$ | -0.051 |
| | (0.022) | (0.048) | (0.032) |
| Quarter fixed effects ($\tau_q$) | yes | yes | yes |
| District fixed effects ($\eta_d$) | yes | yes | yes |
| N | 4,332 | 4,010 | 4,301 |

Table 3 provides baseline estimates of the impact of NREGS on rates of total (column 1), non-farm (column 2), and farm entrepreneurship (column 3), using a differences-in-differences model, which controls for time and district fixed effects. There is no impact of NREGS on total or farm entrepreneurship, but the program results in a 0.13% increase in rural non-farm entrepreneurship. This is consistent with the idea that the program increases entrepreneurship on the extensive margin in the transformational non-farm sector over the subsistence farm sector (Prediction 2 of Section 3). Notes: Each observation corresponds to a quarter-district. The entrepreneurship outcome is defined at the household level. Household responses are aggregated to the quarter-district level in a given year, using the survey weights. Standard errors (in parentheses) are clustered at the district level. *** p<0.01, ** p<0.05, * p<0.1.

## Table 4: Effect of NREGS on Total Entrepreneurship

| | Outcome: Total entrepreneurship | | | | |
|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) |
| NREGS ($\beta$) | -0.020 | 0.007 | -0.005 | 0.003 | -0.004 |
| | (0.031) | (0.030) | (0.033) | (0.033) | (0.040) |
| Post indicator ($\delta$) | -0.050** | -0.142 | -0.275 | -0.222 | -0.579 |
| | (0.022) | (0.088) | (0.171) | (0.179) | (0.677) |
| Poverty and demographics | | yes | yes | yes | yes |
| Labor force characteristics | | yes | yes | yes | yes |
| Rural development measures | | | yes | yes | yes |
| Real per capita NDP | | | | yes | yes |
| Development index components | | | | | yes |
| Quarter fixed effects ($\tau_q$) | yes | yes | yes | yes | yes |
| District fixed effects ($\eta_d$) | yes | yes | yes | yes | yes |
| N | 4,332 | 4,332 | 4,284 | 4,263 | 3,381 |

Table 4 estimates the impact of NREGS on rates of total (rural) entrepreneurship, using a differences-in-differences model. Column 1 presents the baseline results; columns 2-5 extend the baseline regressions, controlling for a range of district characteristics. There is no impact of NREGS on total entrepreneurship, consistent with the results in Table 3. Notes: Each observation corresponds to a quarter-district. The entrepreneurship outcome is defined at the household level. Household responses are aggregated to the quarter-district level in a given year, using the survey weights. Standard errors (in parentheses) are clustered at the district level. *** p<0.01, ** p<0.05, * p<0.1.

## Table 5: Effect of NREGS on Non-farm Entrepreneurship

| | Outcome: Non-farm entrepreneurship | | | | |
|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) |
| **NREGS ($\beta$)** | **0.126**$^{*}$ | **0.176**$^{**}$ | **0.164**$^{**}$ | **0.161**$^{**}$ | **0.170**$^{*}$ |
| | (0.066) | (0.067) | (0.072) | (0.073) | (0.089) |
| Post indicator ($\delta$) | -0.240$^{***}$ | -0.324$^{*}$ | -0.216 | -0.216 | -2.467 |
| | (0.048) | (0.196) | (0.543) | (0.545) | (1.971) |
| Poverty and demographics | | yes | yes | yes | yes |
| Labor force characteristics | | yes | yes | yes | yes |
| Rural development measures | | | yes | yes | yes |
| Real per capita NDP | | | | yes | yes |
| Development index components | | | | | yes |
| Quarter fixed effects ($\tau_q$) | yes | yes | yes | yes | yes |
| District fixed effects ($\eta_d$) | yes | yes | yes | yes | yes |
| N | 4,010 | 4,010 | 3,970 | 3,949 | 3,155 |

Table 5 estimates the impact of NREGS on rates of (rural) non-farm entrepreneurship, using a differences-in-differences model. Column 1 presents the baseline results; columns 2-5 extend the baseline regressions, controlling for a range of district characteristics. In line with Prediction 2 (of Section 3), NREGS increases extensive non-farm entrepreneurship by 18%, equivalent to about 3 percentage points, consistent with the results in Table 3.
Notes: Each observation corresponds to a quarter-district. The entrepreneurship outcome is defined at the household level. Household responses are aggregated to the quarter-district level in a given year, using the survey weights. Standard errors (in parentheses) are clustered at the district level. *** $p<0.01$, ** $p<0.05$, * $p<0.1$.

**Table 6: Effect of NREGS on Farm Entrepreneurship**

| | Outcome: Farm entrepreneurship | | | | |
|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) |
| NREGS ($\beta$) | -0.024 | -0.010 | -0.036 | -0.028 | -0.036 |
| | (0.046) | (0.045) | (0.049) | (0.049) | (0.060) |
| Post indicator ($\delta$) | -0.051 | -0.138 | -0.427* | -0.348 | -0.565 |
| | (0.032) | (0.127) | (0.250) | (0.252) | (1.000) |
| Poverty and demographics | | yes | yes | yes | yes |
| Labor force characteristics | | yes | yes | yes | yes |
| Rural development measures | | | yes | yes | yes |
| Real per capita NDP | | | | yes | yes |
| Development index components | | | | | yes |
| Quarter fixed effects ($\tau_q$) | yes | yes | yes | yes | yes |
| District fixed effects ($\eta_d$) | yes | yes | yes | yes | yes |
| N | 4,301 | 4.301 | 4,253 | 4,237 | 3,364 |

Table 6 estimates the impact of NREGS on rates of (rural) farm entrepreneurship, using a differences-in-differences model. Column 1 presents the baseline results; columns 2-5 extend the baseline regressions, controlling for a range of district characteristics. In all specifications, the impact of NREGS is negative, consistent with the hypothesis that farm entrepreneurs would substitute away from subsistence self-employment towards stable employment in the program, but these estimates are not significant.

Notes: Each observation corresponds to a quarter-district. The entrepreneurship outcome is defined at the household level. Household responses are aggregated to the quarter-district level in a given year, using the survey weights. Standard errors (in parentheses) are clustered at the district level. *** $p<0.01$, ** $p<0.05$, * $p<0.1$.

## Table 7: Effect of NREGS on Share of Household Members in Entrepreneurship

| | **Panel A: Total** | | **Panel B: Non-farm** | | **Panel C: Farm** | |
|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) |
| NREGS ($\beta$) | -0.028 | 0.005 | -0.035 | -0.013 | -0.024 | 0.007 |
| | (0.024) | (0.029) | (0.034) | (0.043) | (0.025) | (0.032) |
| Post indicator ($\delta$) | -0.079*** | -0.439 | -0.091*** | -0.837 | -0.084*** | -0.195 |
| | (0.017) | (0.448) | (0.026) | (0.560) | (0.018) | (0.455) |
| Poverty and demographics | | yes | | yes | | yes |
| Labor force characteristics | | yes | | yes | | yes |
| Rural development measures | | yes | | yes | | yes |
| Real per capita NDP | | yes | | yes | | yes |
| Development index components | | yes | | yes | | yes |
| Quarter fixed effects ($\tau_q$) | yes | yes | yes | yes | yes | yes |
| District fixed effects ($\eta_d$) | yes | yes | yes | yes | yes | yes |
| N | 4,330 | 3,380 | 3,988 | 3,145 | 4,295 | 3,362 |

Table 7 estimates the impact of NREGS on the intensive margin of entrepreneurship, measured by the share of working age household members engaged in the entrepreneurial activity, using a differences-in-differences model. The results are not statistically significant for the combined sectors (Panel A), or when the sectors are disaggregated into non-farm (Panel B) and farm (Panel C) entrepreneurship.

Notes: Each observation corresponds to a quarter-district. The entrepreneurship outcome is defined at the household level. Household responses are aggregated to the quarter-district level in a given year, using the survey weights. Standard errors (in parentheses) are clustered at the district level. *** p<0.01, ** p<0.05, * p<0.1.

**Table 8: Effect of NREGS on Time Spent in Main Entrepreneurial Activity**

| | Panel A: Total | | Panel B: Non-farm | | Panel C: Farm | |
|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) |
| NREGS ($\beta$) | -0.017 | 0.000 | -0.034 | -0.030 | 0.001 | 0.008 |
| | (0.020) | (0.023) | 0.029 | (0.041) | (0.025) | (0.029) |
| Post indicator ($\delta$) | -0.041*** | 0.381* | -0.039* | 0.241 | -0.057*** | 0.509** |
| | (0.016) | (0.206) | (0.023) | (0.429) | (0.020) | (0.226) |
| Poverty and demographics | | yes | | yes | | yes |
| Labor force characteristics | | yes | | yes | | yes |
| Rural development measures | | yes | | yes | | yes |
| Real per capita NDP | | yes | | yes | | yes |
| Development index components | | yes | | yes | | yes |
| Quarter fixed effects ($\tau_q$) | yes | yes | yes | yes | yes | yes |
| District fixed effects ($\eta_d$) | yes | yes | yes | yes | yes | yes |
| N | 4,327 | 3,380 | 3,967 | 3,129 | 4,276 | 3,352 |

Table 8 estimates the impact of NREGS on another measure of intensive entrepreneurship, the share of time working age household members spend engaged in the main entrepreneurial activity, using a differences-in-differences model. The results are not statistically significant for the combined sectors (Panel A), or when the sectors are disaggregated into non-farm (Panel B) and farm (Panel C) entrepreneurship. Combined with Table 7 the results indicate that NREGS has little impact on intensive entrepreneurship, differing from Prediction 3 of Section 3.

Notes: Each observation corresponds to a quarter-district. The entrepreneurship outcome is defined at the household level. Household responses are aggregated to the quarter-district level in a given year, using the survey weights. Standard errors (in parentheses) are clustered at the district level. *** p<0.01, ** p<0.05, * p<0.1.

**Table 9: Pre-NREGS Falsification Test**

| | **Total** (1) | **Non-farm** (2) | **Farm** (3) |
|---|---|---|---|
| NREGS ($\beta$) | 0.052 | 0.070 | 0.048 |
| | (0.040) | (0.080) | (0.054) |
| Post indicator ($\delta$) | 0.102 | 0.046 | 0.049 |
| | (0.064) | (0.128) | (0.091) |
| Poverty and demographics | yes | yes | yes |
| Labor force characteristics | no | no | no |
| Rural development measures | no | no | no |
| Real per capita NDP | no | no | no |
| Development index components | yes | yes | yes |
| Quarter fixed effects ($\tau_q$) | yes | yes | yes |
| District fixed effects ($\eta_d$) | yes | yes | yes |
| N | 2,372 | 2,291 | 2,367 |

Table 9 presents the results from a falsification test: I estimate the impact of NREGS on rates of total, non-farm, and farm entrepreneurship in years when the program was not implemented, 1999-2004. There is no impact of NREGS on entrepreneurship outside of the program era. Note that some controls are excluded from the regression. These controls are from the 2001 Indian Census and were interacted with a post-NREGS indicator to capture district trends that might also be correlated with rates of entrepreneurship during the program period, 2004-2007. In Table 9, I consider the period 1999-2004, and so excluded controls from 2001. However, I have conducted the falsification test for all entrepreneurial, non-farm entrepreneurial, and farm entrepreneurial households with the entire set of controls (from 2001) and find that the results are the same.

Notes: Each observation corresponds to a quarter-district. The entrepreneurship outcome is defined at the household level. Household responses are aggregated to the quarter-district level in a given year, using the survey weights. Standard errors (in parentheses) are clustered at the district level. *** $p<0.01$, ** $p<0.05$, * $p<0.1$.

**Table 10: Urban Falsification Test**

| | Outcome: Non-farm entrepreneurship | | | | |
|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) |
| NREGS ($\beta$) | 0.019 | 0.013 | -0.011 | 0.010 | -0.010 |
| | (0.050) | (0.051) | (0.059) | (0.060) | (0.072) |
| Post indicator ($\delta$) | -0.073** | -0.134 | -0.441 | -0.309 | 0.765 |
| | (0.035) | (0.149) | (0.362) | (0.378) | (0.815) |
| Poverty and demographics | | yes | yes | yes | yes |
| Labor force characteristics | | yes | yes | yes | yes |
| Rural development measures | | | yes | yes | yes |
| Real per capita NDP | | | | yes | yes |
| Development index components | | | | | yes |
| Quarter fixed effects ($\tau_q$) | yes | yes | yes | yes | yes |
| District fixed effects ($\eta_d$) | yes | yes | yes | yes | yes |
| N | 4,046 | 4,046 | 4,008 | 3,894 | 3,235 |

Table 10 presents the results from a falsification test: I estimate the impact of NREGS on rates of non-farm entrepreneurship in urban areas, where the program was not implemented. I consider only the non-farm sector (as there is no farm sector in urban areas). The program has no effect in urban districts. Combined with the results from Table 9, it is clear NREGS has no effect on entrepreneurship in years or areas in which it was not implemented.
Notes: Each observation corresponds to a quarter-district. The entrepreneurship outcome is defined at the household level. Household responses are aggregated to the quarter-district level in a given year, using the survey weights. Standard errors (in parentheses) are clustered at the district level. *** $p<0.01$, ** $p<0.05$, * $p<0.1$.

Table 11: Controlling for Pre-program Trend in Entrepreneurship

| | **Total** | **Non-farm** | **Farm** |
|---|---|---|---|
| | (1) | (2) | (3) |
| **NREGS ($\beta$)** | 0.016 | **0.182**$^{*}$ | -0.021 |
| | (0.037) | (0.097) | (0.056) |
| Post indicator ($\delta$) | 0.372 | -2.226 | 1.033 |
| | (0.555) | (1.856) | (1.000) |
| Entrepreneurship trend, 1999-2004 | -0.293$^{***}$ | -0.340$^{***}$ | -0.283$^{*}$ |
| | (0.107) | (0.098) | (0.157) |
| Poverty and demographics | yes | yes | yes |
| Labor force characteristics | yes | yes | yes |
| Rural development measures | yes | yes | yes |
| Real per capita NDP | yes | yes | yes |
| Development index components | yes | yes | yes |
| Quarter fixed effects ($\tau_q$) | yes | yes | yes |
| District fixed effects ($\eta_d$) | yes | yes | yes |
| N | 2,381 | 2,278 | 2,376 |

Table 11 presents the results from a robustness check: I estimate the impact of NREGS on rates of total, non-farm, and farm entrepreneurship while controlling for pre-program (1999-2004) trends in entrepreneurship. The results are consistent with Table 3-6; NREGS increases non-farm entrepreneurship along the extensive margin, but has no impact on farm entrepreneurship.

Notes: Each observation corresponds to a quarter-district. The entrepreneurship outcome is defined at the household level. Household responses are aggregated to the quarter-district level in a given year, using the survey weights. Standard errors (in parentheses) are clustered at the district level. *** p<0.01, ** p<0.05, * p<0.1.

**Table 12: Impact on Extensive Entrepreneurship, Base Period = 1999**

| | **Total** (1) | **Non-farm** (2) | **Farm** (3) |
|---|---|---|---|
| **NREGS ($\beta$)** | 0.057 | **0.358**$^{***}$ | 0.007 |
| | (0.046) | (0.097) | (0.069) |
| Post indicator ($\delta$) | 0.061 | -0.494$^{***}$ | 0.100 |
| | (0.070) | (0.166) | (0.105) |
| Poverty and demographics | yes | yes | yes |
| Labor force characteristics | no | no | no |
| Rural development measures | no | no | no |
| Real per capita NDP | no | no | no |
| Development index components | yes | yes | yes |
| Quarter fixed effects ($\tau_q$) | yes | yes | yes |
| District fixed effects ($\eta_d$) | yes | yes | yes |
| N | 2,393 | 2,298 | 2,388 |

Table 12 presents the results from a robustness check against possible mean reversion: I estimate the impact of NREGS on rates of total, non-farm, and farm entrepreneurship using 1999 as the base period instead of 2004. The results are consistent with Tables 3-6.

Notes: Each observation corresponds to a quarter-district. The entrepreneurship outcome is defined at the household level. Household responses are aggregated to the quarter-district level in a given year, using the survey weights. Standard errors (in parentheses) are clustered at the district level. *** p<0.01, ** p<0.05, * p<0.1.

Table 13: Dosage Response

| | Total (1) | Non-farm (2) | Farm (3) |
|---|---|---|---|
| Phase 1 * Post indicator ($\beta_1$) | -0.018 | 0.098 | -0.041 |
| | (0.042) | (0.104) | (0.065) |
| **Phase 2 * Post indicator ($\beta_2$)** | 0.013 | **0.192**$^{**}$ | -0.019 |
| | (0.046) | (0.092) | (0.071) |
| Post indicator ($\delta$) | -0.539 | -2.355 | -0.517 |
| | (0.673) | (1.907) | (0.993) |
| Poverty and demographics | yes | yes | yes |
| Labor force characteristics | yes | yes | yes |
| Rural development measures | yes | yes | yes |
| Real per capita NDP | yes | yes | yes |
| Development index components | yes | yes | yes |
| Quarter fixed effects ($\tau_q$) | yes | yes | yes |
| District fixed effects ($\eta_d$) | yes | yes | yes |
| N | 3,541 | 3,302 | 3,524 |
| Dosage response, F-test | $F_{1,44} = 0.46$ | $F_{1,442} = 1.05$ | $F_{1,442} = 0.09$ |

Table 13 presents the results from estimating the differential effect, or dosage response, of NREGS Phases 1 and 2 on rates of total, non-farm, and farm entrepreneurship. The results of the F-test in each column indicate there is no statistical difference across Phases 1 or 2.

Notes: Each observation corresponds to a quarter-district. The entrepreneurship outcome is defined at the household level. Household responses are aggregated to the quarter-district level in a given year, using the survey weights. Standard errors (in parentheses) are clustered at the district level. *** p<0.01, ** p<0.05, * p<0.1.

# Do Working Mothers Make Different Health Decisions? Evidence from India's Public Works Program

A. V. Chari* Peter Glick† Edward Okeke† Sinduja V. Srinivasan‡

August 2014

**Abstract**

We are the first to research how India's public works program (NREGS), by increasing the opportunity for female employment, subsequently affects investments in child health and child health outcomes. The net effect of public works programs on child health is unclear given the trade-off women face between working and care-giving. NREGS raises the value of working, reducing the incentive for women to spend time care-giving at home, potentially negatively impacting children's health. But mothers may use their additional income from NREGS to purchase high-quality healthcare, resulting in improved health outcomes for children.

Our results are consistent with the predictions of a maternal labor supply model. Using linear, non-linear, and non-parametric methods, we find that after the implementation of NREGS, women significantly reduce their engagement in time-intensive health inputs in favor of employment, which subsequently increases their use of money-intensive health inputs. As outside employment options become more lucrative, women in NREGS areas delay their first prenatal visit, reduce the total number of visits received, and spend less time breastfeeding children, compared to women in non-NREGS areas. Subsequently, women in program areas increase their utilization of money-intensive health inputs, and are significantly more likely to deliver in healthcare facilities than women in untreated areas. The net effect of decreased time-intensive health inputs and increased money-intensive inputs on child health is ambiguous.

*University of Sussex.

†Rand Corporation.

‡Pardee RAND Graduate School. Email: sinduja@rand.org

We thank Krishna Kumar, Shanthi Nataraj, and participants in the Labor & Population Seminar at the RAND Corporation for valuable comments and suggestions. This work was supported by the JL Foundation. Any remaining errors are our own.

# 1 Introduction

Increasingly, countries are implementing programs designed to enhance women's livelihoods and earnings. These programs include micro-credit schemes that provide small loans to women to enable them to start their own businesses, as well as government programs that directly provide employment to women (Das, 2010). Of particular interest are these employment generation programs, social safety nets that provide employment to participants, which have been implemented successfully in many countries including Argentina, Ethiopia, India, and South Africa (Subbarao et al., 2013). A main objective of such public works programs is to reduce the striking gender inequities that exist in developing countries: women are less likely than men to attend school, less likely to be employed, and when employed, earn significantly less than men (World Bank, 2011). In many countries, women lack control over resources and, therefore, the ability to determine their own fortunes. In some countries women do not have the right to conduct business or travel without their husband's permission (Duflo, 2012). By providing employment opportunities for women, public works programs can reduce gender inequalities: women who earn have more control over household finances, and subsequently direct more resources towards the benefit of children, especially girls (World Bank, 2011).

The effect of employment generation programs targeting women on the utilization of child health inputs and child health outcomes has not been previously studied. Evaluations of such programs have focused on the effects on female labor force participation and on household consumption (e.g., Azam, 2012; Gilligan et al., 2009). How public works programs, via increased maternal employment, affect child welfare is not immediately evident, especially given the dual role women occupy, as workers outside the household and caregivers at home. In many countries housework and childrearing are considered the primary responsibility of women. A recent paper analyzing time use data from 23 countries showed that women devote significantly more time than men to housework and child care, with differences ranging from about 50 percent more in Cambodia to about 300 percent more in Italy (Berniell and Sánchez-Páramo, 2011). Because of the greater time constraints that women face, labor force participation and household activities, including child care, may come into conflict, with potentially important implications for the welfare of young children.

Our goal is to understand how increased maternal employment affects the utilization of child health inputs, which will ultimately impact child health and welfare. To analyze this issue, we apply the neoclassical model of labor supply to the context of women participating in a public works program. A growing number of studies from various countries show that when women have greater control over household income, children experience greater benefits largely due to increased expenditure on nutrients, health, and housing (Thomas, 1990; Doss, 2006; Rubalcava et al., 2009). Women can use earned income to purchase high-quality health services, which may have positive impacts on child health. However, employment comes at the expense of time spent on child care activities such as breastfeeding or the preparation of nutritious meals, which may have negative consequences for child health. For example, maternal employment has been associated with shorter duration of breastfeeding and with negative effects on child nutrition (Arifeen et al., 2001; Glick and Sahn, 1998).

Thus, this paper analyzes how India's National Rural Employment Guarantee Scheme (NREGS), by increasing maternal employment, impacts the utilization of purchased health inputs versus investment in time-intensive inputs, and how accessibility of care may moderate these effects. We then examine the impact on child health outcomes, and the factors that moderate this relationship. Understanding how employment programs impact the use of child health inputs and child welfare is critical for the design of future programs that aim to draw women into the labor force.

NREGS in India is the largest direct employment program in the world and has increased work opportunities for millions of poor rural women as well as men. To date, the program has provided employment to nearly 50 million households in India (about 36 percent of the rural labor force). Additionally, it is the largest program to specifically target women (Ravi and Engler, 2012). We are conducting the first rigorous evaluation of the effects of this program on the provision of maternal inputs and on child health, and attempt to inform policy by delineating the contextual factors that may diminish or enhance these impacts.

This study makes an important contribution to the literature on the link between maternal employment, the utilization of health inputs, and child health. This link has been studied within developing and developed countries (Desai et al., 1989; Blau and Grossberg, 1992; Glick and Sahn, 1998; Waldfogel, 2002; Anderson et al., 2003; Gordon et al., 2007; Baker and Milligan, 2008), but the findings are mixed. Identifying the true causal effect of maternal employment has proved challenging because maternal employment is a voluntary choice, and as such, may be correlated with confounders, some of which are hard to measure. In addition, reverse causality is a significant problem as child health itself impacts the decision of a mother to work. We improve on this literature by leveraging a quasi-experimental setting: the rollout of NREGS can be treated as a natural experiment that exogenously increases employment opportunities and thereby work participation on the part of women, permitting a credible identification of the effect of maternal employment on the use of health services and child outcomes. We employ a differences-in-differences approach that takes advantage of nationally representative longitudinal data and compares pre- and post-rollout outcomes in treated relative to untreated households. To our knowledge, this is the first rigorous evaluation of the effects of a large-scale public employment program on child health inputs and outcomes.

Our results are consistent with the predictions of a maternal employment model. We find that in response to the implementation of NREGS, women significantly reduce their engagement in time-intensive health inputs in favor of employment, which subsequently increases their use of money-intensive health inputs. As their outside employment options become more lucrative, women in NREGS (treated) areas delay their first prenatal visit, reduce the total number of visits received, and spend less time breastfeeding children, compared to women in non-NREGS areas. Subsequently, women in treated areas increase their utilization of money-intensive health inputs, and are significantly more likely to deliver in healthcare facilities than women in untreated areas. The net effect of decreased time-intensive health inputs and increased money-intensive inputs on child health is ambiguous.

The paper proceeds as follows. Section 2 provides a description of the National Rural Employment Guarantee Scheme. In Sections 3 and 4, we discuss the theoretical framework and describe the data underlying the econometric approach of Section 5. In Section 6, we present and discuss our results. Section 7 concludes.

## 2 National Rural Employment Guarantee Scheme

The National Rural Employment Guarantee Scheme (NREGS) is the largest employment guarantee scheme in the world. Any rural household is eligible to participate in the program. Each enrolled household is guaranteed a maximum of 100 days of paid labor per year, to be allocated across the adult household members in any fashion throughout the year. Employed individuals work on public works projects including land development and rural connectivity (e.g., road construction). The NREGS minimum wage is set at 60 rupees per day (approximately $1.12), with equal wages for men and women.[1] The program was implemented in phases across rural India (see Figure 1).

By law, at least one-third of NREGS beneficiaries should be women. In 2011, the average share of NREGS workdays actually allocated to women was 49.5 percent (Dasgupta and Sudarshan, 2011). Consistent with this explicit program focus, Azam (2012) estimates that NREGS has raised the labor force participation of women by 5.6 percent in treated districts, relative to control districts, but has not significantly impacted the labor force participation of men. Khera and Nayak (2009) find that fewer than 30 percent of women working in NREGS reported being engaged in or looking for wage work prior to the program, highlighting the success NREGS has had on drawing women into the labor force. Local availability of the work, the regularity of working hours, the non-exploitative nature, and the social acceptability of the work were cited as important factors that induced participation among women who had not considered working before.

With our data, we can establish that NREGS is also associated with increased maternal employment (Figure 2). Panel 2(a) presents trends in maternal employment within rural (NREGS) areas in the fourth quarter of 2007 and all quarters of 2008, after NREGS was implemented in Phase 1 and 2 districts. The program is associated with significantly higher (p-value < 0.01) rates of maternal employment in treated Phase 1 and 2 districts (red line) compared to untreated Phase 3 districts (blue line).[2] Further, after Phase 3 was implemented in 2008, we see that maternal employment increased in these districts. Finally, we show that in urban areas, which did not receive NREGS, trends in maternal employment are not associated with the program (Panel 2(b)). Having established that NREGS promotes maternal employment (only) in treated areas, we proceed to examine how maternal employment affects investments in maternal and child health and, subsequently, children's health outcomes.

[1]Recently, NREGS increased the employment ceiling to 150 days per year and raised the daily minimum wage to 100 rupees ($1.50). However, these changes took effect after the period we are studying, 2004-2008.

[2]The survey asks mothers if they worked in the last 7 days. As the survey was administered starting in December 2007, we cannot identify trends in maternal employment prior to the implementation of NREGS Phases 1 and 2. See Section 4 for more information on the timing of NREGS implementation relative to the survey administration.

## 3 Model of Maternal Labor Supply

Our theoretical framework adapts the neoclassical labor supply model. In this context, a mother chooses between investing in time-intensive health inputs at home and earning outside the home to purchase money-intensive health inputs for her children. Time is a finite resource and has an opportunity cost. In general, the value of time is likely to be greater for individuals who have the option of engaging in income-generating activities. An increase in the availability of work opportunities for women increases the opportunity cost of time, making it more costly to invest in time-intensive activities. For a woman who is employed, the cost of taking a child to be immunized is larger relative to one who is unemployed, because her time is more expensive (time spent traveling to the clinic is time lost from working and earning). However, the additional income earned by a working mother may be sufficient to purchase compensatory inputs; e.g., the family may now be able to hire a care-giver for the child. These two effects, income and the opportunity cost of time, work in opposite directions. In contrast, a father's employment is expected to largely work through the income channel, because in developing countries the father typically contributes little time towards childrearing, especially in the case of young children (Evans, 1995; Anandalakshmy, 1994).

### 3.1 Maternal Decision

Mothers must decide how many hours to work versus how much time they will devote to care-giving at home. They face the following utility function, which we assume is well-behaved (continuous, twice-differentiable, quasi-concave).

$$U = U(t, m) \tag{1}$$

where $t$ is the time-intensive maternal health input, such as prenatal care, and $m$ is the money-intensive health input, such as giving birth in a hospital. Although we present a model with a single time-intensive and money-intensive input, the results are generalizable to input vectors. We also recognize health inputs will often have a time and money component; we analyze unidimensional inputs for ease of exposition.

Women face an income constraint given by

$$I = V + \underbrace{w \cdot h}_{E} \geq p \cdot m \tag{2}$$

where income $I$ is the sum of non-labor income $V$ and earned income $E$, and $p$ is the price of $m$. Since we are examining the household decision from the perspective of the woman, consider $V$ to be the money income derived from male employment and any household endowments or assets. A woman's earned income is the product of the wage rate, $w$, and the number of hours a woman worked, $h$.

The corresponding constraint for the total hours worked ($h$) and the total hours spent on the time-intensive health input ($t$) is

$$T \geq h + t \tag{3}$$

where $T$ is the total amount of time available. Time-use data indicate that Indian women, especially those in rural areas, have a very limited amount of free time, usually fewer than 100 minutes each day (OECD, 2011; Sivakami, 2010). Subsequently, we do not include leisure in Equation 3, to more accurately represent the constraints women face. This abstraction does not affect the main implications of the model.

Women maximize Equation 1 subject to the constraints given by Equations 2 and 3.

$$\max_{m,t,h} U(m,t) - \lambda (pm - wh - V) - \mu (h + t - T) \tag{4}$$

with first-order conditions (FOCs) as follows:

$$FOC_h : \lambda w = \mu \tag{5}$$

$$FOC_m : \frac{\partial U}{\partial m} = \lambda p \tag{6}$$

$$FOC_t : \frac{\partial U}{\partial t} = \mu \tag{7}$$

$$FOC_\lambda : p \cdot m = w \cdot h + V \tag{8}$$

$$FOC_\mu : h + t = T \tag{9}$$

Combining Equations 5 and 7, we find that the shadow price for an additional unit of the time-intensive health input $t$ is related to the wage rate $w$. Women give up $\lambda w$ for each additional hour of care-giving at home.

Equations 6-9 yield the Marshallian demand functions health inputs $m$ and $t$, as a function of the exogenous prices, wage rate, and unearned income:

$$\begin{aligned} m^* &= D_m\left(p, w, V\right) \\ t^* &= D_t\left(p, w, V\right) \end{aligned} \tag{10}$$

and $h^* = T - t^*$.

Note that if the wage is zero (a mother has no outside option), this means the time-intensive input is, essentially, free. That is, when the wage is zero or very low we have a corner solution where women will choose not to participate in the labor force, instead devoting all their time to the time-intensive health input. Any income will be used to purchase money-intensive input $m$. We assume an interior solution in the next section, where we consider how the demands for the health inputs change when the outside option for mothers becomes more attractive, i.e., wages increase.

### 3.2 Comparative Statics: Impact of Wage Increases[3]

Impact on Hours Worked ($h$)

When the wage increases, working becomes a more attractive option. This is shown using the Slutsky decomposition:

$$\frac{dh\left(p, w, V\right)}{dw} = \underbrace{\left.\frac{dh}{dw}\right|_{U=\bar{U}}}_{1} + h \underbrace{\frac{dh}{d\left(wh + V\right)}}_{2} > 0 \tag{11}$$

Term 1 of Equation 11 represents the substitution effect, which is unambiguously positive: as the wage increases, the opportunity cost of remaining at home increases, so mothers will increase the time spent in the labor force. However, increases in the wage also increase real income, allowing women to work fewer hours: Term 2 represents this income effect, which is (likely) negative. The

[3]See the Theoretical Appendix for derivations.

substitution effect will outweigh the income effect at relatively low wages, such as those associated with a public works program. Thus we expect Equation 11 to be positive in this context. Indeed, as discussed in Section 2 and shown in Figure 2(a), mothers increased their labor force participation after NREGS was implemented and the wage increased.

Impact on Time-Intensive Input ($t$)

As the time constraint (Equation 9) is binding, and we have shown that mothers will work more when wages increase, investment in the time-intensive input must consequently decrease. In this case, the Slutsky equation is

$$\frac{dt\,(p,w,V)}{dw} = \underbrace{\left.\frac{dt}{dw}\right|_{U=\bar{U}}}_{1} + \underbrace{(T-h^{*})\,\frac{dt}{d\,(wh+V)}}_{2} < 0 \tag{12}$$

where Terms 1 and 2 are the substitution and income effects, respectively. As the wage increases, the time-intensive input becomes relatively expensive, and mothers will reduce their demand for $t$ (substitution effect). The increases in the wage also increase real income, allowing women to work fewer hours, thus increasing their demand for $t$ (income effect). We hypothesize that in a public works program, where wages are low, the former will outweigh the latter. Thus we expect that in NREGS areas, mothers will decrease their utilization of time-intensive health inputs, operationalized by prenatal care and breastfeeding duration (discussed in more detail in Section 4).

Impact on Money-Intensive Input ($m$)

Although the absolute price of the money-intensive is constant (in this context), the relative price decreases as the wage increases, and the time-intensive input becomes more costly. Thus we expect women to substitute away from $t$ towards $m$. Further, at higher wages, more income is available to purchase the money-intensive input; the income effect is positive. The Slutsky decomposition takes the form

$$\frac{dm\,(p,w,V)}{dw} = \underbrace{\left.\frac{dm}{dw}\right|_{U=\bar{U}}}_{1} + \underbrace{m\,\frac{dm}{d\,(wh+V)}}_{2} > 0 \tag{13}$$

As the substitution and the income effects are both positive, empirically we expect that mothers in program districts will increase their use of money-intensive inputs, operationalized by delivery in a healthcare facility.

Impact on Child Health

Changes in the utilization of child health inputs will impact children's health. For example, lack of prenatal care and exclusive breastfeeding are known to be associated with poor child health (Poma 1999; Lara and Pullum 2005). However, access to higher quality healthcare (via increased employment and income) can improve child health. This is ultimately an empirical issue that we bring to the data: we consider the effect of NREGS on infant and neonatal mortality rates. Further, the impact of NREGS will be mediated by contextual factors including ease of access to health facilities (distance), and the availability (and quality) of substitutes for the mother's time (formal or informal caregivers). We discuss how we account for these moderators in our empirical strategy (Section 5).

## 4 Data

To answer the research questions posed, we use publicly available data from the 2007-2008 round of India's District Level Household and Facility Survey (DLHS). DLHS is a nationally representative survey of 720,320 urban and rural households, spread across 601 districts and 30,028 villages and urban wards. DLHS collects data on reproductive and child health. It covers all states and union territories of India (except Nagaland) and follows a multi-stage stratified systematic sampling design. The DLHS has rich information on household characteristics including access to formal health services. All ever-married women ages 15-49 in sampled households were interviewed, providing retrospective information on all pregnancies since January 1, 2004.

### 4.1 Timing of DLHS and NREGS

As described above, NREGS was implemented in three phases. Figure 3 depicts the phased rollout of the program relative to the time period covered by the survey data. India has a total of 615 districts (601 of which are covered by DLHS). Phase 1 was implemented in February 2006, with 200 treated districts receiving NREGS and 415 districts remaining untreated. In April 2007, 130 districts were added in Phase 2, resulting 330 treated districts and 285 comparison districts. Phase 3 was implemented in April 2008, and all districts were treated. The survey data spans January 2004-February 2009 and covers Phases 1, 2, and 3.

Under-developed districts received the program first.[4] Table 1 compares district, household, and mother characteristics across the three NREGS phases. Phase 1 and 2 districts have lower household expenditures and literacy rates compared to Phase 3; rates of labor force participation are similar across all three phases. Households are of similar size across the phases, but the rate of poverty is lowest in Phase 3 districts (28 percent of households hold a poverty card, compared to 39 percent in Phase 1 and 34 percent in Phase 2). The differences across the phases are apparent when comparing mother characteristics. Though mothers are of similar age, those in Phase 1 and 2 districts get married one year earlier compared to Phase 3 districts. Mothers in Phase 3 districts have 1.5 more years of education than in Phase 1 districts. Phase 1 and 2 mothers have experienced a greater number of pregnancies/births than in Phase 3.

### 4.2 Variable Definitions

Below we provide a description of the variables used in our analysis. The Data Appendix provides additional details on variable definitions and management.

Health inputs

We utilize data on (1) Prenatal care (the timing of the first visit, and the number of visits) and (2) Exclusive breastfeeding (initiation and duration) as measures of time-intensive health inputs.[5] Our measure of a money-intensive input is (3) Delivery location (institutional vs. non-institutional deliveries). The data on outcomes (1) and (3) come from the last pregnancy, and the data on outcome (2) come from the last surviving child. Table 2 shows that women in treated (NREGS) districts delay prenatal care, receive fewer visits, exclusively breastfeed for fewer months, and are less likely to deliver in an institution than women in untreated districts.

Child health

[4] We describe how we address any selection issues in our empirical strategy (Section 5).
[5] Note that in India, prenatal care is free or highly subsidized, so utilizing prenatal care is a time-intensive activity.

Our measure of child health is mortality. We distinguish between neonatal mortality (death within one month after birth) and infant mortality (death before the child's first birthday); mortality information is known for all pregnancies since January 2004. Table 2 shows that the neonatal and infant mortality rates are significantly higher in NREGS districts than in untreated areas.

Moderator variables

Apart from understanding how NREGS impacts maternal and child outcomes, we are interested in identifying factors that moderate this effect. Moderator variables include: (1) access to a health facility, measured as distance in kilometers from the village and, (2) substitutes for maternal care, measured by household composition, using the age and gender of all household members. The bottom panel of Table 2 shows that treated and untreated districts are similar with respect to accessibility of care (distance to health facility) and household composition.

Control variables

The DLHS contains rich information on a range of individual, household, and village level characteristics, including the child's age and gender; the mother's age at birth, educational history, and pregnancy history; father's educational history[6]; the household religion, caste, and composition, measured by the share of adults. We describe these variables in detail in the Data Appendix, and define their role in our empirical specifications in Section 5 below.

## 5 Empirical Strategy

Since NREGS was implemented in under-developed districts first, simple cross-sectional comparisons of child outcomes (post-rollout) across treated and untreated districts are likely to yield misleading estimates of the effect of NREGS on the outcomes of interest because of underlying differences between the two groups. We address this issue by implementing the differences-in-differences (DID) estimator.

The DID estimator compares the change in the outcome before and after NREGS implementation in treated districts to the change in the outcome in untreated districts over the same time period. The major advantage of DID is that it allows for comparison of the same unit before and

[6]Since we cannot control for pre-program household wealth, we use the father's educational attainment as a proxy.

after the treatment. Thus we can control for any fixed unobservable differences between treated and untreated units. At the same time, by comparing the change in outcomes for treated and untreated districts, we can control for secular time trends affecting the outcome. Implementing the DID estimator requires longitudinal data. While DLHS is not a longitudinal survey, it collects retrospective data on all births between January 2004 and December 2007. We can observe births that occurred prior to NREGS implementation and those that occurred after. Therefore we implement a DID strategy that treats the district as the fixed unit for which we have pre- and post-period observations.

In practice, we estimate the DID estimator using the following regression model[7]:

$$y_{ijvdt} = \pi_1 treatment_{it} + \beta_1 X_{ijvt} + \delta_d + \nu_m + \tau_y + e_{ijvdt} \tag{14}$$

where $y_{ijvdt}$ is an outcome variable associated with child $i$ of mother $j$ in village $v$ in district $d$ at time $t$. The indicator $treatment_{it}$ takes the value 1 if child $i$ is in the treated group at time $t$, and is 0 otherwise. Everyone is untreated prior to 2006 ($treatment_{it}$ is 0). Because of the phased rollout of the program, 200 districts were treated in February 2006, an additional 130 districts became treated in April 2007, and the remaining 285 districts received treatment in April 2008 (in these months and years $treatment_{it}$ is 1 for those districts). $X_{ijvt}$ is a vector of child, mother, household, and village characteristics (details provided in the Data Appendix). We account for month effects with $\nu_m$ and year effects with $\tau_y$; $\delta_d$ is a district level fixed effect to compare outcomes over time within the same district. We specify the unobserved error term as $e_{ijvdt}$. The coefficient of interest is $\pi_1$, the DID estimate of the intent-to-treat effect. In the context of a maternal employment model, we expect $\pi_1$ to be negative when $y_{ijvdt}$ measures a time-intensive health input, and positive for a money-intensive input.

Validity of the DID estimator requires that treatment and control groups would have maintained the same relative trends in outcomes in the absence of treatment. We discuss the validity of this parallel trends assumption for each outcome (depicted by Figures 4-8) in Section 6 below.

[7] We first implement linear regression models for ease of interpretation. See Section 6.5 for a discussion of the results when estimating non-linear models.

## 5.1 What is the effect of NREGS on the provision of maternal inputs?

To estimate the effects of NREGS on the provision of maternal inputs, we estimate specification 14 above for each of the maternal inputs observed in the data: (1) Prenatal care, (2) Breastfeeding duration. and (3) Delivery location. For outcomes (1) and (2), we expect $\pi_1$ to be negative, as women substitute away from time-intensive inputs towards work, as hypothesized by the labor supply framework. We expect $\pi_1$ to be positive when estimating the impact on the money-intensive input of institutional delivery: increased earnings allow women to purchase better healthcare.

## 5.2 To what extent is the effect of NREGS on the provision of maternal inputs moderated by the accessibility of care?

To answer this question, we estimate a variant of Equation 14. Now the treatment indicator is interacted with a measure of a household's access to care, operationalized as the distance (in kilometers) to the nearest health facility. We estimate the following regression:

$$y_{ijvdt} = \pi_1 treatment_{it} + \pi_2 treatment_{it} * distance_{ijv} + \beta_1 X_{ijvt} + \delta_d + \nu_m + \tau_t + e_{ijvdt} \tag{15}$$

where $y_{ijvdt}$ denotes a particular health input, e.g, number of prenatal visits for child $i$, and $distance_{ijv}$ represents the distance to the nearest health facility. In this specification, $\pi_2$ tells us if the effect of NREGS on the provision of maternal input $y$ is mediated by the distance to the nearest health facility. We expect if the facility is inaccessible, women may not be willing to forgo earnings to spend additional time obtaining healthcare. Thus the sign on $\pi_2$ should be negative, i.e. larger distances negatively impact the program's effect on care seeking.

## 5.3 How does maternal employment affect child health, and what factors moderate this relationship?

To research this question, we estimate the causal effect of NREGS using two measures of child health: neonatal and infant mortality. As discussed earlier, the effect of the program on child health is an empirical issue because the beneficial effects due to increased income generated by the program may be counteracted by the negative effects due to a reduction in time-intensive maternal inputs. The effect on child health will depend on the quality of available substitutes. In developing countries, other household members usually serve as informal caregivers in the absence of the mother. Young children in a household with many females who can serve as substitute caregivers are therefore likely to experience smaller negative effects than a similar household with fewer women. To test this possibility, we will estimate the following variants of Equation 14:

$$y_{ijvdt} = \pi_1 treatment_{it} + \beta_1 X_{ijvt} + \delta_d + \nu_m + \tau_t + e_{ijvdt} \tag{16}$$

$$y_{ijvdt} = \pi_1 treatment_{it} + \pi_2 treatment_{it} * composition_i + \beta_1 X_{ijvt} + \delta_d + \nu_m + \tau_t + e_{ijvdt} \tag{17}$$

where $y_{ijvdt}$ is now a measure of child health, and $composition_i$ is a measure of household composition. We consider the following measures of composition: (1) Share of females older than 15 in the household, (2) Share of females older than 18 in the household, and (3) Share of females ages 15 to 18 in the household. The coefficient of interest is $\pi_2$, which tells us how each of these household composition variables influences the effect of the program on child health. We expect that the third measure will be the most important, as young women ages 15-18 may be taken out of school to care for younger children so their mothers can participate in NREGS (which is available only to adults). One limitation of this approach is that we have data on household composition only at the time of the survey, which may not accurately represent household composition in previous years.

# 6 Results & Discussion

In this section we report and discuss our results of the impact of NREGS on the provision of maternal inputs (Figures 4-7, Tables 3-6), any factors that may moderate this effect (Table 7), and the subsequent impact of NREGS on children's health outcomes (Figure 8, Tables 8-9). All specifications include child, mother, father, and household controls; month and year fixed effects to control for any seasonality (e.g., increased difficulty in accessing hospitals during the rainy season) and secular time trends; and district fixed effects to account for unobservable time-invariant district characteristics that may be correlated with the outcomes, which is particularly important in this context as NREGS was implemented in the least developed districts first.

## 6.1 Impact of NREGS on Time-intensive Inputs

### 6.1.1 Prenatal Care

Prenatal care is a cost-effective method to reduce maternal and child mortality and morbidity (Jowett, 2000). Receiving care earlier during pregnancy as well as receiving more care in total are associated with positive maternal and child health outcomes (WHO and UNICEF, 2003). In rural India, the average share of women receiving any prenatal care is approximately 70 percent, and states with higher rates of prenatal care utilization generally have lower rates of infant mortality (IIPS, 2007).

Within the labor supply model, implementation of a workfare program such as NREGS raises a woman's marginal cost of time. As expressed in Equation 12, any activity limiting the time available to earn, such as the utilization of prenatal care, will be negatively impacted. We consider the impact of NREGS on (1) when women first received prenatal care, measured by the month (within the pregnancy) of the first prenatal care visit, and (2) the total amount of care received, measured by the number of prenatal care visits. Data were collected for the most recent pregnancy. Examining each outcome in turn thus provides information about how women may alter their utilization of prenatal care when faced with increased time constraints due to participation in NREGS, allowing for the identification of channels that may subsequently impact child health.

We first analyze how NREGS affects the timing of the first prenatal care visit. The labor supply model predicts a negative impact of the program. Women are able to work in the early months of pregnancy (especially the physical work NREGS might require), resulting in a high opportunity cost of any activity that reduces the time spent in employment, such as a prenatal care visit. As the pregnancy continues, women will likely face a decreased ability to undertake taxing physical work. This reduces the cost of time and increases the marginal benefit of the first prenatal visit.

Figure 4 presents the trend in the timing of prenatal care (among women who received at least one visit), across treated and untreated districts, before and after NREGS implementation. The figure provides evidence in favor of the parallel trends assumption underlying the differences-in-differences specification. However, the naïve estimate of the treatment is small; NREGS is associated with delaying prenatal care by approximately 0.01 months.[8] The regression estimate of the impact of the program on prenatal care timing is larger (Column 1 of Table 3): we find that women in NREGS districts delay their first prenatal care visit by 0.07 months (p-value < 0.01), or about 2 days, compared to women in non-NREGS districts. This result is consistent with the hypothesis that women in program areas delay receiving prenatal care, likely because the marginal cost of obtaining care is high early in the pregnancy. We note that this estimate is averaged over the entire pregnancy; what remains unclear is whether the impact of NREGS differs according to the month of the pregnancy.

To explore this issue further, we attempt to determine when in the pregnancy NREGS has the greatest impact. To do so, we construct a set of nested indicator variables for when prenatal care was initially received: an indicator equal to 1 if the woman received the first visit in the $1^{st}$ month or later; an indicator equal to 1 if the woman received the first visit in the $2^{nd}$ or later, and so on up to the woman receiving the first visit in the $9^{th}$ month of the pregnancy. We then estimate a series of linear probability models for these indicators. The hypothesis of the maternal labor supply model suggests that the net cost of the first visit will decrease with each month of the pregnancy. Thus we expect that NREGS will reduce the propensity of women to receive care in early months of the pregnancy, when it is costly to give up work, and increase the probability of receiving care in later months of the pregnancy, when the opportunity cost of work is low and the benefit of obtaining care is high.

Columns 2-10 of Table 3 report the regression results for this series of linear probability models. The results confirm our hypothesis: NREGS increases the probability of a woman's first visit occurring in the last trimester ($7^{th}$, $8^{th}$, and $9^{th}$ months) by 1 percentage point (p-value < 0.01). This is consistent with the idea that women find the first prenatal care visit less costly as the pregnancy progresses, due to the decreased opportunity cost of working and an increased benefit of care. Note that the effect of NREGS on the probability of receiving a visit in the $1^{st}$ month or later (Column 2 of Table 3) is approximately zero. We interpret this result to mean that NREGS does not affect the propensity of women to obtain prenatal care, but rather the timing.

[8] The naïve estimate is derived without accounting for any controls. It compares the difference in the average outcome $y$ in the treated group across the pre-treatment and post-treatment periods to the difference in the average outcome in the untreated group across the same time periods: $\pi_{\text{naïve}} = \left(\bar{y}^{post}_{treated} - \bar{y}^{pre}_{treated}\right) - \left(\bar{y}^{post}_{untreated} - \bar{y}^{pre}_{untreated}\right)$.

We then examine the effect of NREGS on the total number of prenatal care visits a woman receives. Our theoretical framework predicts the impact of NREGS will be negative, as more visits increase the total time cost and reduce the time available for working (and earning). The results discussed above indicate that NREGS delays when women first receive prenatal care; in this context, we anticipate women in program areas will reduce the total number of prenatal care visits.

Figure 5 depicts the trends in the number of prenatal visits (for women receiving prenatal care) across treated and untreated districts. As with the timing of prenatal care, there is evidence supporting the parallel trends assumption. The naïve DID estimate of NREGS on the amount of care received is zero visits. The regression results are larger: Column 1 of Table 4 reports that in NREGS areas, women receive 0.05 fewer visits (p-value < 0.01) over the course of a typical pregnancy than women in untreated areas.[9] Similar to the timing outcome, this estimate is averaged over all prenatal visits; below we consider if the impact of NREGS differs according to the number of visits obtained.

To determine where in the prenatal visit trajectory the impact of NREGS is the greatest, we construct another set of indicator variables for the number of prenatal visits received. We define an indicator equal to 1 if the woman received at least 1 prenatal visit; an indicator equal to 1 if the woman received at least 2 prenatal visits, and so on up to the woman receiving at least 9 prenatal visits, which corresponds to the $95^{th}$ percentile of the visit distribution. We then estimate another set of linear probability models. The hypothesis of the maternal employment model suggests that each additional prenatal visit becomes more costly with respect to time. Women may perceive the marginal benefit of each visit to diminish, and if the time required to complete each visit remains constant, this increases the relative cost of subsequent visits. Thus we expect to see little or no impact of NREGS on earlier visits, and a negative impact for later visits.

Columns 2-10 of Table 4 report the regression results for this series of linear probability models. We first note the program has no impact on obtaining any visit; the effect of NREGS on the probability of receiving at least one visit (Column 2 of Table 4) is approximately zero. This is consistent with the results in Table 3: NREGS does not affect the propensity of women to obtain prenatal care. The remaining results confirm our hypothesis: there is little impact of NREGS on

[9] We control for the timing of the first prenatal visit in all specifications presented in Table 4.

the probability of obtaining 5 visits or fewer, but NREGS reduces the probability of obtaining at least 6 or 7 visits by 1 percentage point (p-value $< 0.01$). This supports the idea that women find the marginal visit less beneficial when they have already received a greater number of visits and when the opportunity cost of obtaining the marginal visit is high. Note that this negative impact diminishes, in magnitude (and statistical significance), through the rest of the visit trajectory (Columns 9 and 10 of Table 4), likely because the propensity to obtain more than 8 visits is small in both treated and untreated districts, so NREGS has little effect.

The results in Tables 3 and 4 are consistent with the predictions of a model of maternal employment. With respect to prenatal care outcomes, the substitution effect is dominant for women in NREGS districts. The results indicate that when given an employment opportunity and increased earnings potential, women reveal a preference to substitute away from time-intensive activities that require extended absences from work, such as prenatal care. After onset of the program, women delay the timing of their first visit and reduce the total number of prenatal care visits.

#### 6.1.2 Breastfeeding

Here we consider the impact of NREGS on breastfeeding. Breast milk contains vital nutrients for the newborn, and an increased duration of breastfeeding is associated with a reduced risk of infant infection, malnutrition, and mortality, especially in developing countries (Office of the Surgeon General (US), 2011; Gupta and Khanna, 1999). The World Health Organization (2005)recommends exclusive breastfeeding for a minimum of six months. From 2003 to 2006, about 26 percent of women in rural India initiated breastfeeding, and exclusively breastfed for a median duration of 2 months (IIPS, 2007).

Within our framework of maternal labor supply, breastfeeding is a time-intensive input. Equation 12 predicts we should expect a workfare program, by providing employment opportunities to mothers, to reduce the duration of (exclusive) breastfeeding, as the time a mother takes to feed the child is time spent away from working. However, there are some reasons why a generic workfare program, and why NREGS in particular, may not negatively impact breastfeeding time. First, women may work through the pregnancy and reduce the time spent on other time-intensive inputs (as Section 6.1.1 suggests they do), in anticipation of an inability to work immediately and for some time after delivery. That is, women might be saving for a limited form of maternity leave. If so, then a program such as NREGS might actually increase the duration of exclusive breastfeeding, though it is a time-intensive activity. Further, NREGS is supposed to provide some childcare at project sites (Ministry of Rural Development, 2008). This could allow women in program districts to work and breastfeed at the same time, also contributing to a longer breastfeeding duration (or at least not a shorter duration).

Figure 6 presents trends in the average duration of exclusive breastfeeding across treated and untreated districts, before and after program implementation. There is evidence of the parallel trends assumption prior to the onset of NREGS. Further, we observe a sharp decrease in the duration of breastfeeding in treated areas after they received the program, suggesting the main prediction of the maternal labor supply model is borne out. The naïve differences-in-differences estimate is -0.14 months, indicating women in program areas stop exclusively breastfeeding earlier than mothers in untreated districts.

The regression results for the impact of NREGS on the duration of exclusive breastfeeding, reported in Table 5, are consistent with Figure 6. We first consider (in Column 1) whether the program has an impact on the initiation of breastfeeding (analogous to identifying if NREGS affects the propensity of women to obtain prenatal care); we find the program has no effect. In Column 2, we estimate the impact of NREGS on the total number of months mothers exclusively breastfeed. Compared to untreated areas, women in treated districts reduce total breastfeeding duration by 0.16 months ($p < 0.01$), equivalent to 5 days. This is consistent with a model of maternal labor supply: in response to increased wages, women substitute away from the time-intensive activity of breastfeeding to work instead.

The estimate in Column 2 of Table 5 does not provide information on when women in NREGS districts stop breastfeeding. As in Section 6.1.1, we construct a series of nested indicator variables to better understand the breastfeeding trajectory in treated versus untreated areas. We define indicators equal to 1 if the woman exclusively breastfed for longer than 0.5 months, 1 month, 2 months, 3 months, 6 months, 12 months, 18 months, and 24 months, and estimate a set of linear probability models. The hypothesis of the labor supply model suggests that each additional month of breastfeeding becomes more costly with respect to time, as women recover from delivery and are able to work. In this context, we expect the program will reduce the propensity to breastfeed at time points further along the trajectory. However, if women do save for maternity leave or the NREGS work sites allow for working and breastfeeding, this could mitigate or even reverse the result expected from a typical maternal employment framework.

The results from estimating this series of linear probability models are presented in Columns 3-10 of Table 5, and are consistent with the hypothesis of a standard model of maternal employment. As expected, the propensity to breastfeed 0.5 months after delivery (Column 3) is similar among women in treated and untreated districts, when women face a diminished ability to undertake work outside the household. However, once they can undertake work, women in NREGS areas are less likely to breastfeed by 1-2 percentage points at 1, 2, 3, and 6 months after the child is born. This is consistent with the idea that mothers in program districts face high opportunity costs once they recover completely, and so substitute away from care-giving at home. Note that the difference in the propensity to breastfeed fades 6 months after delivery (Column 8-10), likely because few women in India breastfeed for that long.

The maternal employment model suggests NREGS should have little impact on the propensity to breastfeed immediately after delivery because the opportunity cost of the mother's time is small or zero. If NREGS is associated with a significant reduction in the propensity of women to breastfeed just after childbirth, this suggests that maternal employment is not the (dominant) channel through which the program affects breastfeeding duration and/or our treatment measure is correlated with unobservable socio-economic characteristics (e.g., regional differences in attitudes towards breastfeeding) that affect the outcome, but are not accounted for in our specifications. This provides a falsification test for our hypothesis.

We implement this falsification test by considering the impact of NREGS on the propensity of women to feed infants colostrum, which is thick breast milk produced immediately after childbirth for only a few days (Costello and Orsin, 2003; Orsin et al., 2002). Since movement (and thus work) is restricted for the first few days after delivery (at least), the opportunity cost of time is zero and NREGS should not impact the likelihood of women giving colostrum. Column 11 of Table 5 confirms the prediction of our falsification test: the propensity of women to feed their newborns colostrum is not correlated with the implementation of NREGS.

### 6.2 Impact of NREGS on Money-intensive Inputs

#### 6.2.1 Delivery

In this section we examine one channel through which women can use their earnings to invest in money-intensive health inputs: we research the effect of NREGS on the probability of delivering in a healthcare facility (public or private). Having skilled professionals attend the childbirth is considered critical in ensuring survival of the mother and newborn (Bajpai et al., 2013; UNFPA and Ipea, 2007; Bhutta et al. 2005). In India, however, the costs of institutional delivery are often prohibitive, especially in rural areas (Mohanty and Srivastava, 2013; Kesterton et al., 2010). From 2004-2008 the mean out-of-pocket expenditure for delivery was $37 in a public facility, equivalent to 5 percent of per capita GDP for this period (World Bank, 2014; Mohanty and Srivastava, 2013).[10] High costs contribute to low rates of institutional delivery in rural India; the average share of rural women who delivered their youngest child in a health facility between 2000 and 2005 was under 30 percent (IIPS, 2007).

Within the context of the maternal employment framework, the expected impact of a workfare program (NREGS) on the propensity for institutional delivery is positive. However, although NREGS increases female earnings, and thus the budget constraint of the household, institutional delivery rates may not correspondingly increase. For example, if women have little bargaining power in the household, or other expenditures such as food or clothing are more acute, then additional earnings will not be used for delivery in a healthcare facility.

Figure 7 depicts the trends in the rates of institutional delivery across treated and untreated districts, before and after the implementation of NREGS, providing evidence in favor of the parallel trends assumption. Further, we observe a sharp increase in institutional delivery rates in treated districts after the implementation of NREGS. Correspondingly, the naïve differences-in-differences estimate is large: women in NREGS areas increase the propensity of institutional delivery by 6 percentage points. Table 6 reports the regression results on the impact of NREGS on rates of institutional delivery for the most recent pregnancy, which are smaller than the estimates in Figure 7. We find the program results in a 1 percentage point increase in the probability of women to deliver in either a government or private healthcare facility (Column 1). Figure 7 and

[10] 2004 prices.

Table 6 are consistent with the prediction of a model of maternal employment: women substitute away from care-giving at home towards employment in NREGS, and subsequently use their earnings from NREGS to finance institutional delivery costs.

Combining these results with those from Section 6.1.1, the evidence presented is consistent with the labor supply framework presented in Section 3. Women in NREGS districts forgo time-intensive health investments during pregnancy in favor of earning more income, allowing investment in money-intensive inputs, such as higher quality healthcare at the time of childbirth.

### 6.3 Moderated Impact of NREGS on Maternal Inputs

The evidence presented so far suggests that mothers in NREGS districts reduce time-intensive maternal inputs (prenatal care and breastfeeding) in favor of employment, to subsequently purchase income-intensive inputs (institutional delivery). We now examine if the program effect on maternal inputs is moderated by the accessibility of healthcare, measured by the distance (in kilometers) to the nearest healthcare facility. As distance increases, the added travel time increases the cost of both time-intensive and income-intensive maternal inputs.[11] If a woman's opportunity cost of time is low, she may be willing to travel further to obtain care. Post-NREGS however, each additional kilometer of travel required correspondingly increases the earnings a woman must forgo to receive care. Thus we expect that increased distance will negatively impact the effect of NREGS on the provision of maternal inputs that require travel to a healthcare facility: prenatal care and institutional delivery.

Table 7 reports results from the estimation of Equation 15, how distance to a healthcare facility moderates the effect of NREGS on the utilization of health inputs that depend on the accessibility of care: the timing of the first prenatal visit (Column 1), the number of prenatal visits (Column 2), the probability of an institutional delivery (Column 3), and the probability of delivery in a private institution (Column 4). The distance measure is an indicator variable that is equal to 1 if the nearest healthcare facility is greater than 5 kilometers from the village.[12] All specifications account for child, mother, father, and household characteristics, as well as month, year, and district

[11] Kesterton et al. (2010) and Magadi et al. (2000) find that greater distances are associated with significantly reduced rates of prenatal care and institutional delivery in developing countries.

[12] We consider different distance cutoffs, as well we as continuous and categorical measures of distance; the results are consistent with those presented in Table 7.

fixed effects.[13]

In all columns of Table 7, the estimated impact of NREGS on each health input is similar in magnitude (and statistical significance) to the effects estimated in Tables 3-6. However, the interaction term (NREGS*distance) is not significant in any specification. The evidence suggests that accessibility of care has little impact on the utilization of health inputs in NREGS districts.

### 6.4 Impact of Maternal Employment on Child Health

In Sections 6.1-6.2 we found that increasing maternal employment opportunities results in women decreasing their use of time-intensive inputs and increasing utilization of money-intensive inputs. We now consider how this increased maternal employment, proxied for by NREGS, affects child health, measured by neonatal and infant mortality. As discussed in Section 3, *a priori* the impact of increasing maternal employment on child health outcomes is ambiguous: a reduction in time-intensive health inputs could negatively impact health, while increasing the utilization of money-intensive inputs likely improves health.

Figures 8(a) and 8(b) respectively depict trends in the neonatal and infant mortality rates before and after NREGS was implemented. The parallel trends assumption across treated and untreated areas is validated by the figures, but there is little difference in the mortality rates post-NREGS. Indeed the naïve differences-in-differences estimate is zero percentage points. This is consistent with the regression estimates of Table 8 (Columns 1 and 2).

Further, there is little evidence that household composition moderates the impact of NREGS on neonatal or infant mortality, as shown in Table 9. We consider three measures of substitutes for maternal care: (1) Share of females older than 15 (Columns 1 and 4); (2) Share of females older than 18 (Columns 2 and 5); and (3) Share of females ages 15-18 (Columns 4 and 6).

Though the estimates in Tables 8 and 9 are not significant, this does not imply that altering the utilization of health inputs does not affect child health. The lack of precision is likely due to the fact that mortality is a difficult measure to impact; it requires a large sample size and a large effect size. Our estimated effects on maternal inputs and the sample size (in this context) are small, reducing the power of the regressions specified by Equations 16 and 17. If the data were available, we would use weight at birth as another measure of child health.

[13]We also control for the timing of the first prenatal visit in Column 2.

### 6.5 Robustness to Non-linear Specifications

Thus far, for ease of interpretation, we have considered linear models. In this section we consider the robustness of our results to non-linear specifications. We note that in the non-linear context, the common trends assumption is specified differently (Lechner, 2011). Further, the estimate of the interaction (DID) term cannot be interpreted as the treatment effect (Puhani, 2008; Ai and Norton, 2003). In this context, our goal with these alternate specifications is to verify that the predicted change in outcomes due to NREGS is consistent with the estimates from our linear models. All of the results from the non-linear models are similar in sign, magnitude, and statistical significance with the estimates from our linear specifications. There is strong evidence that NREGS, by increasing maternal employment opportunities, reduces the investment in time-intensive health inputs, subsequently allowing women to use their earnings to purchase better healthcare.

#### 6.5.1 Time-intensive Inputs

Prenatal Care

We estimate the impact of NREGS on the timing of the onset of prenatal care using a a survival model: the longer a woman survives in the data, the greater the delay in receiving prenatal care. We first consider the difference in survival times non-parametrically: Figure 9 displays the Kaplan-Meier survival curves across treated and untreated districts. We see the survival curve for women in NREGS districts lies above the survival curve for women in non-NREGS districts; NREGS women take longer to acquire their first prenatal visit. A Cox regression-based test, stratified at the district level, rejects the null hypothesis for equality of the survival curves (p-value < 0.001).[14]

We refine the estimates from the Kaplan-Meier plot and the stratified test with a semi-parametric Cox proportional hazards model, to account for time trends and control for other covariates:

$$h\left(t|\boldsymbol{x}_j\right) = h_0(t)\exp\left(\boldsymbol{x}_j\boldsymbol{\beta}_x\right) \tag{18}$$

[14] As we use the survey probability weights to allow for population inference, we cannot implement the default log-rank test, and implement Stata's Cox regression test instead. The major difference is that the log-rank test is a score test that reports the likelihood ratio statistic and the Cox test allows for use of sample weights and conducts a Wald test based on a robust variance estimator. The stratified test performs the calculations separately for each stratum. See `http://www.stata.com/manuals13/ststtest.pdf` for more details.

This model has the advantage of making no assumptions about the functional form of the baseline hazard $h_0(t)$, which, if misspecified, will result in biased estimates of $\boldsymbol{\beta}_x$. Just as we allowed for distinct district intercepts in the linear models, we implement a stratified Cox model, where the baseline hazard differs for each district $d$[15]:

$$h\left(t|\boldsymbol{x}_j\right) = h_{0d}(t)\exp\left(\boldsymbol{x}_j\boldsymbol{\beta}_x\right) \tag{19}$$

Table 10 reports the impact of NREGS on the hazard of receiving the first prenatal care visit (Column 1). The coefficient (0.98) is statistically significant, indicating that women in NREGS districts face a hazard rate for the first prenatal visit that is 2 percent smaller than the hazard for women in non-NREGS districts. These non-parametric and semi-parametric results are consistent both in sign and statistical significance with the result from estimation of the linear model in Table 3.

With the total number of prenatal care visits, we have a non-negative and discrete outcome, and so estimate a Poisson model. Consistent with the results from Table 4, there is a statistically significant and negative impact of NREGS on the incidence-rate ratio (IRR) of prenatal care visits (Column 2 of Table 10). The net effect of NREGS, a decrease of 0.04 visits, is also similar to our previous result. In addition, we test for any over-dispersion via an auxiliary regression (Cameron and Trivedi, 2010); the results are reported at the bottom of Column 2 in Table 10.[16] The degree of over-dispersion in Column 2 is minimal (-0.05), and not statistically different from zero.

Breastfeeding Duration

To estimate the effect of NREGS on the time a mother spent breastfeeding her youngest surviving child, we use the survival framework again. In this case, the longer the survival time, the longer the duration of breastfeeding. Figure 10 presents the non-parametric results: the Kaplan-Meier survival curves across treated and untreated districts. The survival curve for women in NREGS districts lies below the survival curve for women in non-NREGS districts, indicating that

[15] Although the Cox model does not impose a functional form, it does assume a constant hazard rate. Since the hazard of obtaining prenatal care may increase over time, we are also working to properly specify a Weibull model, which allows for a monotonically increasing or decreasing hazard.

[16] The Poisson model assumes the mean and variance are equal (there is no over-dispersion). Formally, we consider $\mathrm{Var}\left(y|\mathbf{X}\right) = \mathbb{E}(y|\mathbf{X}) + \alpha^2\mathbb{E}(y|\mathbf{X})$, the variance under the negative binomial model. We test the null hypothesis that $\alpha \neq 0$.

women in program areas do decrease the duration of breastfeeding. A Cox regression-based test, stratified at the district level, indicates this difference in duration is significant, rejecting the null hypothesis for equality of the survival curves (p-value < 0.001).

Column 3 of Table 10 reports the estimates from a semi-parametric Cox model estimating the impact of NREGS on the hazard of breastfeeding.[17] The coefficient of 1.07 indicates that women in treatment districts face a hazard rate for breastfeeding that is approximately 10 percent greater than the hazard faced by women in untreated districts. The results from the non-parametric and semi-parametric models are consistent in sign and statistical significance with the results of the linear model presented in Table 5. We also conduct a non-linear version (logit) of the colostrum falsification test discussed in Section 6.1.2. The results are consistent with Table 5: Column 4 of Table 10 indicates the propensity of women to feed infants colostrum is not correlated with the implementation of NREGS.

#### 6.5.2 Money-intensive Inputs

Institutional Delivery

To identify the impact of NREGS on rates of institutional delivery for the most recent pregnancy, we estimate a logit model; Column 5 of Table 10 reports these results. Consistent with the results in Table 6, we find there is a statistically and substantially significant effect of the program on the propensity of women to deliver in either a government or private healthcare facility: an increase of 1.6 percentage points.

### 6.6 Falsification Tests

Finally, we subject the estimates from our linear models to a falsification test. As NREGS was implemented only in rural areas, there should be no program effect in urban India. Thus we re-estimate 14, comparing the same treated and untreated districts, but limiting the regression sample to urban areas.

Table 11 presents the results from this falsification test. With respect to the investment in time-intensive inputs, the results are as expected: there is no impact of NREGS on the utilization of

[17] In this case, the hazard may decrease over time. Again, we are also considering a Weibull specification.

prenatal care (Columns 1 and 2), or on the propensity and duration of breastfeeding (Columns 3 and 4) in urban areas. However, there is a positive and statistically significant of NREGS on the urban institutional delivery rate (Column 5): the program is associated with an increase of 2 percentage points.

It may be that our results for institutional delivery are somewhat fragile. Another possibility however, may be related to NREGS; hence the observed effect. Ravi et al. (2012), utilizing the Harris-Todaro framework, provide evidence that NREGS has reduced rural-urban migration. As employment opportunities in rural areas increased, rural citizens no longer needed to migrate to cities for employment. Consequently, urban unemployment rates also fell after NREGS was implemented. Our result in Column 5 of Table 11 may be reflecting this phenomenon. If NREGS stemmed the migration of individuals who would have been under- or unemployed, and would not been able to afford an institutional delivery, then our estimate is simply reflecting this change in the population composition.

## 7 Conclusion

We use a nationally representative dataset to test if women, in response to the increased employment opportunities provided by India's public works program NREGS, change the mix of child health inputs utilized, and how this impacts child welfare. Consistent with the predictions of the maternal labor supply model, women reduce investment in time-intensive inputs (prenatal care and breastfeeding) in favor of employment. Facing greater earned incomes, mothers subsequently increase their utilization of money-intensive health inputs (institutional delivery). The net effect on child health is unclear. These results are robust to alternate specifications. In ongoing work, we attempt to estimate the impact of NREGS on alternative measures of child health (e.g., sickness due to diarrhea, fever, and cough). We are also interested in understanding if NREGS has a lasting impact on child investments, e.g. education. Afridi et al. (2012) consider this issue in only one state, but given our national results on child health, it is crucial to understand subsequent investments on a larger scale.

The effects we discuss above are distributed over all women in NREGS districts: the intent-to-treat effect. To understand the impact on treated women, we scale the estimates by the probability a woman enters the labor force post-program, which is 5 percent (Azam, 2012). Thus after the implementation of NREGS, treated women delay their first prenatal visit by 1.5 months, reduce the total number of visits by 1, increase the probability of institutional delivery by 20 percentage points, and reduce the time they spend breastfeeding by 3.2 months. While these are basic extrapolations, they give a better sense of how women who participate in NREGS might alter their utilization of health inputs in response to employment opportunities.

Our findings have broad policy implications. First, as women are willing to forgo care-giving at home in favor of using their earnings to purchase health services, this suggests that such services should be subsidized (further) for low-income women. There is a clear need for affordable, high-quality healthcare amongst low-income women, at least in India. Second, to ensure women are still accessing time-intensive health inputs, policymakers can consider ways to reduce the time cost of such services. Mobile health units providing basic prenatal care are a possibility, and could serve many remote villages inexpensively. Third, and perhaps most importantly, we provide evidence that public works programs can have effects beyond increasing labor force participation. Thus a comprehensive public works policy might provide mobile health units at work sites, in addition to drawing women into the labor force, ultimately allowing them to determine their own fortunes.

## References


Afridi, Farzana, Abhiroop Mukhopadhyay, Soham Shaoo (2012). Female Labour Force Participation and Child Education in India: The Effect of the National Rural Employment Guarantee Scheme. Discussion Paper No. 6593, Institute for the Study of Labor (IZA).

Ai, Chunrong, Edward C. Norton (2003). Interaction terms in logit and probit models. *Economics Letters*, 80(2003): 123-129.

Azam, Mehtabul (2012). The Impact of Indian Job Guarantee Scheme on Labor Market Outcomes: Evidence from a Natural Experiment. Discussion Paper No. 6548, Institute for the Study of Labor (IZA).

Anandalakshmy, S. (ed.) (1994). The Girl Child and the Family: An Action Study. Department of Women and Child Development, Ministry of Human Resource Development, Government of India, New Delhi.

Anderson, Patricia M, Kristin F. Butcher, Philip B. Levine (2003). Maternal Employment and Overweight Children. *Journal of Health Economics*, 22(3): 477-504.

Arifeen S, Robert E. Black, Gretchen Antelman, Abdullah Baqui, Laura Caulfield, Stan Becker (2001). Exclusive breastfeeding reduces acute respiratory infection and diarrhea deaths among infants in Dhaka slums. *Pediatrics*, 108(4): E67.

Bajpai, Nirupam, Ravindra Dholakia, Megan Towle (2013). Increasing the Availability of Skilled Birth Attendance in Rural India. Working Paper No. 9, Columbia Global Centers, South Asia, Columbia University.

Baker, Michael, Kevin Milligan (2008). Maternal Employment, Breastfeeding and Health: Evidence from Maternity Leave Mandates. *Journal of Health Economics*, 27(4): 871-887.

Becker, Gary (1965). A Theory of the Allocation of Time. *Economic Journal*, 75(299): 493-517.

Berniell, María Inés and Carolina Sánchez-Páramo (2011). Closing the Access Gap: Recent Advances in Female Labor Force Participation. Background paper for the World Development Report 2012.

Bhutta, Zulfiqar A., Gary L. Darmstadt, Babar S. Hasan, Rachel A. Haws (2005). Community-Based Interventions for Improving Perinatal and Neonatal Health Outcomes in Developing Countries: A Review of the Evidence. *Pediatrics*, 115(12): 519-617.

Blau, Francine, Adam Grossberg (1992). Maternal Labor Supply and Children's Cognitive Development. *Review of Economics and Statistics*, 74(3): 474-481.

Cameron, A. Colin, Pravin K. Trivedi. *Microeconometrics Using Stata*. College Station, Texas: Stata Press, 2009.

Costello, Anthony M. de L., David Orsin (2003). Micronutrient Status during Pregnancy and Outcomes for Newborn Infants in Developing Countries. *Journal of Nutrition*, 133(5): 1757S-1764S.
Das, Sanjit Kumar (2010). Expansion of Micro-financing through Swarnajayanti Gram Swarojgar Yojana: Experience in West Bengal. *Economic Affairs*, 55(2): 180-186.

Dasgupta, Sukti, Ratna Sudarshan (2011). Issue in Labor Market Inequality and Women's Participation in India's National Rural Employment Guarantee Program. International Labor Office, Geneva, Working Paper 98.

Desai, Sonalde, P. Lindsay Chase-Lansdale, Robert T. Michael (1989). Mother or Market? The Effects of Maternal Employment on the Intellectual Ability of 4-Year-Old Children. *Demography*, 26(4): 546-561.

Doss, Cheryl R. (2006). The Effects of Intrahousehold Property Ownership on Expenditure Patterns in Ghana. *Journal of African Economies*, 15(1): 149-180.

Duflo, Esther (2012). Women's Empowerment and Economic Development. *Journal of Economic Literature*, 50(4): 1051-1079.

Evans, Judith L. (1995). Men in the Lives of Children. *Coordinator's Notebook*, 16: 1-16.

Fox, John. *Applied Regression Analysis and Generalized Linear Models*. Thousand Oaks, California: Sage Publications, 2008.

Gilligan, D., John Hoddinott & Alemayehu Seyoum Taffesse (2009). The Impact of Ethiopia's Productive Safety Net Programme and its Linkages. *The Journal of Development Studies*, 45(10): 1684-1706.

Glick, Peter, David Sahn (1998). Maternal Labor Supply and Child Nutrition in West Africa. *Oxford Bulletin of Economics and Statistics*, 60(3): 325-355.

Gordon, Rachel, Robert Kaestner, and Sanders Korenman (2007). The Effects of Maternal Employment on Child Injuries and Infectious Disease. *Demography*, 44(2): 307-333.

Gupta, Arun, Kuldeep Khanna (1999). Economic value of breastfeeding in India. *The National Medical Journal of India*, 12(3): 123-127.

Handa, S. John A. Maluccio (2010). Matching the gold standard: Comparing experimental and non-experimental evaluation techniques for a geographically targeted program. *Economic Development and Cultural Change*, 58(3): 415-447.

Heckman, James, Hidehiko Ichimura, Petra E. Todd (1997). Matching as an Econometric Evaluation Estimator: Evidence from Evaluating a Job Training Program. *Review of Economic Studies* 64(4): 605-654.

Heckman, James, Hidehiko Ichimura, Jeffrey Smith, Petra Todd (1998). Characterizing selection bias using experimental data. *Econometrica*, 66(5): 1017-1098.

International Institute for Population Sciences (IIPS) (2010). *District Level Household and Facility Survey (DLHS-3), 2007-08: India*. Mumbai: IIPS.

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

Jowett, Matthew (2000). Safe motherhood interventions in low-income countries: an economic justification and evidence of cost-effectiveness. *Health Policy*, 53(3): 201-228.

Kesterton, Amy J., John Cleland, Andy Sloggett, Carine Ronsmans (2010). Institutional delivery in rural India: the relative importance of accessibility and economic status. *BMC Pregnancy and Childbirth*, 10(30).

Khera, Reetika, Nandini Nayak (2009). Women workers and perceptions of the National Rural Employment Guarantee Act. *Economic and Political Weekly*, 44(43): 49-57.

Kolenikov, Stanislav (2010). Resampling variance estimation for complex survey data. *The Stata Journal*, 10(2): 165-199.

Lara, George J., Thomas W. Pullum (2005). Infant Mortality in Egypt: Exploring the Role of Prenatal Care, and Implications for Public Policy. Presented at Population Association of America, 2005: Philadelphia, PA.

Lechner, Michael (2011). The Estimation of Causal Effects by Difference-in-Difference Methods. Discussion Paper No. 2010-28, Universität St. Gallen.

Magadi, Monica Akinyi, Nyovani Janet Madise, Roberto Nascimento Rodrigues (2000). Frequency and timing of antenatal care in Kenya: explaining the variations between women of different communities. *Social Science and Medicine*, 51: 551-561.

Ministry of Rural Development (2008). The National Rural Employment Guarantee Act 2005 (NREGA), Operational Guidelines, $3^{rd}$ edition.

Mohanty, Sanjay K., Akanksha Srivastava (2013). Out-of-pocket expenditure on institutional delivery in India. *Health Policy and Planning*, 28(3): 247-262.

OECD(2011). *Society at a Glance 2011 - Asia/Pacific Edition*, OECD Publishing.

Office of the Surgeon General (US), Centers for Disease Control and Prevention (US), Office on Women's Health (US) (2011). The Surgeon General's Call to Action to Support Breastfeeding. Rockville (MD): Office of the Surgeon General (US), 2011. The Importance of Breastfeeding. Online: `http://www.ncbi.nlm.nih.gov/books/NBK52687/`. Accessed March 27, 2014.

Orsin, David, Kirti M Tumbahangphe, Dej Shrestha, Natasha Mesko, Bhim P Shrestha, Madan K Manandhar, Hilary Standing, Dharma S Manandhar, Anthony M de L Costello (2002). *BMJ*. 325(1063).

Poma, P.A. (1999). Effect of prenatal care on infant mortality rates according to birth-death certificate files. *Journal of the National Medical Association*, 91(9): 515-520.

Puhani, Patrick A. (2008). The Treatment Effect, the Cross Differences, and the Interaction Term in Nonlinear "Differences-in-Differences" Models. Discussion Paper No. 2478, Institute for the Study of Labor (IZA).

Ravi, Shamika, Monika Engler (2009). Workfare in Low Income Countries: An effective way to fight poverty? The case of India's NREGS. Working paper.

Ravi, Shamika, Kapoor, Mudit, Ahluwalia, Rahul (2012). The Impact of NREGS on Urbanization in India. Presented at the Northeast Universities Development Consortium Conference, 2012.

Rosenbaum, Paul R., Donald B. Rubin (1983). The Central Role of the Propensity Score in Observational Studies for Causal Effects. *Biometrika*, 70(1): 41-55.

Rubalcava, Luis, Graciela Teruel, Duncan Thomas (2009). Investments, Time Preferences, and Public Transfers Paid to Women. *Economic Development and Cultural Change*, 57(3): 507-538.

Shah, Anup (2013). Poverty Facts and Stats. Global Issues (Website). Accessed August 22, 2013.

Sivakami, M. (2010). Are poor working mothers in South India investing less time in the next generation? *Asian Population Studies*, 6(1): 99-120.

StataCorp. 2013. *Stata Statistical Software: Release 13*. College Station, TX: StataCorp LP.

Subbaro, Kalanidhi, Carlo del Ninno, Colin Andrews, Claudia Rodriguez-Alas (2013). Public Works as a Safety Net: Design, Evidence and Implementation. Washington, D.C.: World Bank.

Smith, Jeffrey A., Petra E. Todd (2005): Does Matching Overcome Lalonde's Critique of Nonexperimental Estimators? *Journal of Econometrics*, 125(1-2): 305-353.

Thomas, Duncan (1990). Intra-household Resource Allocation: An Inferential Approach. *Journal of Human Resources*, 25(4): 635-664.

UNFPA, Ifpea (2007). Potential Contributions to the MDG Agenda from the Perspective of the ICPD: A Reference Guide to Evidence for Policy Dialogue in the LAC Region. United Nations Population Fund. Online: `http://www.unfpa.org/public/publications/pid/2599`, accessed March 25, 2014.

Waldfogel, Jane, Wen-Jui Han, and Jeanne Brooks-Gunn (2002). The Effects of Early Maternal Employment on Child Cognitive Development. *Demography*, 39(2): 369-392.

WHO & UNICEF (2003). Antenatal Care in Developing Countries: Promises, achievement and missed opportunities: An analysis of trends, levels and differentials, 1990-2001. Department of Reproductive Health and Research, World Health Organization, Geneva.

World Bank (2014). World Development Indicators. Online: `http://data.worldbank.org/indicator/NY.GDP.PCAP.CD?page=1`, accessed March 25, 2014.

World Bank (2011). World Development Report 2012: Gender Equality and Development. Washington, D.C.: World Bank.

World Health Organization (2005). The World Health Report, 2005: Make every mother and child count. World Health Organization, Switzerland. Online: http://www.who.int/whr/2005/whr2005_en.pdf. Acccessed March 27, 2014.

# Theoretical Appendix

We provide proofs of the results derived in Section 3. First we totally differentiate Equations 5, 6, and 8:[18]

$$\underbrace{\begin{bmatrix} U_{hh} & U_{hm} & w \\ U_{mh} & U_{mm} & -p \\ w & -p & 0 \end{bmatrix}}_{\bar{H}} \begin{bmatrix} dh \\ dm \\ d\lambda \end{bmatrix} = \begin{bmatrix} \lambda & 0 & 0 \\ 0 & \lambda & 0 \\ -h & m & -dV \end{bmatrix} \begin{bmatrix} dw \\ dp \\ dV \end{bmatrix} \tag{20}$$

where $U_{xy} = \frac{\partial^2 U}{\partial x \partial y}$ and $\bar{H}$ is the bordered Hessian.

We can solve for $dh$ using Cramer's rule:

$$dh = \frac{\begin{vmatrix} \lambda dw & U_{hm} & w \\ \lambda dp & U_{mm} & -p \\ -hdw + mdp - dV & -p & 0 \end{vmatrix}}{|\bar{H}|} \tag{21}$$

Consider the total effect of a change in the wage $w$ on the number of hours $h$ a mother works, holding price and endowed income constant ($dp = dV = 0$). Equation 21 becomes:

$$dh = \frac{\begin{vmatrix} \lambda dw & U_{hm} & w \\ 0 & U_{mm} & -p \\ -hdw & -p & 0 \end{vmatrix}}{|\bar{H}|} = \frac{\lambda p^2 dw + hdw\,(pU_{hm} + wU_{mm})}{|\bar{H}|} \tag{22}$$

$$\Rightarrow \frac{dh}{dw} = \underbrace{\frac{\lambda p^2}{|\bar{H}|}}_{1} + \underbrace{h\frac{pU_{hm} + wU_{mm}}{|\bar{H}|}}_{2} \tag{23}$$

We need to determine the sign of Terms 1 and 2 in Equation 23.

Term 1

First consider the impact of a change in wage on hours, holding utility, or real income, constant.

[18] After imposing the time constraint and recognizing that $\frac{\partial U}{\partial t} = -\frac{\partial U}{\partial h}$.

Thus $dp = dV = 0$ and $hdw = 0$. Equation 21 becomes:

$$dh|_{U=\bar{U}} = \frac{\begin{vmatrix} \lambda dw & U_{hm} & w \\ 0 & U_{mm} & -p \\ 0 & -p & 0 \end{vmatrix}}{|\bar{H}|} = \frac{\lambda p^2 dw}{|\bar{H}|} \tag{24}$$

$$\Rightarrow \left.\frac{dh}{dw}\right|_{U=\bar{U}} = \underbrace{\frac{\lambda p^2}{|\bar{H}|}}_{1} > 0 \tag{25}$$

We see that Term 1 of Equation 23 is the substitution effect: the impact of wages on hours holding utility constant. Since we have assumed a well-behaved utility function, the determinant of the bordered Hessian, given by $|\bar{H}|$, must be positive. The numerator of Term 1 is also positive. Thus when holding real income and prices constant, women work more when wages increase.

Term 2

Now consider the impact on hours when wages and prices are constant, but non-labor income changes. That is, $dw = dp = 0$, but $dV$ is non-zero. Equation 21 is now:

$$dh = \frac{\begin{vmatrix} 0 & U_{hm} & w \\ 0 & U_{mm} & -p \\ -dV & -p & 0 \end{vmatrix}}{|\bar{H}|} = \frac{dV\left(pU_{hm} + wU_{mm}\right)}{|\bar{H}|} \tag{26}$$

$$\Rightarrow h\frac{dh}{dV} = \underbrace{h\frac{pU_{hm} + wU_{mm}}{|\bar{H}|}}_{2} < 0 \tag{27}$$

Thus Term 2 of Equation 23 is the income effect. Most likely the income effect is negative[19]: as wages increase, women face higher incomes, and so can afford to work fewer hours to purchase the same amount of goods and services. That is, working is an inferior good.

[19] Our assumptions about the utility function imply that $\bar{H} > 0$ and $U_{mm} < 0$. Term 2 will be positive if and only if $U_{hm} > 0$ and $pU_{hm} > |wU_{MM}|$.

Combining all terms together, Equation 23 can be written as:

$$\frac{dh\left(p,w,V\right)}{dw} = \underbrace{\left.\frac{dh}{dw}\right|_{U=\bar{U}}}_{1} + h\underbrace{\frac{dh}{d\left(wh+V\right)}}_{2} > 0 \tag{11 derived}$$

which is the result from Section 3. At low wages, such as those in NREGS, we expect the substitution effect (Term 1) to outweigh the income effect (Term 2), and the net effect of an increase in wages on hours worked will be positive. Equations 12 and 13 can be derived similarly.

# Data Appendix

## Sampling Design

The unit of analysis in all regressions is the individual (child or mother). To account for the sampling design of DLHS, we set the data as survey data, defining villages as the primary sampling units (PSUs) and districts as the strata. Doing so accounts for the hierarchical nature of the data and improves the estimation of the standard errors (Cameron and Trivedi, 2009). All regressions use the sampling weights, allowing for population inference.

## District Borders

During the years in which NREGS was implemented, many Indian districts merged or split, altering district borders. To track consistent geographical units over time, we converted district borders to match the 2001 Indian Census. In 2001, India had 593 districts; DLHS was administered in 584 of these districts. We dropped 28 districts where the new district and the original district received NREGS in different phases. We dropped another 8 districts where NREGS was not administered. Thus our core sample consisted of 548 districts.

## Variable Definition and Management

The sample is limited to women who gave birth at least once since January 1, 2004 living in rural areas (where NREGS was implemented): 185,000 mothers and 247,909 children.

### Prenatal Care

Month of first prenatal care visit: The survey question referred to the most recent pregnancy since January 1, 2004, that resulted in a live or still birth. The variable is defined as the month of the pregnancy in which the woman first received a prenatal check-up.

Number of visits: The survey question referred to the most recent pregnancy since January 1, 2004, if this pregnancy resulted in a live or still birth. The variable is defined as the number of prenatal care check-ups a woman received for this pregnancy. It is coded as (0) if the woman did not receive any prenatal care.

**Breastfeeding**

Initiation: The survey question referred to the youngest surviving child for the most recent pregnancy since January 1, 2004. The variable is coded as (1) if the woman initiated breastfeeding (at any time) for the child; it is coded as (0) if the woman did not breastfeed at all.

Duration: The survey question referred to the youngest surviving child for the most recent pregnancy since January 1, 2004. The variable is defined as the number of months the woman exclusively breastfed her child.

Colostrum: The survey question referred to the youngest surviving child for the most recent pregnancy since January 1, 2004. It is coded as (1) if the woman fed the child colostrum, which is secreted immediately after childbirth for a few days, and (0) otherwise.

**Delivery**

Institutional delivery: The survey question referred to the most recent pregnancy since January 1, 2004, that resulted in a live or still birth. The variable is defined as the delivery location for this last pregnancy. It is coded as (1) if the delivery was in a healthcare facility (government or private), and (0) otherwise.

Private facility delivery: The survey question referred to the most recent pregnancy since January 1, 2004, that resulted in a live or still birth. The variable is defined for women whose last delivery took place in a healthcare facility. It is coded as (1) if the delivery was in private facility, and (0) if the delivery was in a public facility; births in other locations were not included in this sample.

**Child Mortality**

Neonatal mortality: The survey question referred to all children born since January 1, 2004. The variable is coded as (1) if the child died within 28 days of birth, and (0) otherwise.

Infant mortality: The survey question referred to all children born since January 1, 2004. The variable is coded as (1) if the child died within 12 months of birth, and (0) otherwise.

**Moderators**

Distance to health facility: We define an indicator to represent the distance to the nearest health facility (government or private). The variable is coded as follows: (0) if a health facility is located within the village or within 5 kilometers; and (1) if the distance to the health facility is more than

5 kilometers.

Household composition: We define three variables to capture substitutes for maternal care: (1) *Female adult share* is the share of female household members aged 18 or above; (2) *Female adult and young adult share* is the share of female household members aged 15 or above; and (3) *Young female share* is the share of female household members aged 15 to 18.

**Child Controls**

Age: The child's age in months.

Gender: The child's gender, coded as (1) if the child is male, (0) if the child is female.

**Mother Controls**

Age at birth: The mother's age (in years) at the time of birth for a particular child.

Marital status: The mother's marital status. It is coded as (1) if the mother reported being married, and (0) otherwise (e.g., widowed or separated).

Education: The highest grade/standard the mother reported as having passed.

Number of pregnancies, pre-2004: The total number of pregnancies the mother reported, prior to 2004.

Number of births, pre-2004: The total number of births (live and still) the mother reported, prior to 2004.

**Household Controls**

Father's education: The highest grade/standard the father reported as having passed. This is our best proxy for pre-program household income or socio-economic status. Adult share: The share of household members aged 18 or above.

Religion: The religion of the household, coded as (1) if the household reported being Hindu, and (0) otherwise.

Caste: This variable captures membership in historically oppressed groups. It is coded as (1) if the household reported being in a Scheduled Caste; (2) if the household reported being in a Scheduled Tribe, and (0) otherwise.

**Village Controls**

Village development index: We create an index, ranging from 0 to 11, that measures a village's development, e.g. if it has a health facility, a school, etc. Each village receives a score equal to the number of services available within the village.

Village exposure index: We create an index, ranging from 0 to 19, that measures a village's exposure to other government programs. Each village receives a score equal to the number of programs with beneficiaries in the village.

**Table 1**
**Sample characteristics across NREGS Phases 1, 2, and 3**

| | Phase 1 | | Phase 2 | | Phase 3 | |
|---|---|---|---|---|---|---|
| | Mean | (Std. dev.) | Mean | (Std. dev.) | Mean | (Std. dev.) |
| **District** | | | | | | |
| Household expenditure (Rs) | 2,476 | (630) | 2,945 | (794) | 6,503 | (48,640) |
| Male literacy rate | 69.54 | (10.94) | 72.93 | (11.04) | 78.32 | (9.61) |
| Female literacy rate | 44.29 | (12.79) | 50.74 | (15.5) | 57.68 | (13.86) |
| Workforce participation rate | 42.10 | (6.71) | 40.01 | (6.95) | 40.34 | (7.14) |
| *N (Districts)* | *187* | | *122* | | 259 | |
| **Household** | | | | | | |
| Hindu | 0.80 | (0.40) | 0.75 | (0.43) | 0.73 | (0.44) |
| No. members | 5.28 | (2.51) | 5.37 | (2.68) | 5.26 | (2.52) |
| Hold poverty card | 0.39 | (0.49) | 0.34 | (0.47) | 0.28 | (0.45) |
| *N (Households)* | *67,257* | | *42,182* | | *63,586* | |
| **Mother** | | | | | | |
| Current age | 26.23 | (5.74) | 26.58 | (5.86) | 26.4 | (5.44) |
| Age at marriage | 16.69 | (3.27) | 16.72 | (3.81) | 18.02 | (3.7) |
| Highest grade passed | 3.07 | (4.02) | 3.43 | (4.16) | 4.68 | (4.43) |
| No. pregnancies, pre-2004 | 1.84 | (2.06) | 1.87 | (2.15) | 1.58 | (1.92) |
| No. births, pre-2004 | 1.68 | (1.92) | 1.69 | (2.01) | 1.4 | (1.77) |
| *N (Mothers)* | *71,616* | | *45,109* | | *68,071* | |

Table 1 compares district, household, and mother characteristics across the three phases of NREGS. As depicted in Figure 3, Phase 1 was implemented in February 2006, Phase 2 in April 2007, and Phase 3 in April 2008, with least developed districts receiving the program first. Phase 1 and 2 districts have lower household expenditures and literacy rates compared to Phase 3. However, rates of labor force participation are similar across all three phases. Households are of similar size across the phases, but the rate of poverty is lowest in Phase 3 districts (28 percent of households hold a poverty card, compared to 39 percent in Phase 1 and 34 percent in Phase 2). The differences across the phases are apparent when comparing mother characteristics. Though mothers are of similar age in all phases, those in Phase 1 and 2 districts get married one year earlier compared to Phase 3 districts. Mothers in Phase 3 districts have 1.5 more years of education than in Phase 1 districts. Phase 1 and 2 mothers have experienced a greater number of pregnancies/births than in Phase 3. We address possible selection issues due to this non-random rollout in our empirical strategy described in Section 5.

**Table 2**
**Comparing outcomes and moderators across treated and untreated districts**

| | Treated | Untreated | Difference |
|---|---|---|---|
| **Time-intensive inputs** | | | |
| Months to first prenatal visit | 3.85 | 3.59 | 0.26*** |
| Total number of prenatal visits | 3.36 | 3.81 | -0.45*** |
| Breastfeeding duration (months) | 2.89 | 3.14 | -0.25*** |
| **Money-intensive inputs** | | | |
| Institutional delivery rate | 37.39 | 41.09 | -3.7*** |
| **Child health measures** | | | |
| Neonatal mortality rate | 2.97 | 2.51 | 0.45*** |
| Infant mortality rate | 4.19 | 3.71 | 0.48*** |
| **Moderators** | | | |
| Distance to health facility (km) | 1.2 | 1.24 | -.03 |
| Household composition: Females $\geq$ 18 | 30.51 | 30.61 | -.11 |
| Household composition: Females $\geq$ 15 | 33.21 | 33.31 | -.1 |
| Household composition: Females 15-18 | 4.04 | 3.95 | .09 |

Table 2 shows that women in treated (NREGS) districts delay prenatal care, receive fewer prenatal visits, exclusively breastfeed for fewer months, and are less likely to deliver in an institution than women in untreated districts (these differences are significant at a p-value < 0.01). The table also shows that neonatal and infant mortality rates are significantly higher in NREGS districts than in untreated areas. However, in the bottom panel we see that treated and untreated districts are similar with respect to accessibility of care (distance to health facility) and household composition.

**Table 3**
**Impact of NREGS on timing of first prenatal visit**

| | (1) Month of Visit | (2) ≤1 Mth | (3) ≤2 Mths | (4) ≤3 Mths | (5) ≤4 Mths | (6) ≤5 Mths | (7) ≤6 Mths | (8) ≤7 Mths | (9) ≤8 Mths | (10) ≤9 Mths |
|---|---|---|---|---|---|---|---|---|---|---|
| **NREGS** | 0.07*** | -0.00 | -0.00 | 0.00 | 0.00 | 0.00 | 0.01 | 0.01** | 0.01** | 0.01** |
| | (0.022) | (0.002) | (0.004) | (0.005) | (0.005) | (0.005) | (0.005) | (0.005) | (0.005) | (0.005) |
| Child gender = male | 0.04*** | -0.00*** | -0.00 | 0.00 | 0.00 | 0.00 | 0.00** | 0.01*** | 0.01*** | 0.01*** |
| | (0.010) | (0.001) | (0.002) | (0.002) | (0.002) | (0.002) | (0.002) | (0.002) | (0.002) | (0.002) |
| Mother's age at birth | -0.00 | 0.00*** | 0.00*** | 0.00*** | 0.00*** | 0.00*** | 0.00*** | 0.00*** | 0.00*** | 0.00*** |
| | (0.001) | (0.000) | (0.000) | (0.000) | (0.000) | (0.000) | (0.000) | (0.000) | (0.000) | (0.000) |
| Mother married | 0.10** | 0.00 | 0.03*** | 0.04*** | 0.06*** | 0.06*** | 0.06*** | 0.05*** | 0.05*** | 0.05*** |
| | (0.045) | (0.003) | (0.007) | (0.010) | (0.010) | (0.010) | (0.010) | (0.010) | (0.009) | (0.009) |
| Yrs. education, mother | 0.03*** | 0.00*** | 0.01*** | 0.02*** | 0.02*** | 0.02*** | 0.02*** | 0.02*** | 0.02*** | 0.02*** |
| | (0.002) | (0.000) | (0.000) | (0.000) | (0.000) | (0.000) | (0.000) | (0.000) | (0.000) | (0.000) |
| No. pregnancies, pre-2004 | 0.06*** | 0.00*** | 0.01*** | 0.01*** | 0.01*** | 0.02*** | 0.02*** | 0.02*** | 0.02*** | 0.02*** |
| | (0.010) | (0.001) | (0.002) | (0.002) | (0.002) | (0.002) | (0.002) | (0.002) | (0.002) | (0.002) |
| No. births, pre-2004 | -0.13*** | -0.01*** | -0.02*** | -0.04*** | -0.04*** | -0.05*** | -0.05*** | -0.04*** | -0.05*** | -0.05*** |
| | (0.011) | (0.001) | (0.002) | (0.002) | (0.002) | (0.002) | (0.002) | (0.002) | (0.002) | (0.002) |
| Yrs. education, father | 0.03*** | 0.00*** | 0.00*** | 0.01*** | 0.01*** | 0.01*** | 0.01*** | 0.01*** | 0.01*** | 0.01*** |
| | (0.002) | (0.000) | (0.000) | (0.000) | (0.000) | (0.000) | (0.000) | (0.000) | (0.000) | (0.000) |
| Share of HH ≥18 yrs. | 0.02 | -0.00 | 0.00 | 0.00 | -0.01 | -0.00 | -0.00 | 0.00 | 0.00 | 0.00 |
| | (0.023) | (0.002) | (0.004) | (0.005) | (0.005) | (0.005) | (0.005) | (0.005) | (0.005) | (0.005) |
| Hindu | 0.06*** | -0.00 | -0.01** | -0.00 | 0.01 | 0.03*** | 0.02*** | 0.02*** | 0.01*** | 0.01*** |
| | (0.021) | (0.002) | (0.003) | (0.004) | (0.004) | (0.004) | (0.004) | (0.004) | (0.004) | (0.004) |
| Scheduled Caste | 0.01 | -0.00*** | -0.02*** | -0.03*** | -0.03*** | -0.01*** | -0.01*** | -0.01*** | -0.01*** | -0.01*** |
| | (0.015) | (0.001) | (0.002) | (0.003) | (0.003) | (0.003) | (0.003) | (0.003) | (0.003) | (0.003) |
| Scheduled Tribe | -0.20*** | -0.00*** | -0.03*** | -0.06*** | -0.06*** | -0.06*** | -0.07*** | -0.07*** | -0.07*** | -0.07*** |
| | (0.023) | (0.001) | (0.003) | (0.005) | (0.005) | (0.005) | (0.005) | (0.005) | (0.005) | (0.005) |
| Constant | 2.16*** | 0.01 | 0.05*** | 0.25*** | 0.36*** | 0.46*** | 0.49*** | 0.53*** | 0.54*** | 0.54*** |
| | (0.065) | (0.005) | (0.010) | (0.014) | (0.014) | (0.014) | (0.014) | (0.014) | (0.014) | (0.014) |
| N (Mothers) | 140,443 | 140,443 | 140,443 | 140,443 | 140,443 | 140,443 | 140,443 | 140,443 | 140,443 | 140,443 |
| R-squared | 0.134 | 0.118 | 0.166 | 0.219 | 0.238 | 0.256 | 0.253 | 0.258 | 0.257 | 0.257 |

Table 3 reports the results of the impact of NREGS on the timing of prenatal care, for a woman's most recent pregnancy. The treatment, NREGS status, is coded as follows: (1) for women who were exposed to the program for all 9 months of the pregnancy and (0) for women who were exposed for 0 months of the pregnancy. Women who were exposed to the program for some portion of the pregnancy are not included in the regression sample. In Column 1 the dependent variable is the month of the pregnancy in which a woman first received prenatal care. In Columns 2-10, the dependent variable is an indicator for having received the first prenatal visit within $x$ months of the pregnancy, where $x$ is given by the column heading. We find that NREGS delays the first visit by 0.07 months, or 2 days (Column 1). Columns 2-10 estimate a series of linear probability models to identify when in the pregnancy NREGS has the greatest impact. Consistent with a maternal labor supply model, the program increases the probability of obtaining the first prenatal care visit in months 7, 8, and 9 (the last trimester) by 1 percentage point, when the opportunity cost of time is lower. Notes: All specifications include year, month, and district fixed effects. Robust standard errors (in parentheses) are clustered at the village level. ** $p < 0.05$, *** $p < 0.01$.

**Table 4**
**Impact of NREGS on number of prenatal visits**

| | (1) Total Visits | (2) ≥1 Visit | (3) ≥2 Visits | (4) ≥3 Visits | (5) ≥4 Visits | (6) ≥5 Visits | (7) ≥6 Visits | (8) ≥7 Visits | (9) ≥8 Visits | (10) ≥9 Visits |
|---|---|---|---|---|---|---|---|---|---|---|
| **NREGS** | -0.05*** | -0.00 | -0.00 | -0.01** | -0.01 | -0.01 | -0.01** | -0.01** | -0.00 | -0.00** |
| | (0.017) | (0.000) | (0.003) | (0.004) | (0.004) | (0.003) | (0.003) | (0.003) | (0.002) | (0.002) |
| Child gender = male | 0.01 | -0.00 | 0.00 | -0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | -0.00 |
| | (0.008) | (0.000) | (0.001) | (0.002) | (0.002) | (0.002) | (0.001) | (0.001) | (0.001) | (0.001) |
| Mother's age at birth | 0.01*** | 0.00 | 0.00*** | 0.00*** | 0.00*** | 0.00*** | 0.00*** | 0.00*** | 0.00*** | 0.00*** |
| | (0.001) | (0.000) | (0.000) | (0.000) | (0.000) | (0.000) | (0.000) | (0.000) | (0.000) | (0.000) |
| Mother married | 0.08** | -0.00 | 0.01** | 0.02** | 0.01 | 0.01** | 0.01 | 0.01 | 0.00 | 0.00 |
| | (0.033) | (0.000) | (0.005) | (0.008) | (0.008) | (0.007) | (0.007) | (0.005) | (0.004) | (0.003) |
| Yrs. education, mother | 0.06*** | 0.00 | 0.00*** | 0.01*** | 0.01*** | 0.01*** | 0.01*** | 0.01*** | 0.00*** | 0.00*** |
| | (0.001) | (0.000) | (0.000) | (0.000) | (0.000) | (0.000) | (0.000) | (0.000) | (0.000) | (0.000) |
| No. pregnancies, pre-2004 | 0.07*** | 0.00 | 0.00 | 0.01*** | 0.01*** | 0.01*** | 0.01*** | 0.01*** | 0.01*** | 0.00*** |
| | (0.009) | (0.000) | (0.001) | (0.002) | (0.002) | (0.002) | (0.001) | (0.001) | (0.001) | (0.001) |
| No. births, pre-2004 | -0.11*** | -0.00 | -0.01*** | -0.01*** | -0.02*** | -0.02*** | -0.01*** | -0.01*** | -0.01*** | -0.01*** |
| | (0.009) | (0.000) | (0.001) | (0.002) | (0.002) | (0.002) | (0.002) | (0.001) | (0.001) | (0.001) |
| Yrs. education, father | 0.01*** | 0.00 | 0.00*** | 0.00*** | 0.00*** | 0.00*** | 0.00*** | 0.00*** | 0.00*** | 0.00*** |
| | (0.001) | (0.000) | (0.000) | (0.000) | (0.000) | (0.000) | (0.000) | (0.000) | (0.000) | (0.000) |
| Share of HH ≥18 yrs. | 0.02 | 0.00 | 0.00 | -0.00 | 0.01 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 |
| | (0.018) | (0.000) | (0.003) | (0.004) | (0.004) | (0.004) | (0.003) | (0.003) | (0.002) | (0.002) |
| Hindu | 0.02 | 0.00 | -0.00 | -0.00 | 0.01 | 0.01** | 0.01** | 0.00** | 0.00 | -0.00 |
| | (0.015) | (0.000) | (0.002) | (0.003) | (0.003) | (0.003) | (0.003) | (0.002) | (0.002) | (0.001) |
| Scheduled Caste | -0.09*** | -0.00 | -0.00 | -0.01*** | -0.02*** | -0.02*** | -0.01*** | -0.01*** | -0.01*** | -0.00*** |
| | (0.011) | (0.000) | (0.002) | (0.003) | (0.002) | (0.002) | (0.002) | (0.002) | (0.001) | (0.001) |
| Scheduled Tribe | -0.18*** | -0.00 | -0.01*** | -0.03*** | -0.04*** | -0.03*** | -0.03*** | -0.02*** | -0.01*** | -0.00*** |
| | (0.016) | (0.000) | (0.002) | (0.004) | (0.004) | (0.003) | (0.003) | (0.002) | (0.002) | (0.001) |
| Constant | -0.07 | -0.00 | -0.02*** | 0.01 | -0.01 | -0.02 | -0.01 | -0.01 | -0.00 | 0.00 |
| | (0.049) | (0.000) | (0.007) | (0.011) | (0.011) | (0.010) | (0.009) | (0.008) | (0.006) | (0.005) |
| N (Mothers) | 140,127 | 140,127 | 140,127 | 140,127 | 140,127 | 140,127 | 140,127 | 140,127 | 140,127 | 140,127 |
| R-squared | 0.689 | 1.000 | 0.782 | 0.497 | 0.458 | 0.431 | 0.397 | 0.363 | 0.358 | 0.355 |

Table 4 reports the results of the impact of NREGS on the number of prenatal care visits, for a woman's most recent pregnancy. The treatment, NREGS status, is coded as follows: (1) for women who were exposed to the program for all 9 months of the pregnancy and (0) for women who were exposed for 0 months of the pregnancy. Women who were exposed to the program for some portion of the pregnancy are not included in the regression sample. In Column 1 the dependent variable is the total number visits. In Columns 2-10, the dependent variable is an indicator for having received at least $x$ visits, where $x$ is given by the column heading. We find that NREGS reduces the total number of visits by 0.05 visits (Column 1). Columns 2-10 estimate a series of linear probability models to identify where in the prenatal care trajectory NREGS has the greatest impact. Consistent with a model of maternal labor supply, the program reduces the probability of obtaining at least 6 or 7 visits by 1 percentage point, with smaller effects at 8 and 9 visits, likely because the marginal cost of time increases with each additional visit. Notes: All specifications include year, month, and district fixed effects, and control for the timing of the first prenatal visit. Robust standard errors (in parentheses) are clustered at the village level. $^{**}$ $p < 0.05$, $^{***}$ $p < 0.01$.

# Table 5
## Impact of NREGS on breastfeeding

| | (1) Initiation | (2) Duration (Months) | (3) $\geq$0.5 Mths | (4) $\geq$1 Mth | (5) $\geq$2 Mths | (6) $\geq$3 Mths | (7) $\geq$6 Mths | (8) $\geq$12 Mths | (9) $\geq$18 Mths | (10) $\geq$24 Mths | (11) Colostrum |
|---|---|---|---|---|---|---|---|---|---|---|---|
| **NREGS** | 0.00 | -0.16$^{***}$ | -0.00 | -0.01$^{**}$ | -0.02$^{***}$ | -0.02$^{***}$ | -0.01$^{***}$ | 0.00 | -0.00 | -0.00 | 0.01 |
| | (0.003) | (0.038) | (0.003) | (0.003) | (0.004) | (0.005) | (0.005) | (0.002) | (0.002) | (0.002) | (0.003) |
| Child gender = male | -0.00$^{**}$ | -0.02 | -0.00 | -0.00$^{**}$ | -0.01$^{***}$ | -0.01$^{***}$ | 0.00 | 0.00$^{**}$ | 0.00$^{***}$ | 0.00$^{**}$ | -0.00 |
| | (0.002) | (0.020) | (0.002) | (0.002) | (0.002) | (0.002) | (0.002) | (0.001) | (0.001) | (0.001) | (0.002) |
| Child age (months) | -0.01$^{***}$ | -0.11$^{***}$ | -0.00$^{***}$ | -0.01$^{***}$ | -0.01$^{***}$ | -0.01$^{***}$ | -0.03$^{***}$ | -0.04$^{***}$ | -0.04$^{***}$ | -0.04$^{***}$ | 0.01$^{***}$ |
| | (0.000) | (0.030) | (0.000) | (0.000) | (0.000) | (0.000) | (0.000) | (0.000) | (0.000) | (0.000) | (0.000) |
| Mother's age at birth | 0.00$^{***}$ | 0.01 | 0.00 | 0.00$^{**}$ | 0.00$^{***}$ | 0.00$^{***}$ | 0.00 | -0.00 | -0.00 | -0.00 | 0.00$^{***}$ |
| | (0.000) | (0.003) | (0.000) | (0.000) | (0.000) | (0.000) | (0.000) | (0.000) | (0.000) | (0.000) | (0.000) |
| Mother married | 0.01$^{**}$ | 0.11 | 0.01 | 0.01 | 0.02 | 0.02 | 0.01 | -0.00 | 0.00 | 0.00 | 0.01 |
| | (0.006) | (0.102) | (0.007) | (0.007) | (0.009) | (0.010) | (0.011) | (0.005) | (0.004) | (0.004) | (0.008) |
| Yrs. education, mother | 0.00$^{***}$ | 0.01 | 0.00 | 0.00 | 0.00 | 0.00$^{***}$ | 0.00$^{***}$ | -0.00$^{***}$ | -0.00$^{***}$ | -0.00$^{***}$ | 0.01$^{***}$ |
| | (0.000) | (0.003) | (0.000) | (0.000) | (0.000) | (0.000) | (0.000) | (0.000) | (0.000) | (0.000) | (0.000) |
| No. pregnancies, pre-2004 | -0.00$^{***}$ | -0.08$^{***}$ | -0.00 | -0.00$^{***}$ | -0.01$^{***}$ | -0.01$^{***}$ | -0.00 | 0.00 | 0.00 | 0.00 | -0.00 |
| | (0.002) | (0.021) | (0.002) | (0.002) | (0.002) | (0.003) | (0.002) | (0.001) | (0.001) | (0.001) | (0.002) |
| No. births, pre-2004 | 0.00 | 0.06$^{**}$ | 0.00 | 0.00 | 0.00 | 0.01$^{**}$ | 0.00 | -0.00 | -0.00$^{**}$ | -0.00$^{**}$ | -0.00 |
| | (0.002) | (0.024) | (0.002) | (0.002) | (0.002) | (0.003) | (0.003) | (0.001) | (0.001) | (0.001) | (0.002) |
| Yrs. education, father | 0.00$^{***}$ | -0.00 | 0.00 | 0.00 | -0.00 | -0.00 | -0.00 | -0.00$^{***}$ | -0.00$^{***}$ | -0.00$^{***}$ | 0.00$^{***}$ |
| | (0.000) | (0.003) | (0.000) | (0.000) | (0.000) | (0.000) | (0.000) | (0.000) | (0.000) | (0.000) | (0.000) |
| Share of HH $\geq$18 yrs. | -0.00 | 0.01 | -0.00 | 0.00 | 0.01 | 0.01 | 0.00 | 0.00 | -0.00 | 0.00 | 0.00 |
| | (0.004) | (0.045) | (0.003) | (0.004) | (0.005) | (0.005) | (0.005) | (0.002) | (0.002) | (0.002) | (0.004) |
| Hindu | 0.00 | 0.05 | 0.00 | 0.01 | 0.01 | 0.01$^{**}$ | -0.00 | -0.00 | -0.00 | -0.00 | -0.01 |
| | (0.004) | (0.037) | (0.003) | (0.003) | (0.004) | (0.004) | (0.005) | (0.002) | (0.002) | (0.002) | (0.004) |
| Scheduled Caste | 0.01$^{***}$ | 0.13$^{***}$ | 0.00 | 0.00 | 0.01$^{***}$ | 0.02$^{***}$ | 0.02$^{***}$ | 0.01$^{***}$ | 0.01$^{***}$ | 0.00$^{***}$ | 0.01$^{***}$ |
| | (0.003) | (0.028) | (0.002) | (0.003) | (0.003) | (0.004) | (0.004) | (0.002) | (0.001) | (0.001) | (0.003) |
| Scheduled Tribe | 0.02$^{***}$ | 0.09$^{**}$ | 0.01$^{**}$ | 0.01$^{**}$ | 0.01$^{***}$ | 0.02$^{***}$ | 0.00 | 0.00 | 0.00$^{**}$ | 0.00$^{**}$ | -0.01 |
| | (0.004) | (0.040) | (0.003) | (0.003) | (0.004) | (0.004) | (0.005) | (0.002) | (0.002) | (0.002) | (0.004) |
| Constant | 1.32$^{***}$ | 10.12$^{***}$ | 1.13$^{***}$ | 1.17$^{***}$ | 1.25$^{***}$ | 1.34$^{***}$ | 1.75$^{***}$ | 2.00$^{***}$ | 2.00$^{***}$ | 2.00$^{***}$ | 0.46$^{***}$ |
| | (0.021) | (1.579) | (0.011) | (0.012) | (0.016) | (0.019) | (0.025) | (0.025) | (0.025) | (0.025) | (0.025) |
| N (Mothers) | 163,332 | 116,073 | 120,952 | 120,952 | 120,952 | 120,952 | 120,952 | 120,952 | 120,952 | 120,952 | 119,442 |
| R-squared | 0.471 | 0.132 | 0.066 | 0.088 | 0.147 | 0.183 | 0.164 | 0.360 | 0.418 | 0.447 | 0.104 |

Table 5 reports the results of the impact of NREGS on breastfeeding initiation and duration, for the youngest surviving child. The treatment, NREGS status, is coded as follows: (1) for women who were exposed to the program at some time during the pregnancy and (0) for women who were not exposed to the program at any time during the pregnancy. In Column 1 the dependent variable is an indicator for breastfeeding initiation; in Column 2, the dependent variable is the duration of breastfeeding (in months). In Columns 3-10, the dependent variable is an indicator for having breastfed at least $x$ months, where $x$ is given by the column heading. We find that NREGS does not impact the propensity to breastfeed (Column 1), but it reduces the total duration by 0.16 months (Column 2). Columns 3-10 estimate a series of linear probability models to identify where in the breastfeeding trajectory NREGS has the greatest impact. Consistent with the predictions of a model of maternal labor supply, the program reduces the propensity to breastfeed at 1, 2, 3, and 6 months, after mothers have recovered from delivery and the opportunity cost of time is high. Column 11 presents results from a falsification test: NREGS should not impact the propensity to breastfeed immediately after delivery (measured by colostrum), since the opportunity cost of time is low; the results are consistent with this hypothesis. Notes: All specifications include year, month, and district fixed effects. Robust standard errors (in parentheses) are clustered at the village level. $^{**}$ $p < 0.05$, $^{***}$ $p < 0.01$.

**Table 6**
**Impact of NREGS on institutional delivery**

| | (1) Institutional Delivery |
|---|---|
| **NREGS** | $0.01^{***}$ |
| | (0.004) |
| Child gender = male | $0.01^{***}$ |
| | (0.002) |
| Mother's age at birth | $0.00^{***}$ |
| | (0.000) |
| Mother married | 0.01 |
| | (0.009) |
| Yrs. education, mother | $0.02^{***}$ |
| | (0.000) |
| No. pregnancies, pre-2004 | $0.02^{***}$ |
| | (0.002) |
| No. births, pre-2004 | $-0.05^{***}$ |
| | (0.002) |
| Yrs. education, father | $0.01^{***}$ |
| | (0.000) |
| Share of HH $\geq$18 yrs. | 0.00 |
| | (0.004) |
| Hindu | 0.00 |
| | (0.004) |
| Scheduled Caste | $-0.03^{***}$ |
| | (0.003) |
| Scheduled Tribe | $-0.10^{***}$ |
| | (0.004) |
| Constant | $0.13^{***}$ |
| | (0.012) |
| N (Mothers) | 164,948 |
| R-squared | 0.294 |

Table 6 reports the results of the impact of NREGS on institutional delivery rates, for the most recent pregnancy. The treatment, NREGS status, is coded as follows: (1) for women who were exposed to the program at some time during the pregnancy and (0) for women who were not exposed to the program at any time during the pregnancy. In Column 1 the dependent variable is an indicator for institutional delivery (public or private). Consistent with the labor supply framework, we find that women in NREGS districts use their income to purchase money-intensive health inputs: the propensity to deliver in a health facility increases by 1 percentage point. Notes: All specifications include year, month, and district fixed effects. Robust standard errors (in parentheses) are clustered at the village level. $^{**}$ $p < 0.05$, $^{***}$ $p < 0.01$.

**Table 7**
**Moderated impact of NREGS on maternal inputs**

| | (1) Month of Visit | (2) Total Visits | (3) Institutional Delivery |
|---|---|---|---|
| **NREGS** | 0.07*** | -0.05*** | 0.01** |
| | (0.025) | (0.017) | (0.004) |
| **NREGS x Distance** | 0.21 | -0.11 | 0.02 |
| | (0.208) | (0.105) | (0.033) |
| Distance to facility > 0km | -0.14 | 0.07 | -0.00 |
| | (0.108) | (0.059) | (0.018) |
| Child gender = male | 0.05*** | 0.01 | 0.01*** |
| | (0.012) | (0.008) | (0.002) |
| Child age | | | |
| Mother's age at birth | -0.00 | 0.01*** | 0.00*** |
| | (0.002) | (0.001) | (0.000) |
| Mother married | 0.15*** | 0.04 | 0.01 |
| | (0.053) | (0.033) | (0.010) |
| Yrs. education, mother | 0.04*** | 0.05*** | 0.02*** |
| | (0.002) | (0.001) | (0.000) |
| No. pregnancies, pre-2004 | 0.07*** | 0.06*** | 0.02*** |
| | (0.011) | (0.008) | (0.002) |
| No. births, pre-2004 | -0.14*** | -0.10*** | -0.05*** |
| | (0.012) | (0.009) | (0.002) |
| Yrs. education, father | 0.03*** | 0.01*** | 0.01*** |
| | (0.002) | (0.001) | (0.000) |
| Share of HH $\geq$18 yrs. | 0.03 | 0.02 | 0.00 |
| | (0.026) | (0.018) | (0.005) |
| Hindu | 0.06** | 0.03 | 0.01** |
| | (0.025) | (0.014) | (0.004) |
| Scheduled Caste | -0.01 | -0.06*** | -0.04*** |
| | (0.018) | (0.011) | (0.003) |
| Scheduled Tribe | -0.25*** | -0.15*** | -0.09*** |
| | (0.026) | (0.016) | (0.005) |
| Constant | 1.98*** | -0.29*** | 0.07*** |
| | (0.075) | (0.049) | (0.014) |
| N (Mothers) | 118,286 | 118,193 | 139,685 |
| R-squared | 0.134 | 0.620 | 0.220 |

Table 7 considers how the impact of NREGS on the utilization of health inputs is moderated by the accessibility of care (measured as distance to the health facility). We find little evidence that distance to the clinic mediates the effect of NREGS on the use of time-intensive or money-intensive inputs. Notes: See notes to Tables 3-6 for the definition of NREGS status for each health input. All specifications include year, month, and district fixed effects. In Column 2, we also control for the timing of the first prenatal visit. Robust standard errors (in parentheses) are clustered at the village level. $^{**}$ $p < 0.05$, $^{***}$ $p < 0.01$.

**Table 8**
**Impact of NREGS on child mortality**

| | (1) Neonatal Mortality | (2) Infant Mortality |
|---|---|---|
| **NREGS** | 0.00 | 0.00 |
| | (0.001) | (0.001) |
| Child gender = male | 0.00*** | 0.00*** |
| | (0.001) | (0.000) |
| Mother's age at birth | -0.00 | -0.00 |
| | (0.000) | (0.000) |
| Child age (months) | -0.02*** | -0.03*** |
| | (0.000) | (0.000) |
| Mother married | -0.00 | -0.00 |
| | (0.002) | (0.002) |
| Yrs. education, mother | -0.00 | -0.00*** |
| | (0.000) | (0.000) |
| No. pregnancies, pre-2004 | 0.00*** | 0.00*** |
| | (0.001) | (0.001) |
| No. births, pre-2004 | -0.01*** | -0.01*** |
| | (0.001) | (0.001) |
| Yrs. education, father | -0.00 | -0.00*** |
| | (0.000) | (0.000) |
| Share of HH $\geq$18 yrs. | -0.00 | -0.00 |
| | (0.001) | (0.001) |
| Hindu | 0.00 | 0.00 |
| | (0.001) | (0.001) |
| Scheduled Caste | 0.00 | 0.00*** |
| | (0.001) | (0.001) |
| Scheduled Tribe | 0.00 | 0.00** |
| | (0.001) | (0.001) |
| Constant | 1.05*** | 1.51*** |
| | (0.011) | (0.008) |
| N (Children) | 216,932 | 216,932 |
| R-squared | 0.472 | 0.651 |

Table 8 considers how NREGS, by providing maternal employment opportunities that result in decreased use of time-intensive inputs and increased use of money-intensive inputs, affects child health outcomes. The treatment, NREGS status, is coded as follows: (1) for women who were exposed to the program at some time during the pregnancy and (0) for women who were not exposed to the program at any time during the pregnancy. We find that the neonatal mortality rate (Column 1) and the infant mortality rate (Column 2), are not affected by NREGS. Notes: All specifications include year, month, and district fixed effects. Robust standard errors (in parentheses) are clustered at the village level. ** $p < 0.05$, *** $p < 0.01$.

**Table 9**
**Moderated impact of NREGS on child health**

| | Neonatal Mortality | | | Infant Mortality | | |
|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) |
| **NREGS** | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 |
| | (0.002) | (0.002) | (0.001) | (0.002) | (0.002) | (0.001) |
| **NREGS x HH comp.** | -0.00 | -0.00 | -0.01 | -0.00 | 0.00 | -0.01 |
| | (0.004) | (0.004) | (0.007) | (0.004) | (0.004) | (0.007) |
| HH composition | 0.00 | 0.00 | 0.00 | 0.00 | -0.00 | 0.00 |
| | (0.002) | (0.002) | (0.003) | (0.002) | (0.002) | (0.003) |
| Child gender = male | $0.00^{***}$ | $0.00^{***}$ | $0.00^{***}$ | $0.00^{***}$ | $0.00^{***}$ | $0.00^{***}$ |
| | (0.001) | (0.001) | (0.001) | (0.000) | (0.000) | (0.000) |
| Child age | $-0.02^{***}$ | $-0.02^{***}$ | $-0.02^{***}$ | $-0.03^{***}$ | $-0.03^{***}$ | $-0.03^{***}$ |
| | (0.000) | (0.000) | (0.000) | (0.000) | (0.000) | (0.000) |
| Mother's age at birth | -0.00 | -0.00 | -0.00 | -0.00 | -0.00 | -0.00 |
| | (0.000) | (0.000) | (0.000) | (0.000) | (0.000) | (0.000) |
| Mother married | -0.00 | -0.00 | -0.00 | -0.00 | -0.00 | -0.00 |
| | (0.002) | (0.002) | (0.002) | (0.002) | (0.002) | (0.002) |
| Yrs. education, mother | -0.00 | -0.00 | -0.00 | $-0.00^{***}$ | $-0.00^{***}$ | $-0.00^{***}$ |
| | (0.000) | (0.000) | (0.000) | (0.000) | (0.000) | (0.000) |
| Yrs. education, father | -0.00 | -0.00 | -0.00 | $-0.00^{***}$ | $-0.00^{***}$ | $-0.00^{***}$ |
| | (0.000) | (0.000) | (0.000) | (0.000) | (0.000) | (0.000) |
| Share of HH $\geq$18 yrs. | -0.00 | -0.00 | -0.00 | -0.00 | -0.00 | -0.00 |
| | (0.001) | (0.001) | (0.001) | (0.001) | (0.001) | (0.001) |
| Hindu | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 |
| | (0.001) | (0.001) | (0.001) | (0.001) | (0.001) | (0.001) |
| Scheduled Caste | 0.00 | 0.00 | 0.00 | $0.00^{***}$ | $0.00^{***}$ | $0.00^{***}$ |
| | (0.001) | (0.001) | (0.001) | (0.001) | (0.001) | (0.001) |
| Scheduled Tribe | 0.00 | 0.00 | 0.00 | $0.00^{**}$ | $0.00^{**}$ | $0.00^{**}$ |
| | (0.001) | (0.001) | (0.001) | (0.001) | (0.001) | (0.001) |
| Constant | $1.05^{***}$ | $1.05^{***}$ | $1.05^{***}$ | $1.51^{***}$ | $1.51^{***}$ | $1.51^{***}$ |
| | (0.011) | (0.011) | (0.011) | (0.008) | (0.008) | (0.008) |
| N (Children) | 216,932 | 216,932 | 216,932 | 216,932 | 216,932 | 216,932 |
| R-squared | 0.472 | 0.472 | 0.472 | 0.651 | 0.651 | 0.652 |
| Composition measure | Female $\geq$15 | Female $\geq$18 | Female 15-18 | Female $\geq$15 | Female $\geq$18 | Female 15-18 |

Table 9 considers whether any factors might moderate the impact of NREGS on child health. Specifically, we examine if substitutes for maternal care, measured by household composition, are important in this context. We find that the share of females older than 15 (Columns 1 and 4), older than 18 (Columns 2 and 5), or between 15 and 18 years (Columns 3 and 6) do not mediate the impact of the program on neonatal or infant mortality rates. Notes: See notes to Table 8 for the definition of NREGS status. All specifications include year, month, and district fixed effects. Robust standard errors (in parentheses) are clustered at the village level. $^{**}$ $p < 0.05$, $^{***}$ $p < 0.01$.

**Table 10**
**Robustness: Impact of NREGS on maternal inputs, non-linear models**

| | (1) Month of Visit | (2) Total Visits | (3) Breastfeeding Duration (Months) | (4) Colostrum | (5) Institutional Delivery |
|---|---|---|---|---|---|
| **NREGS** | 0.98** | 0.98*** | 1.07*** | 1.05 | 1.10*** |
| | (0.008) | (0.006) | (0.010) | (0.035) | (0.025) |
| Child gender = male | 0.98*** | 1.00 | 1.01 | 0.99 | 1.05*** |
| | (0.004) | (0.003) | (0.005) | (0.018) | (0.013) |
| Child age | | | 1.01 | 1.04*** | |
| | | | (0.008) | (0.003) | |
| Mother's age at birth | 1.00*** | 1.00*** | 1.00*** | 1.01*** | 1.03*** |
| | (0.001) | (0.000) | (0.001) | (0.003) | (0.002) |
| Married | 1.00 | 1.02 | 0.97 | 1.10 | 1.03 |
| | (0.018) | (0.013) | (0.021) | (0.084) | (0.056) |
| Yrs. education, mother | 1.00** | 1.02*** | 1.00*** | 1.06*** | 1.12*** |
| | (0.001) | (0.001) | (0.001) | (0.003) | (0.002) |
| No. pregnancies, pre-2004 | 0.98*** | 1.03*** | 1.02*** | 0.98 | 1.15*** |
| | (0.004) | (0.003) | (0.005) | (0.017) | (0.014) |
| No. births, pre-2004 | 1.02*** | 0.95*** | 0.99** | 0.99 | 0.69*** |
| | (0.004) | (0.003) | (0.006) | (0.019) | (0.009) |
| Yrs. education, father | 1.00*** | 1.01*** | 1.00 | 1.01*** | 1.05*** |
| | (0.001) | (0.000) | (0.001) | (0.003) | (0.002) |
| Share of HH $\geq$18 yrs. | 0.99 | 1.01 | 0.99 | 1.02 | 1.01 |
| | (0.009) | (0.007) | (0.011) | (0.041) | (0.028) |
| Hindu | 0.99 | 1.01** | 0.99 | 0.94 | 1.02 |
| | (0.007) | (0.005) | (0.010) | (0.033) | (0.025) |
| Scheduled Caste | 0.98*** | 0.96*** | 0.96*** | 1.09*** | 0.82*** |
| | (0.006) | (0.004) | (0.007) | (0.030) | (0.015) |
| Scheduled Tribe | 1.04*** | 0.94*** | 0.98** | 0.94 | 0.57*** |
| | (0.009) | (0.006) | (0.009) | (0.034) | (0.015) |
| Baseline | | 0.00*** | | 0.54** | 0.29*** |
| | | (0.000) | | (0.146) | (0.047) |
| **NREGS net effect** | | -0.044*** | | 0.005 | 0.016*** |
| | | (0.016) | | (0.003) | (0.004) |
| Overdispersion | | -.051 | | | |
| | | (0.136) | | | |
| N | 140,443 | 140,127 | 116,073 | 119,275 | 163,741 |
| Model | Cox | Poisson | Cox | Logit | Logit |

Table 10 shows that the results from our linear models are robust to non-linear specifications. Consistent with the labor supply model, mothers face increased wages after the implementation of NREGS, leading to reductions in the utilization of time-intensive inputs (Columns 1-3) and increased use of money-intensive inputs (Column 5). Notes: See notes to Tables 3-6 for definition of NREGS status for each category of health input. All specifications include year, month, and district fixed effects. In Column 2, we also control for the timing of the first prenatal visit. Robust standard errors (in parentheses) are clustered at the village level. ** $p < 0.05$, *** $p < 0.01$.

## Table 11
## Falsification: Impact of NREGS on maternal inputs, urban areas

| | (1) Month of Visit | (2) Total Visits | (3) Breastfeeding Initiation | (4) Breastfeeding Duration (Months) | (5) Institutional Delivery |
|---|---|---|---|---|---|
| **NREGS** | 0.06 | 0.04 | 0.01 | -0.11 | 0.02*** |
| | (0.044) | (0.048) | (0.006) | (0.071) | (0.008) |
| Child gender = male | -0.01 | -0.01 | 0.00 | 0.03 | 0.01** |
| | (0.017) | (0.021) | (0.003) | (0.037) | (0.004) |
| Child age | | | -0.01*** | 0.01 | |
| | | | (0.001) | (0.046) | |
| Mother's age at birth | -0.00** | 0.03*** | 0.00 | -0.01 | 0.01*** |
| | (0.002) | (0.003) | (0.000) | (0.006) | (0.001) |
| Married | -0.01 | 0.19** | -0.00 | 0.01 | -0.01 |
| | (0.068) | (0.084) | (0.013) | (0.149) | (0.016) |
| Yrs. education, mother | 0.00 | 0.07*** | 0.00*** | 0.00 | 0.02*** |
| | (0.003) | (0.003) | (0.001) | (0.006) | (0.001) |
| No. pregnancies, pre-2004 | 0.05*** | 0.07*** | -0.00 | -0.03 | 0.02*** |
| | (0.017) | (0.021) | (0.003) | (0.042) | (0.004) |
| No. births, pre-2004 | -0.09*** | -0.16*** | -0.00 | 0.03 | -0.06*** |
| | (0.020) | (0.023) | (0.004) | (0.046) | (0.004) |
| Yrs. education, father | 0.01 | 0.03*** | 0.00*** | -0.00 | 0.01*** |
| | (0.003) | (0.003) | (0.001) | (0.007) | (0.001) |
| Share of HH $\geq$18 yrs. | 0.03 | 0.01 | 0.01 | -0.10 | -0.00 |
| | (0.039) | (0.048) | (0.007) | (0.080) | (0.009) |
| Hindu | -0.00 | -0.02 | 0.01 | 0.11** | 0.00 |
| | (0.024) | (0.029) | (0.005) | (0.048) | (0.006) |
| Scheduled Caste | 0.01 | -0.07** | -0.00 | -0.00 | -0.05*** |
| | (0.028) | (0.034) | (0.005) | (0.056) | (0.007) |
| Scheduled Tribe | 0.04 | -0.21*** | 0.01 | 0.14 | -0.07*** |
| | (0.052) | (0.055) | (0.008) | (0.126) | (0.012) |
| Constant | 2.71*** | -0.24 | 1.21*** | 4.04 | 0.31*** |
| | (0.099) | (0.125) | (0.045) | (2.409) | (0.023) |
| N | 35,087 | 34,833 | 38,484 | 28,460 | 38,770 |
| R-squared | 0.091 | 0.625 | 0.513 | 0.132 | 0.342 |

Table 11 presents the results from a falsification test: estimating the impact of NREGS on maternal inputs, but limiting the regression sample to urban areas, where the program was not implemented. In Columns 1-4, the results as expected: there is no impact of NREGS on the utilization of time-intensive health inputs. However, there is a statistically significant effect of NREGS on urban institutional delivery rates; this result may be related to decreased rates of rural-urban migration. Notes: See notes to Tables 3-6 for definition of NREGS status for each category of health input. All specifications include year, month, and district fixed effects. In Column 2, we also control for the timing of the first prenatal visit. Robust standard errors (in parentheses) are clustered at the village level. ** $p < 0.05$, *** $p < 0.01$.

**Figure 1**
**Map of NREGS phased implementation**

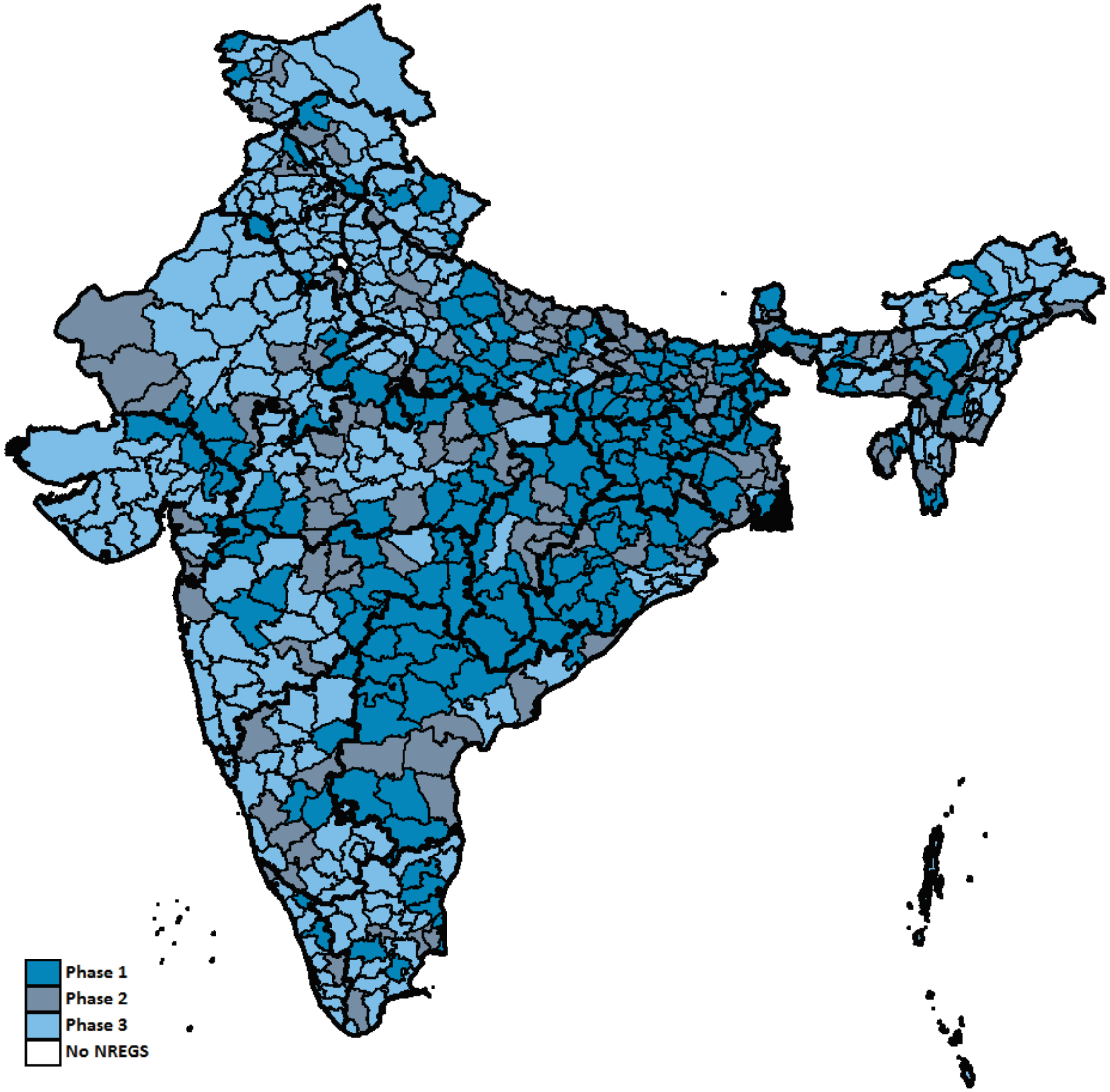
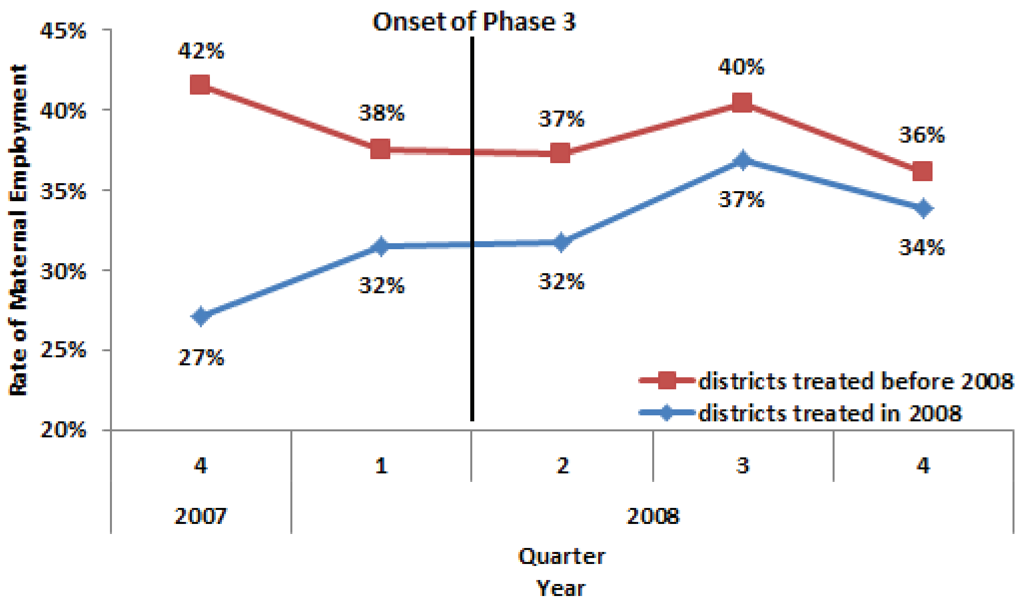


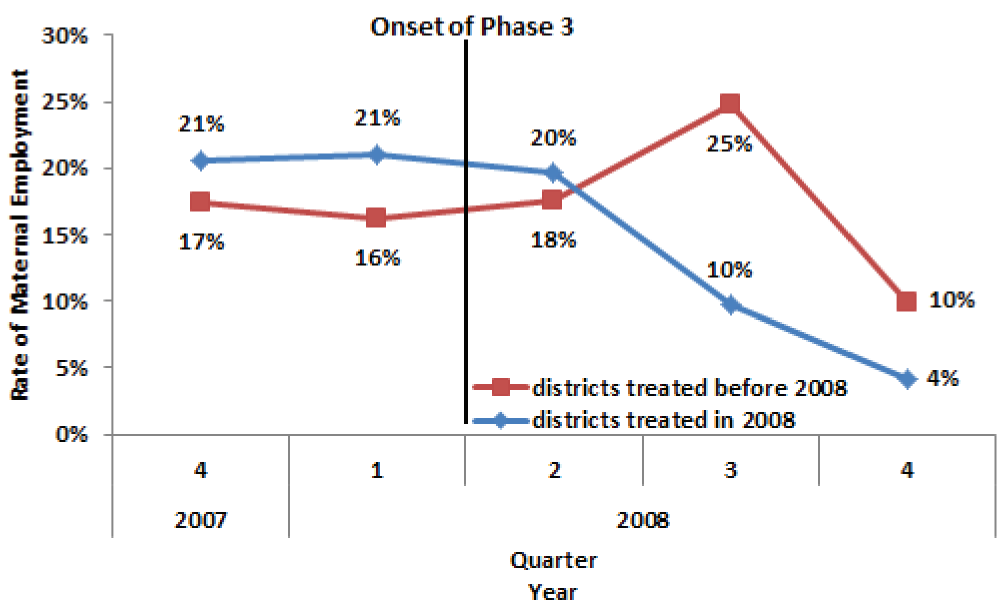

Figure 1 presents a map of the phased implementation of the National Rural Employment Guarantee Scheme (NREGS) in India. In 2006 (Phase 1, dark blue), NREGS was implemented in the 200 least developed districts. In 2007 (Phase 2, gray), the program was implemented in another 130 districts. The remaining 285 districts received the program in 2008 (Phase 3, light blue).

**Figure 2**
**Rate of maternal employment across treated and untreated districts**

(a) **Rural areas**




(b) **Urban areas**




Figures 2(a) and 2(b) depict the rate of maternal employment in the past 7 days across treated and untreated districts, after controlling for time-invariant district characteristics and seasonality in employment. Panel 2(a) presents trends for rural areas, where Phase 1 and 2 districts (treated) received NREGS prior to survey administration; Phase 3 districts (untreated) received the program during the survey period. NREGS is associated with greater maternal employment. Over the survey period, treated districts (red line, square markers) have significantly higher (p-values $< 0.01$) rates of maternal employment compared to untreated areas (blue line, diamond markers), with the gap closing once Phase 3 was implemented in 2008. Panel 2(b) presents trends for urban areas, which never received NREGS. In untreated districts, maternal employment is actually higher prior to Phase 3 implementation, and decreases after the onset of Phase 3, indicating that NREGS is not associated with maternal employment rates in areas not targeted by the program.

**Figure 3**
**Timeline of DLHS and NREGS implementation**

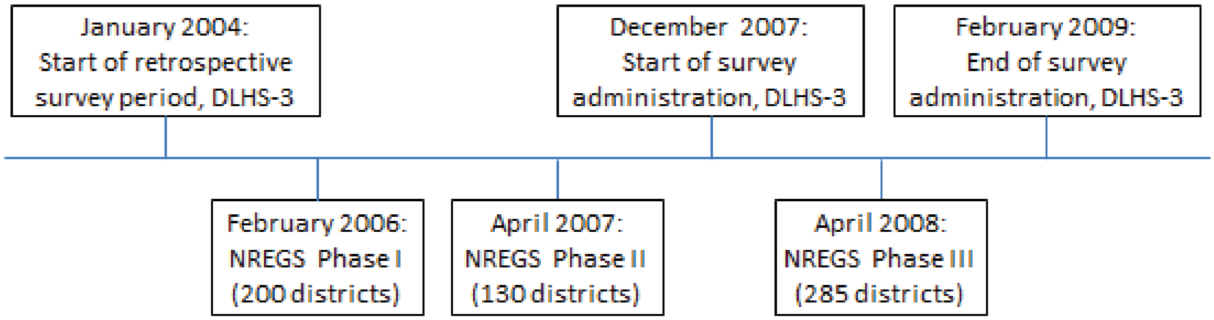
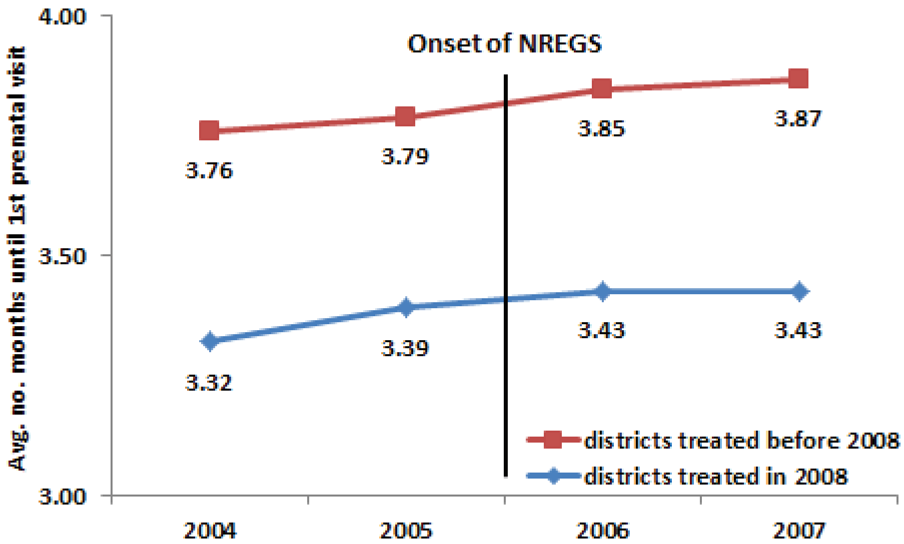



Figure 3 depicts the timing of the District Level Household Survey (DLHS), relative to the phases of NREGS.

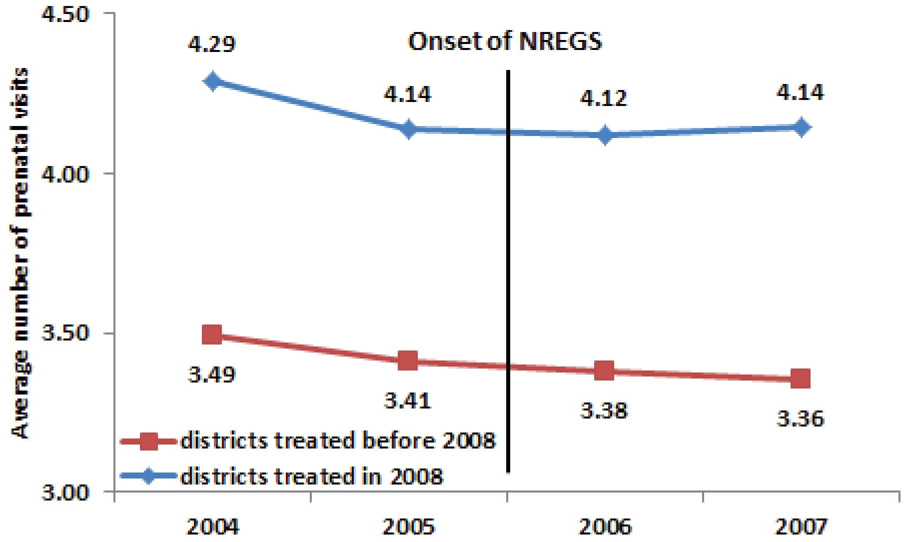
**Figure 4**
**Time until first prenatal visit: Trends in treated and untreated districts**




| | ■ districts | ◆ districts |
|---|---|---|
| pre-NREGS | 3.77 | 3.36 |
| post- NREGS | 3.86 | 3.43 |
| **Naïve DID estimate** | | **0.01 months** |

Figure 4 depicts the trends in the average number of months before a woman receives her first prenatal care visit, across treated (red line, square markers) and untreated (blue line, diamond markers) districts, before and after the implementation of NREGS. The figure provides evidence in favor of the parallel trends assumption underlying the differences-in-differences specification. The naïve estimate of the treatment is small: NREGS is associated with delaying prenatal care by approximately 0.06 months. This result is consistent with our model of labor supply: when women face increased employment opportunities, they substitute away from time-intensive health inputs.

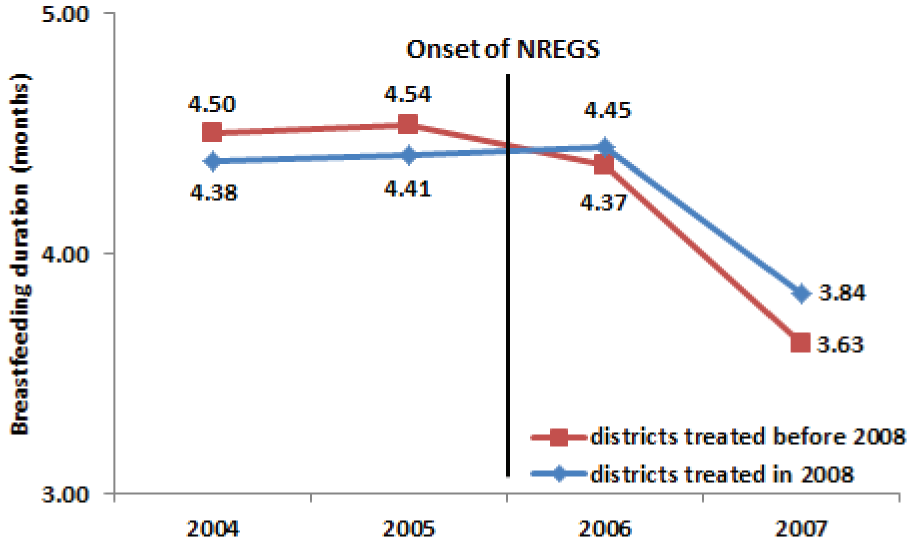
**Figure 5**
**Average number of prenatal visits: Trends in treated and untreated districts**




| | ■ districts | ◆ districts |
|---|---|---|
| Δ pre-NREGS | 3.45 | 4.21 |
| Δ post- NREGS | 3.37 | 4.13 |
| **Naïve DID estimate** | | **0.00 visits** |

Figure 5 depicts the trends in the average number of prenatal visits, across treated (red line, square markers) and untreated (blue line, diamond markers) districts, pre- and post-NREGS. As with the timing of prenatal care (Figure 4), there is evidence supporting the parallel trends assumption. The naïve DID estimate of NREGS on the amount of care received is -0.12 visits, also consistent with the prediction of a model of maternal employment: mothers reduce investment in time-intensive health inputs when the opportunity cost of time increases.

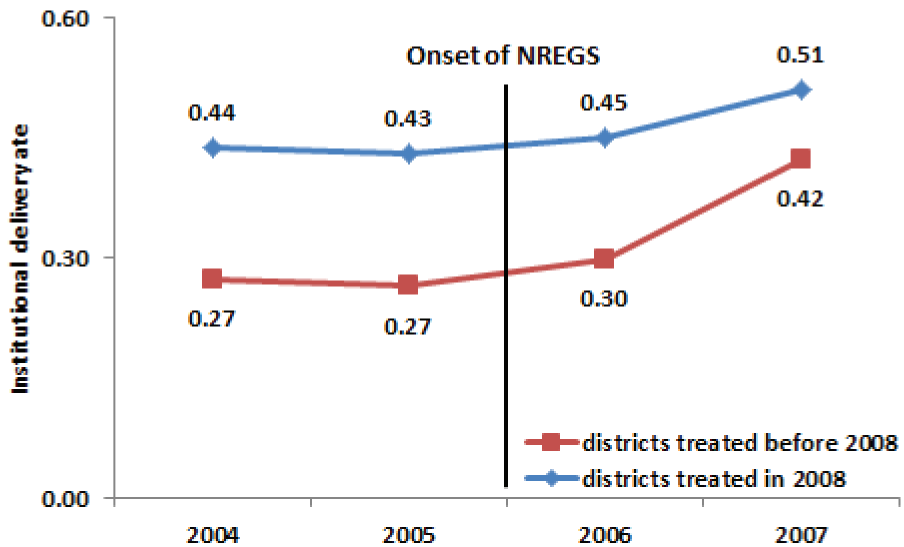
**Figure 6**
**Breastfeeding duration: Trends in treated and untreated districts**




| | ■ districts | ◆ districts |
|---|---|---|
| pre-NREGS | 4.52 | 4.00 |
| post- NREGS | 4.40 | 4.14 |
| **Naïve DID estimate** | | **-0.27 months** |

Figure 6 depicts the trends in breastfeeding duration, across treated (red line, square markers) and untreated (blue line, diamond markers) districts, before and after NREGS. There is evidence of the parallel trends assumption prior to the onset of NREGS. Further, we observe a sharp decrease in the duration of breastfeeding in treated areas after they received the program, suggesting the main prediction of the maternal labor supply model is borne out. The naïve differences-in-differences estimate is -0.14 months, indicating women in program areas stop exclusively breastfeeding earlier than mothers in untreated districts.

**Figure 7**
**Rate of institutional delivery: Trends in treated and untreated districts**

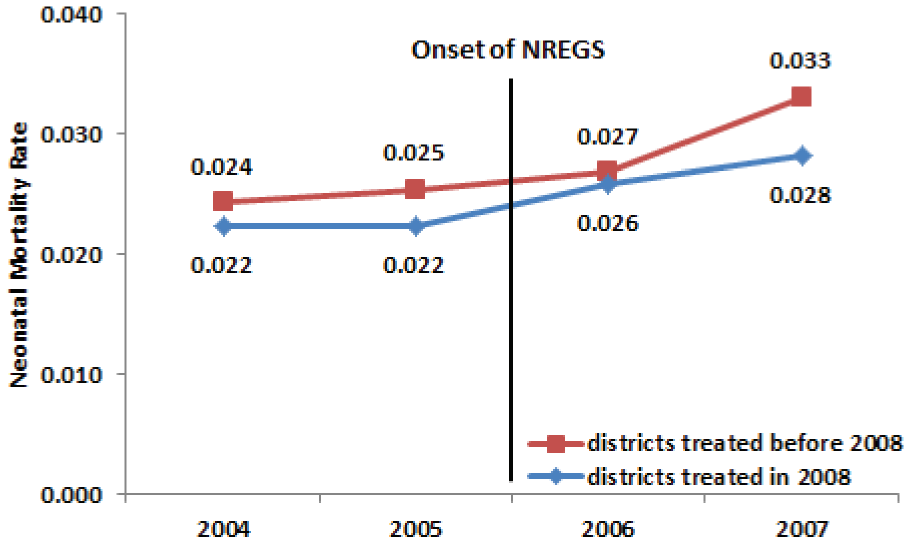



| | ■ districts | ◆ districts |
|---|---|---|
| pre-NREGS | 27 | 43 |
| post- NREGS | 36 | 48 |
| **Naïve DID estimate** | | **5 percentage points** |

Figure 7 depicts the pre- and post-NREGS trends in the rates of institutional delivery, across treated (red line, square markers) and untreated (blue line, diamond markers) districts, providing evidence in favor of the parallel trends assumption. Further, we observe a sharp increase in institutional delivery rates in treated districts after the implementation of NREGS. Correspondingly, the naïve differences-in-differences estimate is large: women in NREGS areas increase their propensity of institutional delivery by 6 percentage points.

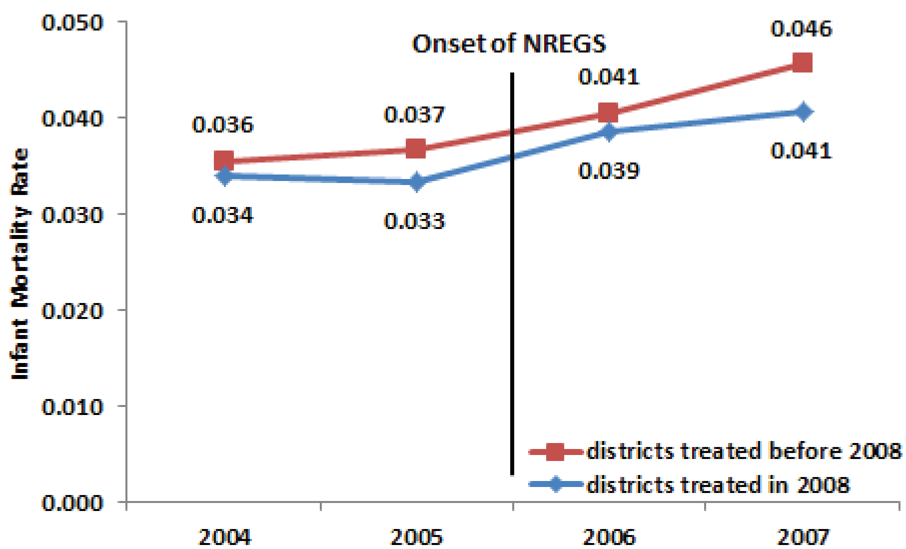
**Figure 8**
**Neonatal and infant mortality rates: Trends in treated and untreated districts**

(a) **Neonatal mortality**




| | ■ districts | ◆ districts |
|---|---|---|
| pre-NREGS | 2 | 3 |
| post- NREGS | 3 | 3 |
| **Naïve DID estimate** | | **0 percentage points** |

(b) **Infant mortality**




| | ■ districts | ◆ districts |
|---|---|---|
| pre-NREGS | 4 | 3 |
| post- NREGS | 4 | 4 |
| **Naïve DID estimate** | | **0 percentage points** |

Figures 8(a) and 8(b) depict, respectively, the pre- and post-NREGS trends in neonatal and infant mortality rates, across treated (red line, square markers) and untreated (blue line, diamond markers) districts. There is little evidence suggesting that NREGS, operating through the channel of maternal employment, affects child health as the naïve DID estimates are close to zero.

**Figure 9**
**Timing of first prenatal visit: Kaplan-Meier survival curve**

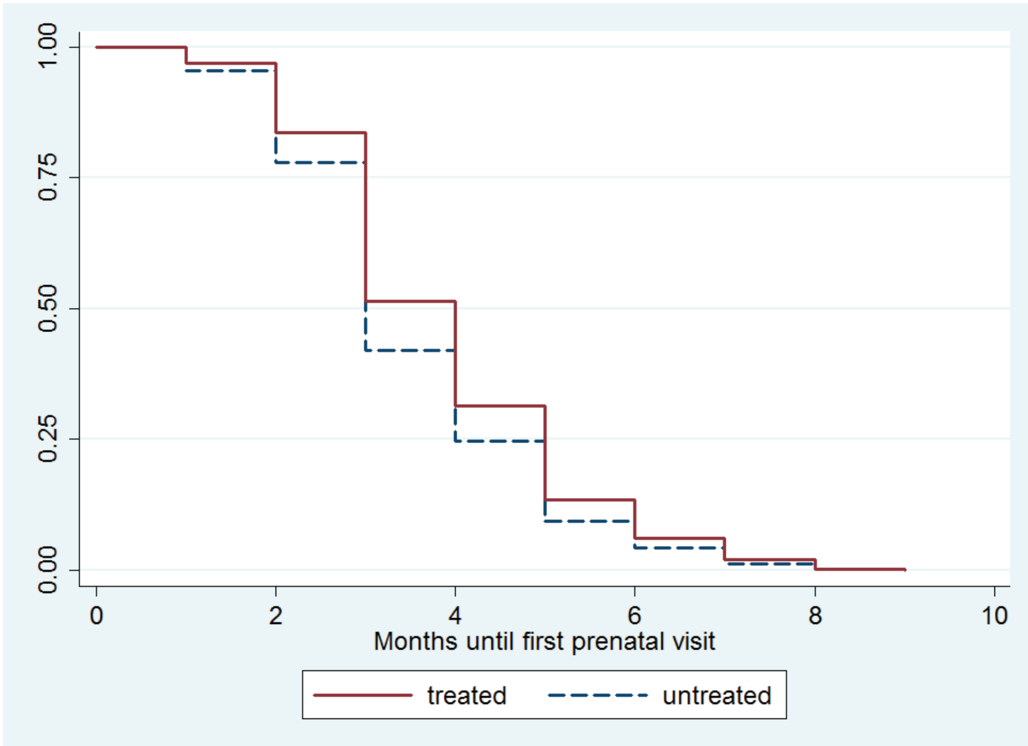



Figure 9 depicts the Kaplan-Meier survival curves for the timing of a woman's first prenatal visit for the most recent pregnancy. We see the (solid red) survival curve for women in NREGS districts lies above the (dashed blue) survival curve for women in non-NREGS districts; NREGS women take longer to acquire their first prenatal visit. A stratified Cox regression-based test (absorbing time-invariant district characteristics) rejects the null hypothesis for equality of the survival curves at a p-value < 0.001.

**Figure 10**
**Breastfeeding duration: Kaplan-Meier survival curve**

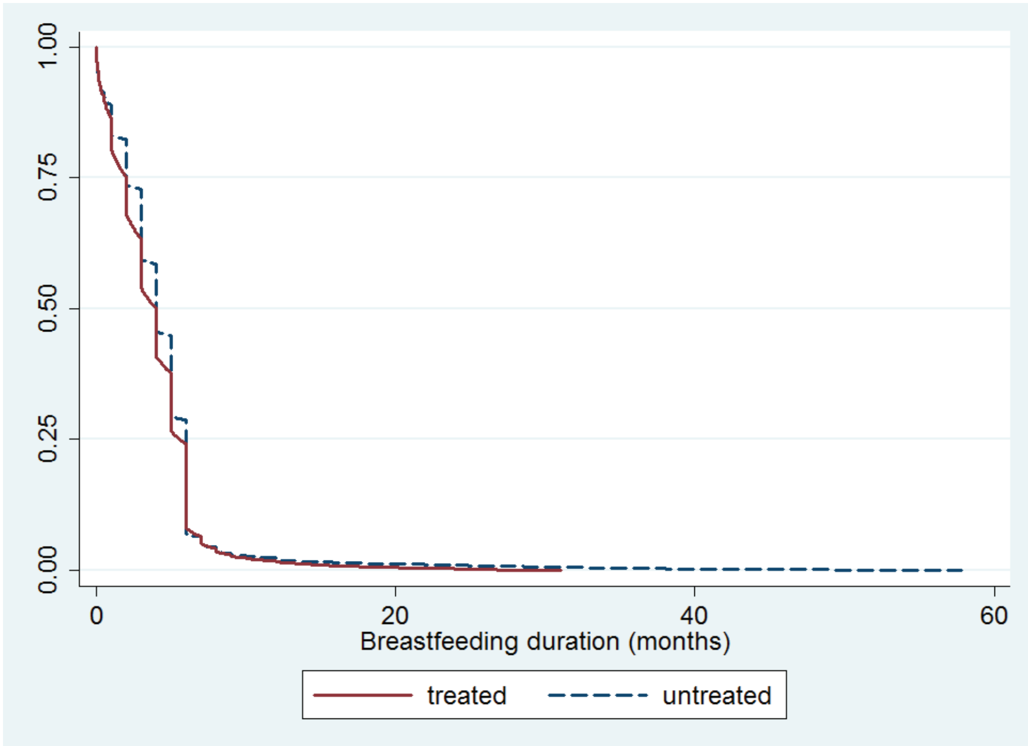



Figure 10 depicts the Kaplan-Meier survival curves for the duration of breastfeeding for the youngest surviving child. The (solid red) survival curve for women in NREGS districts lies below the (dashed blue) survival curve for women in non-NREGS districts, indicating that women in program areas do decrease the duration of breastfeeding. A stratified Cox regression-based test (absorbing time-invariant district characteristics) rejects the null hypothesis for equality of the survival curves at a p-value < 0.001.

# Male Earnings Inequality and Female Marital Outcomes: Evidence from India

A. V. Chari* Annemie Maertens† Sinduja Srinivasan‡

August 2014

**Abstract**

We provide the first evidence, for a developing country, of the causal effect of widening income inequality on the age of marriage of women. We extend the logic of the job search model. In this case, greater income inequality makes high-earning men relatively more attractive candidates than before, and increases the willingness of a woman to wait for such candidates, thus delaying her marriage. We utilize data from India to show that this mechanism does indeed have appear to have increased the average age at marriage. Further, it seems that women are using the extended search duration to acquire more education. We are able exploit the richness of the data to rule out a number of alternative explanations for our results. The results are not due to (i) Men searching longer in the marriage market in response to greater female earnings inequality, (ii) Regional or caste-base social norms, (iii) Men searching longer in the labor market (reducing the gender ratio in the marriage market), (iv) Earnings dispersion proxying for educational premia, and hence encouraging women to stay in school longer, or (v) Women's families needing more time to afford greater dowries.

*University of Sussex. Email: a.chari@sussex.ac.uk
†University of Sussex. Email: a.maertens@sussex.ac.uk
‡Pardee Rand Graduate School. Email: sinduja@rand.org
Acknowledgements: We are grateful to seminar participants at the University of Sussex, PACDEV, the PRGS Brown Bag, and to the Cement participants (Adriana Kugler, Laura Schechter, S. Anukriti, Rachel Heath, Melanie Khamis and Shing-Yi Wang) for comments and feedback. Any remaining errors are our own.

# 1 Introduction

The phenomenon of women marrying young is widespread and is of acute significance in developing countries. In addition to its role in the fertility transition, studies have found that a women's age at marriage correlates strongly with a number of health outcomes. Early marriage has been associated with low contraceptive use, miscarriages, multiple unwanted pregnancies, domestic violence, depression, and even increased HIV risk (Bruce, 2003; Clark, 2004; Nour, 2006; Raj et al., 2009; Santhya et al., 2010). Data from various Demographic and Health Surveys (DHS) conducted between 2003 and 2005 reveal that as many as 27.1% of women in India between the ages of 15 and 19 are married (the legal minimum age is 18 years). The issue is more severe in other developing countries, with corresponding figures of 46.1% in Bangladesh, 42% in Chad, 32.9% in Malawi, 50.4% in Mali, 38.2% in Mozambique and 31.7% in Nigeria.

However, the age at which women marry has been gradually increasing, although there is some heterogeneity in these trends across countries in this regard (Singh and Samara, 1996; Harwood-Lejeune, 2000; Choe et al., 2002; Heaton et al., 2002; Rashad and Osman, 2003; Westoff, 2003; Jensen and Thornton, 2003). In India, from 1980-2005, women's age at marriage increased by 4.3 years (Figure 1). The trend towards later marriage has been hypothesized to reflect factors such as access to education. However, Figure 2 indicates that in the case of India, more education cannot fully explain the increase of the marriage age. From 1980-2005, married women acquired an additional 1.5 years of education, which explains less than half the increase in the marriage age over the same period. Other explanations for delayed marriage include improved labor market opportunities, demographic change, urbanization, and exposure to mass media (see, for example, Jensen and Thornton, 2003; Mensch et al., 2005).

A less obvious mechanism, and one that has not been evaluated in the developing country context, is that increasing income inequality may delay marriage decisions. This mechanism may be understood in the context of a search model in the marriage market as in Loughran (2000 & 2002). In this model, a woman who is on the marriage market samples marriage offers from prospective spouses, evaluating them in terms of their income. In deciding whether to accept a particular offer, she must decide whether the benefits from waiting for a better offer outweigh the costs of staying unmarried during this period. Widening income inequality makes high-earning men relatively more attractive candidates than before, and increases the willingness of the woman to wait for offers from such candidates, thereby delaying her marriage. Loughran (2002) and Gould and Paserman (2003) provide evidence that this mechanism has significantly reduced marriage rates in the United States over the last few decades.

Utilizing nationally representative household survey data from India, we provide the first evidence for such a phenomenon in the context of a developing country. The role of income inequality in delaying marriage is particularly relevant in the Indian setting because the turn towards economic liberalization taken during the 1990s appears to have increased income and consumption inequality, both over time as well as geographically (Banerjee and Piketty, 2005; Deaton and Dreze, 2002; Sundaram and Tendulkar, 2003a & 2003b; Sen and Himanshu, 2005; Pal and Ghosh, 2007). The available evidence indicates that this growing inequality has been largely driven by an increase in incomes at the upper tails of the income distribution, while incomes of other groups have stagnated. Banerjee and Piketty (2005) use income tax data to show that incomes of the top 1% in India increased by 50% in the 1990s, and that their share of total income nearly doubled.

To evaluate the link between earnings inequality and marriage in India, we relate (age-specific) marriage rates of women of marriageable age to male earnings inequality in the relevant marriage market, which we define as all out-of-school unmarried males of marriageable age belonging to the same caste (social group) and residing in the same state. This definition of the marriage market takes into account the fact that inter-caste marriages in India are rare, and that ethno-linguistic differences between states limit the geographical boundaries of the marriage market. This natural segmentation in India provides substantial variation in the marriage market experiences of women, and we are able to utilize this variation to establish the effect of earnings inequality on marriage while controlling for spatial and inter-caste socio-economic characteristics that may influence marriage decisions.

Consistent with the predictions of the marital search model, we find that women in marriage markets characterized by higher male earnings inequality have significantly lower (age-specific) rates of marriage. Men, on the other hand, do not delay their marriage as a response to higher female earnings inequality. While Loughran (2002) and Gould and Paserman (2003) examine the effect of income inequality on marriage rates, we are able to go a step further and consider the extent of time marriage is postponed. We can do so because our data contain a subsample of married women for whom we know the age at marriage. The results indicate that a doubling of the earnings of men at the 90th percentile relative to median male earnings reduces the (age-specific) propensity of women to get married, resulting in an average delay of marriage by nearly 2 years.

Marriage in India (and many other developing countries) is strongly linked with education. In the majority of cases, schooling is discontinued immediately upon cohabitation, if not much earlier. Hence, an important benefit of postponing marriage in the developing country setting is that it can increase female educational attainment (Brien and Lillard, 1994; Lindstrom and Brambila Paz, 2001; Fields and Ambrus, 2008; Maertens, 2013). We find this to be the case: women in marriage markets with greater male earnings inequality also obtain more education. This additional education is concentrated at the point in the educational trajectory at which most women find themselves when they are on the marriage market, i.e., greater male earnings inequality results in increased rates of matriculation and high-school graduation among women, but does not correlate either with the probability of receiving zero education or the probability of completing primary school. We also show that greater earnings inequality does not affect the educational attainment or school enrollment of children who are too young to be on the marriage market. These results confirm that male earnings inequality impacts female educational attainment only at the time of marital search and not before.

We conclude that there is a strong link between male earnings inequality and female age at marriage, and that this relationship plausibly reflects the causal effect of the former on marital search duration. In turn, the increase in search duration appears to improve educational attainment, resulting in significantly higher rates of matriculation and high-school completion. The results in this paper suggest the strong possibility that widening income inequality in India may have had some salutary effects working through the marriage market.

The paper proceeds as follows. Section 2 outlines a model of marital search to motivate the empirical analysis. Section 3 describes the data and the variables used in the analysis. Sections 4 presents the empirical strategy; in Section 5 we report and discuss the baseline results, and explore alternative hypotheses. In Section 6, we consider an extension: the impact of rising male income inequality on the educational attainment of women. We conclude our discussion in Section 7.

## 2 A model of marriage market search

This is a partial equilibrium model, that focuses attention on how changes in the characteristics of the male earnings distribution affect the marital search duration of women. This section presents a model that is used to derive predictions about how the duration of a woman's marriage market search is related to the mean, variance and upper-tail characteristics of the male earnings distribution. In substance, the model is no different from a standard labor market search model. For simplicity of exposition, therefore, we present a version of the model that embodies some functional form assumptions, while referring readers to Mortensen (1986) for the derivation of the results in the more general case.

We model the marriage market search as an extension of a labor market search model (see among others, McCall (1970)), in which an infinitely-lived woman sequentially samples marriage offers from a known distribution of offers, before finally accepting an offer. As in Loughran (2002), we assume that women prefer high-earning males over low-earnings males and that this is the only dimension of groom heterogeneity, so that the male earnings distribution characterizes the distribution of offers. However, the empirical analysis allows the marriage market to be structured horizontally along another important dimension: caste.

The asymmetric nature of the model fits into the Indian context. Traditionally, when a female child reached the age of menarche, the child was taken out of school and arrangements for her marriage were made (Caldwell et al., 1983). Nowadays, while parents in India have more leeway, they do become anxious to marry off their daughter once she has reached menarche, partly to avert any unwanted pregnancies (see also Srinivas, 1984). Once a daughter is considered of marriageable age, inquiries are sent across the (caste-based) social network. Suitable candidates present themselves, and out of this set, a selection is made. Maertens (2013), using data from three villages in India, finds that families meet up to 5 candidates in this process. While the prospective spouse is usually only met on the day of the wedding (65% of the women in our sample first met their husband at the wedding), 42% of the brides do have a say in the selection process. From the point of view of the men, Maertens (2013) finds that education and settling down (e.g., getting a regular job or establishing a household) receives preference over marriage; so usually only out-of-school men could be considered as potential candidates, and only men with certain income and wealth prospects would be considered desirable matches.

## 2.1 Set-up

Suppose $F$ is the cumulative distribution function corresponding to the male earnings distribution; let $f$ denote the associated density function. For simplicity, we assume that the woman receives exactly one offer in each period and that this offer is an independent and random draw from $F$. We denote the earnings associated with this offer by $x$. The woman may either accept or reject this proposal. If she chooses to reject it, she will have to wait one more period, during which she will receive a utility of $c$, which is the per-period utility from remaining single, and may be negative for some women. Recognizing the iterative nature of the problem we follow a traditional dynamic programming framework.

Denote by $V_{accept}(x)$ the lifetime value obtained by accepting offer $x$. Assuming that the woman is risk-neutral, we can write:

$$V_{accept}(x) = \frac{x}{1-\beta} \tag{1}$$

where $\beta$ is the discount rate.

Denote by $V_{reject}$ the value of rejecting offer $x$. Because of the recursive nature of the problem, we can write:

$$V_{reject} = c + \beta \mathbb{E}\left[\max\left\{\frac{x}{1-\beta}, V_{reject}\right\}\right] \tag{2}$$

The form of the solution to this optimization problem can be shown to be the following: there is a reservation earnings level, denoted by $R$, above which the woman will accept any offer, and below which she will reject any offer and instead choose to wait another period. We can therefore re-write (2) as:

$$\begin{aligned} V_{reject} &= c + \beta\left[\int_R^\infty \frac{x}{1-\beta} dF(x) + \int_{-\infty}^R V_{reject}\, dF(x)\right] \\ &= c + \frac{\beta}{1-\beta}\int_R^\infty x dF(x) + \beta F(R) V_{reject} \end{aligned} \tag{3}$$

We also note that if the offer exactly equals the reservation level, $R$, the woman will be indifferent between accepting and rejecting, i.e. we must have:

$$V_{reject} = V_{accept}(R) = \frac{R}{1-\beta} \tag{4}$$

Combining (3) and (4), we can write, after some algebra:

$$[r+1-F(R)]R = rc + \int_{R}^{\infty} x dF(x) \tag{5}$$

where $r = \frac{1-\beta}{\beta}$.

This equation may be solved to obtain the reservation earnings level, $R$. Associated with $R$, one can define the per-period probability of "escape" from the marriage market, i.e., the probability of getting married, $q$:

$$q = P(x > R) = 1 - F(R) \tag{6}$$

Lower values of $q$ imply greater search duration (in expectation), or equivalently, higher age at marriage. The expected number of periods of search before an offer is accepted can be shown to be $1/q$.[1]

## 2.2 Comparative Statics

We now present the key comparative statics of interest. We note that the outcome of interest is the expected search duration, which is inversely related to $q$. The comparative static results presented here are derived by assuming that $F$ is $N(0,1)$. This allows us to simplify the algebra. It can be verified however that the results obtained here are fully general, and are not dependent on the normality assumption (see Mortensen, 1986).

First, we consider the effect of changes in the per-period utility ($c$) on the per-period escape probability ($q$), while holding fixed the earnings distribution ($F$). Note that in this case it is sufficient to know the effect on $R$ to infer the effect on $q$ (this follows from Equation (6) above). For different levels of $c$, the reservation earnings will be different. Applying the Implicit Function Theorem to equation (5) and using (6), one can show that higher levels of $c$ correspond to higher

[1]Denote search duration by T. Then, $\mathbb{E}(T) = 1 * P(T=1) + 2 * P(T=2) + 3 * P(T=3)... = q[1 + (2(1-q) + 3(1-q)^2 + ...]$
Denote $S = 1 + 2x + 3x^2 + 4x^3 + ...$with $x = 1 - q$. We can show that $S = 1/q^2$ and hence $\mathbb{E}(T) = 1/q$.

values of $R$ and hence lower values of $q$:

$$\frac{dR}{dc} = \frac{r}{r+q} > 0 \tag{7}$$

This result says that women who obtain little utility (or great disutility) from remaining single are less likely to turn down offers.

Next, we examine how changes in the earnings distribution, $F$, affect search duration. A complication in this case is that when $F$ changes, the effect on search duration cannot automatically be inferred from the change in $R$.

We begin by considering the effect of a marginal rightward shift in the earnings distribution, while holding constant the variance, i.e. the mean of the distribution increases. Set $F(x) = \Phi(\frac{x-\mu}{\sigma})$ where $\Phi$ denotes the standard normal distribution function. Again, applying the Implicit Function Theorem to equation (5) utilizing (6) yields (see also Appendix):

$$\frac{dR}{d\mu} = \frac{q}{r+q} < 1 \tag{8}$$

And, using (6):

$$\frac{dq}{d\mu} = \left(1 - \frac{dR}{d\mu}\right)\phi(R) > 0 \tag{9}$$

where $\phi$ denotes the standard normal density function. Thus, a rightward shift in the offer distribution will increase the reservation level, $R$. However, $R$ increases less than one-for-one with mean $\mu$, and in turn this implies that the per-period escape probability increases, and so the expected age at marriage must fall. Because $r$ is likely to be quite small relative to $q$, we expect this effect to be small.

Next, we consider a mean-preserving spread in the offer distribution. Because of our assumption of normality, this is equivalent to an increase in the standard deviation of the offer distribution, while holding constant the mean. It is important to note that except in this special case, the equivalence does not hold. Mortensen (1986) provides the comparative static result in the general case of a mean-preserving spread starting from an arbitrary distribution.

Applying the Implicit Function Theorem to equation (5) using (6) yields (see also Appendix):

$$\frac{dR}{d\sigma} = \frac{\phi(R)}{r+q} > 0 \tag{10}$$

And, using (6):

$$\frac{dq}{d\sigma} = \left[R - \frac{dR}{d\sigma}\right]\phi(R) \tag{11}$$

The increase in earnings dispersion, $\sigma$, unambiguously increases the reservation level, $R$. However, this time, the direction of the effect on $q$ is ambiguous, depending on the term in square brackets in Equation (11).

To understand these results, note first that for any initial value of $R$, the expected earnings of individuals above $R$ increases with a mean-preserving spread of $F$. This implies that the value of rejecting an offer equal to $R$ has increased, which in turn will cause the woman to revise $R$ upwards (because initially, the value of accepting an offer equal to $R$ was equal to the value of rejecting it). Thus the reservation level unambiguously rises for all women, regardless of their initial value of $R$.

Next, note that the mean-preserving spread increased the probability mass to the right of $R$ (i.e. the escape probability, $P(x > R)$) in the case of women for whom $R$ was above the median, while decreasing $P(x > R)$ in the case of women for whom $R$ was below the median. For the latter group, as $R$ rises in response to the mean-preserving spread, $P(x > R)$ falls further, whereas for initial reservation values above the median, the net effect of the increase in $R$ on $P(x > R)$ is ambiguous, and depends on how large the increase in $R$ is. For this reason, the mean-preserving spread has a theoretically ambiguous effect on search duration for women whose initial reservation levels were above the median.

Nevertheless, it is likely that for most values of $R$, the net effect on search duration is positive, i.e. $q$ falls (see also Burdett and Ondrich (1985) for discussion on this point). Figure 3 illustrates this in the context of our particular functional form assumptions, by graphing $R - \frac{dR}{d\sigma}$ against different values of $R$ (where we have assumed a discount rate of $\beta = 0.95$). The figure also graphs the standard normal cumulative distribution function. The figure shows that $R - \frac{dR}{d\sigma}$ is indeed negative over a substantial range of values of $R$, accounting for a large proportion of the probability mass. It is only in the far right tail of the distribution, high values of $R$, that $R - \frac{dR}{d\sigma}$ (and hence

$\frac{dq}{d\sigma}$) turns positive. Nonetheless, the question of whether increasing earnings dispersion increases search duration and the age at marriage is ultimately an empirical one, and will be tested in the data.

We have so far considered a mean-preserving spread as an example of rising inequality. However, we are unlikely to observe mean-preserving spreads in the data. The available evidence also indicates that for a number of English-speaking countries as well as India and China, recent increases in income inequality are largely being driven by increases in income for people in the upper tails of the income distribution (Atkinson, Piketty, and Saez, 2011). In the empirical analysis, we will accordingly separate out the impact of increases in upper-tail incomes (relative to median incomes) from the impact of reductions in lower-tail incomes.

While analytical proofs are difficult, it is possible to intuitively predict the direction of impact in these cases. Consider an increase in income at the 90th percentile, relative to the median income. For women with reservation levels below the 90th percentile (this includes the majority of women), this would increase the value of rejecting an offer equal to $R$. This would result in an increase in $R$ and (most likely) an increase in search duration. However, a reduction in income at the 10th percentile relative to the median income would have no effect on reservation levels and search durations for women whose initial reservation levels were above the 10th percentile (this likely includes the majority of women). These asymmetric predictions can be tested in the data.

## 3 Data

**Overview of the data:** We use the India Human Development Survey 2005 (IHDS) to analyze the relationship between earnings inequality and marriage in India. The IHDS is a nationally representative survey of 41,554 rural and urban Indian households in all twenty-eight states and five union territories.[2] IHDS administered household, school, village, and medical facility surveys, collecting poverty, health, employment, economic, and social data. The breadth and depth of IHDS permits extensive analysis of human and economic development in India.

**Definition of the marriage market:** We define a set of spatially non-overlapping marriage markets, in which we exploit variation in the earnings distributions across markets to analyze

[2] Lakshwadeep and the Andaman and Nicobar Islands were excluded.

how earnings inequality impacts female marriage. This variation arises naturally from two facts: (1) Women tend to marry within their own caste[3], and earnings distributions tend to be very different across castes, and (2) The cultural and linguistic differences between the states of India result in marriage markets being circumscribed by state boundaries, and socio-economic differences across the states result in substantial variation in the earnings distributions. In India, about 95% of women marry within their caste (see Panel A, Figure 4). Further, after marriage, women settle very close to their place of birth: nearly 90% of married women, live within 5 hours traveling distance of their families (see Panel B, Figure 4)[4], suggesting that women marry within state, and most likely, districts.[5] The caste composition of the households in our sample is as follows: 5% Brahmins, 16% other high castes, 38% Other Backward Castes, 22% Dalits (Scheduled Castes), 6% Adivasis (Scheduled Tribes) and 12% Muslims (and a small minority of Sikh, Jain and Christians). Thus, the marriage market of a woman of "marriageable age"(as defined in Section 4) is defined as all unmarried men ages 18 to 35, who are from her same caste and state, and who are not currently enrolled in school.

**Description of the earnings variables:** The 2005 IHDS elicited detailed information on the employment, occupation, and earnings of individuals in the household, including the annual earnings of individuals from employment outside the home, the net income the household received from farm and business activity, and the contribution of each household individual to the farm or business, in terms of days worked in the year. From these data, we construct a comprehensive measure of annual individual earnings, which accounts for earnings from paid employment for work done outside the home and the individual's share of net income from household farm and business ventures. To obtain this share, we assign household farm and business income to each household member in proportion to the number of days an individual spent on that activity during the year (relative to the total number of days household members devoted to the activity). In our sample, about 55% of unmarried men earn income from employment solely outside the household; 34% of men receive farm income, and 13% receive some business income. In our empirical analysis we assess the robustness of the results using a narrower definition of earnings

[3]In India, caste is equivalent to a social or ethnic group. Households provide information on their caste similar to survey respondents in the US providing interviewers with race/ethnicity information.

[4]Traveling distance refers to travel by road or train, not plane.

[5]A district is an administrative division in India similar to a US county.

that only includes wage and salaried employment (from outside the home).

We use the sample weights in the IHDS to construct the earnings distribution corresponding to each marriage market. We consider only the earnings of unmarried males between ages 18 and 35, who are not currently enrolled in school. From the earnings distribution we compute (i) Mean earnings, (ii) Standard deviation of earnings, (iii) Various percentiles of the distribution, including the 90th, 50th, and 10th. In our sample, men earn close to Rs. 30,000 (approximately $680) per year, with those at the 90th percentile earning 23 times as much as men at the 10th percentile. We also derive measures of the mean and dispersion of the earnings distributions for unmarried women ages 18 to 35 who are not in school, which we utilize to test the robustness of our results.

**Sample of women:** Our analysis uses two overlapping samples of women. The first sample is derived from the household roster, which provides the age and marital status of each household member. We will refer to this as the "full sample". The second sample is based on the sample of ever-married women (one such woman was randomly chosen from each surveyed household), who were administered a separate questionnaire. For these women, we know their marriage histories, educational attainment, and anthropometric measurements. We refer to this sample as the "ever-married sample". Table 1 introduces this ever-married sample of 33,482 women and their households. The ever-married woman is (on average) 33 years old (standard deviation 8 years), was 17.53 years old at the time of her marriage (standard deviation 3.79), and has 2.29 children living with her in the household (standard deviation 1.44). The average number of household members is 5.2 (standard deviation 2.5), and the number of children between the ages of 0 and 14 years is 1.6 (standard deviation 1.6). The highest education level among adults (over 21 years) in the household is 7.5 years (standard deviation 5.1 years). The monthly consumption per capita is, on average, Rs. 953 (standard deviation Rs. 1,024) and the yearly income is estimated at Rs. 53,922 (standard deviation Rs. 83,284). To measure household possessions and housing quality, we construct a household asset index, which ranges from 0 to 30. In the ever-married sample, households, on average, have an asset level of 12.3 (standard deviation 6.3).

## 4 Empirical Strategy

Our baseline regressions investigate the effects of male income dispersion on female marriage using a simple specification. We begin by looking at how male earnings inequality affects the probability of being married conditional on age, using the full sample of women obtained from the household roster. We employ the following linear probability model:[6]

$$y_{ics} = \alpha + \beta_1\sigma_{cs} + \beta_2\mu_{cs} + \gamma Age_{ics} + \eta_c + \eta_s + u_{ics} \tag{12}$$

where $y_{ics}$ is an indicator that takes the value 1 if individual $i$ belonging to caste $c$ and residing in state $s$ is married; $\sigma_{cs}$, the measure of earnings dispersion, is the standard deviation of male earnings in the marriage market defined by caste $c$ and state $s$; $\mu_{cs}$ denotes the mean of the male earnings distribution in the marriage market; $Age_{ics}$ denotes the age of the individual (thereby capturing differences in $c$, the utility of staying single, by age); $\eta_c$ denotes a caste fixed effect; $\eta_s$ denotes a state fixed effect; $u_{ics}$ is an unobserved error term. The inclusion of caste fixed effects allows us to control for marital norms and socio-economic characteristics that differ across castes. The inclusion of state fixed effects controls for regional differences in socio-economic conditions that may impact the marriage decisions of women, as well as for regional differences in marriage norms. The standard errors are corrected by clustering at the marriage market level.

The main coefficient of interest is $\beta_1$, which represents the effect of a one unit increase in male earnings dispersion on an individual's probability of being married. The marriage market search hypothesis, as outlined in Section 2, suggests that $\beta_1$ should be negative, i.e. greater dispersion in male earnings will result in a lower per-period probability of getting married, and $\beta_2$ should be positive.

One concern with the analysis above is that the earnings dispersion may be changing rapidly over time. As a result, the proportion of married women in a given marriage market may not truly reflect the effects of the current earnings dispersion, especially if many of the married women were married several years prior to the survey. To mitigate this concern, we limit the regression sample to women between the ages of 15 and 30.

[6] We opted for a linear probability model instead of a probit as the latter failed to converge due to the large number of fixed effects in several specifications. Where convergence was achieved, we obtained very similar results (results are available on request).

While the analysis based on specification (12) can tell us how the male earnings dispersion affects the probability of being married conditional on age, it does not directly indicate the extent to which marriage is delayed as a consequence of increased earnings dispersion. When examining age at marriage, there is a censoring issue: for a woman who is unmarried at the time of the survey, we only know that her age at marriage is at least as great as her current age. As Glick and Sahn (2002) note, the typical approach to addressing this censoring issue is to estimate duration models. We employ the following specification:

$$h\left(t\right) = h_0(t)\exp\left(\alpha + \beta_1\sigma_{cs} + \beta_2\mu_{cs} + \eta_c + \eta_s\right) \tag{13}$$

where $h\left(t\right)$ is the hazard of marriage for a woman $i$ belonging to caste $c$ and residing in state $s$ at time/age $t$, conditional on remaining unmarried until $t$. In Equation 13, the baseline hazard $h_0(t)$ takes the Weibull functional form, which allows the probability of marriage to increase with age, reflecting the fact that nearly every women in India gets married by age 30. Further, the Weibull functional form allows us to easily transform coefficients from the proportional hazard metric to the failure-time metric, allowing for easier interpretation of the impact of increased male income inequality on female age at marriage. The measures of male earnings dispersion in each marriage market, ($\sigma_{cs}$ and $\mu_{cs}$), are the same as in Equation 12; Equation 13 also accounts for caste effects.[7] We estimate a stratified regression at the state level, which allows the both the scale and shape of the hazard function to vary by state, analogous to the state fixed effects included in Equation 12. In this specification (13), we expect $\beta_1$ to be negative, indicating that greater earnings inequality results in a later age at marriage (the hazard of marriage decreases). In line with the theoretical model presented in Section 2, we expect the coefficient $\beta_2$ to be positive.

We want to use specification (13) to examine whether women in higher-inequality markets get married later than women in lower-inequality markets. Some women in the sample married many years ago, when earnings inequality may have been very different, and current income inequality cannot explain their age at marriage. Further, older women who married recently may have been

[7]Though we recognize that social norms regarding marriage might differ across the years, Equation (13) does not include the current age of the married woman as a control. Age reflects search duration in the marriage market, i.e., in high inequality markets, married women will be, on average, older than in low-inequality markets, or age itself is an outcome variable of the treatment “inequality”. For this reason, the age of the individual is a “bad control”, in the language of Angrist and Pischke (2009), because its inclusion can bias the estimate of the causal effect of male earnings inequality.

on the marriage market for a long time and current earnings inequality may not really explain their late marriages. Our solution to these issues is to restrict the regression sample to women between the ages of 15 and 30, who are more likely to have married recently or be on the marriage market currently.

In Section 2, we noted the caveat that the typical earnings distribution is not normal, and that therefore an increase in the standard deviation while holding the mean constant does not necessarily constitute a mean-preserving spread (recall that the comparative static results in Section 2 pertained to a mean-preserving spread). For this reason, the effect of an increase in the standard deviation in the data may not necessarily correspond to the theorized effect of a mean-preserving spread. We also noted that increases in income inequality in a number of countries, including India, have largely been due to increases in the incomes of highest-earning groups. It is therefore clearly important to consider finer measures of income inequality. Following Loughran (2002), we construct measures of earnings inequality using differences between earnings at different points in the earnings distribution. Specifically, we consider the difference between earnings at the 90th and 50th percentiles, and the difference between earnings at the 50th and the 10th percentile. In Section 2 we theorized that, for most women, increases in income at the upper tail of the income distribution are likely to increase the duration of marital search, whereas reductions in income at the lower tail of the income distribution should not affect search duration. We now test these hypotheses by examining how the percentile-based differences in earnings affect female marital outcomes, while controlling for the median earnings. The regression specifications are:

$$y_{ics} = \alpha + \beta_1(e^{90} - e^{50}) + \beta_2(e^{50} - e^{10}) + \beta_3 e^{50} + \gamma Age_{ics} + \eta_c + \eta_s + u_{ics} \tag{14}$$

$$h\left(t\right) = h_0(t)\exp\left(\alpha + \beta_1(e^{90} - e^{50}) + \beta_2(e^{50} - e^{10}) + \beta_3 e^{50} + \eta_c + \eta_s\right) \tag{15}$$

where $e^{90}$, $e^{50}$, and $e^{10}$ denote the 90th, 50th and 10th percentiles, respectively, of the male earnings distribution in the relevant marriage market. We expect $\beta_1$ to be negative in Equations 14 and 15, while we expect $\beta_2$ to be small and not significantly different from zero in both specifications.

# 5 Results

## 5.1 Effect of male income inequality on female marriage

Table 2 reports the results of the baseline regressions 12 and 13. Column 1 reports the results from estimating Equation 12, while Column 2 reports the results from estimating Equation 13. An increase in the standard deviation of the male earnings distribution by one unit (approximately 10,000 Rupees) reduces the probability of a woman being married by 0.5 percentage points (Column 1). The same change in $\sigma_{ics}$ reduces the hazard of marriage by about 3 percent. It should be recalled that unless the change in $\sigma_{ics}$ represents a mean-preserving spread, the theoretical effect on search duration is ambiguous. For this reason, the results of these regressions should be interpreted with caution. An increase in the mean of the earnings distribution has a small and statistically insignificant effect in both samples. Again, this is consistent with our expectation that while an increase in the mean should technically reduce search duration (and hence increase the probability of being married), in practice the effect is likely to be very small. The coefficient on age in Column 1 indicates the probability of being married increases by 7 percentage points every year. This validates our choice of the Weibull model: the estimate of the ancillary parameter $p$ is significantly greater than one (Column 2), indicating that the hazard of marriage does increase with age. The coefficients on the caste indicators in Column 1 indicate that the probability of being married (conditional on age) is lowest for Brahmins (the omitted caste category), with the probability increasing as we move from other high-caste to OBC, Dalits, and Adivasis. This is consistent with the caste coefficients in Column 2: Brahmin women experience the lowest hazard (age) of marriage, and subsequently face the longest survival time. Their survival time is 1.05 greater than high-caste women.[8] The age-gap increases across the other caste categories, with the largest gap occurring between Brahmins and Adivasis (the survival time is 1.1 times greater for Brahmin women). Note that we restricted our analysis to Hindus only, and did not include Muslims, Sikh, Jain and Christians in the analysis.

We now turn to finer measures of earnings inequality. Table 3 presents the estimates from Equations 14 and 15, Column 1 and Column 2, respectively. Because of the inherent difference in

[8] The Weibull model allows for easy transformation from the proportional hazards (PH) metric to the accelerated failure-time (AFT) metric: $\beta_{AFT} = \frac{-\beta_{PH}}{p}$. The survival time ratio is then calculated as $\exp(\beta_{AFT})$.

scales between the measures of upper-tail and lower-tail earnings inequality, we have converted these measures into Z-scores in order to facilitate a comparison of the coefficients on these variables.[9] The results in Table 3 paint a consistent picture. A unit increase in standardized upper-tail earnings inequality lowers the probability of getting married (at any given age) by 1.6 percentage points, and correspondingly results in the hazard of marriage decreasing by about 7 percent, increasing the age at marriage by a factor of 1.02, or about 0.35 years. However, an increase in lower-tail inequality has a small and statistically insignificant effect on both marriage propensities as well as the age at marriage.

### 5.2 Testing alternative hypotheses

We have established that in high-inequality markets, women tend to postpone marriage. By including state and caste fixed effects in our regressions in Table 3, we attempt to control for unobservable differences between women in high- and low-inequality markets that may arise from socio-economic conditions specific to regions or societal groups. Nonetheless, the main threat to identification of the causal effects remains that earnings inequality may be correlated with unobserved confounding factors that may also affect marriage. In this section, we address this potential omitted variable bias.[10]

We first show that the effects found in Table 3 are not symmetric, i.e., men's probability of being married does not relate to women's earnings inequality (Table 4). We then demonstrate that women in marriage markets with high inequality are on average not different from women in low inequality markets, in terms of characteristics such as the age at which they reached menarche (Table 5). As this variable does not represent a short-term decision that should be affected by earnings inequality, it is suitable for testing for differences in observable characteristics between women in high- and low-inequality markets. Finally, we provide evidence that the results are not due to (i) Increased duration of labor market search on the part of men as a response to greater dispersion, (ii) Male earnings dispersion proxying for expected or current female earnings, or (iii) Financial constraints faced women's families (Tables 6A and 6B).

[9] Define, for instance, $Z_{90-50} = \frac{(X_{90}-X_{50})-\mu_{X_{90}-X_{50}}}{\sigma_{X_{90}-X_{50}}}$.

[10] We also subject the results in Table 3 to a range of robustness tests, using alternate earnings measures and alternate regression samples. The impact of male income inequality on female marital outcomes remains the same. These results are available upon request.

It is possible that the results presented in Table 3 could reflect increased male search duration due to increased female earnings inequality; if male earnings dispersion is correlated with female earnings dispersion, this would appear as a negative correlation between the female marriage age and male income inequality. To investigate this possibility, we re-estimate Equation 14, now considering the effect of female income inequality on the probability of male marriage, in a sense a falsification test of our original specification. Table 4 presents the results. We find that is no symmetric effect: the coefficients for the measures of the upper-and lower-tail of the female earnings distribution are both small and statistically insignificant. Thus men are not searching longer due to increased dispersion in female earnings.

We note that woman in high-and low-inequality markets could be different from each other, resulting in their different marriage patterns. In this section, we verify whether women in high- and low-inequality markets are similar in observable characteristics and whether inclusion of these characteristics changes our results. Specifically, we look at the correlation between the earnings inequality measures and an outcome that is not a decision variable, namely the age at which a woman attained menarche, which are reported for the ever-married women sample. Because this variable does not represent a short-term decision that should be affected by earnings inequality, it is suitable for testing for differences in observable characteristics between women in high- and low-inequality markets. Age at menarche does reflect childhood nutrition, however, so may be correlated with earnings inequality if the latter is correlated with unobserved socio-economic characteristics. This variable is of particular interest because women typically get married only after they attain menarche (indeed, in our sample of ever-married women, age at menarche is a very strong predictor of age at marriage). If individuals in marriage markets with greater earnings inequality also happen to be associated with poor childhood nutrition (which would act to delay menarche), this might result in a spurious positive correlation between earnings inequality and age at marriage.

Table 5 presents the results of examining whether male earnings inequality predicts age at menarche (Column 1). The results are negative: the coefficients on the earnings inequality variables are small and statistically insignificant, confirming that women in high inequality markets are not different from women in low inequality markets in terms of observable characteristics. In accordance with these results, including age at menarche as a control in regression 15 does not significantly affect the estimated effect of male earnings inequality on age at marriage, as shown in Column 2 of Table 5.

Table 6A reports the regression results of Equation 14, with the following additional controls included: male/female sex-ratio (defined as the number of unmarried (eligible) men to the number of unmarried (eligible) women) (Column 2), female earnings dispersion (Column 3), and mean expense the bride's family usually incurs for a wedding (Column 4). Table 6B reports the corresponding regression results for estimation of Equation 15.

The inclusion of sex-ratio is motived as follows. It is possible that increasing the dispersion of the male earnings distribution may extend the labor market search of men and thereby delay their own entry into the marriage market (or make them ineligible for marriage). This may reduce the number of available men in the marriage market, possibly delaying female marriage. Alternatively, it is plausible that men only join the marriage market when they reach a minimum earnings level. An increased dispersion of earnings, could increase the number of men on the market in this case, affecting the probability of marriage for women. Finally, it might be possible that men in high-inequality areas move to areas with low-inequality, hence reducing the number of available men in high-inequality areas. Column 2 in Tables 6A and 6B show that inclusion of sex-ratio has little impact on the coefficients of the earnings inequality variables. We note however that the sex-ratio is also likely to be endogenous and its coefficient should not be interpreted causally.

It may also be the case that inequality in the upper-tail of the male earnings distribution captures the skill/education premia, and these returns to higher education may motivate women or men to pursue higher education (delaying their marriage). To address this we control for measures of female earnings dispersion (Column 3 of Tables 6A and 6B). We note, however, that the female earnings variables are potentially endogenous. Nonetheless, our intention is not to interpret the coefficients on these variables, but to examine whether their inclusion affects the coefficients of the male earnings inequality variables. The results are qualitatively the same as the baseline results of Table 3, although the estimated effect of earnings inequality on the age at marriage is now larger than in previous specifications.

We also consider the possibility that increased male earnings dispersion at the upper-tail may actually induce matches to form more quickly if women and their families are making offers to secure high-earning men. However, the realization of the match, i.e. the marriage, may be delayed as it takes time to afford the wedding, especially if more attractive candidates ask for greater dowries, thereby resulting in a spurious positive correlation between earnings inequality and age at marriage. To address this, we control for the average expenditure the bride's family incurs for a wedding. The inclusion of wedding expenses has little impact on the coefficients on the earnings inequality measures (Column 4 of Tables 6A and 6B). Again we note the caveat that marriage expenditures may be endogenous and its coefficient should not be interpreted causally.

# 6 Extension: Effect of male income inequality on female education

## 6.1 Empirical specifications

In an extension, we hypothesize that to the extent increased male inequality delays marriage, it may also have an impact on female educational attainment by keeping women in school longer. To test this hypothesis, we employ the following specification:

$$EduYears_{ics} = \alpha + \beta_1(e^{90} - e^{50}) + \beta_2(e^{50} - e^{10}) + \beta_3 e^{50} + \eta_c + \eta_s + u_{ics} \tag{16}$$

where we use $EduYears_{ics}$, the completed educational attainment of ever-married women, as the dependent variable to determine how many extra years of education are added by greater earnings inequality. As in previous specifications, $e^{90}$, $e^{50}$, and $e^{10}$ denote the 90th, 50th and 10th percentiles of the male earnings distribution in the relevant marriage market, and we account for caste and state effects.

Next, we attempt to determine where in the educational trajectory the additional years (if any) accrue. To do this, we construct a set of nested indicator variables for different levels of educational attainment: the first takes the value 1 if the woman received zero years of education and 0 otherwise; the second takes the value 1 if the woman did not complete primary school (i.e. received fewer than 5 years of education); the third takes the value 1 if the woman did not matriculate at middle school (i.e. received fewer than 8 years of education); the fourth takes the value 1 if the woman did not matriculate at high school (i.e. received fewer than 10 years of education); the fifth takes the value 1 if the woman did not complete high school (i.e. received fewer than 12 years of education); the sixth takes the value 1 if the woman did not complete her college education (i.e. received fewer than 15 years of education). We then estimate linear probability models on the ever-married sample, using these indicators as dependent variables:

$$level_{ics} = \alpha + \beta_1(e^{90} - e^{50}) + \beta_2(e^{50} - e^{10}) + \beta_3 e^{50} + \eta_c + \eta_s + u_{ics} \tag{17}$$

where $level_{ics}$ is an indicator for each level in the education trajectory, and $e^{90}$, $e^{50}$, and $e^{10}$ are our measures of male income inequality.

The hypothesis of the marital search model suggests that increased search duration due to widening earnings inequality should largely impact educational attainment at the later stages of the educational trajectory because that is when women are likely to be on the marriage market, i.e. it may affect rates of high-school matriculation and completion (and even possibly college education), but is unlikely to move women from zero to positive levels of education or to raise primary school completion rates.[11] In a sense, this provides another falsification test for our hypothesis. If greater earnings inequality is associated with a significant proportion of women completing lower levels of education, then this would suggest either that marital search duration is not the channel by which earnings inequality affects education and/or that our measures of earnings inequality are correlated with unobservable socio-economic characteristics that affect education (and possibly marriage) not accounted for in our current specifications.

We extend this logic to consider if male income inequality affects the age at marriage for women with no or little education. If in fact women are acquiring more education prior to entering the marriage market, then women with little educational attainment should not respond to changes in the male income distribution. We test this hypothesis using a variation of Equation 15: we interact the male income inequality measures, $(e^{90} - e^{50})$ and $(e^{50} - e^{10})$, with an indicator for women's educational attainment. In one version the indicator is equal to 1 if a woman has no education (0 years); in another version the indicator is equal to 1 if a woman did not complete primary school. These variations allow us to better identify the order of events: whether women acquire more education in response to increased search duration, as the marriage search model suggests, or if educational attainment occurs before entering the marriage market. Within the our

[11] We do not entirely rule out the possibility, however, because there is a small percentage of women who are enrolled in school in their late teens, but have only completed four or five years of education. For these individuals an increase in search duration may result in primary school completion.

search framework, we anticipate that even among women with no or little education, their age at marriage will increase in response to greater male earnings dispersion.

Finally, we implement a falsification test that examines the schooling of girls who are not yet of marriageable age. If, for this group, greater earnings inequality is associated with more years of education and/or a greater propensity to be enrolled in school, then this would again suggest either that the channel of influence is not marital search duration or that earnings dispersion is correlated with unobserved socio-economic characteristics that affect education. To implement this test, we estimate regression specification 14, but now consider (i) Completed years of education and (ii) Current school enrollment status for unmarried girls under the age of 10 (who are not likely to be in the marriage market).

### 6.2 Results: Effect on female education

To estimate the effect of earnings inequality on female educational attainment, we restrict ourselves to the ever-married sample, as these women have completed their education and are no longer in school. Table 7 reports the regression results for the educational attainment variables. The dependent variables in Columns 1-7 are, respectively, (i) Completed years of education, (ii) Indicator for zero years of education, (iii) Indicator for having obtained fewer than 5 years of education, (iv) Indicator for having obtained fewer than 8 years of education, (v) Indicator for having obtained fewer than 10 years of education, (vi) Indicator for having obtained fewer than 12 years of education, and (vii) Indicator for having obtained fewer than 15 years of education. A doubling of 90th percentile earnings relative to the median increases completed years of education by 0.6 years. Turning to the specific attainment variables, we find that earnings inequality has small and statistically insignificant effects on lower-level educational attainment, but it has a significant effect on high-school completion and college matriculation rates. In particular, the probability of completing high-school (Column 6) increases by 8 percentage points and the probability of entering college increases by 6 percentage points (Column 7).

One also may hypothesize that women get married later because they decided to get more education. That is, the total search duration remains the same, but women enter the marriage market later, after acquiring more education. We present results excluding this possibility in Table

8, by examining the impact of increased earnings dispersion on the age at marriage for women who completed their educational attainment prior to entering the marriage market. We find that even women who have never been to school (who have zero years of education) or who have very little education (did not complete primary school) delay marriage to a greater extent as women with higher educational attainment. Women with zero years of education experience a 18 percent lower risk of marriage: they delay marriage by a factor of 1.33 (Column 1). Women who have not completed primary school face a 13 percent lower hazard, delaying marriage by a factor of 1.25 (Column 2).

Table 9 examines the effect of earnings inequality on educational attainment and current school enrollment of unmarried children below the age of 10. Because these children are plausibly not on the marriage market, earnings inequality should not affect these educational outcomes. The results in Table 9 confirm this hypothesis: neither current school enrollment nor completed years of education are correlated with male earnings inequality for this sample.

These results are consistent with the hypothesis that greater earnings inequality affects female marriage by increasing search duration, and therefore impacts educational attainment at the point of the educational trajectory that women find themselves at when they enter the marriage market.

## 7 Conclusion

We use a nationally representative dataset from India to test whether women delay their marriage as a response to increased earnings inequality of men. In line with the predictions from a marital search model, we find that increases in upper tail inequality delay marriage, while increases in lower tail inequality have no significant effect. There is a corresponding effect on educational attainments, with women in high-inequality markets being more likely to obtain higher education. Marriage is delayed even for women who attained no education, ruling out the hypothesis that greater inequality may delay marriage by causing women to seek higher education, i.e. the direction of effect runs from marriage to education, not the other way around.

In particular, we find that a unit increase in standardized upper-tail earnings inequality lowers the propensity to get married (at any given age) by 1.6 percentage points, and correspondingly results in a delay of marriage by about 0.35 years. Comparing these results with the US context: Gould and Paserman (2003) using US census data from 1970 till 1990 find that higher male inequality in a city lowers the marriage rate of women, and that this effect is not unimportant: increasing male inequality explains about 30 percent of the marriage rate decline for women of the last few decades. Loughran (2002), using the same data, finds that rising within-group make wage inequality accounts for anywhere between 5 and 35 percent of the decline in the age-specific female propensity to marry between 1970 and 1990, the strongest results being for women age 22-28 with less than a college education.

The richness of the data allows us to test and rule out a number of alternative hypotheses for these findings. The results are not due to (i) Men searching longer in the marriage market in response to greater female earnings inequality, (ii) Regional or caste-base social norms, (iii) Men searching longer in the labor market (reducing the gender ratio in the marriage market), (iv) Earnings dispersion proxying for educational premia, and hence encouraging women to stay in school longer, or (v) Women's families needing more time to afford greater dowries. While we have little information on the woman's natal family in the data, e.g. we cannot control for variables such as natal family composition and wealth, we do establish that women in high and low-inequality markets are not different with respect to some observable characteristics: height and age at menarche.

There are clear long-term implications of the phenomenon we have documented. The acquisition of more education due to increased search duration on the marriage market might increase the future income stream of women, contributing to their economic well-being. In addition, greater educational attainment may improve their bargaining power in the household, allowing women to direct resources towards children's, especially girls', education and health, thus affecting the skills and productivity of the future workforce.

## References


[1] Angrist, Joshua and J-S Pischke. 2009. Mostly Harmless Econometrics. Princeton University Press.

[2] Atkinson, Anthony B., Thomas Piketty and Emmanuel Saez, 2011. "Top Incomes in the Long Run of History."Journal of Economic Literature 49, no. 1: 3-71.

[3] Banerjee, Abhijit, and Thomas Piketty. 2005. "Top Indian Incomes, 1922–2000."The World Bank Economic Review 19, no. 1: 1-20.

[4] Brien, Michael J., and Lee A. Lillard. 1994. "Education, Marriage, and First Conception in Malaysia."Journal of Human Resources 29, no. 4: 1167-1204.

[5] Bruce, Judith. 2003. "Married Adolescent Girls: Human Rights, Health, and Developmental Needs of a Neglected Majority."Economic and Political Weekly, October 11: 4378-4380.

[6] Burdett, K. and J. Ondrich. "How Changes in Labor Demand Affect Unemployed Workers," Journal of Labor Economics, 3 (January): 1-10.

[7] Caldwell, J.C., Reddy, P.H. & Caldwell, P. 1983. "Causes of Marriage Change in South India." Population Studies 37(3), 343-361

[8] Choe, Kim, Sidney B. Westley, and Robert D. Rutherford. 2002. "Tradition and Change in Marriage and Family Life". In: The Future of Population in Asia. Honolulu, HI: East-West Center.

[9] Clark, Shelley. 2004. "Early Marriage and HIV Risks in Sub-Saharan Africa."Studies in family planning 35, no. 3: 149-160.

[10] Deaton, Angus, and Jean Dreze. 2002. "Poverty and Inequality in India: a Re-Examination."Economic and Political Weekly, September 7: 3729-3748.

[11] Deaton, Angus, and Valerie Kozel. 2005. "Data and Dogma: The Great Indian poverty Debate."The World Bank Research Observer 20, no. 2: 177-199.

[12] Field, Erica, and Attila Ambrus. 2008. “Early Marriage, Age of Menarche, and Female Schooling Attainment in Bangladesh.”Journal of Political Economy 116, no. 5: 881-930.

[13] Glick, Peter J., and David E. Sahn. 2002. “Are Africans Practicing Safer Sex? Evidence from Demographic and Health Surveys for Either Countries. ”Economic Development and Cultural Change 56, no. 2: 397-439.

[14] Gould, Eric D., and M. Daniele Paserman. 2003. “Waiting for Mr. Right: Rising Inequality and Declining Marriage Rates.” Journal of Urban Economics 53, no.2: 257-281.

[15] Harwood-Lejeune, Audrey. 2001. “Rising age at Marriage and Fertility in Southern and Eastern Africa.”European Journal of Population/Revue Européenne de Démographie 17, no. 3: 261-280.

[16] Heaton, Tim B., Renata Forste, and Samuel M. Otterstrom. 2002. “Family Transitions in Latin America: First Intercourse, First Union and First Birth.”International Journal of Population Geography 8, no. 1: 1-15.

[17] Jensen, Robert, and Rebecca Thornton. 2003. “Early Female Marriage in the Developing World.”Gender & Development 11, no. 2: 9-19.

[18] Lindstrom, David P., and Carlos Brambila Paz. 2001. “Alternative Theories of The Relationship of Schooling and Work to Family Formation: Evidence from Mexico.”Biodemography and Social Biology 48, no. 3-4: 278-297.

[19] Loughran, David S. 2000. “Does Variance Matter? The Effect of Rising Male Wage Inequality on Female age at First Marriage.” RAND Working Paper.

[20] Loughran, David S. 2002. “The Effect of Male Wage Inequality on Female Age at First Marriage.”Review of Economics and Statistics 84, no. 2: 237-250.

[21] Maertens, Annemie. 2013. “Social Norms and aspirations: age of Marriage and Education in Rural India.” World development 47. No.C : 1-15.

[22] McCall, John J. 1970. “Economics of Information and Job Search.” Quarterly Journal of Economics 84, no.1: 113-126.

[23] Mensch, Barbara S., Susheela Singh and John B. Casterline. 2005. "Trends in the Timing of First Marriage among Men and Women in the Developing World," in Cynthia B. Lloyd, Jere R. Behrman, Nelly P. Stromquist, and Barney Cohen (eds.), The Changing Transitions to Adulthood in Developing Countries: Selected Studies, Washington, DC: National Academies Press, 118–171.

[24] Mortensen, Dale T. 1986. "Job Search and Labor Market Analysis," in Handbook of Labor Economics. O. Ashenfelter and R. Layard, eds. Amsterdam: North Holland, 849–920.

[25] Nour, Nawal M. 2006. "Health Consequences of Child marriage in Africa."Emerging infectious diseases 12, no. 11: 1644-1649.

[26] Parthapratim Pa, and Jayati Ghosh. 2007. "Inequality in India: A Survey of Recent Trends." Working Papers 45, United Nations, Department of Economics and Social Affairs.

[27] Raj, Anita, Niranjan Saggurti, Donta Balaiah, and Jay G. Silverman. 2009. "Prevalence of Child Marriage and its Effect on Fertility and Fertility-Control Outcomes of Young Women in India: a Cross-Sectional, Observational Study."The Lancet 373, no. 9678: 1883-1889.

[28] Rashad, Hoda and Osman, Magued. 2003. "Nuptiality in Arab countries: Changes and implications."In Nicholas Hopkins (eds) The New Arab Family. Cairo: The American University in Cairo Press: 20-50.

[29] Santhya, K. G., Usha Ram, Rajib Acharya, Shireen J. Jejeebhoy, Faujdar Ram, and Abhishek Singh. 2010. "Associations between Early marriage and Young Women's Marital and Reproductive Health Outcomes: Evidence from India."International Perspectives on Sexual & Reproductive Health 36, no. 3: 132-139.

[30] Srinivas, M. N. 1984. Some Reflections on Dowry. New Delhi: Oxford University Press.

[31] Sundaram, Krishnamurty, and Suresh D. Tendulkar. 2003A. "Poverty in India in the 1990s: Revised results for all-India and 15 major states for 1993-94."Economic and Political Weekly, December 13-19: 4865-4872.

[32] Sundaram, Krishnamurty, and Suresh D. Tendulkar. 2003B. "Poverty in India in the 1990s: An Analysis of Changes in 15 Major States." Economic and Political Weekly, April 5: 1385-1393.

[33] Westoff, Charles F. 2003. Trends in Marriage and Early Childbearing in Developing Countries. DHS Comparative Reports No. 5. Calverton, Maryland: ORC Macro.

# Appendix (Proofs)

We provide outlines for the proofs of the results cited in Section 2.2. We showed that the reservation earnings $R$ could be obtained by solving the equation:

$$[r+1-F(R)]R = rc + \int_R^\infty x dF(x) \tag{18}$$

where $r = \frac{1-\beta}{\beta}$.

We now assume that the $F$ distribution is normal with mean $\mu$ and standard deviation $\sigma$. We can now apply the standard formula for the mean of a truncated normal distribution to write:

$$\begin{aligned} \int_R^\infty x dF(x) &= P(x > R) \cdot E(x|x > R) = P(x > R) \cdot \left[\mu + \sigma \frac{\phi\left(\frac{R-\mu}{\sigma}\right)}{1 - \Phi\left(\frac{R-\mu}{\sigma}\right)}\right] \\ &= \mu\left[1 - \Phi\left(\frac{R-\mu}{\sigma}\right)\right] + \sigma\phi\left(\frac{R-\mu}{\sigma}\right) \end{aligned} \tag{19}$$

Thus, we have, from (18):

$$\left[r + 1 - \Phi\left(\frac{R-\mu}{\sigma}\right)\right] R = rc + \mu\left[1 - \Phi\left(\frac{R-\mu}{\sigma}\right)\right] + \sigma\phi\left(\frac{R-\mu}{\sigma}\right) \tag{20}$$

We now assume that the initial (starting) distribution of $F$ is $N(0,1)$. This simplifies some of the algebra. Using Equation (20) above, and noting that $\phi'(z) = -z\phi'(z)$, one can obtain the following derivatives:

$$\frac{dR}{dc} = \frac{r}{q} > 0 \tag{21}$$

$$\frac{dR}{d\mu} = \frac{q}{r+q} < 1 \tag{22}$$

$$\frac{dR}{d\sigma} = \frac{\phi(R)}{r+q} > 0 \tag{23}$$

These were the comparative static results presented in Section 2.2.

**Figure 1: Average Age at Marriage (in India), 1980-2005**

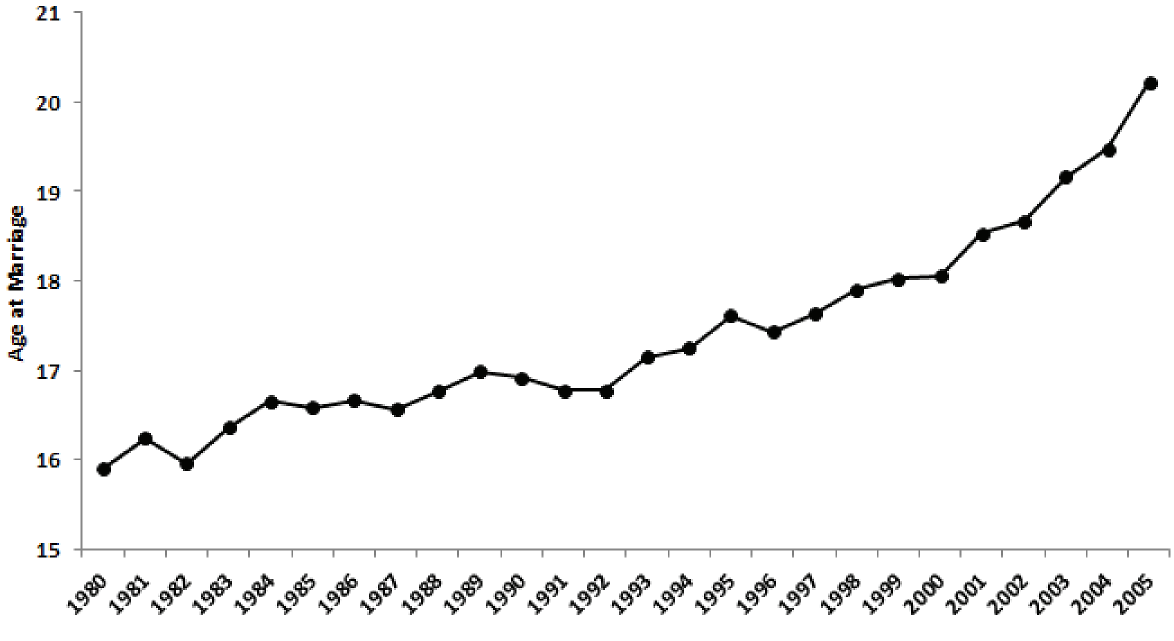



Figure 1 depicts the average age at marriage for women in India, from 1980-2005. Over this period, we see an increase in the marriage age from 16.25 years to 20.36 years.
*Source*: Authors' calculation, IHDS data.

**Figure 2: Completed Years of Education, 1980-2005**

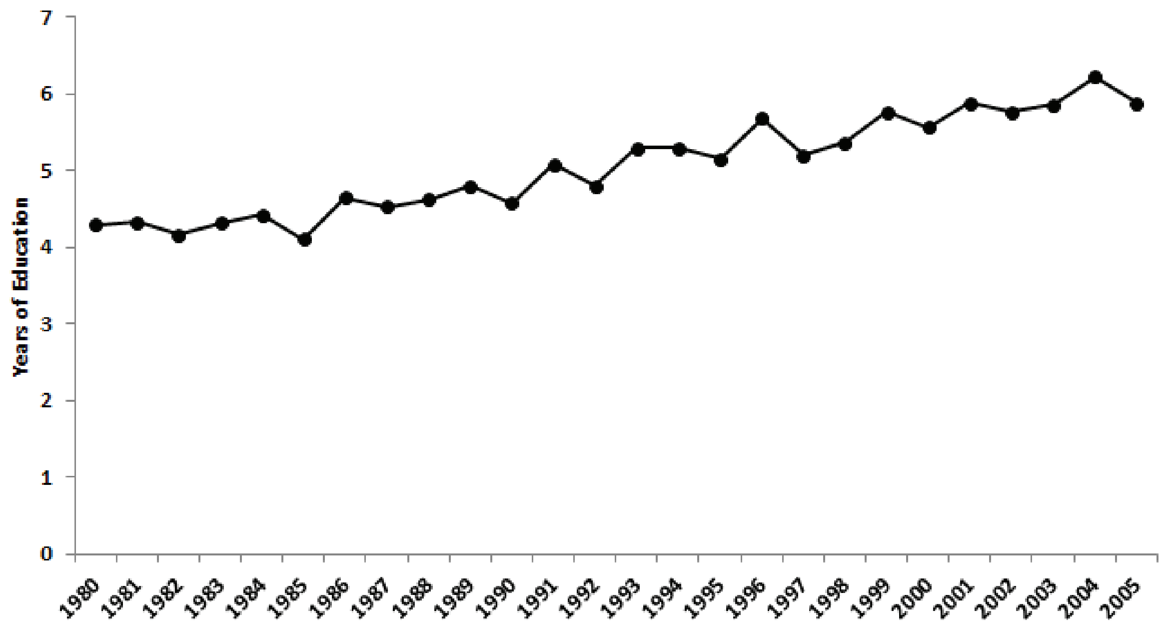



Figure 2 depicts the average years of educational attainment amongst married women in India, from 1980-2005. Over this period, we see an increase in the years of education from 4.65 to 6.47. We note that this increase in educational attainment does not fully explain the four year increase in the marriage age of women depicted in Figure 1.
*Source*: Authors' calculation, IHDS data.

**Figure 3: Functional form of $R - \frac{dR}{d\sigma}$ as a function of R (reservation level) (in red)**

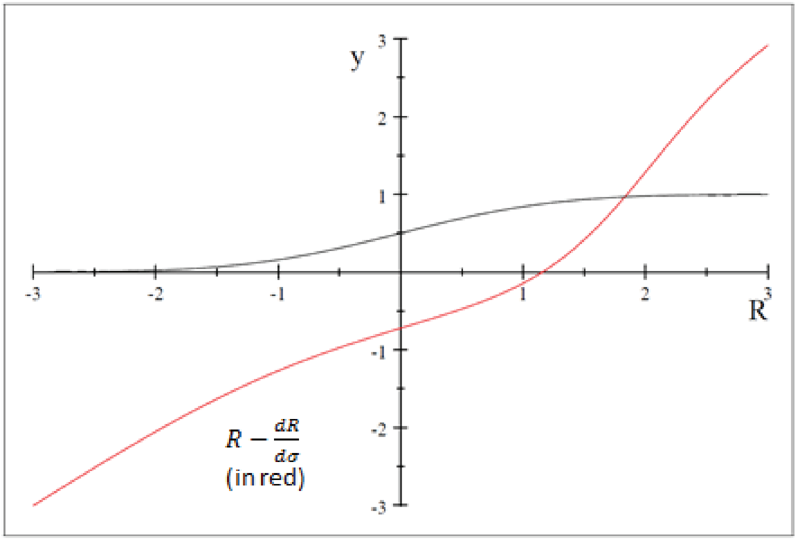



Figure 3 graphs $R - \frac{dR}{d\sigma}$ against different values of $R$ (where we have assumed a discount rate of $\beta = 0.95$). The figure also graphs the standard normal cumulative distribution function. The figure shows that $R - \frac{dR}{d\sigma}$ is negative over a substantial range of values of $R$, implying that increasing earnings dispersion most likely increases search duration and the age at marriage.

**Figure 4: Marriage Market Characteristics**

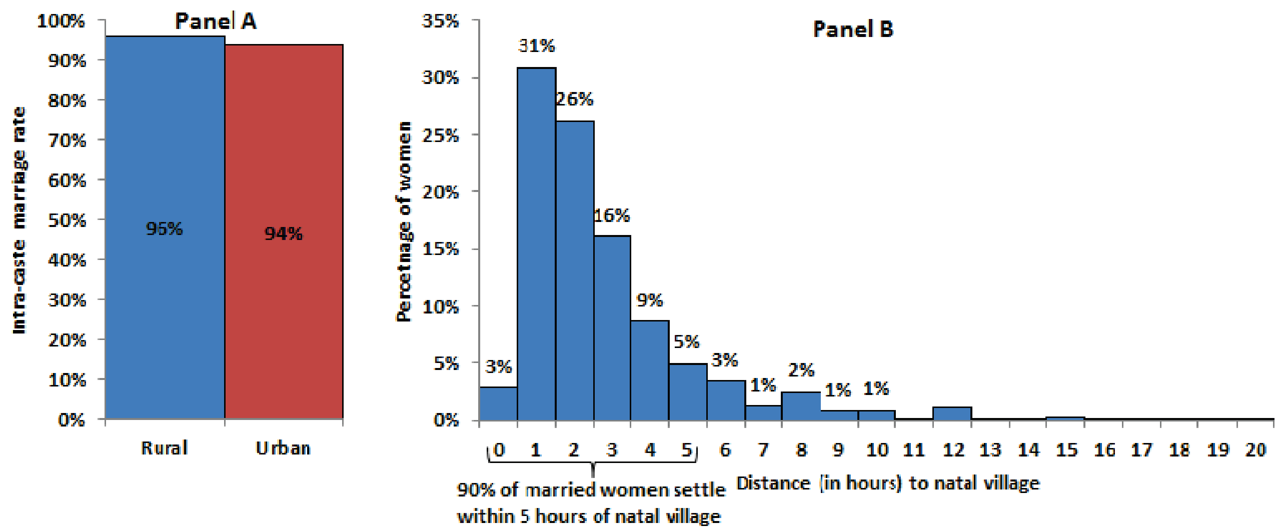



Figure 4 depicts the marriage market in India. Panel A shows that the intra-caste marriage rate is high: 96% in rural areas, and 94% in urban areas. Panel B shows that after marriage, women in India settle very close to their place of birth. (Traveling distance refers to travel by road or train, not plane.) Combined, this provides justification to define the marriage market according to a woman's caste and geographical location.
*Source*: Authors' calculation, IHDS data.

**Table 1: Introducing the "ever-married" women sample**

| | Mean | Std. Dev. |
|---|---|---|
| Age of ever-marrried woman (years) | 33.0 | 8.0 |
| Age at marriage of ever-married woman (years) | 17.53 | 3.79 |
| Number of children of ever-married woman | 2.29 | 1.44 |
| Number of household members | 5.2 | 2.5 |
| Number of children (0-14 years) in the household | 1.6 | 1.6 |
| Years of education (adult $>$ 21 years) | 7.5 | 5.1 |
| Monthly Consumption per capita (in Rs) | 953 | 1,024 |
| Income (yearly, in Rs) | 53,922 | 83,284 |
| Asset index (range 0 to 30) | 12.3 | 6.3 |

Table 1 presents descriptive statistics on the ever-married sample of women. The household is defined as the persons who live under the same roof and share the same kitchen for 6+ months. Asset index: IHDS asked a series of questions about what goods the household owned and about the quality of the housing. This index sums 30 dichotomous items measuring household possessions and housing quality. The age of marriage is noted for women who are married just once. The number of children only includes the children of the woman living in the household.
*Source*: Authors' calculation, IHDS data.

**Table 2: Effect of male earnings inequality (standard deviation and mean) on female marriage**

| | (1) Probability of marriage | (2) Age at marriage |
|---|---|---|
| Standard deviation of male earnings | -0.005$^{***}$ | -0.031$^{***}$ |
| | (0.001) | (0.011) |
| Mean of male earnings | -0.005 | 0.066 |
| | (0.008) | (0.046) |
| Age of woman (years) | 0.068$^{***}$ | |
| | (0.001) | |
| (Other) High-caste indicator | 0.048$^{***}$ | 0.261$^{***}$ |
| | (0.013) | (0.085) |
| OBC indicator | 0.078$^{***}$ | 0.489$^{***}$ |
| | (0.014) | (0.079) |
| Dalit indicator | 0.104$^{***}$ | 0.594$^{***}$ |
| | (0.016) | (0.085) |
| Adivasi indicator | 0.114$^{***}$ | 0.680$^{***}$ |
| | (0.021) | (0.109) |
| Constant | -0.946$^{***}$ | -20.032$^{***}$ |
| | (0.032) | (0.873) |
| Ancillary parameter | | |
| ln($p$) | | 1.804$^{***}$ |
| | | (0.046) |
| State fixed effects included? | Yes | Yes |
| N (women) | 25,549 | 24,888 |
| R-squared | 0.451 | |

Table 2 reports the results of the baseline regressions. Column 1 reports the results from estimating the impact of male income inequality on female marital status, using a Linear Probability Model. Column 2 reports the results from estimating a hazard model (Weibull specification) to understand the impact on the age at marriage. We find that an increase in the standard deviation of the male earnings distribution by one unit (10,000 Rupees) reduces the probability of a woman being married by 0.5 percentage points (Column 1), and reduces the hazard of marriage by 3 percent (age at marriage increases) (Column 2). Notes: The excluded caste category is Brahmins; Standard errors are clustered by marriage market; $^{*}$ $p < 0.1$, $^{**}$ $p < 0.05$, $^{***}$ $p < 0.01$.

**Table 3: Effect of male earnings inequality (upper and lower tails) on female marriage**

| | (1) Probability of marriage | (2) Age at marriage |
|---|---|---|
| Difference in male earnings: $90^{th}$-$50^{th}$ percentile | -0.016*** | -0.067* |
| | (0.006) | (0.038) |
| Difference in male earnings: $50^{th}$-$10^{th}$ percentile | -0.007 | -0.033 |
| | (0.009) | (0.057) |
| Male earnings at $50^{th}$ percentile | 0.014 | 0.132** |
| | (0.013) | (0.060) |
| Age of woman (years) | 0.068*** | |
| | (0.001) | |
| (Other) High-caste indicator | 0.043*** | 0.204*** |
| | (0.014) | (0.078) |
| OBC indicator | 0.108*** | 0.551*** |
| | (0.016) | (0.075) |
| Dalit indicator | 0.138*** | 0.656*** |
| | (0.016) | (0.067) |
| Adivasi indicator | 0.156*** | 0.759*** |
| | (0.021) | (0.094) |
| Constant | -1.017*** | -20.129*** |
| | (0.031) | (0.901) |
| Ancillary parameter | | |
| $\ln(p)$ | | 1.809*** |
| | | (0.045) |
| State fixed effects included? | Yes | Yes |
| N (women) | 25,550 | 24,889 |
| R-squared | 0.451 | |

Table 3 reports the effect of the difference in male earnings between the $90^{th}$ and $50^{th}$ percentile and the difference between the $50^{th}$ and the $10^{th}$ percentile on women's marital outcomes. Column 1 reports the results from estimating the impact on female marital status, using a Linear Probability Model. Column 2 reports the results from estimating a hazard model (Weibull specification) to understand the impact on the age at marriage. Because of the inherent difference in scales between the measures of upper-tail and lower-tail earnings inequality, we converted these measures into Z-scores ( $Z \text{ score} = \frac{x-\mu}{\sigma}$ ) in order to facilitate a comparison of the coefficients on these variables. The results in Table 3 paint a consistent picture. A unit increase in standardized upper-tail earnings inequality lowers the propensity to get married (at any given age) by 1.6 percentage points, and correspondingly results in a 7 percent reduction in the hazard of marriage (equivalent to a delay of approximately 0.35 years). However, an increase in lower-tail inequality has a small and statistically insignificant effect on both marriage propensity and the age at marriage. Notes: The excluded caste category is Brahmins; Standard errors are clustered by marriage market; $^{*}$ $p < 0.1$, $^{**}$ $p < 0.05$, $^{***}$ $p < 0.01$.

**Table 4: Effect of female earnings inequality on male marriage**

| | (1) Probability of marriage |
|---|---|
| Difference in female earnings: $90^{th}$-$50^{th}$ percentile | -0.001 |
| | (0.004) |
| Difference in female earnings: $50^{th}$-$10^{th}$ percentile | -0.004 |
| | (0.008) |
| Female earnings at $50^{th}$ percentile | $0.020^{*}$ |
| | (0.012) |
| Age of man (years) | $0.047^{***}$ |
| | (0.000) |
| High-caste indicator | $0.030^{**}$ |
| | (0.012) |
| OBC indicator | $0.080^{***}$ |
| | (0.010) |
| Dalit indicator | $0.112^{***}$ |
| | (0.011) |
| Adivasi indicator | $0.123^{***}$ |
| | (0.011) |
| Constant | $-0.798^{***}$ |
| | (0.017) |
| State fixed effects included? | Yes |
| N (men) | 37,841 |
| R-squared | 0.546 |

Table 4 examines the effect of female earnings inequality on the probability of male marriage using a Linear Probability Model, and shows that there is no symmetric effect: the coefficients for the measures of the upper- and lower-tail of the female earnings distribution are both small and statistically insignificant. Notes: The excluded caste category is Brahmins; Standard errors are clustered by marriage market; $^{*}$ $p < 0.1$, $^{**}$ $p < 0.05$, $^{***}$ $p < 0.01$. See also notes to Table 3.

**Table 5: Do pre-marriage market characteristics correlate with male earnings dispersion?**

| | (1) Age at menarche | (2) Age at marriage |
|---|---|---|
| Difference in male earnings: $90^{th}$-$50^{th}$ percentile | -0.012 | -0.061$^{*}$ |
| | (0.024) | (0.035) |
| Difference in male earnings: $50^{th}$-$10^{th}$ percentile | -0.052 | -0.007 |
| | (0.041) | (0.058) |
| Male earnings at $50^{th}$ percentile | 0.093 | 0.147$^{***}$ |
| | (0.064) | (0.055) |
| High-caste indicator | 0.270$^{***}$ | 0.138$^{*}$ |
| | (0.087) | (0.074) |
| OBC indicator | 0.355$^{***}$ | 0.469$^{***}$ |
| | (0.114) | (0.073) |
| Dalit indicator | 0.416$^{***}$ | 0.551$^{***}$ |
| | (0.110) | (0.062) |
| Adivasi indicator | 0.526$^{***}$ | 0.681$^{***}$ |
| | (0.135) | (0.093) |
| Age at menarche | | -0.056$^{***}$ |
| | | (0.017) |
| Constant | -28.722$^{***}$ | -23.411$^{***}$ |
| | (3.591) | (1.165) |
| Ancillary parameter | | |
| $\ln(p)$ | 2.345$^{***}$ | 2.032$^{***}$ |
| | (0.130) | (0.049) |
| State fixed effects included? | Yes | Yes |
| N (women) | 22,754 | 22,752 |

Table 5 examines whether male earnings inequality predicts a woman's age at menarche (Column 1) and whether including this variable into the analysis of age at marriage alters the results (Column 2). The coefficients on the earnings inequality variables are small and statistically insignificant, confirming that women in high inequality markets are not different from women in low inequality markets in terms of observable characteristics, and that including age at menarche as a control in the regression does not significantly affect the estimated effect of male earnings inequality on the age at marriage. Notes: The excluded caste category is Brahmins; Standard errors are clustered by marriage market; $^{*}$ $p < 0.1$, $^{**}$ $p < 0.05$, $^{***}$ $p < 0.01$. See also notes to Table 3.

**Table 6: Testing alternative hypotheses**

**(a) Probability of marriage**

| | (1) | (2) | (3) | (4) |
|---|---|---|---|---|
| Male earnings: $90^{th}$-$50^{th}$ | -0.016*** | -0.014** | -0.021*** | -0.013** |
| | (0.006) | (0.005) | (0.007) | (0.006) |
| Male earnings: $50^{th}$-$10^{th}$ | -0.007 | 0.002 | -0.012 | -0.005 |
| | (0.009) | (0.008) | (0.009) | (0.009) |
| Male earnings: $50^{th}$ | 0.014 | 0.002 | 0.010 | 0.011 |
| | (0.013) | (0.011) | (0.013) | (0.013) |
| Age of woman (years) | 0.068*** | 0.068*** | 0.068*** | 0.068*** |
| | (0.001) | (0.001) | (0.001) | (0.001) |
| Male:Female ratio | | 0.044*** | | |
| | | (0.011) | | |
| Female earnings: $90^{th}$-$50^{th}$ | | | 0.012** | |
| | | | (0.006) | |
| Female earnings: $50^{th}$-$10^{th}$ | | | -0.006 | |
| | | | (0.010) | |
| Female earnings: $50^{th}$ | | | 0.017 | |
| | | | (0.014) | |
| Wedding expenditure | | | | -0.003 |
| | | | | (0.002) |
| Constant | -1.017*** | -1.067*** | -1.018*** | -0.962*** |
| | (0.031) | (0.033) | (0.032) | (0.047) |
| Caste fixed effects included? | Yes | Yes | Yes | Yes |
| State fixed effects included? | Yes | Yes | Yes | Yes |
| N (women) | 25,550 | 25,530 | 25,550 | 25,550 |
| R-squared | 0.451 | 0.451 | 0.451 | 0.451 |

**(b) Age at marriage**

| | (1) | (2) | (3) | (4) |
|---|---|---|---|---|
| Male earnings: $90^{th}$-$50^{th}$ | -0.067* | -0.067* | -0.125*** | -0.067* |
| | (0.038) | (0.038) | (0.040) | (0.035) |
| Male earnings: $50^{th}$-$10^{th}$ | -0.033 | -0.026 | -0.083 | -0.033 |
| | (0.057) | (0.061) | (0.058) | (0.058) |
| Male earnings: $50^{th}$ | 0.132** | 0.129* | 0.096 | 0.132** |
| | (0.060) | (0.066) | (0.061) | (0.062) |
| Male:Female ratio | | 0.189* | | |
| | | (0.114) | | |
| Female earnings: $90^{th}$-$50^{th}$ | | | 0.118*** | |
| | | | (0.029) | |
| Female earnings: $50^{th}$-$10^{th}$ | | | 0.131** | |
| | | | (0.056) | |
| Female earnings: $50^{th}$ | | | -0.149* | |
| | | | (0.082) | |
| Wedding expenditure | | | | -0.000 |
| | | | | (0.012) |
| Constant | -20.129*** | -20.279*** | -20.398*** | -20.126*** |
| | (0.901) | (0.879) | (0.763) | (0.953) |
| Ancillary parameter | | | | |
| ln($p$) | 1.809*** | 1.810*** | 1.823*** | 1.809*** |
| | (0.045) | (0.045) | (0.040) | (0.045) |
| Caste fixed effects included? | Yes | Yes | Yes | Yes |
| State fixed effects included? | Yes | Yes | Yes | Yes |
| N (women) | 24,889 | 24,870 | 24,889 | 24,889 |

Table 6 reports the regression results from estimating the impact on marital status, using a Linear Probability Model (6a) and from estimating a hazard model to understand the impact on the age at marriage (6b), with the following additional controls included: male/female sex-ratio (defined as the number of unmarried (eligible) men to the number of unmarried (eligible) women) (Column 2), female earnings dispersion (Column 3), and the average amount the bride's family spends for a wedding (by marriage market) (Column 4). The results are qualitatively the same as the baseline results of Table 3. Notes: Standard errors are clustered by marriage market; * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. See also notes to Table 3.

**Table 7: The effect of male earnings inequality on female educational attainment (ever-married sample)**

| | (1) Years of education | (2) No education | (3) Under 5 years | (4) Under 8 years | (5) Under 10 years | (6) Under 12 years | (7) Under 15 years |
|---|---|---|---|---|---|---|---|
| Difference in male earnings: $90^{th}$-$50^{th}$ percentile | 0.607* | -0.013 | -0.016 | -0.052 | -0.042 | -0.076*** | -0.057*** |
| | (0.313) | (0.026) | (0.027) | (0.035) | (0.029) | (0.027) | (0.021) |
| Difference in male earnings: $50^{th}$-$10^{th}$ percentile | 0.270 | -0.052 | -0.035 | 0.042 | -0.009 | -0.057 | 0.010 |
| | (0.590) | (0.047) | (0.049) | (0.066) | (0.060) | (0.058) | (0.053) |
| Male earnings at $50^{th}$ percentile | -0.656 | 0.124** | 0.104** | 0.068 | 0.065 | 0.041 | -0.099* |
| | (0.541) | (0.051) | (0.048) | (0.047) | (0.045) | (0.057) | (0.056) |
| High-caste indicator | -1.901** | 0.138** | 0.118** | 0.170** | 0.206*** | 0.063 | 0.080 |
| | (0.776) | (0.057) | (0.057) | (0.066) | (0.072) | (0.091) | (0.096) |
| OBC indicator | -4.446*** | 0.319*** | 0.293*** | 0.328*** | 0.466*** | 0.296*** | 0.179 |
| | (0.931) | (0.075) | (0.076) | (0.087) | (0.086) | (0.112) | (0.109) |
| Dalit indicator | -5.535*** | 0.366*** | 0.374*** | 0.460*** | 0.581*** | 0.363*** | 0.191* |
| | (0.882) | (0.077) | (0.077) | (0.082) | (0.079) | (0.108) | (0.105) |
| Adivasi indicator | -6.507*** | 0.577*** | 0.524*** | 0.482*** | 0.538*** | 0.287* | 0.164 |
| | (1.322) | (0.107) | (0.105) | (0.111) | (0.113) | (0.153) | (0.132) |
| Constant | 12.097*** | -0.242* | -0.161 | 0.002 | 0.148 | 0.515*** | 0.918*** |
| | (1.374) | (0.128) | (0.120) | (0.120) | (0.115) | (0.161) | (0.160) |
| State fixed effects included? | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| N (women) | 627 | 627 | 627 | 627 | 627 | 627 | 627 |
| R-squared | 0.201 | 0.139 | 0.130 | 0.164 | 0.176 | 0.167 | 0.232 |

Table 7 reports the regression results for the educational attainment variables. The dependent variables in Columns 1-7 are, respectively, (i) Completed years of education, (ii) Indicator for zero years of education, (iii) Indicator for having obtained fewer than 5 years of education, (iv) Indicator for having obtained fewer than 8 years of education, (v) Indicator for having obtained fewer than 10 years of education, (vi) Indicator for having obtained fewer than 12 years of education, and (vii) Indicator for having obtained fewer than 15 years of education. Columns 2-7 estimate a Linear Probability Model. We find that a doubling of 90th percentile earnings relative to the median increases completed years of education by 0.6 years and that the probability of completing high-school increases by 8 percentage points and the probability of entering college increases by 6 percentage points. Notes: The excluded caste category is Brahmins; Standard errors are clustered by marriage market; * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. See also notes to Table 3.

**Table 8: Do women with no or little education experience delayed marriage?**

| | Age at marriage | |
|---|---|---|
| | (1) | (2) |
| Difference in male earnings: $90^{th}$-$50^{th}$ percentile | 0.003 | 0.014 |
| | (0.034) | (0.034) |
| Difference in male earnings: $50^{th}$-$10^{th}$ percentile | -0.003 | -0.004 |
| | (0.054) | (0.053) |
| Male earnings at $50^{th}$ percentile | 0.059 | 0.068 |
| | (0.062) | (0.060) |
| Female no education indicator | 0.762*** | |
| | (0.073) | |
| Female less than primary indicator | | 0.800*** |
| | | (0.057) |
| Male $90^{th}$-$50^{th}$ * No education | -0.180* | |
| | (0.100) | |
| Male $50^{th}$-$10^{th}$ * No education | 0.102 | |
| | (0.072) | |
| Male $90^{th}$-$50^{th}$ * Less than primary | | -0.148** |
| | | (0.074) |
| Male $50^{th}$-$10^{th}$ * Less than primary | | 0.061 |
| | | (0.056) |
| Constant | -20.175*** | -20.286*** |
| | (1.020) | (1.000) |
| Ancillary parameter | | |
| $\ln(p)$ | 1.815*** | 1.819*** |
| | (0.050) | (0.049) |
| **Net effect: $90^{th}$-$50^{th}$** | -0.177* | -0.135* |
| | (0.103) | (0.078) |
| State fixed effects included? | Yes | Yes |
| N (women) | 24,889 | 24,889 |

Table 8 examines whether women who have no or little education still delay marriage in response to greater male earnings dispersion. We report the results for a variation of the baseline age at marriage regression: the male income inequality measures are interacted with indicators for low levels of female educational attainment (Column 1: no education, Column 2: did not complete primary). Women with zero years of education delay marriage by a factor of 1.33 (Column 1) and women who have not completed primary school delay marriage by a factor of 1.25 (Column 2). This rules out the hypothesis that greater inequality may delay marriage by causing women to seek higher education, i.e. the direction of effect runs from marriage to education, not the other way around. Notes: The excluded caste category is Brahmins; Standard errors are clustered by marriage market; * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. See also notes to Table 3.

This product is part of the Pardee RAND Graduate School (PRGS) dissertation series. PRGS dissertations are produced by graduate fellows of the Pardee RAND Graduate School, the world's leading producer of Ph.D.s in policy analysis. The dissertation has been supervised; reviewed; and approved by the faculty committee composed of Krishna Kumar (Chair), Shanthi Nataraj, and Peter Glick.

PARDEE RAND GRADUATE SCHOOL

www.rand.org

RGSD-342