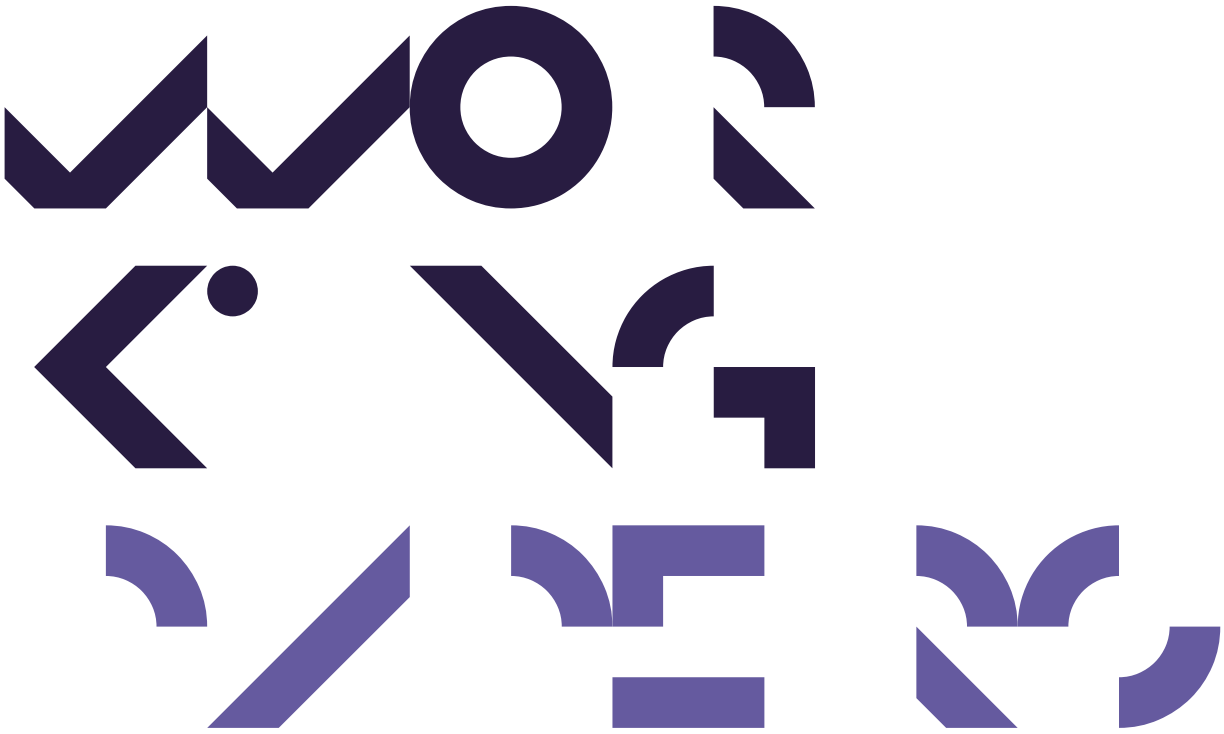


Economics Working Papers

2014-24

Social Networks and Health Knowledge in India: Who You Know or Who You Are?

Niels-Hugo Blunch and Nabanita Datta Gupta

AARHUS
UNIVERSITY
BUSINESS AND SOCIAL SCIENCES
DEPARTMENT OF ECONOMICS AND BUSINESS

Social Networks and Health Knowledge in India: Who You Know or Who You Are?[†]

Niels-Hugo Blunch
Washington and Lee University & IZA
blunchn@wlu.edu

Nabanita Datta Gupta
Aarhus University & IZA
ndg@econ.au.dk

This Version: October 23, 2014

JEL Classifications: I12, I14, I15.
Keywords: Health knowledge, caste, religion, social networks, India

Abstract:
Addressing several methodological shortcomings of the previous literature, this paper explores the relationship among health knowledge and caste and religion and a number of important mediating factors in India, estimating causal impacts through a combination of instrumental variables and matching methods. The results indicate the presence of a substantively large caste and religion health knowledge gap in the context of proper treatment of diarrhea in children favoring high caste women relative to low caste and Muslim women. We also provide evidence that while observed individual characteristics such as education and access to social networks explain part of the gap, a substantial part of the health knowledge gap is left unexplained. All groups have greater health knowledge in urban than in rural areas, but the gap is even wider in urban than in rural areas. Additionally, high caste women benefit more in terms of health knowledge from having health networks than women from other groups; except if the health person is of the same caste/religion, in which case low caste and Muslim women sometimes benefit by as much as double that of high caste women, or even more. It may therefore not be enough to give individuals access to high quality networks if caste and religion-related gaps in health knowledge are to be reduced; such networks also have to be homophilous, to have the maximum effect. Improved treatment from and confidence in the medical profession is found to be part of the mechanism linking health social network formation with improved health knowledge about the treatment of diarrhea in children.

[†] We thank Keera Allendorf, Ritwik Banerjee, Christian Bjørnskov, Tor Eriksson, Dylan Fitz, Art Goldsmith, Joseph Guse, Tim Lubin, Martin Paldam, Howard Pickett, Claus Pörtner, Nishith Prakash, and Gautam Rao; and conference participants at the Population Association of America Annual Meetings (2014, Boston), the 9th Annual Conference on Economic Growth and Development (New Delhi), the 51st Annual Meetings of the Public Choice Society (Charleston), the 3rd Danish Academic Economists in North America Annual Meetings (Durham), the 2nd Liberal Arts Colleges Development Conference (Colgate), and the 2nd IHDS User Conference (New Delhi); and seminar participants at Aarhus University, Seattle University, and Washington and Lee University for helpful comments and suggestions. Remaining errors and omissions are our own. The data were kindly provided by the Inter-university Consortium for Political and Social Research, Ann Arbor, MI, on behalf of Sonalde Desai, Reeve Vanneman, and the National Council of Applied Economic Research, New Delhi. The findings and interpretations are those of the authors and should not be attributed to any of the aforementioned individuals or organizations, however.

1. Introduction

While diarrhea is a common phenomenon in developed and developing countries alike, in developing countries it can have devastating effects. Each year, about 760,000 children under five are killed due to diarrhea, making it the second leading cause of death for children in that age group (WHO, 2013)[1]. And yet, many of these deaths are preventable: the main treatment for diarrhea is increased fluids, especially water, possibly combined with oral rehydration salts. Despite this simple remedy the death toll due to diarrhea increases year after year. Why? Because frequently mothers (and fathers) just do not know that increased water intake is what needs to be done upon the onset of diarrhea in their children[2]. Understanding better what determines health knowledge in the context of diarrhea in children therefore seems to be a health issue of utmost policy relevance.

Socioeconomic differences across different groups in society would seem to give rise to differences in health knowledge across these groups, also. In India, the institution of caste has been heavily researched and found, together with the institution of religion, to be associated with differentials in unemployment and poverty (Das, 2012), as well as directly in health outcomes (Borooah, 2012; Borooah et al, 2012). Health knowledge, arguably, is one of the main inputs in the production of health outcomes, thus motivating the policy relevance of understanding how health knowledge is produced. In particular, marginalized groups in society can be expected to benefit even more from improvements in their health knowledge than groups that are already (relatively) well off.

Addressing several methodological shortcomings of the previous literature, this paper seeks to improve our understanding of what underlies this lack of health knowledge specifically regarding diarrhea in children and their increased water intake as the correct treatment by addressing six research questions regarding this specific type of health knowledge and the relationship with caste and religion in India: (1) Does a ("raw") health knowledge gap exist between high caste women and women from low castes and other disadvantaged groups, such as Muslims, in India?; (2) Can observable characteristics help explain such an observed caste and religion health knowledge gap—in other words, does the gap "disappear" as potentially important mediating factors are introduced in the caste/religion-health knowledge relationship such as education, household income and social networks? (3) Where is health knowledge produced? In particular, what are the relative returns to what seems to be two of the main factors producing health knowledge—apart from social network

[1] Diarrhea is particularly dangerous for very small children, due to their immune systems not being fully developed yet, so that onset of diarrhea is much more likely to lead to death here (as compared to older children and adults).

[2] In a descriptive study based on Indian NFHS data (3rd wave, 2005-06), only 26% of children with diarrhea received oral rehydration treatment or increased fluids, as recommended, and 26 percent received no treatment at all. Sixteen percent received antibiotics, which are not recommended for treating most childhood diarrhea (Sharma, 2008).

access—namely education and information exposure, in terms of health knowledge?; (4) When considering specifically the importance of social networks as a means of producing health knowledge, which aspect of network formation is more important—access, quality, or homophily of networks?; (5) What can we learn from both an overall and a detailed decomposition of the observed caste and religion gap in health knowledge?; and (6) What are some potential components of the mechanism underlying the observed impact of health social network formation on health knowledge?

Estimation of raw caste and religion health knowledge gaps and overall and detailed earnings decompositions leads to six main results: (1) education and social network access are all strongly associated with increased health knowledge; (2) the presence of a substantively large health knowledge caste gap favoring high caste women; (3) evidence that the endowments and the returns to these endowments increase the health knowledge gaps—indicating that high caste women have higher education and better access to social networks but also higher returns to these characteristics; (4) network homophily frequently is much more important for low caste women than either network access or network quality—unlike what is the case for high caste/religion women; (5) while observed individual characteristics explain part of the gaps, a substantial part of the health knowledge gap is left unexplained; and (6) improved treatment from and confidence in the medical profession is found to be part of the mechanism linking health social network formation with improved health knowledge about the treatment of diarrhea in children.

The remainder of this paper is structured as follows. The next section motivates studying the inter-linkages among health knowledge, caste and religion, and social networks in India. This is followed by a description of the data in section three, while section four discusses the empirical strategy and related issues. Section five presents the results for the raw caste and religion gaps in health knowledge, as well as OLS, IV/2SLS, and matching estimations of the determinants of health knowledge and decompositions of the caste and religion gaps in endowments/observables and returns/unobservables—with all decompositions pursued both at the aggregate level, as well as tracing the determinants to their component parts, such as education, social network access, and information exposure. Finally, section six concludes and discusses policy implications and provides suggestions for further research.

2. Institutional Background, Caste and Religion in India and Conceptual Framework

India has the second highest rate of diarrhea mortality among children in the world (after Afghanistan). In 2010, 13% of all child deaths in India were due to diarrhea (UNICEF). As mentioned earlier, socioeconomic factors, malnourishment, poor sanitation (high rates of open defecation) and drinking

contaminated water are some of the major reasons for diarrhea in South Asian countries. In terms of caste differences in mortality, compared to forward caste Hindus, the average age at death in India – after imposing all the controls - was 4.9 years lower for *Adivasis*[3], 7.1 years lower for *Dalits*[4], and 6.1 years lower for Muslims (Borooah, 2010). Another study focusing on women only found that mortality of Dalit women shows that the average age at death for Dalit women (39.5 years) is 14.6 years less than the average age at death for higher caste women (54.1 years) (Borooah et al 2012). There are hardly any studies looking at caste differences in child health and mortality. One exception is a study showing a mortality divide by caste in India, particularly among children and the elderly (Subramanium et al, 2006). Using the dataset examined here, we show that children of the lower castes are more vulnerable to diarrhea mortality and severity:

Table 1 depicts the cause of death for the most recent non-birth related death in the household, focusing at infants 2 years of age or less. While the overall sample size is quite low, so that the results should be interpreted with caution, the results here are indicative of (1) diarrhea being a leading cause of death in India; and (2) low caste and Muslim children being particularly susceptible to dying from diarrhea as compared to other reasons, relative to high caste children.[5]

Table 1. Cause of Death for Recently Died Children 2 Years of Age or Less (Percent)

	Full sample:	*High:*	*Dalit:*	*Adivasi:*	*OBC:*	*Muslim:*
Accident	1.9	8.7	0.0	0.0	3.2	0.0
Fever	39.9	32.2	47.7	63.5	34.4	23.2
Diarrhea	11.6	0.0	16.8	14.7	5.0	23.3
Cancer	0.7	0.0	0.0	0.0	1.7	0.0
Heart attack	0.3	0.0	0.0	0.0	0.8	0.0
TB	0.5	7.4	0.0	0.0	0.0	0.0
Other	45.1	51.7	35.5	21.8	54.9	53.5
N	123	11	26	15	48	23

Notes: Calculations incorporate sampling weights.
Source: 2004/05 India Human Development Survey (IHDS).

When it comes to diarrhea, not only diarrhea as a cause of death is relevant, however—also the incidence as well as the severity and treatment of diarrhea are relevant, as well as important for policy

[3,4] Adivasis: indigenous tribal peoples living in remote areas also denoted as Schedule Tribes; Dalits: considered the untouchables, denoted as Schedule Castes.

[5] Reflecting also the low overall sample size here, there were no children from Sikh, Jain, or Christian households in this sub-sample.

makers—by indicating some of the mechanisms leading to diarrheal mortality, as well as to the differential diarrheal mortality across caste and religion. In particular, it is preferable not have blood in the diarrhea, which could be seen as a more severe form of diarrhea. Table 2 depicts these factors, again using the current dataset and focusing on children 2 years of age or less. From the table, low caste and Muslim children are generally worse off than children from high caste, Sikh/Jain or Christian backgrounds in terms of incidence, while the gaps in severity are not as pronounced. Perhaps surprisingly, Adivasi children experience a lower incidence than even high caste children—though this may be explained by Adivasis traditionally living in hills and forests, so that access to wild berries and fruits and the use of herbal remedies may possibly yield some protection against diarrhea in children.

Table 2. Incidence and Severity of Diarrhea in Children 2 Years of Age or Less (Share)

	High caste	*Dalit*	*Adivasi*	*OBC*	*Muslim*	*Sikh/Jain*	*Christian*
Diarrhea incidence							
Any diarrhea	0.108	0.134	0.072	0.150	0.135	0.087	0.048
N	1,976	2,407	993	3,620	1,575	129	101
Diarrhea Severity							
Blood in diarrhea	0.187	0.205	0.162	0.187	0.150	0.100	0.243
N	209	316	69	478	192	11	5

Notes: Calculations incorporate sampling weights.
Source: 2004/05 India Human Development Survey (IHDS).

To understand the origins of some of the differences above, it is necessary to review the role played by caste in Indian society. Indian society has been stratified according to an elaborate and rigid system of occupational specialization for thousands of years. The system of caste has resulted in the practice of extreme forms of prejudice against and the complete exclusion of certain groups from opportunities for advancement. Caste discrimination is constitutionally illegal in India and starting in 1950, several government schemes have been implemented to improve the labor market and social conditions of the low caste. Yet, poor economic outcomes for the low caste persist. There are 5 major caste groups, which, over time have branched out into around 3000 subcastes or *jatis*. The most marginalized groups are the Scheduled castes (the "untouchables" or Dalits) and the Scheduled tribes (Adivasis, or indigenous peoples). Other Backward Castes (OBC) are placed higher in the hierarchy but still suffer economic disadvantage. Caste is endogamous and rarely can be changed. It is generally revealed by the family name, village location, occupation or even dress and custom. Since caste is mainly associated with Hinduism, conversion to another religion or name changing is in principle a way to

escape from one's caste of birth. The act of conversion or caste-free names, however, signals a prior low caste affiliation, and caste segmentation has even found its way into the other religions of India.

Why should caste be important for health and health knowledge? First of all, low caste individuals are disproportionately poor and have lower levels and quality of education. In 2004/5, poverty rates among Dalits and Adivasis were 46% and 37% compared to 23% among non-SCs/STs (Chin and Prakash, 2011). Even though the educational attainment of the low caste has improved since the turn of the century, it lags behind that of upper caste Hindus (Borooah and Iyer, 2005). According to Grossman's model for the demand for health and health care (Grossman, 1972), educated individuals both have greater allocative efficiency and greater productive efficiency when it comes to making investments in their health. Kenkel (1991), studying the US, empirically shows that the more educated possess greater levels of health knowledge and thereby enjoy an informational advantage.[6] Although our model controls for educational attainment, it does not capture the quality of the education obtained. Educational quality may be poorer for the lower castes either because they may be debarred from or cannot afford attending the same schools as the upper caste (for instance, they are much less likely to attend private schools) or when attending the same schools often have an inferior educational experience due to discrimination and unequal treatment by the teacher or co-students (Ramachandran and Naorem, 2013).

Because of extreme prejudicial notions of contamination and loss of purity when encountering the low caste, the high-caste have traditionally forced them to live on the outskirts of the villages. This means lower *access* to resources such as clean water[7]. Not only would they be more likely to drink contaminated water and lack toilets in their own homes, the age-old forced specialization in human waste remover occupations bring them and their families in close and regular contact with open latrines and dry pit toilets exposing them to viral and bacterial infections that can lead to diarrhea. By virtue of their isolation, the low caste also have lower access to health care facilities and thereby less contact with professional staff from whom health knowledge is obtained, and therefore they are exposed to less information. Even when they do establish contact, extreme discrimination excludes them from receiving the right treatment and knowledge.

[6] Altindag et al. (2011) using two waves of the NLSY79 find evidence consistent with Kenkel (1991), but do not find that allocative efficiency is the primary reason for why education affects health knowledge.

[7]The lower castes do not have rights to drink from common water sources such as public wells and taps out a fear of impurity and pollution, although this practice is on the decline. On the other hand, the donation of water is seen as a virtue in Hinduism, so a dichotomy exists. While the lower castes may be donated clean water, they do not have property rights to it (Pradhan and Meinzen-Dick, 2010).

A further consequence of low education and isolation is that the social networks of the low caste are of poorer *quality*—they would have greater social distance to members of the medical community, especially those in higher positions such as doctors. The same quality difference in networks would arise if there is network *homophily*—individuals interacting mainly with other individuals like themselves for example due to trust and (lack of) self-esteem—simply because too few members of the low caste are represented in these professions. In our empirical framework, we will be able to disentangle the access, quality, and homophily effects.

Though most of the focus on socioeconomically disadvantaged groups in India has been related to caste, many of the previous observations carry over to other disadvantaged groups—characterized, for example, by their religious affiliation. This is associated with differentials in unemployment and poverty (Das, 2012), as well as directly in health outcomes (Borooah, 2012; Borooah et al, 2012).

It is by now fairly well established that health knowledge plays an important and sometimes independent (from maternal education) role in child health (Glewwe,1999; Kovsted et al., 2003) in a developing country context. One of the lessons from this literature is that health knowledge can be endogenous to child health outcome, say if parents of sick children make special efforts to gain knowledge, then this will impart a downward bias to the effect of health knowledge. Surprisingly little literature exists on health knowledge differences and the role of social networks. A few explorations in the economics of risky behaviors have shown evidence of contagions effects of social networks for smoking (Cutler and Glaser, 2007), and obesity (Christakis and Fowler, 2007). These effects arise even after accounting for homophily, i.e. individuals selecting people who resemble themselves to be part of their network. One recent U.S. study exploits random assignment of undergraduates to residence halls which either offer or do not offer vaccination clinics to test whether it influences the share of a resident's friends (as determined from online social networks) who choose to get vaccinated (Rao, Mobius and Rosenblat, 2007). The results showed that when 10% of a student's social network was exposed to medicinal information, this raised their own probability of getting vaccinated by 8.3 percentage points, and that 85% of the increase could be explained by an enhanced positive belief of the benefits of getting vaccinated. Another US study finds that among female employees of a large medical organization, there were strong social multiplier effects (about 2 to 3 times) in response to a co-worker or neighbor's breast cancer screening (Banerjee, Cohen-Cole and Zanella, 2007). As the context of the first study is influenza shots in the setting of Harvard University and the second a large health establishment in the U.S., containing highly educated and knowledgeable individuals, it is not clear whether a similar finding would carry over for health knowledge of mothers in India.

The study that comes closest to ours is Patra et al (2013) who link mothers' health knowledge to childhood morbidity (defined as if a child 0-59 months suffered from either fever, diarrhea, cold or cough in the month preceeding the survey), medical care and medical care expenditure in India using IHDS 2005 data. They first model the determinants of health knowledge[8] in a multinomial regression (No knowledge/low knowledge, Medium knowledge, High knowledge). The determinants of health knowledge are mother's background characteristics, age categories, education in levels, rural/urban, wealth index, caste and religious groupings. Next, the same covariates plus child gender and health knowledge levels are used to determine child short-term morbidity, use of modern medical treatments, and median medical expenditure levels. Their main findings relating to health knowledge are that the socioeconomic indicators show considerable impact on mothers' health knowledge and that a mother's health knowledge is a stronger predictor of short-term morbidities among children than the other background characteristics. Furthermore, they find that the prevalence rate of any short-term morbidity (fever, cough or diarrhea) among children is inversely related to the level of health knowledge of their mothers. Controlling for socioeconomic status, mother's health knowledge is a critical factor in determining the type of treatment received among children for their short-term morbidities and on spending on medical care for child morbidities.

This study gives some of the first evidence on the importance of maternal health knowledge for child short term morbidities in India. However, it does not explore the importance of social networks or mass media exposure for producing health knowledge. The study also pools together different morbidities (fever, diarrhea, cold or cough) and creates in addition a composite index of health knowledge, thereby not allowing for deeper insight into how diarrhea health knowledge is generated.

Based on this brief review of relevant studies in the area and the main issues regarding caste pertinent to this study, we develop the analytical framework.[9] In addition to bringing greater clarity of the issues and mechanisms involved when examining the production of health knowledge, this will also have empirical implications—including the specification of explanatory variables in a regression analysis, as well as potentially important issues that may affect the empirical analysis such as endogeneity and unobserved heterogeneity.

The starting point for the analysis is an extension of the health production model (Grossman, 1972). We consider a mother who obtains utility from child health and allow the human capital effects

[8] Health knowledge in their setting is an index composed of 7 health knowledge questions in the survey covering knowledge of water purity (2 questions), knowledge of the dangers of a wood-burning stove, knowledge of the benefits of breastfeeding during early pregnancy, knowledge of handwashing after defecation and knowledge on hydration following diarrhea.

[9] This model and the associated discussion is an adapted version of the expanded Grossman (1972) model developed in Blunch (2006).

to also potentially come from health knowledge, rather than only from education per se. Further, the human capital effects run from the mother's education and health knowledge to the child's health, rather than from one's own education to one's own health.

Specifically, we consider a two-person household consisting of a mother and a child in which the mother has preferences over the child's health (Z_H) and other commodities (Z_O). Alternatively, the set-up may be viewed as a multi-person household, where the focus is on the linkages of maternal health knowledge, schooling, labor supply and child health investments and the resources of other household members enter the model through contributions of earned and unearned income.[10] The relative preferences between the two commodities are affected by three factors: schooling (S), caste or religion (C),[11] and other important characteristics both observable and unobservable (δ)—giving rise to the following utility function:

$$U = u(Z_H, Z_O; S, C, \delta) \quad (1)$$

The utility function is assumed to exhibit the usual desirable properties; most importantly it is assumed to be quasi-concave and increasing in its arguments.

The household maximizes utility subject to a set of technological, budget and time constraints. The technological constraints are given by the two main production functions, f_H and f_O, in which outputs of child health and all other goods depend on inputs of a market good (X), and the mother's time (T) and are conditional on the mother's schooling, (S), caste or religion (C), and all other important characteristics, both observable and unobservable (δ). Note caste or religion may enter production functions because of the social and physical distance of lower caste individuals to inputs into the production process (for example, water and health workers) and since discrimination may have feedback effects in terms of constraining investments. For child health production some of these additional factors captured by δ may be further specified—namely as social health network formation (N), mass media exposure (M), the child's initial, unobserved health endowment, η, and a community specific health related variable, B, which includes health infrastructure, treatment practices, and the local disease environment. Health knowledge, which is used in the production of child health, is also itself produced and therefore has its own production function, f_K, in which health knowledge is

[10] Issues related to intrahousehold bargaining over resources are not incorporated here. A large literature, starting with Manser and Brown (1980) and McElroy and Horney (1981), examines issues related to marriage and household decision-making. One of the main results from this literature is that the bargaining power over resources within the household related to for example child health depends on the opportunities outside of the household. To the extent that bargaining power is correlated with mother's skills and schooling, however, the analyses here will at least capture some elements of the bargaining structure within the household. For a review of family economics, including cooperative household models, see also Bergstrom (1996).

[11] For example due to differing norms and traditions related to health and health practice (Das, 2013).

produced from health social networks formation (N) and mass media exposure (M) and time (T_K). These production functions can therefore be expressed as:

$$Z_H = f_H(X_H, T_H; K, S, C, N, M, B, \eta, \delta) \quad (2)$$

$$Z_O = f_O(X_O, T_O; S, C, \delta) \quad (3)$$

$$K = f_K(N, M, T_K; S, C, B, \eta, \delta) \quad (4)$$

The household obtains income from engaging in labor activities, exogenously supplying H hours of labor at the wage rate W, and from income from other unearned sources (O), which is also exogenously given. Income can be spent on market goods such that:

$$WH + O \geq P_H X_H + P_O X_O \quad (5)$$

Lastly, the maximization of (1) is subject to a time constraint:

$$T_H + T_K + T_O + H = T_{max}, \quad (6)$$

where T_{max} is the maximum time available to the mother.

The mother is assumed to choose N, M, T_H, T_K, T_O, X_H, and X_O to maximize utility subject to these constraints. Solving the model yields a series of market goods demands and production time supply functions. Using these to solve for N^*, M^*, and T_K^* and then substituting these into (4) yields the reduced form health knowledge production function:

$$K^* = f_K(W, O, P_H, P_O, K, S, C, N, M, B, \delta, \eta) \quad (7)$$

From this model—even if the direct estimation of (7) *per se* is not the objective of the empirical work that follows[12]—there are several implications, which can help guide the empirical analyses. First, the reduced-form health knowledge demand function (7) points to the variables that should be included in its empirical counterpart. These variables include the mother's schooling, caste or religion, health social network formation, the availability of health care and services, and other background characteristics. Second, the reduced-form health knowledge demand function indicates that schooling, caste (or religion), and network formation all have direct effects on health knowledge demand.

Lastly, the conceptual model outlined above highlights the importance of unobserved heterogeneity for the subsequent analyses. Unobserved preference shifters (δ) and child health endowments (η) are seen to affect health knowledge demand directly. Further, network formation and obtaining information through mass media both reflect a purposeful activity, which causes potential endogeneity (the same can be said to be true for schooling—though this at least is predetermined).[13]

[12] If only because of having mostly either rather crude empirical proxies for the theoretical counterparts of the model (for example maternal wages), or even frequently missing altogether (for example the price of medical care, for people who did not go to the doctor).

[13] As we will also discuss later, we do, however, still attempt to endogenize schooling—though unsuccessfully so: the

These twin issues of unobserved heterogeneity and endogeneity/simultaneity will be addressed in the subsequent empirical analyses.

Treating network effects as predetermined and unidimensional simplified the presentation of the model and helped to illustrate its implications. Conceptually, however, several extensions are warranted. First, what are health social networks exactly? Rather than merely one generic "health social network," there seems to be at least three separate dimensions to health social network formation: access, quality, and homophily. Access here refers to simply having access to any type of health professional outside one's family, whereas quality refers to the level of that network access—for example, is it a health technician, a nurse, or a doctor? The doctor here would seem to indicate the highest quality level. Homophily refers to the notion of "sameness" between a given individual and her network contact. For example, is the individual and her network contact of the same caste or religion? Here, one might expect that quality or homophily of health social networks improves health knowledge more than just access, per se: a higher quality network contact is able to provide better information and an individual would seem both better able and more eager to obtain information from a network contact, who is "similar" to oneself in terms of socioreligious background.

Based on this conceptual framework we will examine whether schooling, mass media exposure, and network formation increase health knowledge as it relates to the correct treatment of diarrhea in children regarding their water intake. In particular, we will further examine whether the proposed three distinct aspects of network formation have differential effects on this specific form of health knowledge.

3. Data

The 2004/05 India Human Development Survey (IHDS) is a comprehensive nationally representative multi-purpose household survey of 41,554 households in 1,503 villages and 971 urban neighborhoods across India. The IHDS survey was produced by the National Council of Applied Economic Research (NCAER), New Delhi, and the University of Maryland and used a multi-stage clustered sampling design, ensuring national representativeness of the survey findings.[14] The survey consisted of several questionnaires, which collected information at different levels of aggregation and/or for different household members. The household questionnaire collected household and individual information such as age, education, information exposure, and access to social networks from a knowledgeable

proposed instruments turn out to be extremely weak, leading to huge standard errors. We, therefore, endogenize only network formation and information exposure and resort to the predeterminedness argument for schooling.

[14] See Desai et al (2010) for more details.

informant, typically the head of the household. A special questionnaire on health and education was given to an ever-married woman, aged 15-49, in the household.

Relevant for the purposes here, for the subsample of ever-married women 15-49 years of age, the survey collected information on health information regarding several health issues. While most of these are not particularly policy relevant,[15] the survey also asked about the correct treatment in children regarding their water intake. Again, considering the adverse outcomes overall for the lower castes and Muslims in terms of incidence, mortality, and, to some extent, the severity of diarrhea established above and the fact that health knowledge of diarrhea is plausibly an important factor affecting these, examining more closely health knowledge specifically regarding diarrhea would seem a worthwhile endeavor—from an academic as well as a public policy point of view. The dependent variable, therefore, is a binary measure of the individual woman's health knowledge in terms of her knowledge of the correct treatment of diarrhea in children in terms of water intake. It is based on the question "When children have diarrhea, do you think that they should be given less to drink than usual, more drink than usual, about the same, or it doesn't matter?" and is coded one if the women responds that more water should be given to a child with diarrhea and zero otherwise.

The two focal explanatory variables are caste and religious affiliation and social network formation, respectively. Caste and religious affiliation is coded as a set of dummy variables, where caste includes the four main groups High Caste (Brahmin and Other Forward Castes), Dalits/Scheduled Castes (SC), Adivasis/Scheduled Tribes (ST), and Other Backward Castes (OBC). Non-Hindus are divided into three groups: Muslims, Christians, and Sikh and Jain. The social network variables examined here are based on the main question, "Among your acquaintances and relatives, are there any who are doctors or nurses or who work in hospitals and clinics?" For this main question we construct a binary measure, which is coded one for "yes" and zero for "no." In addition to this main question, several follow-up questions are asked, which enables exploring additional dimensions of social network access. These include "What does he/she do?" where the possible answers include "Doctors," "Nurses," "Technician," and "Other" and "Is his / her community / jati the same as yours?" where the possible answers include "Different jati" and "Same jati." The responses to these two additional questions are also coded into two binary measures, one which is one for doctor and zero otherwise (thus taking "doctor" as the highest level of health network access) and one which is one for same jati and zero otherwise.[16] In essence, therefore, we measure three distinct aspects of social health network

[15] Including, for example, asking about whether men become physically weak even months after sterilization.

[16] The interpretation of "jati" here is thus not literally as jati in terms of (occupational) sub-caste but rather as main caste or religious community. This interpretation was supported by Sonalde Desai, one of the PIs of the IHDS examined here, in

formation: *access*, *quality* (taking "doctor" as the highest level), and *homophily* of the network (as measured by likeness regarding caste relationship of the two network participants). In the empirical analyses these measures are therefore included separately in the regressions, yielding three separate sets of results as far as the specification of network formation is concerned.

Additional explanatory variables include birth cohorts (constructed as a set of dummy variables spanning five years each), and a set of six dummy variables for the highest level of education completed—spanning no education (the reference) through tertiary. To ensure that any observed caste and religion health knowledge gaps are not primarily due to differences in income and/or poverty, we also include (log) per capita household expenditures (consumption) to proxy for income. This measure is based on a series of 47 questions about household consumption designed to estimate total household consumption.[17] Information exposure is constructed as a dummy based on the question "How often do people in your household watch TV?" with possible responses including "Never," "Sometimes," and "Regularly." A dummy variable is then constructed for "Regular" usage, motivated by the fact that that can be seen as the higher threshold.[18] The remaining explanatory variables include access to health facilities in the community (only collected for rural areas, from the village survey)—specified as a set of dummy variables for availability of a Health Sub-center, Primary Health Center, Community Health Center, Government Maternity Center, Government Communicable Disease Facility (e.g., tuberculosis)—and state fixed effects.

Since the dependent variable is health knowledge, the sample was first conditioned on ever-married females 15-49 years of age, since this was the group for which this information was exclusively collected (32,154 observations). Some variables were missing for some women, so that the final estimation sample consists of 31,537 ever-married women 15-49 years of age. The means and standard deviations for the final estimation samples by caste and religious affiliation are reported in Table A1, Appendix A.

4. Estimation Strategy and Related Issues

The conceptual framework discussed in Section 2 suggests that caste, religious affiliation, educational attainment, information exposure, and access to social networks can directly affect the acquisition of health knowledge and also suggest additional factors that are potentially important for health

remarks given to the audience at the 2014 IHDS User Conference in Delhi (at which the authors were both participants).

[17] See Desai et al (2010) for more details.

[18] These are asked both for women and men in the household overall. We use the information pertaining to the women of the household here.

knowledge and therefore should be included in the empirical specifications.[19] The empirical analysis will examine this relationship, using linear approximations of the health knowledge equation. The natural starting point is estimating the following regression by OLS[20] (i.e. as a Linear Probability Model, LPM):

$$HK_i = \alpha_0 + \alpha_1 CASTE + \alpha_2 EDU_i + \alpha_3 INFEXP_i + \alpha_4 SOCNET_i + \alpha_2 CONTROLS_i + \varepsilon_i, \quad (8)$$

Where HK_i is one if the mother knows that a child with diarrhea requires more water than usual, $CASTE_i$ is a set of dummy variables for caste and religious affiliation, EDU_i is a set of dummy variables for educational attainment; $INFEXP_i$ is a vector containing the two dummy variables for information (newspaper and television); $SOCNET_i$ is a vector of social network access variables; and $CONTROLS_i$ is a vector of all additional controls, including log per capita household consumption and state fixed-effects; and ε_i is an error-term capturing unobservables.

As also discussed in Section 2, several issues need to be addressed pertaining to the estimation of the caste/religion-health knowledge relationship in equation (8)—where the most important arguably is the possible endogeneity of the education, information exposure, and social network access variables. Again, endogeneity has three possible causes: omitted variables, simultaneity, and measurement error, all of which are potentially relevant in this application.

Regarding omitted variables, ability and preferences, for example, are unobserved and at the same time also main determinants of educational attainment, information exposure, and social network access. As a result, the estimated impacts of these variables may be affected by omitted variables bias. Second, simultaneity may be a potential issue, since obtaining health knowledge, information exposure, as well as access to social network access all involve choices on the part of the woman. Regarding measurement error, one important issue is that the variables for social network access are binary measures of access per se, that is, they do not measure the intensity of an individual's network access. Information exposure is also measured as binary variables and are also self-reported, both of which leads to measurement error.

One widely applied approach to deal with endogeneity involves instrumental variables (IVs), by estimating by Two-Stage Least Squares. It is often a daunting task, however, to come up with

[19] At a minimum, if these factors are not included, one may systematically over- or underestimate the strength of the caste/religion-health knowledge relationship.

[20] While there may be some concern about using the linear probability model (LPM) (for example, predicted probabilities may fall outside the (0,1)-range and heteroskedasticity also is present by default), it can be argued that the LPM is a more robust alternative to a logit or probit model and also approximates the response probability well—especially if (1) the main purpose is to estimate the partial effect of a given regressor on the response probability, averaged across the distribution of the other regressors, (2) most of the regressors are discrete and take on only a few values and/or (3) heteroskedasticity-robust standard errors are used in place of regular standard errors (Wooldridge, 2010). All three factors seem to work in favor of the LPM for the purposes of the application here.

variables that are both highly correlated with the potentially endogenous variable(s) and which at the same time may also validly be excluded from the main equation. Arguably, human capital accumulation and skills acquisition depend on the availability of educational institutions, as well as their quality. This has led researchers to follow two main IV strategies in recent years: either using as IVs (1) various combinations of time of year, birth cohort, and/or geographical area of birth dummies to capture variation in institutional factors relevant for human capital accumulations such as compulsory schooling laws or expansion of educational programs (Angrist and Krueger, 1991; Duflo, 2001) or (2) variables for proximity or exposure to educational institutions in the local area (Card, 2001).

Since we do not have available the geographical area of birth, we first explored the second of these approaches for the case of educational attainment. It turned out, however, that the instruments were quite weak, thus leading to the so-called weak instruments problem (Staiger and Stock, 1997). Additionally, it can be argued that education is at least pre-determined, thus at least addressing the simultaneity-part of the potential endogeneity issue.

Turning to information exposure and social network access, where the endogeneity issues seems particularly worrisome due to the likely string simultaneity of these variables vis-à-vis the dependent variable, namely health knowledge, one promising candidate is the share of the population in the area with regular information exposure and social network access (across the different sources of information and types/levels of social networks). The intuition behind this instrument—inspired by Gruber (2005)—is that the more information exposure or social network access there is in an area, the more likely it is that a given woman will be exposed to information from the media and/or gain access to social networks. One concern here is that if the geographical area chosen is too small, endogenity once again become an issue (Dustmann and Preston, 2001; Jensen and Rasmussen, 2011). In this case, for example, there would seem to be more mobility in terms of migration between villages than between districts, so that an IV measure based on the village of residence would be "more endogenous" than an IV measure based on the district of residence. Similarly, an IV measure based on the district of residence is likely to be "less endogenous" than an IV measure based on the village of residence.[21]

These considerations lead us to estimate the health knowledge equation with instrumental variables using the first stage equation:

$$END_i = a_0 + a_1 Z_i + a_2 CONTROLS_i + n_i, \quad (9)$$

[21] Sensitivity analysis (see Results section) reveals that choosing the (potentially endogenous) village level instead yields even stronger results, in substantive as well as statistical terms.

where END_i is a binary, possibly endogenous, variable (information exposure and social network access), Z_i is a vector of instrumental variables based on district shares, and $CONTROLS_i$ is a vector of all additional controls from the second stage regression (primarily included for efficiency), including all other (exogenous) variables. ν_i is an error-term capturing unobservables. The first-stage test for weak instruments then is performed as a joint test on the variables in Z_i (the identifying instruments – that is, excluded from the second-stage regression). The second stage equation (the estimating equation) then includes the predicted values of the potentially endogenous variables from the first stage:

$$HK_i = b_0 + b_1 \hat{END}_i + b_2 CONTROLS_i + g_i, \quad (10)$$

where HK_i measures the health knowledge regarding diarrhea in children of the ith woman; $\hat{END}_i$ is a vector of the fitted values of the potentially endogenous variables from the first-stage equation (9); $CONTROLS_i$ is a vector of all additional (exogenous) controls; and γ_i is an error-term capturing unobservables. Again, since (10) includes predicted variables as regressors, the standard errors must be adjusted accordingly. Further, so as to allow for arbitrary heteroskedasticity, the estimations of (8)-(10) will be carried out using Huber-White standard errors (Huber, 1967; White, 1980). To allow for the possibility that observations are correlated within communities the standard errors are also adjusted for within-cluster correlation (Wooldridge, 2010). Regressions are estimated both for the full sample—with caste and religion variables included among the focal explanatory variables—as well as in separate, stratified (by caste and religious affiliation) regressions. The objective of the latter is to examine whether the impact of the other focal explanatory variable, namely social network access, differs across the different castes/religions. For example, it is possible that high caste women get a higher return from having access to social networks than other women due to their already favorable characteristics.

To help strengthen the social network results we also use matching on observables to estimate the impact of having access to social networks, in all three dimensions. This estimation method goes straight towards constructing the proper counterfactual; that is, what would have happened in the absence of the treatment? Specifically, treatment and control groups are matched on observables[22] and then the treatment effect is estimated as the mean difference between the average effects between the matched samples (Rosenbaum and Rubin, 1983). "Treatment" for the application here refers to having

[22] It here seems prudent to exclude TV watching from the observables used in the matching procedure due to the endogeneity concerns of that variable—so that is also the approach taken here. Sensitivity analysis (see Results section) reveals that this choice is not crucial to the obtained matching results.

any network access or having access to a doctor or having access to a health person of the same jati—so that we pursue three analyses in turn. The counterfactual is approximated by the experiences of a "comparison group" of women who are similar in all respects except social network access. This is achieved using matching, which in practice amounts to using a two-stage approach. In the first stage participants and non-participants are matched, based on their observable characteristics. In the second stage the impact estimate—which corresponds to the estimate of α_1 or α_2 in (8) or β_1 in (10) from the regression case, depending on which "treatment" is considered—is then calculated as the difference in means of health knowledge outcomes between matched participants and non-participants.

There are several different ways to conduct the matching in practice. For the application here we use Mahalanobis covariate matching which is based on the Mahalanobis distance to match participant and treatment individuals[23] and has the attractive property that it reduces differences in covariates within matched pairs in all directions (Imbens, 2004: 15). The match is then based on the closest non-participant ("nearest neighbor"). The estimations implement the robust analytical standard errors proposed by Abadie and Imbens (2006) to correct for possible heteroskedasticity and also impose common support by excluding participant observations whose propensity scores are higher than the maximum or less than the minimum covariate values of the comparison group (as also suggested by Rosenbaum and Rubin, 1983).

After estimating the relationship between health knowledge and its main determinants as expressed by (8) and (10) above, the next step is to decompose the health knowledge gap into its main components using the Blinder-Oaxaca approach (Blinder, 1973; Oaxaca, 1973).

Specifically, the decomposition analysis amounts to examining to which extent the observed health knowledge gaps across caste and religion are attributable to observable/explained and unobservable/unexplained characteristics (the so-called "two-fold decomposition"), where the unexplained part is often treated as discrimination in the literatures on gender and racial earnings gaps.[24] And, to be sure, in the application here it is indeed possible to talk about "discrimination," per se, as being a low caste/religion woman is an intrinsic characteristic. We base the decomposition on the prevailing health knowledge structure of high caste women. This seems reasonable for the application here, since the high caste dominates in the Indian society—and, thus, can be perceived as the non-discriminated-against group—as also revealed by the existence of substantial "raw" health knowledge gaps in the subsequent analysis. Still, while the main analysis here takes the high caste

[23] For details on Mahalanobis matching, see for example Rosenbaum and Rubin (1985).
[24] It is important to note, however, that the "unexplained" part also captures all potential effects of differences in unobserved variables (Jann, 2008).

health knowledge structure as the reference, several different specifications for the baseline specification (also known as the "absence of discrimination" specification) will be pursued in the sensitivity analysis as a robustness check.[25] The standard errors of the decompositions are computed using the Delta method by applying the procedure detailed in Jann (2008), which extends the earlier method developed in Oaxaca and Ransom (1998) to deal with stochastic regressors.

In addition to examining the overall composition of the established health knowledge gaps we also perform detailed decompositions, whereby it is possible to see which explanatory variables contribute the most to the overall health knowledge gap decompositions. An issue here is that while the overall decompositions are always identified, the results for categorical variables in detailed decompositions depend on the choice of the reference category (Oaxaca and Ransom 1999). A possible solution to this problem is to apply the deviation contrast transformation to the estimates before conducting the decomposition (Yun 2005); this is also the approach pursued here. Similar to the OLS regressions, the decomposition estimations also all allow for arbitrary heteroskedasticity (Huber, 1967; White, 1980). So as to condense the wealth of results obtained here—thereby easing the interpretation of the many results—the detailed decompositions are done groupwise, rather than for each individual variable (for example, for education as a whole, rather than separately for each educational level, and so on). Here, too, the focus will be on the case where the high caste structure is taken as the reference, though sensitivity analysis again will consider alternative specifications, as well.

5. Results

This section reviews the main results. Since the (full) results tables are rather large, the results are mainly presented graphically and in excerpt tables, reporting only the main coefficients of interest (in the case of coefficients reported in figures, additional excerpt tables reporting also the standard errors and level of statistical significance of coefficients are placed in the Appendices and referred to in the below).[26]

Question 1: *Does a ("raw") health knowledge gap exist between high caste women and women from the low castes and disadvantaged religious groups in India?*

[25] See Oaxaca (1973), Blinder (1973), Cotton (1988), Reimers (1983), Neumark (1988), and Jann (2008) for different approaches—basically, these differ in the relative weights they attribute to the two groups in the decomposition.
[26] The complete set of results is available upon request.

From Table 3 below two main results stand out. First, raw health knowledge gaps favoring high caste (the reference category, comprised of Brahmin and other forward castes) are found across almost all castes and religious affiliations and for both rural and urban areas. For example, high caste women are about 13.3 percentage-points more likely to know that children require more water when they have diarrhea than Dalit women—and with substantial gaps versus women from other traditionally disadvantaged groups, as well: about 12.4 percentage-points for Adivasi women, and just above 9 percentage-points for both women from Other Backward Castes and Muslim women. Second, however, while the levels of health knowledge are greater in urban areas than in rural areas across almost all caste and religious groups, the gaps are greater, still. For OBCs, in particular, the gap widens from about 9.4 percentage-points in rural areas to about 17 percentage-points in urban areas.

Table 3. Raw Caste and Religion Health Knowledge Gaps: Health Knowledge Regressions Including Only Caste and Religion (OLS/LPM)

	Rural	*Urban*
OBC	-0.094***	-0.169***
	[0.023]	[0.018]
Dalit	-0.133***	-0.140***
	[0.023]	[0.023]
Adivasi	-0.124***	-0.042
	[0.027]	[0.045]
Muslim	-0.093***	-0.112***
	[0.031]	[0.025]
Sikh/Jain	-0.014	0.088**
	[0.049]	[0.035]
Christian	0.220***	-0.171**
	[0.037]	[0.068]
R^2	0.012	0.023
N	19,844	11,693

Notes: High Caste (Brahmin + Other Forward Castes) is the reference category. Robust Huber-White (Huber, 1967; White, 1980) standard errors, adjusted for within-cluster correlation/clustering (Wooldridge, 2010), in brackets under parameter estimates. ***: statistically significant at 1 percent; **: statistically significant at 5 percent; *: statistically significant at 10 percent; ++: statistically significant at 15 percent +: statistically significant at 20 percent.
Source: 2004/05 India Human Development Survey (IHDS).

Question 2: *Can observable characteristics help explain the observed caste and religion health knowledge gap—in other words, does the gap "disappear" as mediating factors are introduced in the caste/religion-health knowledge relationship?*

To start answering the second question, we first need to determine the preferred estimation method—where the main options are OLS/LPM and 2SLS/IV. The results from specification tests indicate that the use of 2SLS/IV is prudent (Table B1 in the Appendix). First, the results from Wu-Hausman tests indicate that information exposure and social network formation are strongly endogenous for two out

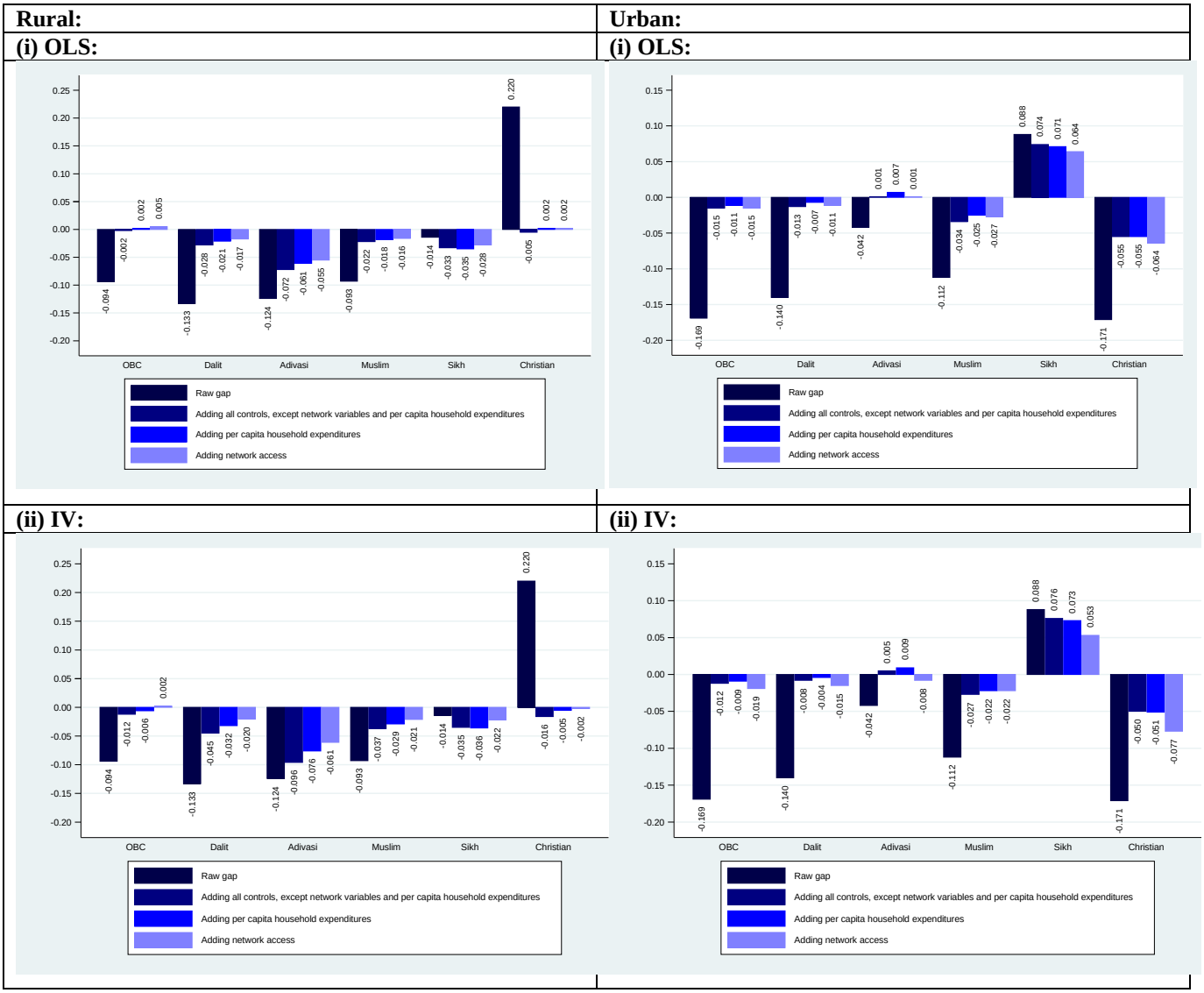
of the three specifications for urban areas and borderline endogenous for the third case (while rural areas show far less indications of endogeneity). Second, the results from the F-tests of the joint significance of the identifying instruments from the first stage of the 2SLS procedure indicate that the identifying instruments are highly correlated with all the potentially endogenous variables—with statistical significance levels of their joint statistical significance of 1 percent or better in all cases, and with high F-statistics, too.[27] It therefore seems prudent to use 2SLS/IV, since this is empirically relevant and will, thus, also address the endogeneity concerns discussed earlier. Nevertheless, we will present OLS results alongside those of 2SLS/IV as benchmarks, since the previous literature overwhelmingly has used OLS and since rural areas showed far less indications of prevalent endogenity (so that OLS/LPM in that case would yield more precise results than 2SLS/IV).

The main impression from the results from building up the models presented in Figure 1[28] is that the caste and religion gap narrows substantially when all explanatory variables (discussed in Section 3) are included in the regressions—for both the (benchmark) OLS regressions and for the (preferred) 2SLS/IV regressions and across both rural and urban areas. Indeed, statistically significant coefficients only remain for Adivasi women in rural areas—who are between about 5.5 and 7.2 percentage-points less likely to be knowledgeable about the correct treatment for diarrhea in children regarding their water intake, depending on the specification for social network access—and for Sikh and Jain women in urban areas, who are between about 5 and 7 percentage-points *more* likely to be knowledgeable about the correct treatment for diarrhea in children regarding their water intake, depending on the specification for social network access. This indicates that the combination of personal characteristics ("endowments") and coefficients ("returns" to these characteristics) are important for explaining the established raw caste health knowledge gaps. And since from the descriptive statistics (Table A1, Appendix A) high caste women are favored over low caste and Muslim women for favorable characteristics such as education, information exposure, and social network access, this already hints at differences in characteristics/endowments potentially becoming important determinants in the subsequent Oaxaca decompositions.

[27] It should be noted that since the first-stage regression is exactly identified, Hansen's (1982) J-test for over-identification is not relevant for this application.

[28] And in Table C1, Appendix C—where additional results, including standard errors and statistical significance of the coefficients presented in Figure 1, can be found.

Figure 1. Caste and Religion Coefficients: Caste and Religion, only (Raw Gap) and for Various Specifications of OLS/LPM and 2SLS/IV



Notes: All specifications include state fixed effects. Additional controls include dummies for caste and religion, educational attainment, age cohort dummies, log per capita household expenditures and dummies for access to health facilities in the community (rural areas, only).
Source: 2004/05 India Human Development Survey (IHDS).

Question 3: *How is health knowledge produced? In particular, what are the relative returns to what seems to be two of the main factors—apart from social network access—in producing health knowledge, namely education and information exposure, in terms of health knowledge?*

In Section 2, we suggested that education, information exposure, and social network access were potentially among the main determinants of health knowledge. Starting with the results for education and information exposure, the evidence from Table 4 confirms this to some extent—namely for education, where both the OLS and 2SLS/IV specifications yield substantively large coefficients,

whereas the returns to information exposure are minimal in both substantive and statistical terms. Further, mostly it is the case that the greater the education, the greater the health knowledge—though at a decreasing rate. Also, the returns to education in terms of increased health knowledge is greater in rural areas than in urban areas—possibly again reflecting that health knowledge is more scarce in rural than in urban areas, so that education is even more effective in producing it relative to what is the case in urban areas (where there are so many other sources available, including more access to health facilities). Health exposure in terms of watching TV regularly only has a marginally statistically significant coefficient (at 15 percent level of significance), though the estimated effect is substantively large: about 12 percentage-points. This is in line with, for example, the findings from the returns to education studies in the labor markets literature when education is endogenized (e.g., Card, 2001).

Table 4. Education and Information Exposure Coefficients from Full Specifications (Access Network Measure)

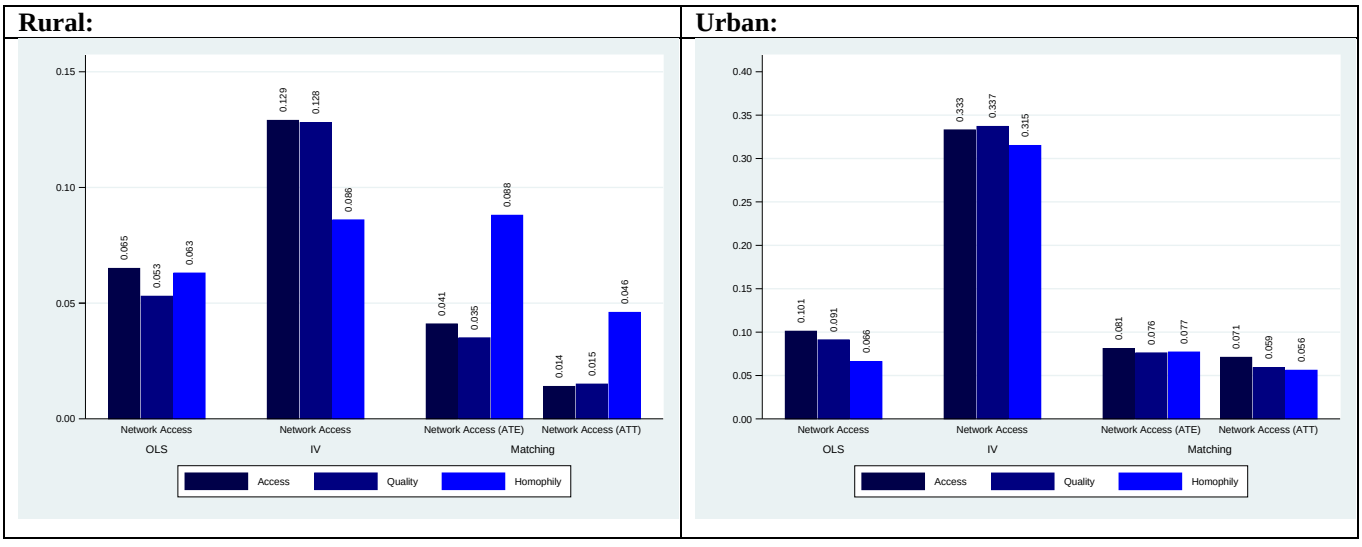
	(1) Access	
	Rural	*Urban*
(i) OLS		
Some education	0.044**	0.068***
	[0.017]	[0.022]
Primary	0.072***	0.085***
	[0.015]	[0.015]
Middle/some sec	0.119***	0.116***
	[0.018]	[0.016]
Higher secondary	0.159***	0.143***
	[0.037]	[0.022]
Tertiary	0.185***	0.163***
	[0.036]	[0.021]
Watches TV regularly	0.009	-0.003
	[0.014]	[0.014]
(ii) IV		
Some education	0.047**	0.053**
	[0.018]	[0.024]
Primary	0.078***	0.057***
	[0.019]	[0.020]
Middle/some sec	0.135***	0.067***
	[0.031]	[0.025]
Higher secondary	0.173***	0.071**
	[0.042]	[0.034]
Tertiary	0.204***	0.083**
	[0.049]	[0.034]
Watches TV regularly	-0.077	0.122++
	[0.098]	[0.085]
N	19,844	11,693

Notes: Robust Huber-White (Huber, 1967; White, 1980) standard errors, adjusted for within-cluster correlation/clustering (Wooldridge, 2010), in brackets under parameter estimates. ***: statistically significant at 1 percent; **: statistically significant at 5 percent; *: statistically significant at 10 percent; ++: statistically significant at 15 percent +: statistically significant at 20 percent. All specifications include state fixed effects and additional controls similar to previous estimations.
Source: 2004/05 India Human Development Survey (IHDS).

Question 4: *Which aspect of social networks is more important for health knowledge: access, quality, or homophily?*

To examine the importance of access to social networks for health knowledge, due to the likely presence of strong multicollinearity among the network access variables, we add each of the three health network access variables separately for a total of three different specifications, for each of OLS, IV and matching (Figure 2 and Table C2, Appendix C). From these results two observations stand out in particular. First, there appears to be a fairly strong relationship (both in statistical and substantive terms) between health knowledge and all three of the health social network access measures. For example, women having access to any health person in rural areas are about 6.5 percentage-points more likely to know about the correct treatment of diarrhea in children regarding their water intake according to the OLS/LPM estimate and about 13 percentage-points according to the IV estimate. Second, the access to health social networks is associated with even greater health knowledge in urban than in rural areas. In turn, this indicates an additional urban bias in health knowledge in addition to the previous finding of the raw caste/religion health knowledge gap being greater in urban than in rural areas. The results further reveals that all three aspects of social network formation are important, and—with a few exceptions—almost to the same degree. Again, these results assume that social network formation has homogenous effects across women from all the different castes and religions considered here—which is not likely to be the case and therefore warrant further examinations, which are pursued in the following.

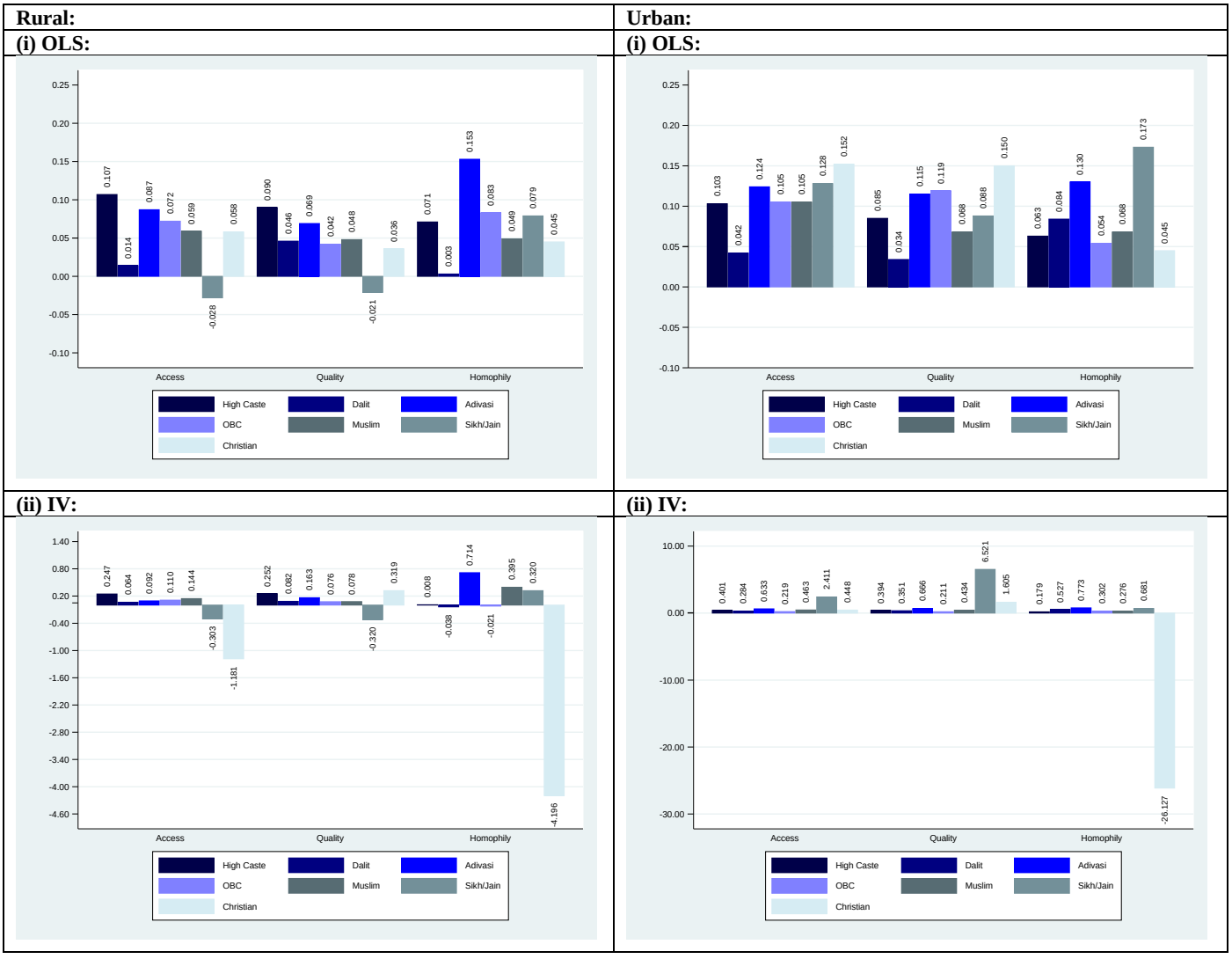
Figure 2. OLS/LPM and 2SLS/IV Health Social Network Coefficients With Three Alternative Specifications of Health Social Network Formation: Main Estimation Samples (With All Explanatory Variables Included)



Notes: All specifications include state fixed effects and additional controls similar to previous estimations.
Source: 2004/05 India Human Development Survey (IHDS).

The possibility that the relationship between health network access and health knowledge is different across caste and religious groups is explored in the analysis underlying Figures 3 and 4 (and Appendix C, Table C3), which reveals the estimated coefficients from regressions stratified by caste and religious group. That is, each coefficient in the figures (and the table) is the network formation coefficient from an individual regression, as per the caste/religion and geographical (rural/urban) designation given in the pertinent figure (and in the table). From these results three main findings stand out. First, high caste—and, to some extent, Christian—women benefit more from social network

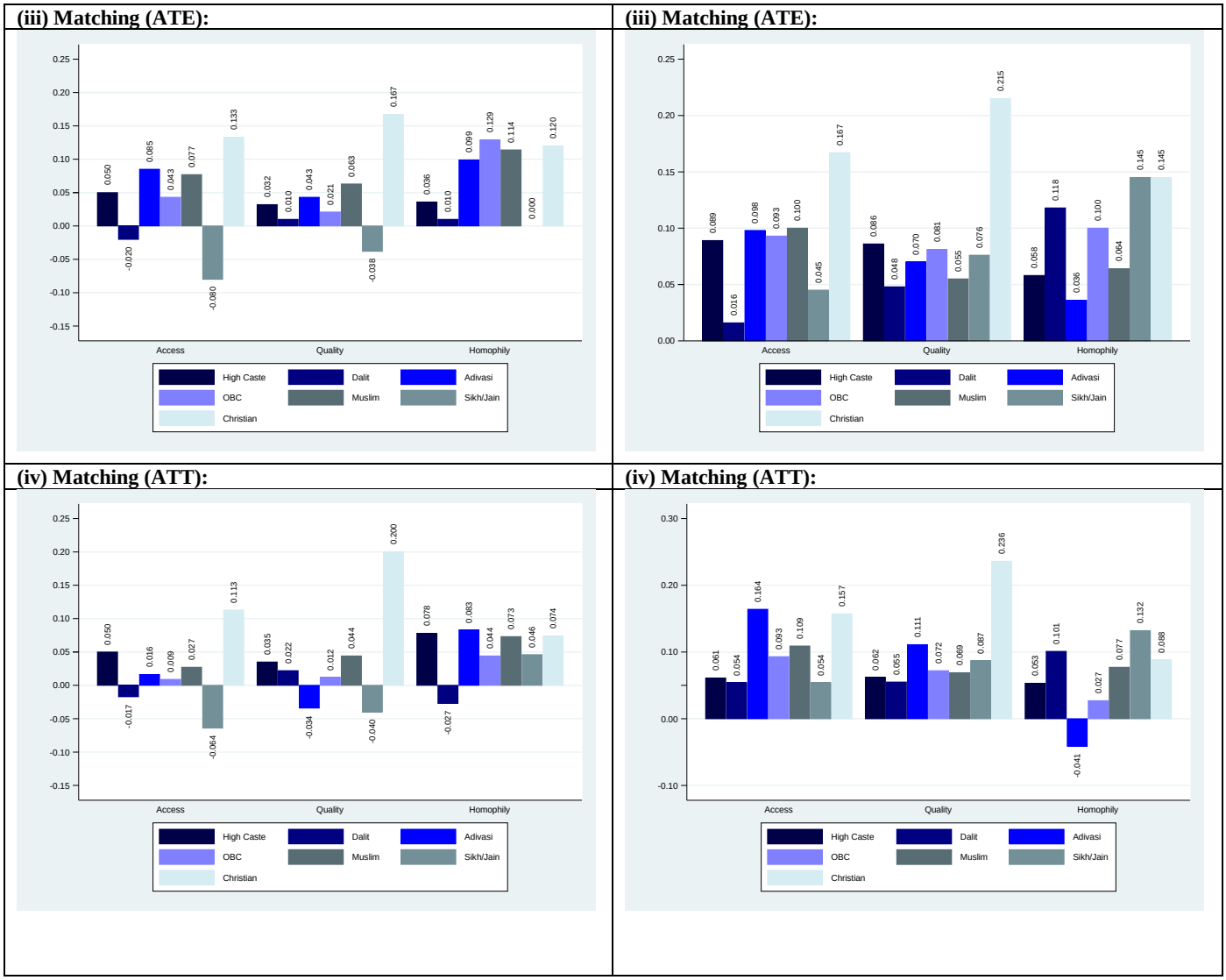
Figure 3. OLS/LPM and 2SLS/IV Health Social Network Coefficients With Three Alternative Specifications of Health Social Network Formation: Stratified Across Caste and Religion



Notes: All specifications include state fixed effects and additional controls similar to previous estimations.
Source: 2004/05 India Human Development Survey (IHDS).

formation overall in terms of their health knowledge than other groups, especially in rural areas. One exception is Adivasi women in urban areas, where the IV coefficients (the preferred specification) are far greater than those of high caste women— and across all three health network categories, and statistically significantly so. Second, however, low caste women frequently benefit far more from knowing somebody from their own jati/religion than from knowing just anybody or even a doctor, both in rural and urban areas—unlike what is the case for high caste/religion women. In turn, this indicates that—contrary to the results for the overall, pooled sample above—homophily of networks sometimes

Figure 4. Mahalanobis Matching Health Social Network Coefficients With Three Alternative Specifications of Health Social Network Formation: Stratified Across Caste and Religion



Notes: All specifications include state fixed effects and additional controls similar to previous estimations.
Source: 2004/05 India Human Development Survey (IHDS).

has much larger effects than the two other aspects of network formation, access and quality, depending on the woman's caste/religion. Third, the returns from social network formation again are mostly far greater in urban than in rural areas.

Question 5: *What can we learn from both an overall and a detailed decomposition of the observed caste and religion gap in health knowledge?*

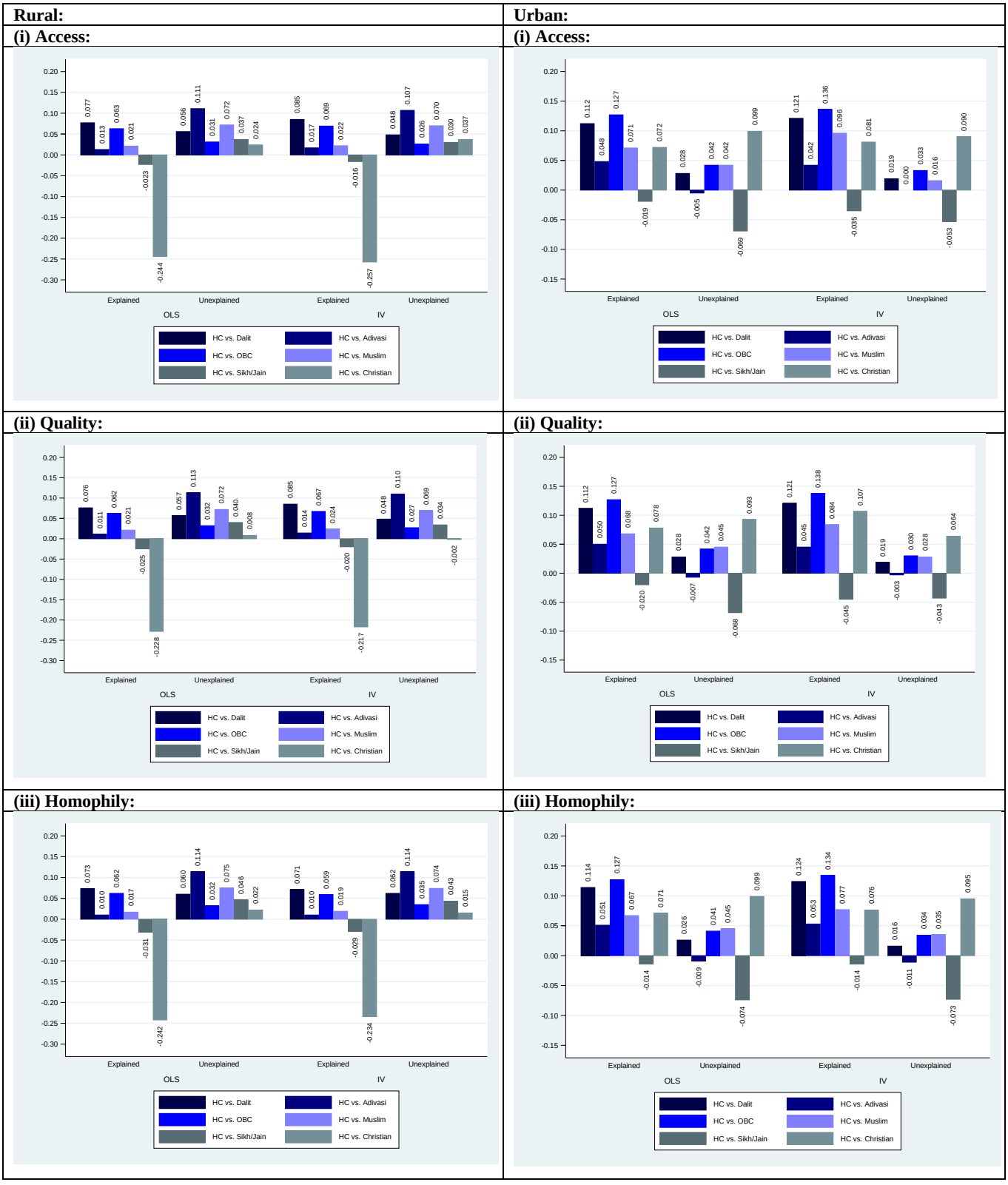
(i) Overall health knowledge gap decompositions:

A couple of results stand out particularly strongly from the results of the overall two-fold decompositions (Figure 5, and Appendix C Table C4). First, the endowments increase the overall caste/religion health knowledge gap in most cases (and that mostly both statistically and substantively significantly so), indicating that high caste women have relatively more favorable observable characteristics—that is, they have more (and possibly also better) education, are more exposed to information relevant for health knowledge production, and have more access to social networks (this will be examined more closely when considering the detailed decompositions in the next sub-section). Second, the returns to these characteristics—the unexplained part of the overall gaps—also increase the gaps in most cases. Third, however, the unexplained part of the gap is mostly much larger in rural than in urban areas—where it is frequently nil, in both substantive and statistical terms. In turn, this indicates that what has traditionally been termed "discrimination" in the labor markets literature on gender and racial pay gaps (the interpretation of the unexplained part in that literature) towards low caste and Muslim women is more prevalent in rural than in urban areas.

Notably, these results are quite robust to whether the decomposition is performed from low caste women's viewpoint (i.e., using high caste endowments and returns) or whether the decomposition is performed from high caste women's viewpoint (i.e., using low caste endowments and returns) or from any of the several other different possibilities of specifying the "absence of discrimination" group in the two-fold decompositions.[29]

[29] Due to space constraints the results are not given here but are available upon request.

Figure 5. Overall Health Knowledge Gap Decompositions With Three Alternative Specifications of Health Social Network Formation



Notes: All specifications include state fixed effects and additional controls similar to previous estimations.
Source: 2004/05 India Human Development Survey (IHDS).

(ii) Detailed health knowledge gap decompositions:

Examining the detailed caste health knowledge decompositions allows us to assess in more detail what the individual components of the overall health knowledge gaps are, in terms of specific (groups of) explanatory variables. From Tables D1 - D6 (Appendix D) the main component of the explained part overall is education—though income (as measured by per capital household consumption) and access to social networks also are important. For the IV (i.e. our preferred) specification health network access is the most important single factor (apart from the state FEs) contributing to the explained part of the health knowledge gap. These results again are quite robust to whether the decomposition is performed from low caste women's viewpoint (i.e., using high caste endowments and returns) or whether the decomposition is performed from high caste women's viewpoint (i.e., using low caste endowments and returns) or from any of the several other different possibilities of specifying the "absence of discrimination" group in the two-fold decompositions.[30]

Question 6: *What are some potential components of the mechanism underlying the observed impact of health social network formation on health knowledge?*

The massive caste gap in health social network formation is one of the obvious explanations of the observed health social impact of health social network formation on health knowledge as far as mechanism is concerned. Indeed, if lower caste/religion individuals generally have much lower access to such networks, and frequently also of lower quality, and of less likelihood of being homophilous networks the marginal benefit when they do can also be expected to be much larger than for individuals from higher castes/religions. Table A1 from the Appendix reveals that, for example, in terms of access alone there are substantial differences in network formation, favoring women from higher castes/ religions—with an access of 40.7 percent for high caste women, 25.8 percent for Dalits, 17.7 percent for Adivasis, 34.1 percent for OBCs, and 25.1 from Muslim women. These gaps persist when moving to the quality and homophily aspects of network formation, where they are even greater, in relative terms. For homophily, for example, network formation among high caste women is 17.5 percent, while it is about only a third of this among Dalit and Adivasi women, at about 6 percent for both groups.

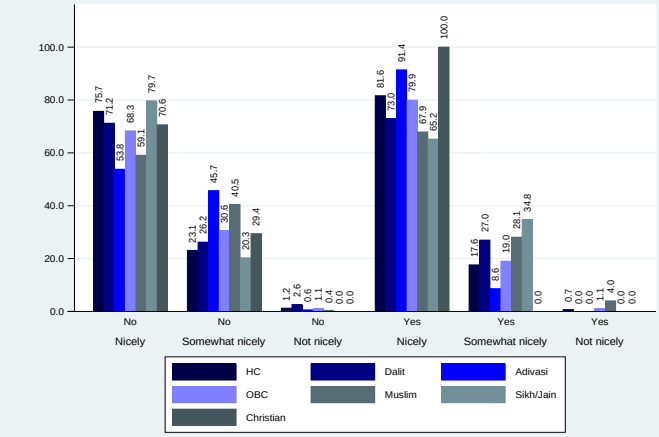
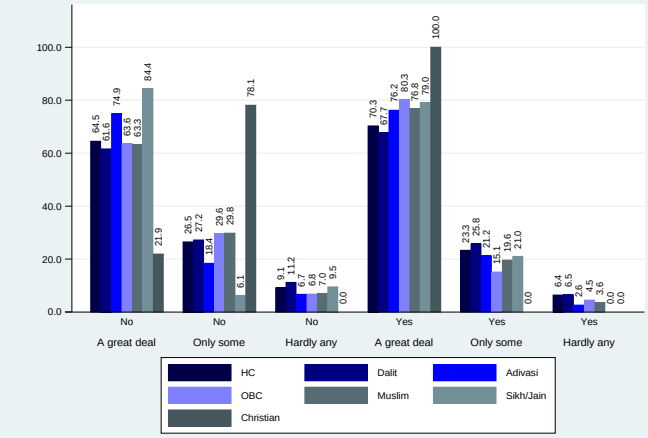
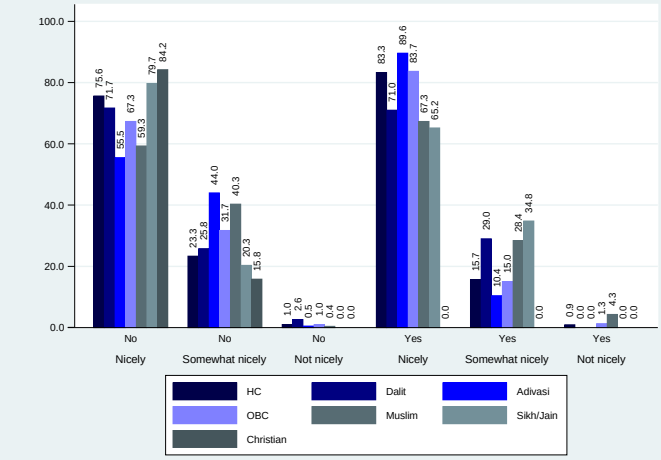
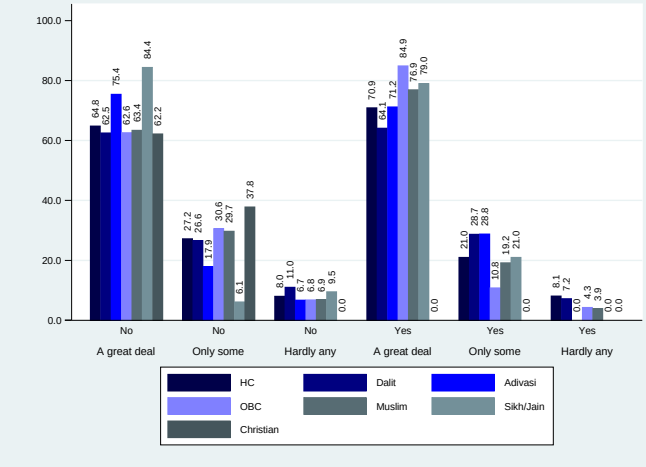
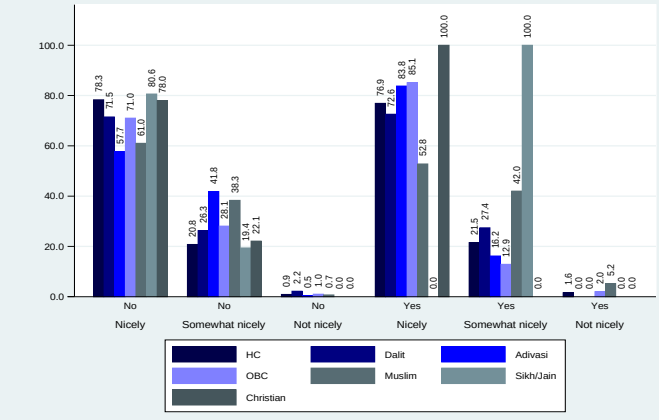
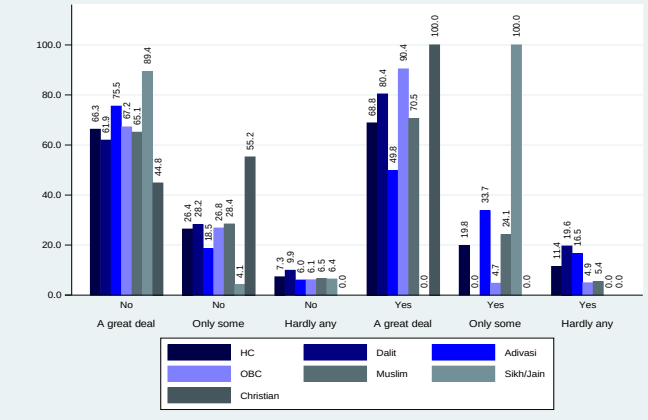
But the data allow further exploration into the potential mechanism behind the observed asymmetric impact of network formation on health knowledge. Among many other things, women are asked about their most recent consultation with a health worker—what it was for, and how they were

[30] Due to space constraints the results from this sensitivity analysis are not given here but are available upon request.

treated. Conditioning on women who went for diarrhea, we next explore how women from different castes and religions were treated on their most recent visit. The thinking here is that access to health social networks likely would give both more self-esteem and more knowledge about what and how to ask medical personnel—all of which is likely to improve how the woman is treated by the health worker, part of which include sharing additional health knowledge with the woman. From Figure 6, left panel (and Table C5, Appendix C) it can be seen that in almost all cases, network formation improves how the woman was treated, in terms of increasing the "Nicely" and decreasing the "Somewhat Nicely" and "Not Nicely" groups. Indeed, in several cases both the absolute and relative increases are much higher for women from the lower castes/religions. An issue here, of course, is that the number of observations is frequently quite small (as can also be seen from Table C5, Appendix C).

Another possible mechanism through which health social network formation might affect health knowledge is through its impact on peoples' confidence in the medical profession. Fortunately, households were also asked about their confidence in hospitals and doctors to provide good treatment. We once again restrict the sample to households, where women consulted a health worker for a diarrhea related illness. From Figure 6, right panel (and Table C5, Appendix C) we see how once again, in agreement with the results in Figure 7, there is an increase in the "A great deal" category and similarly decreases in the "Only some" and "Hardly any" categories whenever the household is engaged in health social network formation. In several cases both the absolute and relative increases are again much higher for women from the lower castes/religions. But again the number of observations frequently being quite small becomes an issue (as can also be seen from Table C5, Appendix C).

Figure 6. Perception of Medical Treatment & Confidence in Medical Institutions: How Woman Was Treated Last Time She Consulted a Health Worker for Diarrhea Related Illness and Confidence in Doctors and Hospitals (Stratified by Caste/Religion and Three Alternative Types of Health Social Network Formation)

How Woman Was Treated:	*Confidence in Institutions:*
(i) Access:	**(i) Access:**
	
(ii) Quality:	**(ii) Quality:**
	
(iii) Homophily:	**(iii) Homophily:**
	

Notes: Calculations incorporate sampling weights.
Source: 2004/05 India Human Development Survey (IHDS).

6. Conclusion

This paper examines the gap in health knowledge as measured by knowledge of the correct treatment of diarrhea in children regarding their water intake across caste and religious groups in India in terms of its prevalence, magnitude and determinants using a recent data set and thereby adds to the emerging literature on caste, religion, and health knowledge.

Estimation of raw caste and religion health knowledge gaps and overall and detailed earnings decompositions leads to six main results: (1) education and social network access are all strongly associated with increased health knowledge; (2) the presence of a substantively large health knowledge caste gap (favoring high caste women); (3) evidence that the endowments and the returns to these endowments increase the health knowledge gaps—indicating that high caste women have higher education and better access to social networks and also higher returns to these characteristics; (4) network homophily frequently is much more important for low caste women than either network access or network quality—sometimes double or even more—unlike what is the case for high caste/religion women; (5) while observed individual characteristics explain part of the gaps, a substantial part of the health knowledge gap is left unexplained, especially in rural areas; and (6) improved treatment from and confidence in the medical profession is found to be part of the mechanism linking health social network formation with improved health knowledge about the treatment of diarrhea in children.

These results have strong policy implications. First, the continued presence of a caste and religion health knowledge gap regarding the correct treatment of diarrhea in children regarding their water intake is likely to lead to continued higher child mortality for children from low caste and Muslim backgrounds. Notably, these are deaths that could have been averted, had the mothers only been taught the arguably simple—and at the same time also relatively cheap—measure of increasing the water intake for their children when experiencing diarrhea. Second, the finding that low caste and Muslim women frequently benefit (much) more than high caste women from homophilous networks indicate that it may not be enough to give individuals access to high quality networks if caste and religion-related gaps in health knowledge are to be reduced; such networks also have to be homophilous, to have the maximum effect. Third, while somewhat contentious, the decomposition results are consistent with the presence of what has traditionally been termed “discrimination” in the labor markets literature on gender and racial pay gaps—but here instead indicate the presence of continued discrimination towards low caste and Muslim women in the context of health knowledge, especially in rural areas. Even if one were to dismiss this interpretation of the decomposition results, however, low caste and Muslim women still appears to be subject to discrimination in terms of their

much lower access to education and social networks. In turn, this points towards the importance of continued attention towards education, institutions and economic policy for decreasing the caste and religion health knowledge gap in India—notably through increased attention towards the education system and public provision of health information campaigns and health networks, possibly through an increase in job reservations favoring lower caste and Muslim health professionals.

References:

Abadie A. and Imbens, G. (2006) "Large Sample Properties of Matching Estimators for Average Treatment Effects," *Econometrica* 74(1): 235-267.

Banerjee R, Cohen-Cole E, and Zanella G. (2007) "Demonstration Effects in Preventive Care," Working Paper No. QAU07-7. Mayo Clinic College of Medicine.

Borooah, Vani and Sriya Iyer (2005) "Vidya, Veda, and Varna: The influence of religion and caste on education in rural India," *The Journal of Development Studies* 41(8): 1369-1404.

Becker, Gary S. (1964) *Human Capital*, Chicago: University of Chicago Press.

Blinder, A.S. (1973) "Wage Discrimination: Reduced Form and Structural Estimates," *Journal of Human Resources* 8: 436-455.

Borooah, Vani K. (2012) Inequality in Health Outcomes in India: The Role of Caste and Religion," in Sukhadeo Thorat and Katherine S. Newman (eds): *Blocked by Caste: Economic Discrimination in Modern India*, New Delhi: Oxford University Press.

Borooah, Vani K., Nidhi Sadana Sabharwal, and Sukhadeo Thora (2012) "Gender and Caste-Based Inequality in Health Outcomes in India," Working Paper, Vol. 6, No. 3, New Delhi: Indian Institute of Dalit Studies.

Card, David (2001) "Estimating the Return to Schooling: Progress on Some Persistent Econometric Problems," *Econometrica* 69(5): 1127-1160.

Chin, Aimee and Nishith Prakash (2011) "The redistributive effects of political reservation for minorities: Evidence from India," *Journal of Development Economics* 96 (2): 265-277.

Christakis, Nicholas A., Fowler, James H. (2007) "The Spread of Obesity in a Large Social Network over 32 Years," *New England Journal of Medicine* 357: 370-379.

Cutler, D.M. and Glaeser, E.L. (2007) "Social Interactions and Smoking," *NBER Working Paper* No. 13477.

Cotton, J. (1988) "On the Decomposition of Wage Differentials," *Review of Economics and Statistics* 70: 236-243.

Das, Smita (2012) "Caste, Ethnicity, and Religion: Linkages with Unemployment and Poverty," in Sukhadeo Thorat and Katherine S. Newman (eds): *Blocked by Caste: Economic Discrimination in Modern India*, New Delhi: Oxford University Press.

Das, Suranjan (2013) "Health Culture of Scheduled Caste: A Case Study of Patni in Cachar District of Assam," International Research Journal of Social Sciences 2(12), 35-41.

Desai, Sonalde, Reeve Vanneman, and National Council of Applied Economic Research, New Delhi (2010). India Human Development Survey (IHDS), 2005 [Computer file].

ICPSR22626-v8. Ann Arbor, MI: Inter-university Consortium for Political and Social Research [distributor], 2010-06-29. doi:10.3886/ICPSR22626.v8

Dustmann, C. and I. Preston (2001) "Attitudes to Ethnic Minorities, Ethnic Context and Location Decisions," *Economic Journal* 111(470): 353-373.

Glewwe, Paul (1999) "Why Does Mother's Schooling Raise Child Health in Developing Countries? Evidence from Morocco," *Journal of Human Resources* 34(1): 124-159.

Grossman, Michael (1972) "On the Concept of Health Capital and the Demand for Health," *Journal of Political Economy* 80(2): 223-55.

Heckman, James J., Lance J. Lochner, and Petra E. Todd (2008) "Earnings Functions and Rates of Return," *Journal of Human Capital* 2(1):1-31.

Huber, P. J. (1967) "The Behavior of Maximum Likelihood Estimates under Nonstandard Conditions," in: Lucien M. Le Cam and Jerzy Neyman (Eds.) *Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability Vol. 1*, Berkeley, CA: University of California Press.

Imbens, Guido W. (2004) "Nonparamteric Estimation of Average Treatement Effects Under Exogeneity: A Review," *The Review of Economics and Statistics* 86(1): 4-29.

Jann, Ben (2008) "The Blinder-Oaxaca decomposition for linear regression models," *The Stata Journal* 8(4): 453-479.

Jensen, Peter and Astrid Würtz Rasmussen (2011) "The effect of immigrant concentration in schools on native and immigrant children's reading and math skills," *Economics of Education Review* 30(6): 1503–1515.

Kenkel, Donald S (1991) "Health Behavior, Health Knowledge, and Schooling," *Journal of Political Economy* 99(2): 287-305.

Kovsted, Jens, Claus C. Pörtner and Finn Tarp (2003) "Child Health and Mortality: Does Health Knowledge Matter?" *Journal of African Economies* 11(4): 542-560.

Mincer, Jacob (1974) *Schooling, Experience and Earnings*, New York: National Bureau of Economic Research.

Neumark, D. (1988) "Employers' Discriminatory Behavior and the Estimation of Wage Discrimination," *Journal of Human Resources* 23: 279-295.

Oaxaca, R. (1973) "Male-Female Wage Differentials in Urban Labor Markets," *International Economic Review* 14: 693-709.

Oaxaca, R.L., Ransom, M.R. (1998) "Calculation of approximate variances for wage Decomposition differentials," *Journal of Economic and Social Measurement* 24: 55-61.

Oaxaca, R.L. and M.R. Ransom (1999) "Identification in Detailed Wage Decompositions," *Review of Economics and Statistics* 81: 154-157.

Patra, Shraboni, Arokiasamy Perianayagam, and Srinivas Goli (2013) "Mother's health knowledge and practice and their linkage with childhood morbidity, medical care and medical care expenditure in India," Paper presented at the 27th IUSSP International Polulation Conference, 26-31 August, 2013, BEXCO, Busan, Korea.

Pradhan, R. and Meinzen-Dick, R. (2010) "Which Rights are Right? Water Rights, Culture and Underlying Values," in Peter G. Brown and Jeremy J. Schmidt (eds): *Water Ethics: Foundational Readings for Students and Professionals*: Washington D.C., Island Press.

Ramachandran,V. and T.Naorem (2013) "What It Means To Be a Dalit or Tribal Child in Our Schools? A Synthesis of a Six-State Qualitative Study," *Economic and Political WEEKLY*, November 2, 2013, Vol. XLVIII no. 44, 43-52.

Rao, Neel, Mobius, Markus M. and Rosenblat, Tanya (2007) "Social Networks and Vaccination Decisions". Federal Reserve Band of Boston Working Paper No. 07-12.

Reimers, C.W. (1983) "Labor Market Discrimination against Hispanic and Black Men," *Review of Economics and Statistics* 65: 570-579.

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

Rosenbaum, P., and D. Rubin (1985) "Constructing a Control Group Using Multivariate Matched Sampling Methods that Incorporate the Propensity," *American Statistician*, 39, 33-38.

Subramanian, S.V., Shailen Nandy, Michelle Irving, Dave Gordon, Helen Lambert, and George Davey Smith (2006) "The Mortality Divide in India: The Differential Contributions of Gender, Caste, and Standard of Living Across the Life Course," Am J Public Health 96(5): 818–825.

White, H. (1980) "A Heteroskedasticity-Consistent Covariance Matrix Estimator and a Direct Test for Heteroskedasticity," *Econometrica* 48(4):817–830.

WHO (2013) "Diarrhoeal Disease," Fact sheet N°330, April 2013, Geneva, Switzerland: World Health Organization. http://www.who.int/mediacentre/factsheets/fs330/en/ (Accessed online July 11, 2014).

Wooldridge, Jeffrey M. (2010) *Econometric Analysis of Cross-Section and Panel Data*, Second Edition, Cambridge, Massachusetts: The MIT Press.

Yun, Myeong-Su (2005) "A Simple Solution to the Identification Problem in Detailed Wage Decompositions," *Economic Inquiry* 43: 766-772.

APPENDIX A: Descriptive Statistics

Table A1. Means and Standard Deviations for Estimation Samples

	Rural							*Urban*						
	High caste	*Dalit*	*Adivasi*	*OBC*	*Muslim*	*Sikh/Jain*	*Christian*	*High caste*	*Dalit*	*Adivasi*	*OBC*	*Muslim*	*Sikh/Jain*	*Christian*
Dependent variable														
Correct treatment of diarrhea in children, re water intake	0.605	0.472	0.481	0.511	0.512	0.591	0.825	0.67	0.53	0.627	0.501	0.557	0.757	0.499
	[0.489]	[0.499]	[0.500]	[0.500]	[0.500]	[0.493]	[0.381]	[0.470]	[0.499]	[0.484]	[0.500]	[0.497]	[0.430]	[0.501]
Explanatory variables														
15-19 years old	0.025	0.05	0.046	0.034	0.043	0.016	0.005	0.013	0.018	0.028	0.016	0.025	0.007	0.012
	[0.156]	[0.218]	[0.210]	[0.181]	[0.202]	[0.127]	[0.068]	[0.112]	[0.134]	[0.165]	[0.127]	[0.157]	[0.082]	[0.107]
20-24 years old	0.136	0.153	0.164	0.159	0.172	0.061	0.068	0.097	0.165	0.13	0.134	0.149	0.093	0.045
	[0.343]	[0.360]	[0.370]	[0.366]	[0.377]	[0.239]	[0.251]	[0.296]	[0.371]	[0.337]	[0.340]	[0.356]	[0.292]	[0.208]
25-29 years old	0.189	0.186	0.186	0.176	0.179	0.144	0.094	0.18	0.193	0.179	0.204	0.2	0.169	0.17
	[0.391]	[0.389]	[0.389]	[0.381]	[0.384]	[0.352]	[0.293]	[0.384]	[0.395]	[0.384]	[0.403]	[0.400]	[0.375]	[0.377]
30-34 years old	0.2	0.197	0.18	0.191	0.179	0.21	0.24	0.196	0.197	0.175	0.194	0.204	0.132	0.234
	[0.400]	[0.398]	[0.385]	[0.393]	[0.383]	[0.408]	[0.428]	[0.397]	[0.398]	[0.380]	[0.395]	[0.403]	[0.340]	[0.424]
35-39 years old	0.198	0.194	0.201	0.196	0.184	0.277	0.276	0.211	0.191	0.21	0.194	0.186	0.238	0.202
	[0.399]	[0.396]	[0.401]	[0.397]	[0.388]	[0.448]	[0.448]	[0.408]	[0.393]	[0.408]	[0.396]	[0.390]	[0.427]	[0.402]
40-44 years old	0.149	0.128	0.138	0.141	0.145	0.164	0.148	0.176	0.146	0.139	0.147	0.147	0.222	0.173
	[0.356]	[0.334]	[0.346]	[0.348]	[0.352]	[0.371]	[0.356]	[0.381]	[0.353]	[0.347]	[0.354]	[0.354]	[0.416]	[0.379]
45-49 years old	0.104	0.092	0.085	0.103	0.099	0.128	0.169	0.128	0.089	0.139	0.111	0.089	0.139	0.164
	[0.305]	[0.289]	[0.278]	[0.304]	[0.299]	[0.334]	[0.376]	[0.334]	[0.285]	[0.346]	[0.315]	[0.285]	[0.347]	[0.371]
No education	0.353	0.684	0.694	0.548	0.633	0.344	0.084	0.122	0.391	0.408	0.253	0.396	0.062	0.067
	[0.478]	[0.465]	[0.461]	[0.498]	[0.482]	[0.476]	[0.278]	[0.328]	[0.488]	[0.492]	[0.435]	[0.489]	[0.242]	[0.250]
Some edu (primary incomplete)	0.076	0.079	0.083	0.084	0.106	0.062	0.042	0.044	0.099	0.066	0.061	0.09	0.021	0.039
	[0.265]	[0.270]	[0.275]	[0.278]	[0.308]	[0.242]	[0.201]	[0.204]	[0.299]	[0.248]	[0.239]	[0.286]	[0.143]	[0.194]
Primary	0.228	0.124	0.118	0.167	0.135	0.231	0.133	0.163	0.184	0.161	0.206	0.222	0.142	0.192
	[0.419]	[0.330]	[0.322]	[0.373]	[0.342]	[0.422]	[0.341]	[0.370]	[0.388]	[0.368]	[0.405]	[0.416]	[0.350]	[0.395]
Middle/some secondary	0.271	0.095	0.091	0.168	0.107	0.299	0.375	0.346	0.254	0.194	0.31	0.209	0.411	0.346
	[0.445]	[0.293]	[0.287]	[0.374]	[0.309]	[0.459]	[0.485]	[0.476]	[0.435]	[0.396]	[0.462]	[0.407]	[0.493]	[0.477]
Higher secondary	0.046	0.013	0.012	0.022	0.015	0.035	0.217	0.116	0.038	0.074	0.08	0.042	0.129	0.178
	[0.210]	[0.113]	[0.107]	[0.145]	[0.122]	[0.184]	[0.413]	[0.320]	[0.190]	[0.263]	[0.272]	[0.200]	[0.336]	[0.383]
Tertiary	0.026	0.005	0.004	0.011	0.004	0.028	0.149	0.208	0.034	0.097	0.09	0.041	0.235	0.179
	[0.159]	[0.069]	[0.060]	[0.106]	[0.063]	[0.166]	[0.357]	[0.406]	[0.182]	[0.297]	[0.286]	[0.199]	[0.425]	[0.384]
Network access	0.407	0.258	0.177	0.341	0.251	0.484	0.316	0.411	0.306	0.368	0.355	0.272	0.563	0.442
	[0.491]	[0.438]	[0.382]	[0.474]	[0.434]	[0.501]	[0.466]	[0.492]	[0.461]	[0.483]	[0.479]	[0.445]	[0.497]	[0.498]
Network quality	0.331	0.199	0.122	0.274	0.203	0.457	0.109	0.319	0.228	0.238	0.268	0.21	0.497	0.274
	[0.471]	[0.399]	[0.328]	[0.446]	[0.403]	[0.499]	[0.312]	[0.466]	[0.420]	[0.426]	[0.443]	[0.408]	[0.501]	[0.447]
Network homophily	0.175	0.06	0.06	0.098	0.11	0.167	0.272	0.194	0.093	0.186	0.15	0.14	0.215	0.269
	[0.380]	[0.237]	[0.237]	[0.297]	[0.313]	[0.374]	[0.446]	[0.395]	[0.290]	[0.389]	[0.357]	[0.348]	[0.412]	[0.444]
Watches TV regularly	0.47	0.248	0.139	0.311	0.231	0.682	0.582	0.823	0.641	0.652	0.707	0.597	0.851	0.763

	[0.499]	[0.432]	[0.346]	[0.463]	[0.422]	[0.467]	[0.494]	[0.382]	[0.480]	[0.477]	[0.455]	[0.491]	[0.357]	[0.426]
Ln (per cap household expenditures)	6.592	6.227	5.926	6.352	6.337	6.937	6.626	7.107	6.641	6.714	6.835	6.575	7.278	7.051
	[0.602]	[0.547]	[0.600]	[0.595]	[0.561]	[0.546]	[0.757]	[0.617]	[0.567]	[0.659]	[0.607]	[0.577]	[0.554]	[0.621]
Health Sub-center in community	0.42	0.489	0.351	0.44	0.433	0.51	0.8							
	[0.494]	[0.500]	[0.478]	[0.496]	[0.496]	[0.501]	[0.401]							
Primary Health Center in community	0.14	0.155	0.118	0.186	0.164	0.206	0.736							
	[0.347]	[0.362]	[0.322]	[0.390]	[0.370]	[0.405]	[0.442]							
Community Health Center	0.032	0.026	0.023	0.02	0.017	0.049	0.194							
	[0.176]	[0.160]	[0.151]	[0.139]	[0.130]	[0.215]	[0.396]							
Government Maternity Center	0.029	0.062	0.017	0.062	0.016	0.035	0.018							
	[0.167]	[0.241]	[0.131]	[0.241]	[0.126]	[0.184]	[0.132]							
Govt. Comm. Disease Facility (e.g., TB)	0.028	0.053	0.026	0.057	0.026	0.038	0.003							
	[0.164]	[0.225]	[0.160]	[0.232]	[0.159]	[0.191]	[0.054]							
Delhi	0.001	0	0	0.001	0	0	0	0.062	0.065	0.008	0.027	0.047	0.079	0.014
	[0.028]	[0.011]	[0.000]	[0.037]	[0.000]	[0.000]	[0.000]	[0.241]	[0.246]	[0.090]	[0.162]	[0.213]	[0.271]	[0.119]
Jammu/Kashmir	0.012	0.003	0.001	0	0.083	0	0.002	0.004	0.002	0.005	0.001	0.025	0.002	0
	[0.108]	[0.055]	[0.029]	[0.013]	[0.276]	[0.000]	[0.042]	[0.064]	[0.046]	[0.073]	[0.025]	[0.155]	[0.040]	[0.000]
Himachal Pradesh	0.029	0.009	0.002	0.002	0.001	0.004	0	0.005	0.003	0.007	0	0.001	0.003	0
	[0.166]	[0.094]	[0.043]	[0.045]	[0.031]	[0.061]	[0.000]	[0.070]	[0.058]	[0.084]	[0.021]	[0.024]	[0.054]	[0.000]
Uttarakhand	0.029	0.033	0	0.011	0.007	0	0	0.019	0.011	0	0.005	0.006	0.011	0
	[0.167]	[0.179]	[0.000]	[0.104]	[0.085]	[0.000]	[0.000]	[0.136]	[0.102]	[0.000]	[0.073]	[0.077]	[0.106]	[0.000]
Punjab	0.011	0.031	0	0.013	0.003	0.82	0.017	0.039	0.054	0.002	0.015	0.002	0.237	0.04
	[0.102]	[0.174]	[0.022]	[0.113]	[0.053]	[0.385]	[0.131]	[0.192]	[0.226]	[0.048]	[0.121]	[0.048]	[0.426]	[0.196]
Haryana	0.039	0.025	0.001	0.015	0.01	0.066	0	0.028	0.013	0.003	0.008	0.009	0.024	0
	[0.193]	[0.157]	[0.033]	[0.122]	[0.102]	[0.249]	[0.000]	[0.164]	[0.115]	[0.052]	[0.090]	[0.095]	[0.153]	[0.000]
Uttar Pradesh	0.136	0.136	0.012	0.146	0.191	0.01	0	0.094	0.069	0.019	0.058	0.203	0.051	0.015
	[0.343]	[0.343]	[0.108]	[0.353]	[0.393]	[0.102]	[0.000]	[0.292]	[0.254]	[0.138]	[0.235]	[0.402]	[0.220]	[0.123]
Bihar	0.057	0.083	0.006	0.128	0.085	0	0.004	0.014	0.016	0.002	0.038	0.031	0.003	0
	[0.231]	[0.275]	[0.076]	[0.334]	[0.279]	[0.000]	[0.063]	[0.116]	[0.126]	[0.049]	[0.192]	[0.173]	[0.059]	[0.000]
Jharkhand	0.029	0.029	0.142	0.031	0.077	0	0	0.022	0.019	0.093	0.028	0.006	0.031	0.003
	[0.167]	[0.168]	[0.349]	[0.175]	[0.266]	[0.000]	[0.000]	[0.147]	[0.137]	[0.291]	[0.165]	[0.080]	[0.175]	[0.056]
Rajasthan	0.063	0.056	0.046	0.045	0.027	0.026	0	0.035	0.031	0.083	0.032	0.057	0.072	0
	[0.244]	[0.231]	[0.209]	[0.207]	[0.161]	[0.161]	[0.000]	[0.184]	[0.172]	[0.277]	[0.175]	[0.233]	[0.258]	[0.000]
Chhattisgarh	0.006	0.016	0.137	0.04	0.003	0	0	0.023	0.014	0.031	0.03	0.002	0.02	0.005
	[0.075]	[0.127]	[0.344]	[0.196]	[0.056]	[0.000]	[0.000]	[0.150]	[0.117]	[0.174]	[0.170]	[0.043]	[0.141]	[0.068]
Madhya Pradesh	0.052	0.046	0.111	0.057	0.008	0.011	0	0.032	0.052	0.127	0.036	0.068	0.07	0.014
	[0.221]	[0.209]	[0.315]	[0.233]	[0.087]	[0.106]	[0.000]	[0.175]	[0.221]	[0.334]	[0.187]	[0.252]	[0.256]	[0.116]
Northeast	0.009	0.006	0.06	0.006	0.007	0	0.019	0.006	0.004	0.125	0.007	0.003	0.008	0
	[0.095]	[0.077]	[0.238]	[0.080]	[0.083]	[0.000]	[0.135]	[0.080]	[0.063]	[0.331]	[0.084]	[0.058]	[0.092]	[0.018]
Assam	0.01	0.006	0.04	0.013	0.105	0	0	0.013	0.014	0.02	0.013	0.017	0.003	0
	[0.100]	[0.076]	[0.197]	[0.115]	[0.307]	[0.000]	[0.000]	[0.113]	[0.119]	[0.139]	[0.114]	[0.129]	[0.054]	[0.000]
West Bengal	0.09	0.147	0.053	0.021	0.215	0.023	0.009	0.14	0.097	0.056	0.032	0.086	0	0.033
	[0.286]	[0.354]	[0.224]	[0.143]	[0.411]	[0.150]	[0.097]	[0.347]	[0.296]	[0.231]	[0.176]	[0.280]	[0.000]	[0.178]

Orissa	0.022	0.045	0.093	0.07	0.002	0	0.006	0.018	0.026	0.043	0.026	0.006	0.002	0.002
	[0.145]	[0.208]	[0.291]	[0.256]	[0.049]	[0.000]	[0.078]	[0.132]	[0.159]	[0.202]	[0.159]	[0.075]	[0.048]	[0.046]
Gujarat	0.075	0.019	0.08	0.034	0.024	0	0	0.097	0.048	0.116	0.067	0.054	0.08	0.034
	[0.264]	[0.136]	[0.272]	[0.181]	[0.153]	[0.000]	[0.000]	[0.296]	[0.213]	[0.321]	[0.250]	[0.225]	[0.271]	[0.181]
Maharashtra, Goa	0.166	0.061	0.111	0.088	0.024	0.005	0.043	0.203	0.188	0.123	0.136	0.11	0.242	0.154
	[0.372]	[0.238]	[0.314]	[0.283]	[0.153]	[0.071]	[0.204]	[0.402]	[0.391]	[0.329]	[0.343]	[0.314]	[0.429]	[0.362]
Andhra Pradesh	0.104	0.13	0.042	0.11	0.04	0	0.007	0.078	0.064	0.032	0.109	0.067	0	0.054
	[0.305]	[0.336]	[0.201]	[0.313]	[0.195]	[0.000]	[0.085]	[0.268]	[0.245]	[0.177]	[0.311]	[0.250]	[0.000]	[0.226]
Karnataka	0.025	0.039	0.056	0.065	0.042	0.034	0.013	0.037	0.054	0.101	0.086	0.078	0.062	0.124
	[0.157]	[0.194]	[0.230]	[0.247]	[0.200]	[0.182]	[0.113]	[0.189]	[0.227]	[0.301]	[0.280]	[0.268]	[0.242]	[0.331]
Kerala	0.033	0.012	0.006	0.021	0.046	0	0.787	0.014	0.014	0.002	0.032	0.057	0	0.234
	[0.178]	[0.110]	[0.077]	[0.142]	[0.210]	[0.000]	[0.410]	[0.116]	[0.119]	[0.048]	[0.177]	[0.231]	[0.000]	[0.424]
Tamil Nadu	0.006	0.067	0	0.081	0.001	0	0.092	0.019	0.142	0	0.213	0.065	0	0.274
	[0.075]	[0.250]	[0.022]	[0.273]	[0.022]	[0.000]	[0.290]	[0.138]	[0.349]	[0.000]	[0.409]	[0.247]	[0.000]	[0.447]
N	3,747	4,491	2,056	7,121	1,877	310	242	3,566	1,915	367	3,604	1,785	228	228

Notes: Values in brackets are robust Huber-White (Huber, 1967; White, 1980) standard deviations.
Source: 2004/05 India Human Development Survey (IHDS).

APPENDIX B: Specification Tests

Table B1. Specification Tests for 2SLS/IV Health Knowledge Regressions: Predictive Power of Identifying Instruments (First Stage); Endogeneity (Second Stage)

	Rural			*Urban*		
	Access	*Quality*	*Homophily*	*Access*	*Quality*	*Homophily*
(i) Joint F-test of predictive power of IVs						
Network formation	463.63	455.48	89.99	447.21	333.25	181.21
	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]
Watches TV regularly	77.41	79.04	80.97	174.31	172.68	165.36
	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]
(ii) Wu (1973)-Hausman (1978) endogeneity test	1.50	1.57	0.72	8.51	6.60	2.32
	[0.224]	[0.208]	[0.488]	[0.000]	[0.001]	[0.098]
N	19,844	19,844	19,844	11,693	11,693	11,693

Notes: Terms in brackets are the p-values of the corresponding test-statistic. The tests employ robust Huber-White (Huber, 1967; White, 1980) standard errors and also adjust for within-cluster correlation/clustering (Wooldridge, 2010). All estimations include state fixed-effects and additional explanatory variables similar to those used for the main estimations.
Source: 2004/05 India Human Development Survey (IHDS).

APPENDIX C: Detailed Results Tables Underlying Estimated Coefficients and Descriptive Analysis of Mechanisms Reported in Figures 1 through 6[31]

Table C1. Caste and Religion Coefficients: Caste and Religion, only (Raw Gap) and for Various Specifications of OLS/LPM and 2SLS/IV

	(1) Raw gap		*(2) Adding all controls, except network variables and per capita household expenditures*		*(3) Adding per capita household expenditures*		*(4) Adding network access*	
	Rural	*Urban*	*Rural*	*Urban*	*Rural*	*Urban*	*Rural*	*Urban*
(i) OLS								
OBC	-0.094***	-0.169***	-0.002	-0.015	0.002	-0.011	0.005	-0.015
	[0.023]	[0.018]	[0.018]	[0.014]	[0.018]	[0.014]	[0.018]	[0.014]
Dalit	-0.133***	-0.140***	-0.028+	-0.013	-0.021	-0.007	-0.017	-0.011
	[0.023]	[0.023]	[0.020]	[0.019]	[0.020]	[0.019]	[0.020]	[0.019]
Adivasi	-0.124***	-0.042	-0.072***	0.001	-0.061**	0.007	-0.055**	0.001
	[0.027]	[0.045]	[0.025]	[0.036]	[0.025]	[0.036]	[0.025]	[0.036]
Muslim	-0.093***	-0.112***	-0.022	-0.034+	-0.018	-0.025	-0.016	-0.027
	[0.031]	[0.025]	[0.027]	[0.023]	[0.027]	[0.023]	[0.027]	[0.023]
Sikh	-0.014	0.088**	-0.033	0.074**	-0.035	0.071**	-0.028	0.064*
	[0.049]	[0.035]	[0.052]	[0.035]	[0.052]	[0.035]	[0.050]	[0.034]
Christian	0.220***	-0.171**	-0.005	-0.055	0.002	-0.055	0.002	-0.064
	[0.037]	[0.068]	[0.046]	[0.060]	[0.043]	[0.060]	[0.042]	[0.059]
R^2	0.012	0.023	0.161	0.193	0.162	0.194	0.165	0.202
(ii) IV								
OBC			-0.012	-0.012	-0.006	-0.009	0.002	-0.019+
			[0.019]	[0.015]	[0.019]	[0.014]	[0.018]	[0.014]
Dalit			-0.045*	-0.008	-0.032+	-0.004	-0.02	-0.015
			[0.025]	[0.019]	[0.023]	[0.019]	[0.022]	[0.019]
Adivasi			-0.096***	0.005	-0.076***	0.009	-0.061**	-0.008
			[0.030]	[0.037]	[0.027]	[0.037]	[0.027]	[0.039]
Muslim			-0.037	-0.027	-0.029	-0.022	-0.021	-0.022
			[0.029]	[0.025]	[0.028]	[0.024]	[0.028]	[0.023]
Sikh			-0.035	0.076**	-0.036	0.073**	-0.022	0.053++
			[0.052]	[0.036]	[0.052]	[0.036]	[0.049]	[0.033]
Christian			-0.016	-0.05	-0.005	-0.051	-0.002	-0.077+
			[0.050]	[0.060]	[0.045]	[0.060]	[0.043]	[0.060]
R^2			0.15	0.19	0.154	0.192	0.158	0.148
N	19,844	11,693	19,844	11,693	19,844	11,693	19,844	11,693

Notes: Robust Huber-White (Huber, 1967; White, 1980) standard errors, adjusted for within-cluster correlation/clustering (Wooldridge, 2010), in brackets under parameter estimates. ***: statistically significant at 1 percent; **: statistically significant at 5 percent; *: statistically significant at 10 percent; ++: statistically significant at 15 percent +: statistically significant at 20 percent. All specifications include state fixed effects. Additional controls include age cohort dummies, and dummies for educational attainment, information exposure, and access to health facilities in the community (rural areas, only).
Source: 2004/05 India Human Development Survey (IHDS).

[31] Additional controls are included as specified in the body paper and in the notes under tables; the full set of results is available upon request.

Table C2. OLS/LPM and 2SLS/IV Health Social Network Coefficients With Three Alternative Specifications of Health Social Network Formation: Main Estimation Samples (With All Explanatory Variables Included)

	(1) Access		*(2) Quality*		*(3) Homophily*	
	Rural	*Urban*	*Rural*	*Urban*	*Rural*	*Urban*
(i) OLS	0.065***	0.101***	0.053***	0.091***	0.063***	0.066***
	[0.015]	[0.014]	[0.017]	[0.015]	[0.017]	[0.015]
(ii) IV	0.129***	0.333***	0.128**	0.337***	0.086	0.315***
	[0.046]	[0.063]	[0.054]	[0.072]	[0.126]	[0.120]
(iii) Matching						
ATE	0.041***	0.081***	0.035***	0.076***	0.088***	0.077***
	[0.010]	[0.011]	[0.013]	[0.012]	[0.015]	[0.014]
ATT	0.014+	0.071***	0.015+	0.059***	0.046***	0.056***
	[0.010]	[0.013]	[0.011]	[0.013]	[0.017]	[0.015]
N	19,844	11,693	19,844	11,693	19,844	11,693

Notes: Robust Huber-White (Huber, 1967; White, 1980) standard errors, adjusted for within-cluster correlation/clustering (Wooldridge, 2010), in brackets under parameter estimates. ***: statistically significant at 1 percent; **: statistically significant at 5 percent; *: statistically significant at 10 percent; ++: statistically significant at 15 percent; +: statistically significant at 20 percent. All specifications include state fixed effects. Additional controls include dummies for caste and religion, age cohort dummies, and dummies for educational attainment, information exposure, log per capita household expenditures and dummies for access to health facilities in the community (rural areas, only).
Source: 2004/05 India Human Development Survey (IHDS).

Table C3. OLS/LPM and 2SLS/IV Health Social Network Coefficients With Three Alternative Specifications of Health Social Network Formation: Stratified Estimation Samples (Across Caste and Religion—With All Explanatory Variables Included)

	Rural							*Urban*						
	High caste	*Dalit*	*Adivasi*	*OBC*	*Muslim*	*Sikh/Jain*	*Christian*	*High caste*	*Dalit*	*Adivasi*	*OBC*	*Muslim*	*Sikh/Jain*	*Christian*
(i) OLS														
Access	0.107***	0.014	0.087**	0.072***	0.059++	-0.028	0.058+	0.103***	0.042++	0.124*	0.105***	0.105***	0.128++	0.152**
	[0.025]	[0.028]	[0.039]	[0.023]	[0.040]	[0.072]	[0.043]	[0.021]	[0.029]	[0.067]	[0.021]	[0.031]	[0.083]	[0.064]
Quality	0.090***	0.046+	0.069++	0.042++	0.048	-0.021	0.036	0.085***	0.034	0.115*	0.119***	0.068*	0.088	0.150++
	[0.028]	[0.033]	[0.043]	[0.026]	[0.045]	[0.070]	[0.045]	[0.023]	[0.033]	[0.068]	[0.022]	[0.035]	[0.081]	[0.097]
Homophily	0.071**	0.003	0.153**	0.083***	0.049	0.079	0.045	0.063***	0.084**	0.130++	0.054**	0.068*	0.173**	0.045
	[0.032]	[0.041]	[0.060]	[0.028]	[0.045]	[0.110]	[0.038]	[0.023]	[0.036]	[0.088]	[0.026]	[0.038]	[0.082]	[0.064]
(ii) IV														
Access	0.247***	0.064	0.092	0.110*	0.144	-0.303	-1.181	0.401***	0.284*	0.633***	0.219***	0.463***	2.411	0.448
	[0.063]	[0.079]	[0.140]	[0.064]	[0.145]	[0.252]	[2.096]	[0.107]	[0.153]	[0.162]	[0.078]	[0.131]	[18.574]	[0.617]
Quality	0.252***	0.082	0.163	0.076	0.078	-0.320+	0.319	0.394***	0.351*	0.666***	0.211**	0.434***	6.521	1.605
	[0.071]	[0.088]	[0.133]	[0.076]	[0.174]	[0.232]	[0.512]	[0.121]	[0.199]	[0.225]	[0.098]	[0.137]	[139.836]	[2.127]
Homophily	0.008	-0.038	0.714**	-0.021	0.395	0.320+	-4.196	0.179	0.527++	0.773***	0.302*	0.276	0.681	-26.127
	[0.155]	[0.365]	[0.354]	[0.176]	[0.339]	[0.237]	[21.378]	[0.195]	[0.321]	[0.218]	[0.164]	[0.281]	[1.109]	[397.804]
(iii) Matching														
Access (ATE)	0.050***	-0.020	0.085*	0.043***	0.077**	-0.080	0.133**	0.089***	0.016	0.098++	0.093***	0.100***	0.045	0.167***
	[0.017]	[0.023]	[0.052]	[0.015]	[0.038]	[0.066]	[0.060]	[0.017]	[0.030]	[0.063]	[0.021]	[0.025]	[0.074]	[0.060]
Quality (ATE)	0.032*	0.010	0.043	0.021+	0.063++	-0.038	0.167***	0.086***	0.048++	0.070	0.081***	0.055*	0.076	0.215*
	[0.017]	[0.029]	[0.047]	[0.016]	[0.041]	[0.068]	[0.030]	[0.019]	[0.031]	[0.058]	[0.024]	[0.029]	[0.076]	[0.110]
Homophily (ATE)	0.036+	0.010	0.099++	0.129***	0.114**	0.000	0.120**	0.058***	0.118***	0.036	0.100***	0.064++	0.145*	0.145**
	[0.026]	[0.039]	[0.067]	[0.027]	[0.049]	[0.104]	[0.060]	[0.022]	[0.046]	[0.075]	[0.024]	[0.044]	[0.077]	[0.067]
Access (ATT)	0.050**	-0.017	0.016	0.009	0.027	-0.064	0.113*	0.061***	0.054++	0.164**	0.093***	0.109***	0.054	0.157**
	[0.021]	[0.023]	[0.041]	[0.017]	[0.036]	[0.073]	[0.064]	[0.020]	[0.034]	[0.066]	[0.025]	[0.035]	[0.089]	[0.074]
Quality (ATT)	0.035*	0.022	-0.034	0.012	0.044	-0.040	0.200***	0.062***	0.055++	0.111+	0.072***	0.069*	0.087	0.236**
	[0.019]	[0.024]	[0.044]	[0.018]	[0.039]	[0.073]	[0.067]	[0.021]	[0.037]	[0.080]	[0.027]	[0.038]	[0.095]	[0.098]
Homophily (ATT)	0.078***	-0.027	0.083++	0.044++	0.073++	0.046	0.074	0.053**	0.101**	-0.041	0.027	0.077*	0.132	0.088
	[0.025]	[0.040]	[0.053]	[0.027]	[0.044]	[0.095]	[0.067]	[0.024]	[0.050]	[0.096]	[0.030]	[0.042]	[0.105]	[0.096]
N	3,747	4,491	2,056	7,121	1,877	310	242	3,566	1,915	367	3,604	1,785	228	228

Notes: Each coefficient in the table is the network access coefficient from an individual regression, as per the caste/religion and geographical (rural/urban) designation given in the row and columns, respectively. Robust Huber-White (Huber, 1967; White, 1980) standard errors, adjusted for within-cluster correlation/clustering (Wooldridge, 2010), in brackets under parameter estimates. ***: statistically significant at 1 percent; **: statistically significant at 5 percent; *: statistically significant at 10 percent. All specifications include state fixed effects. Additional controls include age cohort dummies, and dummies for educational attainment, information exposure, log per capita household expenditures and dummies for access to health facilities in the community (rural areas, only).
Source: 2004/05 India Human Development Survey (IHDS).

Table C4. Overall Health Knowledge Gap Decompositions With Three Alternative Specifications of Health Social Network Formation

		Rural						*Urban*					
		HC vs Dalit	*HC vs Adivasi*	*HC vs OBC*	*HC vs Muslim*	*HC vs Sikh/Jain*	*HC vs Christian*	*HC vs Dalit*	*HC vs Adivasi*	*HC vs OBC*	*HC vs Muslim*	*HC vs Sikh/Jain*	*HC vs Christian*
(1) Access													
(i) OLS	Explained	0.077***	0.013	0.063***	0.021	-0.023	-0.244***	0.112***	0.048**	0.127***	0.071***	-0.019+	0.072***
		[0.014]	[0.022]	[0.013]	[0.018]	[0.053]	[0.032]	[0.013]	[0.019]	[0.015]	[0.013]	[0.014]	[0.024]
	Unexplained	0.056***	0.111***	0.031*	0.072***	0.037	0.024	0.028++	-0.005	0.042**	0.042**	-0.069*	0.099**
		[0.020]	[0.026]	[0.018]	[0.025]	[0.062]	[0.038]	[0.019]	[0.031]	[0.018]	[0.018]	[0.035]	[0.039]
(ii) IV	Explained	0.085***	0.017	0.069***	0.022	-0.016	-0.257***	0.121***	0.042*	0.136***	0.096***	-0.035*	0.081***
		[0.017]	[0.025]	[0.015]	[0.020]	[0.056]	[0.034]	[0.017]	[0.022]	[0.017]	[0.019]	[0.020]	[0.031]
	Unexplained	0.048**	0.107***	0.026+	0.070***	0.03	0.037	0.019	0	0.033*	0.016	-0.053	0.090*
		[0.021]	[0.029]	[0.019]	[0.026]	[0.075]	[0.119]	[0.021]	[0.037]	[0.019]	[0.023]	[0.329]	[0.055]
(2) Quality													
(i) OLS	Explained	0.076***	0.011	0.062***	0.021	-0.025	-0.228***	0.112***	0.050**	0.127***	0.068***	-0.020++	0.078***
		[0.014]	[0.022]	[0.013]	[0.018]	[0.052]	[0.032]	[0.013]	[0.020]	[0.015]	[0.013]	[0.014]	[0.024]
	Unexplained	0.057***	0.113***	0.032*	0.072***	0.04	0.008	0.028++	-0.007	0.042**	0.045**	-0.068*	0.093**
		[0.020]	[0.026]	[0.018]	[0.025]	[0.062]	[0.038]	[0.019]	[0.032]	[0.018]	[0.018]	[0.035]	[0.039]
(ii) IV	Explained	0.085***	0.014	0.067***	0.024	-0.02	-0.217***	0.121***	0.045**	0.138***	0.084***	-0.045**	0.107***
		[0.017]	[0.025]	[0.015]	[0.020]	[0.056]	[0.033]	[0.017]	[0.022]	[0.017]	[0.018]	[0.020]	[0.032]
	Unexplained	0.048**	0.110***	0.027+	0.069***	0.034	-0.002	0.019	-0.003	0.030++	0.028	-0.043	0.064
		[0.022]	[0.029]	[0.019]	[0.026]	[0.075]	[0.052]	[0.021]	[0.037]	[0.020]	[0.022]	[0.960]	[0.099]
(3) Homophily													
(i) OLS	Explained	0.073***	0.01	0.062***	0.017	-0.031	-0.242***	0.114***	0.051***	0.127***	0.067***	-0.014	0.071***
		[0.014]	[0.022]	[0.013]	[0.018]	[0.052]	[0.032]	[0.013]	[0.020]	[0.015]	[0.013]	[0.014]	[0.025]
	Unexplained	0.060***	0.114***	0.032*	0.075***	0.046	0.022	0.026+	-0.009	0.041**	0.045**	-0.074**	0.099**
		[0.020]	[0.026]	[0.018]	[0.025]	[0.061]	[0.038]	[0.019]	[0.032]	[0.018]	[0.018]	[0.035]	[0.040]
(ii) IV	Explained	0.071***	0.01	0.059***	0.019	-0.029	-0.234***	0.124***	0.053***	0.134***	0.077***	-0.014	0.076***
		[0.017]	[0.025]	[0.015]	[0.019]	[0.051]	[0.034]	[0.016]	[0.020]	[0.017]	[0.016]	[0.015]	[0.027]
	Unexplained	0.062***	0.114***	0.035*	0.074***	0.043	0.015	0.016	-0.011	0.034*	0.035*	-0.073*	0.095
		[0.022]	[0.029]	[0.020]	[0.027]	[0.066]	[0.287]	[0.021]	[0.037]	[0.020]	[0.021]	[0.044]	[1.238]
	N	8,238	5,803	10,868	5,624	4,057	3,989	5,481	3,933	7,170	5,351	3,794	3,794

Notes: Decompositions are from low caste and religion females' viewpoint—i.e., using high caste (HC) endowments and returns. Values in brackets are robust Huber-White (Huber, 1967; White, 1980) standard errors. ***: statistically significant at 1 percent; **: statistically significant at 5 percent; *: statistically significant at 10 percent; ++: statistically significant at 15 percent +: statistically significant at 20 percent.
Source: 2004/05 India Human Development Survey (IHDS).

Table C5. How Woman Was Treated Last Time She Consulted a Health Worker for Diarrhea Related Illness: Stratified by Caste and Three Alternative Types of Health Social Network Formation

	HC	*Dalit*	*Adivasi*	*OBC*	*Muslim*	*Sikh/Jain*	*Christian*
(1) Access							
No:							
Nicely	75.7	71.2	53.8	68.3	59.1	79.7	70.6
Somewhat nicely	23.1	26.2	45.7	30.6	40.5	20.3	29.4
Not nicely	1.2	2.6	0.6	1.1	0.4	0.0	0.0
N	176	256	128	331	184	7	5
Yes:							
Nicely	81.6	73.0	91.4	79.9	67.9	65.2	100.0
Somewhat nicely	17.6	27.0	8.6	19.0	28.1	34.8	0.0
Not nicely	0.7	0.0	0.0	1.1	4.0	0.0	0.0
N	109	72	20	167	46	4	3
(2) Quality							
No:							
Nicely	75.6	71.7	55.5	67.3	59.3	79.7	84.2
Somewhat nicely	23.3	25.8	44.0	31.7	40.3	20.3	15.8
Not nicely	1.0	2.6	0.5	1.0	0.4	0.0	0.0
N	201	266	134	356	190	7	8
Yes:							
Nicely	83.3	71.0	89.6	83.7	67.3	65.2	0.0
Somewhat nicely	15.7	29.0	10.4	15.0	28.4	34.8	0.0
Not nicely	0.9	0.0	0.0	1.3	4.3	0.0	0.0
N	84	62	14	142	40	4	0
(3) Homophily							
No:							
Nicely	78.3	71.5	57.7	71.0	61.0	80.6	78.0
Somewhat nicely	20.8	26.3	41.8	28.1	38.3	19.4	22.1
Not nicely	0.9	2.2	0.5	1.0	0.7	0.0	0.0
N	236	311	142	454	210	10	6
Yes:							
Nicely	76.9	72.6	83.8	85.1	52.8	0.0	100.0
Somewhat nicely	21.5	27.4	16.2	12.9	42.0	100.0	0.0
Not nicely	1.6	0.0	0.0	2.0	5.2	0.0	0.0
N	49	17	6	44	20	1	2

Notes: Calculations incorporate sampling weights.
Source: 2004/05 India Human Development Survey (IHDS).

Table C6. Confidence in Hospitals and Doctors to Provide Good Treatment for Households Where Woman Consulted a Health Worker for Diarrhea Related Illness: Stratified by Caste and Three Alternative Types of Health Social Network Formation

	HC	*Dalit*	*Adivasi*	*OBC*	*Muslim*	*Sikh/Jain*	*Christian*
(1) Access:							
No:							
A great deal	64.5	61.6	74.9	63.6	63.3	84.4	21.9
Only some	26.5	27.2	18.4	29.6	29.8	6.1	78.1
Hardly any	9.1	11.2	6.7	6.8	7.0	9.5	0.0
N	177	256	128	330	181	7	4
Yes:							
A great deal	70.3	67.7	76.2	80.3	76.8	79.0	100.0
Only some	23.3	25.8	21.2	15.1	19.6	21.0	0.0
Hardly any	6.4	6.5	2.6	4.5	3.6	0.0	0.0
N	109	71	20	170	47	4	3
(2) Quality:							
No:							
A great deal	64.8	62.5	75.4	62.6	63.4	84.4	62.2
Only some	27.2	26.6	17.9	30.6	29.7	6.1	37.8
Hardly any	8.0	11.0	6.7	6.8	6.9	9.5	0.0
N	202	266	134	355	187	7	7
Yes:							
A great deal	70.9	64.1	71.2	84.9	76.9	79.0	0.0
Only some	21.0	28.7	28.8	10.8	19.2	21.0	0.0
Hardly any	8.1	7.2	0.0	4.3	3.9	0.0	0.0
N	84	61	14	145	41	4	0.0
(3) Homophily:							
No:							
A great deal	66.3	61.9	75.5	67.2	65.1	89.4	44.8
Only some	26.4	28.2	18.5	26.8	28.4	4.1	55.2
Hardly any	7.3	9.9	6.0	6.1	6.5	6.4	0.0
N	237	310	142	455	207	10	5
Yes:							
A great deal	68.8	80.4	49.8	90.4	70.5	0.0	100.0
Only some	19.8	0.0	33.7	4.7	24.1	100.0	0.0
Hardly any	11.4	19.6	16.5	4.9	5.4	0.0	0.0
N	49	17	6	45	21	1	2

Notes: Calculations incorporate sampling weights.
Source: 2004/05 India Human Development Survey (IHDS).

APPENDIX D: Detailed Decomposition Results Using Three Alternative Specifications for Network Formation (OLS/LPM and 2SLS/IV)

Table D1. Detailed Decomposition Results: Network Access (OLS/LPM)

	(1) HC vs Dalit		*(2) HC vs Adivasi*		*(3) HC vs OBC*		*(4) HC vs Muslim*		*(5) HC vs Sikh/Jain*		*(6) HC vs Christian*	
	Explained	*Unexpl.*	*Explained*	*Unexpl.*	*Explained*	*Unexpl.*	*Explained*	*Unexpl.*	*Explained*	*Unexpl.*	*Explained*	*Unexpl.*
(i) Rural												
Network access	0.016***	0.024***	0.025***	0.003	0.007***	0.012	0.017***	0.012	-0.008*	0.065*	0.010**	0.015
	[0.004]	[0.009]	[0.006]	[0.007]	[0.002]	[0.010]	[0.004]	[0.011]	[0.005]	[0.037]	[0.005]	[0.016]
Watches TV regularly	-0.001	-0.014++	-0.002	-0.005	-0.001	-0.003	-0.001	0.003	0.001	0.032	0.001	-0.024
	[0.005]	[0.009]	[0.008]	[0.007]	[0.004]	[0.009]	[0.006]	[0.010]	[0.005]	[0.062]	[0.003]	[0.029]
Ln (per cap household expenditures)	0.021***	0.226+	0.038***	0.532**	0.014***	0.214+	0.015***	0.051	-0.020***	0.017	-0.002	0.555**
	[0.007]	[0.175]	[0.013]	[0.217]	[0.005]	[0.156]	[0.005]	[0.220]	[0.007]	[0.499]	[0.004]	[0.258]
Age	-0.001	-0.009	0	-0.011	0	-0.009	0	-0.01	0.002	-0.054*	0.002	0.021
	[0.002]	[0.009]	[0.002]	[0.010]	[0.001]	[0.009]	[0.002]	[0.010]	[0.003]	[0.029]	[0.005]	[0.017]
Education	0.017**	0.019	0.018**	0.076**	0.011**	0.040**	0.015**	0.026	0	0.056*	-0.029**	-0.066**
	[0.008]	[0.033]	[0.009]	[0.038]	[0.005]	[0.020]	[0.008]	[0.037]	[0.002]	[0.033]	[0.013]	[0.029]
Health facilities access	0.003	0.015	0	-0.007	0.003	0.020+	0.001	0.038**	-0.002	0.002	-0.021	0.070+
	[0.003]	[0.016]	[0.002]	[0.014]	[0.003]	[0.014]	[0.002]	[0.019]	[0.003]	[0.046]	[0.022]	[0.054]
State FEs	0.022**	-0.031*	-0.066***	-0.028	0.029**	-0.038**	-0.025*	0.044*	0.004	-0.157	-0.204***	0.643***
	[0.010]	[0.017]	[0.019]	[0.027]	[0.012]	[0.019]	[0.015]	[0.024]	[0.052]	[0.236]	[0.037]	[0.156]
(ii) Urban												
Knows any health person	0.011***	0.019*	0.004+	-0.008	0.006***	-0.001	0.014***	0	-0.016***	-0.014	-0.003	-0.022
	[0.002]	[0.010]	[0.003]	[0.025]	[0.002]	[0.009]	[0.003]	[0.009]	[0.005]	[0.042]	[0.004]	[0.032]
Watches TV regularly	-0.008**	-0.049**	-0.008*	0.017	-0.005**	-0.032++	-0.010**	-0.044**	0.001	0.014	-0.003+	0.038
	[0.004]	[0.022]	[0.004]	[0.043]	[0.003]	[0.020]	[0.005]	[0.021]	[0.001]	[0.078]	[0.002]	[0.073]
Ln (per cap household expenditures)	0.012*	-0.098	0.010*	0.488+	0.007*	0.074	0.014*	0.097	-0.004++	0.801*	0.001	0.029
	[0.007]	[0.191]	[0.006]	[0.341]	[0.004]	[0.146]	[0.008]	[0.191]	[0.003]	[0.435]	[0.001]	[0.462]
Age	0.003+	0	0.003	0.005	0.001	0.013	0.004++	0.007	-0.003+	0.023	0	-0.036+
	[0.002]	[0.014]	[0.003]	[0.022]	[0.002]	[0.012]	[0.002]	[0.013]	[0.003]	[0.055]	[0.003]	[0.025]
Education	0.045***	0.007	0.040***	0.055**	0.024***	-0.007	0.045***	-0.01	-0.010***	-0.004	-0.006++	0.038
	[0.009]	[0.015]	[0.009]	[0.023]	[0.005]	[0.008]	[0.009]	[0.015]	[0.004]	[0.039]	[0.004]	[0.031]
State FEs	0.049***	0.027*	-0.003	-0.012	0.094***	0.01	0.004	0.008	0.013	0.012	0.082***	0.015
	[0.010]	[0.015]	[0.018]	[0.047]	[0.014]	[0.017]	[0.008]	[0.020]	[0.013]	[0.031]	[0.023]	[0.049]
N	5,481	5,481	3,933	3,933	7,170	7,170	5,351	5,351	3,794	3,794	3,794	3,794

Notes: Decompositions are from low caste and religion females' viewpoint—i.e., using high caste endowments and returns. Values in brackets are robust Huber-White (Huber, 1967; White, 1980) standard errors. ***: statistically significant at 1 percent; **: statistically significant at 5 percent; *: statistically significant at 10 percent; ++: statistically significant at 15 percent +: statistically significant at 20 percent.
Source: 2004/05 India Human Development Survey (IHDS).

Table D2. Detailed Decomposition Results: Network Quality (OLS/LPM)

	(1) HC vs Dalit		*(2) HC vs Adivasi*		*(3) HC vs OBC*		*(4) HC vs Muslim*		*(5) HC vs Sikh/Jain*		*(6) HC vs Christian*	
	Explained	*Unexpl.*	*Explained*	*Unexpl.*	*Explained*	*Unexpl.*	*Explained*	*Unexpl.*	*Explained*	*Unexpl.*	*Explained*	*Unexpl.*
(i) Rural												
Network quality	0.012***	0.009	0.019***	0.003	0.005***	0.013++	0.012***	0.009	-0.011**	0.051+	0.020***	0.006
	[0.004]	[0.007]	[0.005]	[0.006]	[0.002]	[0.009]	[0.004]	[0.010]	[0.005]	[0.035]	[0.006]	[0.006]
Watches TV regularly	-0.001	-0.013++	-0.001	-0.004	-0.001	-0.003	-0.001	0.004	0.001	0.033	0	-0.028
	[0.005]	[0.009]	[0.008]	[0.007]	[0.004]	[0.009]	[0.006]	[0.010]	[0.005]	[0.062]	[0.003]	[0.029]
Ln (per cap household expenditures)	0.022***	0.245+	0.040***	0.527**	0.014***	0.221+	0.015***	0.06	-0.020***	0.035	-0.002	0.537**
	[0.007]	[0.175]	[0.014]	[0.217]	[0.005]	[0.157]	[0.005]	[0.219]	[0.007]	[0.497]	[0.005]	[0.264]
Age	-0.001	-0.008	0	-0.011	0	-0.01	0	-0.01	0.002	-0.054*	0.002	0.02
	[0.002]	[0.009]	[0.002]	[0.010]	[0.001]	[0.009]	[0.002]	[0.010]	[0.003]	[0.029]	[0.005]	[0.018]
Education	0.018**	0.015	0.018**	0.080**	0.011**	0.043**	0.016**	0.027	0	0.056*	-0.030**	-0.065**
	[0.008]	[0.033]	[0.009]	[0.037]	[0.005]	[0.020]	[0.008]	[0.037]	[0.002]	[0.033]	[0.013]	[0.029]
Health facilities access	0.002	0.015	0	-0.008	0.003	0.020+	0.001	0.038**	-0.002	0.002	-0.02	0.067
	[0.003]	[0.016]	[0.002]	[0.014]	[0.003]	[0.014]	[0.002]	[0.019]	[0.003]	[0.046]	[0.022]	[0.054]
State FEs	0.024**	-0.029*	-0.064***	-0.027	0.029**	-0.038**	-0.021++	0.043*	0.006	-0.154	-0.198***	0.643***
	[0.010]	[0.017]	[0.019]	[0.026]	[0.012]	[0.019]	[0.015]	[0.024]	[0.052]	[0.237]	[0.037]	[0.156]
(ii) Urban												
Knows any doctor	0.008***	0.012++	0.007**	-0.007	0.004***	-0.009	0.009***	0.004	-0.015***	-0.001	0.004+	-0.018
	[0.002]	[0.008]	[0.003]	[0.016]	[0.001]	[0.007]	[0.002]	[0.008]	[0.005]	[0.037]	[0.003]	[0.027]
Watches TV regularly	-0.009**	-0.050**	-0.008*	0.022	-0.005**	-0.034*	-0.011**	-0.046**	0.001	0.018	-0.003+	0.036
	[0.004]	[0.022]	[0.004]	[0.044]	[0.003]	[0.020]	[0.005]	[0.021]	[0.001]	[0.077]	[0.002]	[0.072]
Ln (per cap household expenditures)	0.013*	-0.09	0.011*	0.429	0.008*	0.101	0.015*	0.067	-0.005*	0.819*	0.002	0.073
	[0.007]	[0.192]	[0.006]	[0.347]	[0.004]	[0.146]	[0.008]	[0.191]	[0.003]	[0.430]	[0.002]	[0.498]
Age	0.003+	0.001	0.004	0.008	0.001	0.013	0.004++	0.007	-0.004+	0.022	0	-0.036+
	[0.002]	[0.014]	[0.003]	[0.021]	[0.002]	[0.013]	[0.002]	[0.013]	[0.003]	[0.057]	[0.003]	[0.025]
Education	0.046***	0.007	0.042***	0.055**	0.024***	-0.007	0.046***	-0.008	-0.010***	0	-0.006*	0.031
	[0.009]	[0.015]	[0.010]	[0.023]	[0.005]	[0.008]	[0.009]	[0.015]	[0.004]	[0.040]	[0.004]	[0.031]
State FEs	0.050***	0.024++	-0.006	-0.01	0.095***	0.006	0.004	0.008	0.012	0.013	0.082***	0.029
	[0.010]	[0.015]	[0.018]	[0.050]	[0.014]	[0.017]	[0.008]	[0.021]	[0.013]	[0.032]	[0.024]	[0.050]
N	5,481	5,481	3,933	3,933	7,170	7,170	5,351	5,351	3,794	3,794	3,794	3,794

Notes: Decompositions are from low caste and religion females' viewpoint—i.e., using high caste endowments and returns. Values in brackets are robust Huber-White (Huber, 1967; White, 1980) standard errors. ***: statistically significant at 1 percent; **: statistically significant at 5 percent; *: statistically significant at 10 percent; ++: statistically significant at 15 percent +: statistically significant at 20 percent.
Source: 2004/05 India Human Development Survey (IHDS).

Table D3. Detailed Decomposition Results: Network Homophily (OLS/LPM)

	(1) HC vs Dalit		*(2) HC vs Adivasi*		*(3) HC vs OBC*		*(4) HC vs Muslim*		*(5) HC vs Sikh/Jain*		*(6) HC vs Christian*	
	Explained	*Unexpl.*	*Explained*	*Unexpl.*	*Explained*	*Unexpl.*	*Explained*	*Unexpl.*	*Explained*	*Unexpl.*	*Explained*	*Unexpl.*
(ii) Rural												
Network homophily	0.008**	0.004+	0.008**	-0.005++	0.005**	-0.001	0.005**	0.002	0.001	-0.001	-0.007*	0.007
	[0.003]	[0.003]	[0.003]	[0.003]	[0.002]	[0.004]	[0.002]	[0.005]	[0.002]	[0.018]	[0.004]	[0.015]
Watches TV regularly	0	-0.013++	0	-0.004	0	-0.002	0	0.004	0	0.036	0	-0.022
	[0.005]	[0.009]	[0.008]	[0.007]	[0.004]	[0.009]	[0.006]	[0.010]	[0.005]	[0.061]	[0.003]	[0.029]
Ln (per cap household expenditures)	0.023***	0.261++	0.042***	0.588***	0.015***	0.256++	0.016***	0.071	-0.022***	0.07	-0.002	0.585**
	[0.008]	[0.176]	[0.014]	[0.219]	[0.005]	[0.158]	[0.005]	[0.221]	[0.008]	[0.484]	[0.005]	[0.263]
Age	-0.001	-0.009	-0.001	-0.012	0	-0.01	0	-0.01	0.002	-0.052*	0.002	0.019
	[0.002]	[0.009]	[0.002]	[0.010]	[0.001]	[0.009]	[0.002]	[0.010]	[0.003]	[0.029]	[0.005]	[0.018]
Education	0.018**	0.019	0.018**	0.077**	0.011**	0.041**	0.016**	0.026	0	0.052++	-0.030**	-0.065**
	[0.008]	[0.033]	[0.009]	[0.038]	[0.005]	[0.020]	[0.008]	[0.037]	[0.002]	[0.033]	[0.013]	[0.029]
Health facilities access	0.002	0.017	0.001	-0.005	0.003	0.022++	0.001	0.040**	-0.002	0.003	-0.022	0.071+
	[0.003]	[0.016]	[0.002]	[0.014]	[0.003]	[0.014]	[0.002]	[0.019]	[0.003]	[0.045]	[0.022]	[0.055]
State FEs	0.023**	-0.027++	-0.059***	-0.022	0.028**	-0.035*	-0.020+	0.043*	-0.009	-0.332***	-0.182***	0.630***
	[0.010]	[0.017]	[0.019]	[0.025]	[0.012]	[0.019]	[0.015]	[0.024]	[0.052]	[0.088]	[0.036]	[0.156]
(ii) Urban												
Health person is of same jati	0.006***	-0.002	0.001	-0.012	0.003**	0.001	0.003**	-0.001	-0.001	-0.024+	-0.005*	0.005
	[0.002]	[0.004]	[0.002]	[0.014]	[0.001]	[0.005]	[0.001]	[0.006]	[0.002]	[0.017]	[0.003]	[0.021]
Watches TV regularly	-0.009**	-0.052**	-0.008*	0.01	-0.006**	-0.034*	-0.011**	-0.048**	0.001	0.006	-0.003+	0.031
	[0.004]	[0.022]	[0.004]	[0.043]	[0.003]	[0.021]	[0.005]	[0.021]	[0.001]	[0.080]	[0.002]	[0.072]
Ln (per cap household expenditures)	0.016**	-0.049	0.014**	0.565*	0.009**	0.085	0.018**	0.094	-0.006**	0.874**	0.002	-0.124
	[0.007]	[0.189]	[0.006]	[0.339]	[0.004]	[0.146]	[0.008]	[0.190]	[0.003]	[0.432]	[0.002]	[0.476]
Age	0.003+	0.002	0.004	0.008	0.002	0.012	0.004*	0.007	-0.004++	0.025	0	-0.039++
	[0.002]	[0.014]	[0.003]	[0.024]	[0.002]	[0.013]	[0.002]	[0.014]	[0.003]	[0.060]	[0.003]	[0.025]
Education	0.048***	0.003	0.043***	0.055**	0.026***	-0.007	0.048***	-0.01	-0.010***	-0.003	-0.006*	0.035
	[0.009]	[0.015]	[0.010]	[0.023]	[0.005]	[0.008]	[0.009]	[0.015]	[0.004]	[0.041]	[0.004]	[0.031]
State FEs	0.048***	0.026*	-0.001	-0.003	0.094***	0.011	0.004	0.01	0.006	0.005	0.084***	0.024
	[0.010]	[0.015]	[0.017]	[0.049]	[0.014]	[0.017]	[0.008]	[0.021]	[0.012]	[0.030]	[0.024]	[0.048]
N	5,481	5,481	3,933	3,933	7,170	7,170	5,351	5,351	3,794	3,794	3,794	3,794

Notes: Decompositions are from low caste and religion females' viewpoint—i.e., using high caste endowments and returns. Values in brackets are robust Huber-White (Huber, 1967; White, 1980) standard errors. ***: statistically significant at 1 percent; **: statistically significant at 5 percent; *: statistically significant at 10 percent; ++: statistically significant at 15 percent +: statistically significant at 20 percent.
Source: 2004/05 India Human Development Survey (IHDS).

Table D4. Detailed Decomposition Results: Network Access (IV)

	(1) HC vs Dalit		*(2) HC vs Adivasi*		*(3) HC vs OBC*		*(4) HC vs Muslim*		*(5) HC vs Sikh/Jain*		*(6) HC vs Christian*	
	Explained	*Unexpl.*	*Explained*	*Unexpl.*	*Explained*	*Unexpl.*	*Explained*	*Unexpl.*	*Explained*	*Unexpl.*	*Explained*	*Unexpl.*
(i) Rural												
Network access	0.037***	0.047**	0.057***	0.028	0.016***	0.047*	0.039***	0.026	-0.019*	0.266**	0.023**	0.451
	[0.009]	[0.021]	[0.013]	[0.022]	[0.005]	[0.024]	[0.010]	[0.031]	[0.011]	[0.104]	[0.011]	[0.559]
Watches TV regularly	-0.002	-0.048	-0.003	0.056+	-0.002	0.067++	-0.002	-0.076+	0.002	0.701+	0.001	1.03
	[0.023]	[0.045]	[0.035]	[0.040]	[0.017]	[0.046]	[0.025]	[0.053]	[0.022]	[0.526]	[0.012]	[1.345]
Ln (per cap household expenditures)	0.015+	0.235	0.027+	0.151	0.010+	-0.073	0.010+	0.303	-0.014	-1.22	-0.001	-1.714
	[0.011]	[0.252]	[0.021]	[0.329]	[0.007]	[0.237]	[0.008]	[0.351]	[0.011]	[1.235]	[0.003]	[2.959]
Age	-0.001	-0.007	0	-0.012	0	-0.013+	0	-0.006	0.002	-0.045++	0.002	-0.12
	[0.002]	[0.009]	[0.002]	[0.010]	[0.001]	[0.010]	[0.002]	[0.011]	[0.003]	[0.029]	[0.005]	[0.181]
Education	0.014	0.003	0.014	0.117**	0.009	0.069***	0.012	-0.019	0	0.047	-0.023++	-0.178
	[0.012]	[0.041]	[0.012]	[0.053]	[0.007]	[0.024]	[0.011]	[0.054]	[0.002]	[0.054]	[0.015]	[0.155]
Health facilities access	0.003	0.015	0	-0.009	0.004++	0.016	0	0.043**	-0.002	-0.044	-0.021	0.588
	[0.003]	[0.017]	[0.002]	[0.015]	[0.003]	[0.014]	[0.002]	[0.020]	[0.003]	[0.058]	[0.022]	[0.713]
State FEs	0.019*	-0.043**	-0.077***	-0.035	0.032**	-0.039*	-0.036**	0.056*	0.014	-0.757++	-0.237***	0.817
	[0.010]	[0.018]	[0.020]	[0.028]	[0.013]	[0.021]	[0.016]	[0.032]	[0.055]	[0.497]	[0.040]	[0.832]
(ii) Urban												
Knows any health person	0.042***	0.036	0.017	-0.085++	0.023***	0.065*	0.056***	-0.017	-0.061***	-1.131	-0.012	-0.021
	[0.011]	[0.041]	[0.014]	[0.053]	[0.007]	[0.036]	[0.013]	[0.033]	[0.019]	[10.886]	[0.015]	[0.292]
Watches TV regularly	0.037++	0.061	0.035+	0.297++	0.024++	0.133	0.046++	0.03	-0.006	-7.995	0.012	-0.624
	[0.025]	[0.114]	[0.025]	[0.191]	[0.016]	[0.109]	[0.031]	[0.100]	[0.007]	[72.487]	[0.010]	[0.808]
Ln (per cap household expenditures)	-0.021++	-0.21	-0.017++	0.146	-0.012++	-0.308	-0.024++	0.17	0.008++	4.036	-0.003	0
	[0.014]	[0.342]	[0.012]	[0.436]	[0.008]	[0.247]	[0.015]	[0.301]	[0.005]	[33.326]	[0.003]	[1.262]
Age	0	-0.003	0.002	-0.014	0	0.009	0.001	0.009	-0.002	-0.353	0.001	0.007
	[0.003]	[0.014]	[0.003]	[0.024]	[0.002]	[0.013]	[0.003]	[0.013]	[0.003]	[3.246]	[0.003]	[0.066]
Education	0.013	0.014	0.013	0.059*	0.007	-0.002	0.013	-0.019	-0.002	0.204	-0.002	0.054
	[0.016]	[0.019]	[0.014]	[0.032]	[0.008]	[0.009]	[0.016]	[0.019]	[0.004]	[1.993]	[0.003]	[0.058]
State FEs	0.050***	0.031**	-0.007	-0.012	0.096***	0.013	0.005	0.012	0.028*	-0.376	0.083***	-0.128
	[0.010]	[0.016]	[0.020]	[0.052]	[0.014]	[0.017]	[0.009]	[0.020]	[0.014]	[3.362]	[0.024]	[0.163]
N	5,481	5,481	3,933	3,933	7,170	7,170	5,351	5,351	3,794	3,794	3,794	3,794

Notes: Decompositions are from low caste and religion females' viewpoint—i.e., using high caste endowments and returns. Values in brackets are robust Huber-White (Huber, 1967; White, 1980) standard errors. ***: statistically significant at 1 percent; **: statistically significant at 5 percent; *: statistically significant at 10 percent; ++: statistically significant at 15 percent +: statistically significant at 20 percent.
Source: 2004/05 India Human Development Survey (IHDS).

Table D5. Detailed Decomposition Results: Network Quality (IV)

	(1) HC vs Dalit		*(2) HC vs Adivasi*		*(3) HC vs OBC*		*(4) HC vs Muslim*		*(5) HC vs Sikh/Jain*		*(6) HC vs Christian*	
	Explained	*Unexpl.*	*Explained*	*Unexpl.*	*Explained*	*Unexpl.*	*Explained*	*Unexpl.*	*Explained*	*Unexpl.*	*Explained*	*Unexpl.*
(i) Rural												
Network quality	0.033***	0.034*	0.053***	0.011	0.015***	0.048**	0.032***	0.036	-0.032**	0.262***	0.056***	-0.007
	[0.009]	[0.019]	[0.013]	[0.015]	[0.005]	[0.022]	[0.009]	[0.028]	[0.012]	[0.086]	[0.015]	[0.078]
Watches TV regularly	-0.003	-0.052	-0.005	0.049	-0.002	0.069++	-0.003	-0.069+	0.003	0.659	0.002	0.417++
	[0.023]	[0.046]	[0.035]	[0.040]	[0.017]	[0.046]	[0.025]	[0.052]	[0.022]	[0.520]	[0.012]	[0.270]
Ln (per cap household expenditures)	0.015+	0.255	0.028+	0.199	0.010+	-0.079	0.011+	0.24	-0.014+	-1.156	-0.001	0.347
	[0.011]	[0.254]	[0.021]	[0.327]	[0.008]	[0.240]	[0.008]	[0.344]	[0.011]	[1.227]	[0.003]	[0.686]
Age	-0.001	-0.008	-0.001	-0.011	0	-0.014++	0	-0.007	0.002	-0.046++	0.002	-0.013
	[0.002]	[0.009]	[0.002]	[0.010]	[0.001]	[0.010]	[0.002]	[0.011]	[0.003]	[0.029]	[0.005]	[0.036]
Education	0.015	0	0.015	0.112**	0.009	0.075***	0.013	-0.01	0	0.049	-0.024++	-0.100**
	[0.012]	[0.042]	[0.012]	[0.051]	[0.007]	[0.025]	[0.011]	[0.052]	[0.002]	[0.051]	[0.016]	[0.047]
Health facilities access	0.003	0.014	0	-0.01	0.004+	0.015	0	0.042**	-0.001	-0.047	-0.018	0.284*
	[0.003]	[0.017]	[0.002]	[0.015]	[0.003]	[0.014]	[0.002]	[0.020]	[0.003]	[0.058]	[0.022]	[0.163]
State FEs	0.022**	-0.042**	-0.076***	-0.038	0.033***	-0.039*	-0.029*	0.057*	0.023	-0.720++	-0.233***	0.278*
	[0.010]	[0.018]	[0.020]	[0.030]	[0.012]	[0.021]	[0.015]	[0.032]	[0.055]	[0.496]	[0.039]	[0.156]
(ii) Urban												
Knows any doctor	0.036***	0.01	0.032**	-0.065+	0.020***	0.049++	0.043***	-0.008	-0.070***	-3.048	0.018+	-0.332
	[0.010]	[0.037]	[0.013]	[0.047]	[0.007]	[0.030]	[0.011]	[0.029]	[0.022]	[72.373]	[0.014]	[0.616]
Watches TV regularly	0.037++	0.036	0.035+	0.299++	0.024++	0.144+	0.046++	0.012	-0.006	-23.74	0.012	-1.422
	[0.025]	[0.119]	[0.025]	[0.187]	[0.016]	[0.110]	[0.031]	[0.101]	[0.007]	[555.090]	[0.010]	[2.296]
Ln (per cap household expenditures)	-0.022++	-0.078	-0.018++	-0.215	-0.013++	-0.342+	-0.025++	0.122	0.008+	14.046	-0.003	2.319
	[0.014]	[0.384]	[0.012]	[0.457]	[0.008]	[0.259]	[0.016]	[0.313]	[0.006]	[316.788]	[0.003]	[4.578]
Age	0.001	-0.003	0.002	-0.002	0	0.011	0.001	0.013	-0.002	-1.162	0.001	0.084
	[0.003]	[0.014]	[0.003]	[0.015]	[0.002]	[0.013]	[0.003]	[0.014]	[0.003]	[27.245]	[0.003]	[0.194]
Education	0.015	0.015	0.016	0.056*	0.008	-0.001	0.015	-0.015	-0.003	0.858	-0.003	0.034
	[0.016]	[0.019]	[0.014]	[0.032]	[0.008]	[0.009]	[0.016]	[0.020]	[0.004]	[20.131]	[0.003]	[0.080]
State FEs	0.054***	0.019	-0.021	-0.006	0.100***	-0.001	0.004	0.011	0.028*	-0.945	0.081***	-0.214
	[0.010]	[0.016]	[0.020]	[0.066]	[0.014]	[0.017]	[0.009]	[0.021]	[0.015]	[21.928]	[0.025]	[0.368]
N	5,481	5,481	3,933	3,933	7,170	7,170	5,351	5,351	3,794	3,794	3,794	3,794

Notes: Decompositions are from low caste and religion females' viewpoint—i.e., using high caste endowments and returns. Values in brackets are robust Huber-White (Huber, 1967; White, 1980) standard errors. ***: statistically significant at 1 percent; **: statistically significant at 5 percent; *: statistically significant at 10 percent; ++: statistically significant at 15 percent +: statistically significant at 20 percent.
Source: 2004/05 India Human Development Survey (IHDS).

Table D6. Detailed Decomposition Results: Network Homophily (IV)

	(1) HC vs Dalit		*(2) HC vs Adivasi*		*(3) HC vs OBC*		*(4) HC vs Muslim*		*(5) HC vs Sikh/Jain*		*(6) HC vs Christian*	
	Explained	*Unexpl.*	*Explained*	*Unexpl.*	*Explained*	*Unexpl.*	*Explained*	*Unexpl.*	*Explained*	*Unexpl.*	*Explained*	*Unexpl.*
(i) Rural												
Network homophily	0.001	0.003	0.001	-0.042***	0.001	0.003	0.001	-0.043+	0	-0.052	-0.001	1.142
	[0.013]	[0.017]	[0.013]	[0.016]	[0.009]	[0.018]	[0.007]	[0.031]	[0.001]	[0.043]	[0.011]	[5.937]
Watches TV regularly	0	-0.042	0	0.053+	0	0.076*	0	-0.071	0	0.533	0	2.123
	[0.024]	[0.046]	[0.036]	[0.040]	[0.017]	[0.046]	[0.026]	[0.058]	[0.023]	[0.441]	[0.012]	[10.135]
Ln (per cap household expenditures)	0.025**	0.377+	0.046**	0.581*	0.017**	0.042	0.018**	0.509+	-0.024**	-0.641	-0.002	-6.137
	[0.012]	[0.272]	[0.022]	[0.337]	[0.008]	[0.251]	[0.009]	[0.390]	[0.012]	[0.899]	[0.005]	[34.414]
Age	-0.001	-0.007	0	-0.013+	0	-0.013+	0	-0.005	0.002	-0.039*	0.002	-0.31
	[0.002]	[0.009]	[0.002]	[0.010]	[0.001]	[0.009]	[0.002]	[0.011]	[0.003]	[0.023]	[0.005]	[1.612]
Education	0.020*	0.005	0.020*	0.085++	0.012*	0.075***	0.018++	-0.033	-0.001	0.029	-0.033**	-0.276
	[0.012]	[0.046]	[0.012]	[0.057]	[0.007]	[0.025]	[0.011]	[0.058]	[0.003]	[0.042]	[0.016]	[1.021]
Health facilities access	0.002	0.018	0.001	-0.001	0.003	0.018	0.001	0.047**	-0.002	-0.028	-0.021	1.193
	[0.003]	[0.017]	[0.002]	[0.015]	[0.003]	[0.015]	[0.002]	[0.021]	[0.003]	[0.054]	[0.022]	[5.419]
State FEs	0.024**	-0.029++	-0.058***	-0.017	0.027**	-0.023	-0.017	0.055++	-0.004	-0.204+	-0.180***	-0.408
	[0.010]	[0.018]	[0.019]	[0.026]	[0.012]	[0.021]	[0.016]	[0.035]	[0.052]	[0.144]	[0.037]	[3.653]
(ii) Urban												
Health person is of same jati	0.018	-0.032	0.001	-0.110***	0.008	-0.018	0.01	-0.014	-0.004	-0.108	-0.013	7.08
	[0.015]	[0.027]	[0.004]	[0.041]	[0.006]	[0.028]	[0.008]	[0.030]	[0.006]	[0.256]	[0.012]	[121.970]
Watches TV regularly	0.018	0.01	0.017	0.261+	0.011	0.087	0.022	-0.026	-0.003	-0.536	0.006	-19.055
	[0.025]	[0.114]	[0.024]	[0.193]	[0.016]	[0.110]	[0.032]	[0.101]	[0.005]	[0.504]	[0.009]	[326.188]
Ln (per cap household expenditures)	0.003	0.007	0.003	0.583+	0.002	-0.003	0.003	0.19	-0.001	1.179	0	-33.784
	[0.014]	[0.328]	[0.012]	[0.426]	[0.008]	[0.263]	[0.017]	[0.308]	[0.005]	[1.167]	[0.002]	[580.524]
Age	0.002	0.003	0.003	-0.005	0.001	0.012	0.003	0.011	-0.004+	-0.003	0.001	0.39
	[0.003]	[0.014]	[0.003]	[0.035]	[0.002]	[0.013]	[0.003]	[0.014]	[0.003]	[0.058]	[0.003]	[7.276]
Education	0.035**	-0.005	0.031**	0.061**	0.019**	-0.007	0.034**	-0.019	-0.007++	0.002	-0.005+	-0.309
	[0.016]	[0.021]	[0.015]	[0.031]	[0.008]	[0.009]	[0.016]	[0.020]	[0.005]	[0.043]	[0.004]	[6.026]
State FEs	0.048***	0.025++	-0.002	0.053	0.094***	0.01	0.005	0.019	0.004	-0.055	0.088***	-0.19
	[0.010]	[0.016]	[0.018]	[0.050]	[0.014]	[0.017]	[0.008]	[0.021]	[0.013]	[0.093]	[0.024]	[3.957]
N	5,481	5,481	3,933	3,933	7,170	7,170	5,351	5,351	3,794	3,794	3,794	3,794

Notes: Decompositions are from low caste and religion females' viewpoint—i.e., using high caste endowments and returns. Values in brackets are robust Huber-White (Huber, 1967; White, 1980) standard errors. ***: statistically significant at 1 percent; **: statistically significant at 5 percent; *: statistically significant at 10 percent; ++: statistically significant at 15 percent +: statistically significant at 20 percent.
Source: 2004/05 India Human Development Survey (IHDS).

Economics Working Papers

2014-11: Jesper Bagger and Rasmus Lentz: An Empirical Model of Wage Dispersion with Sorting

2014-12: Louise Voldby Beuchert, Maria Knoth Humlum and Rune Vejlin: The Length of Maternity Leave and Family Health

2014-13: Julia Bredtmann, Sebastian Otten and Christian Rulff: Husband's Unemployment and Wife's Labor Supply – The Added Worker Effect across Europe

2014-14: Andrew B. Bernard, Valerie Smeets and Frederic Warzynski: Rethinking Deindustrialization

2014-15: Bo Sandemann Rasmussen: An Interpretation of the Gini Coefficient in a Stiglitz Two-Type Optimal Tax Problem

2014-16: A. R. Lamorgese, A. Linarello and Frederic Warzynski: Free Trade Agreements and Firm-Product Markups in Chilean Manufacturing

2014-17: Kristine Vasiljeva: On the importance of macroeconomic factors for the foreign student's decision to stay in the host country

2014-18: Ritwik Banerjee: On the Interpretation of Bribery in a Laboratory Corruption Game: Moral Frames and Social Norms

2014-19: Ritwik Banerjee and Nabanita Datta Gupta: Awareness programs and change in taste-based caste prejudice

2014-20: Jos Jansen and Andreas Pollak: Strategic Disclosure of Demand Information by Duopolists: Theory and Experiment

2014-21: Wenjing Wang: Do specialists exit the firm outsourcing its R&D?

2014-22: Jannie H. G. Kristoffersen, Morten Visby Krægpøth, Helena Skyt Nielsen and Marianne Simonsen: Disruptive School Peers and Student Outcomes

2014-23: Erik Strøjer Madsen and Yanqing Wu: Globalization of Brewing and Economies of Scale

2014-24: Niels-Hugo Blunch and Nabanita Datta Gupta: Social Networks and Health Knowledge in India: Who You Know or Who You Are?