NBER WORKING PAPER SERIES

MOTHERS' SOCIAL NETWORKS AND SOCIOECONOMIC GRADIENTS OF ISOLATION

Alison Andrew
Orazio Attanasio
Britta Augsburg
Jere Behrman
Monimalika Day
Pamela Jervis
Costas Meghir
Angus Phimister

Working Paper 28049
http://www.nber.org/papers/w28049

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
November 2020, Revised September 2023

We thank Yasmin Al-Ghabra, Cung Truong Hoang, Anusha Guha, and Gianina Morales Suazo for excellent research assistance. This study was funded by the NICHD (grant R01 HD072120), the Cowles Foundation and ISPS at Yale, the Population Studies Center at the University of Pennsylvania, the Leverhulme Trust (RPG-2021-391) and the SIEF program at the World Bank. This project has also received funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation program (grant agreement No 695300-HKADeC-ERC-2015-AdG). In addition, we gratefully acknowledge the support of the Economic and Social Research Council's Centre for the Microeconomic Analysis of Public Policy at IFS (grant reference ES/M010147/1). We bear sole responsibility for all errors and opinions expressed. The views expressed herein are those of the authors and do not necessarily reflect the views of the National Bureau of Economic Research.

Mothers’ Social Networks and Socioeconomic Gradients of Isolation
Alison Andrew, Orazio Attanasio, Britta Augsburg, Jere Behrman, Monimalika Day, Pamela Jervis, Costas Meghir, and Angus Phimister
NBER Working Paper No. 28049
November 2020, Revised September 2023
JEL No. D13,D71,O1,O35

ABSTRACT

Social connections are fundamental to human wellbeing. We examine the social networks of mothers of young children in rural Odisha, India. Gendered norms around marriage, mobility and work likely shape this group’s opportunities to form and maintain ties. We track 2,170 mothers’ networks over four years and find a high degree of isolation. Wealthier women and women from more-advantaged castes and tribes have smaller networks than their less-advantaged peers, primarily because they know fewer women within their own socioeconomic group. There exists strong, but symmetric, homophily by socioeconomic group. Socioeconomic differences are associated with toilet ownership and labor force participation.

Alison Andrew
The Institute for Fiscal Studies
alison_a@ifs.org.uk

Orazio Attanasio
Department of Economics
Yale University
87 Trumbull Street
New Haven, CT 06511
and Institute for Fiscal Studies,
FAIR, BREAD and CEPR
and also NBER
orazio.attanasio@yale.edu

Britta Augsburg
The Institute for Fiscal Studies
7 Ridgmount Street
London WC1E 7AE
britta_a@ifs.org.uk

Jere Behrman
University of Pennsylvania
The Ronald O. Perelman Center for
Political Science
133 South 36th Street
Philadelphia, PA 19104
jbehrman@econ.upenn.edu

Monimalika Day
Ambedkar University
New Delhi
India
monimalika.day@gmail.com

Pamela Jervis
Department of Industrial Engineering
University of Chile
Beauchef 851
Santiago, Chile
and Institute for Fiscal Studies
pjervisr@uchile.cl

Costas Meghir
Department of Economics
Yale University
87 Trumbull Street
New Haven, CT 06511
and IZA
and also NBER
c.meghir@yale.edu

Angus Phimister
Institute for Fiscal Studies
and University College London
angus.phimister.18@ucl.ac.uk

1 Introduction

'To engage in … social interaction[s]' and 'to live with and toward others' are basic capabilities essential for human dignity and freedom (Nussbaum 2011). Social networks and social interactions are crucial for broad aspects of wellbeing and are key drivers of economic outcomes.[1] The role of personal networks for economic outcomes is particularly important in low-income contexts where they often provide informal insurance (as stressed, among others, by Townsend (1994) and Munshi and Rosenzweig (2016)),[2] while also playing a key role in the diffusion of information about technological innovations, as discussed by Banerjee et al. (2013).

In this paper, we describe the social connections, or lack thereof, between younger married women in rural Odisha, India. Social ties with other women may be important in increasing women's support for more gender-equitable norms (Kabeer 1994; Rowlands 1997). The support that these ties provide and the collective action they enable are critical for social and political movements that empower women, both in their homes and in their broader communities (Prillaman 2023; Sanyal 2009).[3] Therefore, isolation may be part of a vicious cycle that entrenches the disadvantages that women face in terms of political representation and their voice and involvement in decision-making in their households and communities. Likewise, since social networks are important transmitters of information (Beaman et al. 2021; Behrman, Kohler, and Watkins 2002; BenYishay and Mobarak 2019; Kohler and Bühler 2001; Kohler, Behrman, and Watkins 2000; 2007), isolation may limit women's knowledge, particularly of heavily gendered subjects, such as sexual and reproductive health or child development (Anukriti et al., 2020; 2023), that are not typically discussed within married couples or within male social networks (Mason and Smith 2000) and which may be particularly pertinent around pregnancy and motherhood. Networks can also be an important means of access to capital, markets and insurance (e.g. Fafchamps and Lund, 2003; Feigenberg, Field and Pande, 2013; Field et al., 2016; Barnhardt, Field and Pande, 2017), implying that isolation may limit women's business endeavors and economic wellbeing.

The important linkages between women's social connections and their freedoms, mental health, empowerment and access to information raise several questions about women's social connections in

[1] On the relationship between networks and mental wellbeing and life satisfaction see Berkman et al. (2000), Cacioppo and Hawkley (2003), De Silva et al. (2007), Fowler and Christakis (2009) and Sawyer, Ayers and Smith (2010).
[2] For example, Ambrus, Mobius, and Szeidl (2014); Ambrus, Gao, and Milan (2022) and Attanasio and Krutikova (2020) analyze the role of networks in providing insurance.
[3] A recent review by Diaz-Martin et al. (2023) found positive effects on women's decision-making in roughly half the evaluations of women's groups they studied.

contexts with strict gender norms. How connected, or isolated, are women on average and how does this change over time? What is the depth of the social connections that women do have? How does connectedness or isolation vary by women's socioeconomic status? What drives socioeconomic gradients in isolation?

In this paper, we address some of these questions, by documenting the social ties of 2,170 married women with young children living in 192 villages of Odisha, India. This group may face particular barriers to building and maintaining social connections with peers. The custom of brides moving into their husbands' households upon marriage coupled with women typically marrying outside of their own communities (patrilocality) means that young women typically lose their adolescent and familial social networks upon marriage. Moreover, strong gender norms that frown upon married women moving freely about their communities or working outside the home mean that married women may find it hard to create and maintain new ties with peers in their new communities (Miller 1982; Chen 1995; Field et al. 2019; Jayachandran 2021), further affected by the presence of a mother-in-law (Anukriti et al., 2020). Restrictions on men and women from different households socializing mean that married women do not have access to their husbands' social networks.

We follow the same women over four years and measure on a yearly basis not only the number of connections they have but also the depth of these connections. We asked up to 12 mothers with children between the ages of 1 and 20 months in each village whether, and how well, they knew each of the other interviewed mothers in the village. On average, we interviewed a (quasi-randomly selected) 45% of mothers with children in this age range within a village. The median mother in our sample knew just 1, or 11%, of the other mothers we asked about despite the other mothers living on average just 237 meters away. Moreover, 39% of mothers reported that they did not know any other mother in our sample. An extrapolation exercise to account for the fact that we only asked about a fraction of other mothers in the village suggests a median within-village peer group size of 3.2. Additional data, collected 7 years after the first wave and 3 years after the last, suggests that other young women living in close proximity, and indeed primarily in the same household, are the *main* social network for young mothers in our sample, suggesting a very small overall network. We also document that in-person interactions are the main form of contact with mobile phone or online communication being very rare. This is similar to other contexts where women of similar ages and circumstances represent an important source of advice and support (Richardson, Barbour, and Bubenzer 1995). Furthermore, this group represents a key margin of network size adjustment. Whereas other components of one's networks, such as familial or caste ties, are fixed, the network of one's peers is more likely shaped by the individual.

The high level of social isolation that women face in rural India has been previously documented in qualitative studies (Crivello et al. 2018; Sanyal 2009) and, more recently, also in quantitative studies (Anukriti et al. 2020, 2023; Kandpal and Baylis 2019). The degree of isolation we find is consistent with these past studies. An important contribution of this work is that the panel nature of the data allows us to document the persistence of this isolation over a four-year period. Further, in addition to data on the existence of connections, we have rich data on the strength of connections and frequency of interactions which enables us to paint a detailed picture of women's social lives. Finally, our study captures social networks around new motherhood, a particularly crucial life stage for both women and their children.

We next describe the socioeconomic gradient of isolation and examine its correlates. We might expect social and economic characteristics, such as caste and poverty, to intersect with gendered norms and restrictions, resulting in differences in the types and strengths of women's social networks by their socioeconomic status (SES). It is, however, not obvious in which direction this gradient would go. For example, mothers from higher-SES households might acquire more social connections if their high status makes them a valuable connection that others seek out and if time devoted to social connections is something that only women from more-advantaged households can afford. Conversely, more-affluent women may be less-'valuable' connections or may benefit less from social connections if, for example, they are less involved in agricultural production and hence can less likely serve as a source of information (Magnan et al. 2015). More-advantaged households may also be able to 'afford' to adhere to more restrictive gendered roles for women. This may lead to women in more-advantaged households facing more restrictions to their mobility because these households may not have to rely on women's work outside the home to meet basic economic needs and can afford amenities such as indoor gas stoves and private toilets. Previous work has found that women from both more-advantaged castes (Boserup 1970) and wealthier households (Chen 1983) face more restrictions than their less-advantaged peers. Many studies have found that, in India, women's participation in paid work outside the home declines rapidly as other sources of household income, including men's earnings, rise (Kapadia 1995; Eswaran, Ramaswami, and Wadhwa 2013; Klasen and Pieters 2015; Mehrotra and Parida 2017; Chatterjee, Desai, and Vanneman 2018). This strong income effect on women's labor force participation is consistent with the idea that women not working is something that households value highly and opt for readily when economically and practically feasible. It has long been noted that in South Asia, women not leaving the home and not being in public spaces often brings households social status (Miller 1982; Chen 1995; Klasen and Pieters 2015). Having concrete reasons to leave home and be in public spaces, either for work or for other needs, may well be crucial for allowing mothers to make and maintain social connections.

In practice, we observe a negative socioeconomic gradient in connectedness: we find that mothers from richer households and those from more advantaged castes and tribes are more isolated than their peers from poorer households and less-advantaged castes and tribes. The additional data we collected suggests that mothers of higher SES do *not* engage in more interaction through mobile phones or online interactions. To the best of our knowledge, this is the first quantitative study to examine the socioeconomic gradient of women's social isolation.

We next analyze the drivers of the SES gradients we observe. We decompose the gradients in three parts: first, the SES composition of villages, second differences by SES in women's propensities to have social connections *within* their SES group; and third, SES differences in women's propensities to have ties *across* SES groups (i.e. differences in the degree of homophily). Our data suggest that the second component is the chief driver of both the caste/tribe and the wealth gradients: higher-SES women are substantially less likely to know the other higher-SES women in their village than lower-SES women are to know the other lower-SES women. The negative relationship between wealth status and connections also holds *within caste/tribe groups*. Social ties *across* SES groups are less common than those within groups, indicating substantial homophily, but this is equally the case for higher- and lower-SES mothers. Village composition can explain about a third of the observed caste/tribe gradient. To the best of our knowledge, this is the first study to offer such a decomposition of the drivers of heterogeneity in network size.

Finally, we examine correlates, or `mediators', of homophily and of SES differences in within- and across-group social ties. We find that the higher rates of toilet ownership amongst higher-SES households are associated with a substantial portion of both the homophily we observe and the lower within-group connectedness of high-SES women (by both wealth and caste/tribe). Toilet ownership means that women are less likely to have to defecate in the open. However, in this context, for the sake of safety, women often form informal groups with whom they travel out of the house to more isolated areas of the village to defecate, which opens up opportunities for social interactions (Patil 2019). Together, we interpret the mediation of isolation with toilet ownership as evidence that actions that households take as they get wealthier may end up worsening women's isolation. A similar conclusion is drawn from an analysis of labor force participation. This paper contributes towards building a new strand of literature on the drivers of women's isolation in contexts with strong gender norms. Recent work has suggested that women's isolation is intensified through living with a mother in law (Anukruti et al., 2020) or geographic relocation

(Barnhardt, Field, Pande, 2017).

Our paper proceeds as follows. In Section 2 we discuss the study context and data, Section 3 documents the features of social networks in this context and Section 4 concludes.

2 Study Context and Data

The setting for this study is rural Odisha, India. The Eastern Indian state is poorer and more rural than the country as a whole, with an income per capita of around US$1,300 and with 33% of the population living below the poverty line in 2018 (Government of Odisha 2018). According to the 2011 census, 17.1% of the population belong to a scheduled caste (SC), which is greater than the national proportion (16.6%), while the proportion belonging to a scheduled tribe (ST), at 22.9%, is far greater than the national average (8.6%).[4,5]

This study uses primary data gathered as part of a randomized control trial (RCT) of an early childhood intervention in 192 villages across three blocks (districts) of Odisha: Balangir (in Balangir district), Soro (Balasore) and Salepur (Cuttack) (see Figure A1 in Appendix A). The study sample consists of a panel of 2,170 mothers with infants between the ages of 1 and 20 months at the time of the first survey (wave 1) in November 2015, with an average of 11 study participants per village.[6]

We generated our sample as follows. We first identified the participants who met the study's eligibility criteria, which were based on the requirements of the early childhood intervention (see Attanasio et al. (2016) for more details). Pregnant women and mothers with children under the age of 2 years were identified through a census of each village in the summer of 2015, just prior to wave 1 of the study. The sample was split into two groups: *target* children, who met the age criteria for the intervention (between 7 and 16 months at the start of the intervention), and *spillover* children, who would be aged just above or just below the age criteria (1-6 and 17-20 months). In villages where there were fewer than eight eligible target children (roughly 38% of villages), all eligible children were selected. Villages with more than eight eligible target children had a median of 15 eligible children. In these villages, one child was selected at random,

[4] Odisha figures from 2011 census taken from https://www.censusindia2011.com/odisha-population.html

[5] Social interactions in Hindu communities in India continue to be influenced by the caste hierarchy. A detailed discussion of this complex system is beyond the scope of this paper; however, several authors have written about how it especially influences the lives of rural women (Byres, Kapadia, and Lerche 1999; Chakravarti and Krishnaraj 2018).

[6] Where the mother was not the primary caregiver of the child, we collected information on both mothers and primary caregivers. This occurred in 8.4% of cases. For cases where the biological mother is still alive, we used her as part of the networks-module; where this was not the case, we replaced her with the primary caregiver.

and that child's seven nearest neighbors were then targeted for enrollment.[7] This geographic approach to sampling was chosen to facilitate intervention implementation. However, since it meant that more-central households were more likely to be selected, a legitimate concern is whether this method may have systematically under-sampled groups with certain socio-economic characteristics who might be more likely to live on the outskirts of villages. If this were the case, it could bias our estimates of connectivity. To check this concern empirically, in Table A1 we assess how representative our chosen sample is of the all those with children in this age range in these villages using the census data. In particular, we use the census data on all eligible target children (those who we estimated would be between 7 and 16 months at the start of the intervention) and regress indicators of whether or not we selected respondents for inclusion on indicators of caste, household size and two proxies of wealth.[8] We see that none of these characteristics appears correlated with whether or not the respondent was selected for inclusion and this continues to hold true once we look *within* villages by including village fixed effects. While these results are reassuring, we show that our main findings are robust to only using the 38% of villages where there were eight or fewer target children and therefore where we didn't perform any within-village selection.[9]

Spillover children were selected by creating a list of all other children under 2 years in the village ordered by average distance from the randomly chosen central target child. A total of four spillover children per village from the ordered list were enrolled in the sample, choosing first children who were closer to age range of intervention eligibility.[10]

This generated an overall sample for wave 1 (target and spillover children combined) of 2,170 children with ages between 1 and 20 months, from between 10 and 12 households in each village (1,401 target, 769 spillover). Since households of target and spillover children are observationally equivalent on key margins (see Table A2 in Appendix A), we make no distinction between the two groups.

Mothers were re-interviewed in three further surveys in November 2016 (wave 2), November 2017 (wave 3) and March 2019 (wave 4). We collected additional data by phone to collect information on wider social

[7] All surplus children (children in the eligible age range who lived further than the first seven children from the central child) were placed on a reserve list and were added to the sample only if one of the targeted households had left the sample area between the census and wave 1 or refused the survey, and were added in order of distance from the central child. This occurred in around 14% of cases.

[8] Because we only collected a very limited set of variables in the census, we only have particularly crude measures of wealth. The only two proxies we have are whether the household has a mobile phone and whether they live in a multistory building.

[9] In particular, Table A3 replicates Table 2 while Table A4 replicates Table 3.

[10] This was done with the following order of priority: up to three 5- to 6-month-old children; up to two 17- to 18-month-old children; all other 5- to 6-month-olds; all other 17- to 18-month-olds; all other 4-month-olds and under; and all 19- to 20-month-olds. The quota of spillover children was filled using this order of priority when spillover children targeted for enrollment refused the survey.

networks in the summer of 2022.[11] Since our aim is to describe the social networks of young mothers in the absence of the treatment, we use data from all 2,170 mothers in wave 1 (pre-treatment) but use data only from the 532 mothers in the 48 control villages when analyzing dynamics in networks.

The characteristics of sample mothers and their households in wave 1 are given in Table 1. Mothers in our sample are young and poor, with an average age of 25 years; 6% of mothers work and the average household per-capita income per day of $0.84 (2019 USD); 93% live below the US$1.90 per day international poverty line. Around 65% of the households hold a ration card, for which only the poorest households are eligible. Households on average live 237m from each other, and constitute around a third of total mothers with children under 30 months in their village. We asked each respondent for the religion and the caste or tribe of the household head, which was then categorized into scheduled caste (SC), scheduled tribe (ST), other backward castes (OBC), dominant caste (Brahmin or Khandayata), or other. Sample households are 92% Hindu and 8% Muslim. Our sample is predominantly SC/ST/OBC (62%) with a significant minority identifying as the upper (dominant) caste (21%). In what follows we categorize SC/ST/OBC households as belonging to a 'disadvantaged caste or tribe'.

Table 1: Sample Characteristics in Wave 1

Variable	Mean	SD	N
Household Economic Characteristics			
Number of household members	5.46	2.36	2,170
HH under $1.90 per day poverty line (2019 USD) (proportion)	0.93	0.25	2,167
HH owns a toilet	0.47	0.50	2,167
HH has a ration card	0.65	0.48	2,164
HH engaged in agriculture	0.68	0.47	2,163
HH main room has dirt floor	0.43	0.50	2,167
HH owns a refrigerator	0.19	0.39	2,166
Household Social Characteristics			
Scheduled caste or tribe + OBC (proportions)	0.62	0.49	2,161
Khandayata or Brahmin	0.21	0.41	2,161
Hindu	0.92	0.28	2,166
Muslim	0.08	0.27	2,166
Mother and Child Characteristics			
Mother age (years)	25.4	4.38	2,162
Years since first child born	3.33	3.69	2,024
Grades of schooling attained	7.38	3.50	2,169
In labor force	0.06	0.24	2,167
Distance from other sample mothers within village (m)	237.4	213.38	2,168

[11] The choice to conduct a phone survey rather than face-to-face was driven by budget considerations as well as the COVID-19 pandemic.

In each survey wave, we collected detailed data on the social network among study participants. Each respondent was asked 'Do you know [NAME]?', for each other survey member in their village.[12] If a respondent answered affirmatively to knowing another participant, we asked a series of follow-up questions relating to the intensity of their relationship. These questions spanned a range of topics such as the duration of the relationship, whether or not they spoke about their children, and whether they could borrow food from this person.[13] These data provide a detailed picture of not only *who knows whom*, but also *how well* they know each other. We additionally collected each household's geographic location using GPS, cross-checking measurements over the multiple survey waves to reduce measurement error.[14]

It is important to consider the implications of our sampling strategy for our network data. Our social networks data are incomplete in two senses. First, in villages where not all mothers are included in the sample, we might underestimate the size of their `peer network'. To gauge this degree of underestimation, we conduct an exercise to extrapolate the patterns of connectivity we see in the partial network to the complete village network of mothers of children aged under 30 months as captured in the census data (for details, see Appendix C). We find that we capture on average one third of the total peer network. In wave 1 mothers self-reported their total peer network size. Our extrapolation is lower than the self-reported total number of known mothers (which is 6.3 on average), in line with literature suggesting that asking about named individuals, rather than aggregate number of connections is seen as more precise (McCormick et al, 2012; Breza et al 2020). In our main analysis, we therefore use the reported number of named individuals, but show robustness to our results using the estimated total network of mothers in Appendix Table A5. Results are in line. The second implication of our sampling strategy for mothers' networks relates to its location-based nature, which implies that our mothers are on average physically closer to each other than would be the case if they were selected at random. This selection might therefore bias upwards our estimates of connectivity in the complete network, implying that the degree of isolation could be underestimated. As discussed above, we show in Table A1 that our sample is representative at the village level along various

[12] In wave 1, this list was populated with the 12 mothers targeted for inclusion in the study on the basis of the census. However, not all these mothers were actually enrolled (due to refusals, incorrect information about the children's ages having been recorded during the census, or the interviewers being unable to relocate the house). This implies that for the dyad-level analysis, where we require characteristics of both the respondent and the mother asked about, we have a smaller analysis sample than for the general analysis of social connectedness. In waves 2–4, the list was populated with the actual study participants.

[13] For full module, see Appendix B.

[14] We primarily used GPS measurements taken at the census carried out at the start of the study. However, in cases where these coordinates suggested that a respondent lived more than 1 kilometer from their nearest neighbor, we manually compared these measures with those taken at later rounds and took the measure that appeared most reasonable.

margins. A limitation of our study remains that we cannot test representativeness in terms of networks. We however do show that our results hold when focusing only on villages where we did not need to do location-based sampling.

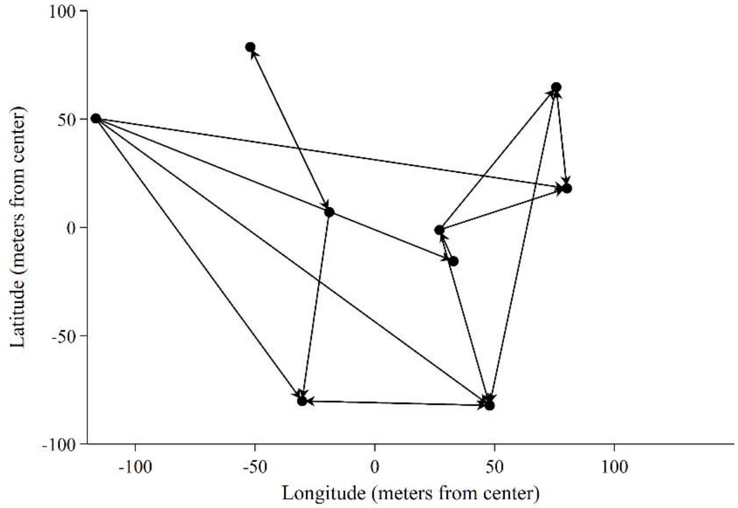
Finally, our network data are incomplete in the sense that we only analyze connections to other mothers of young children. To assess the extent to which this peer-group network reflects the mothers' total social network, we fielded additional data collection when their children were about seven years old. In particular, we asked mothers about how many people they know with whom they can relax and have fun with, and how many they know with whom they can talk when being down or depressed. We also asked them about the frequency of interaction by mobile phone or online (including through text messaging, WhatsApp, Facebook, etc.) and about their agreement with the statement that "[m]ost of my social interactions are with members of my household

Figure 1 shows examples of village networks in wave 1. Figure 1a shows an example sample village where each dot represents a respondent plotted, on the basis of their geographic position in the village, on a Cartesian coordinate system with the village center at (0,0), and each arrow represents a connection from one respondent to another. The direction the arrow points represents the direction of the reported connection. This village is smaller than average, and had five target children and four spillover children identified as part of the census. An advantage of the way we collect network data is that we are able to detect asymmetric or unreciprocated connections, making the captured network more precise (Breza et al 2020). Figure 1a makes clear that many reported connections are unreciprocated (around 48% in wave 1). Given the question we use to form these connections asked about whether the respondent knew the other mother, it is perfectly feasible that some respondents knew who the other mother was or had a brief acquaintance with her but that the connection was not reciprocated. For example, if some women are particularly prominent in the village, they may have many inward connections but themselves know relatively few others. The fact that many connections are unreciprocated highlights a point we make later in the paper, which is that even the connections that do exist (and defined so broadly – just an acquaintance) often appear weak in terms of how well individuals report knowing one another.

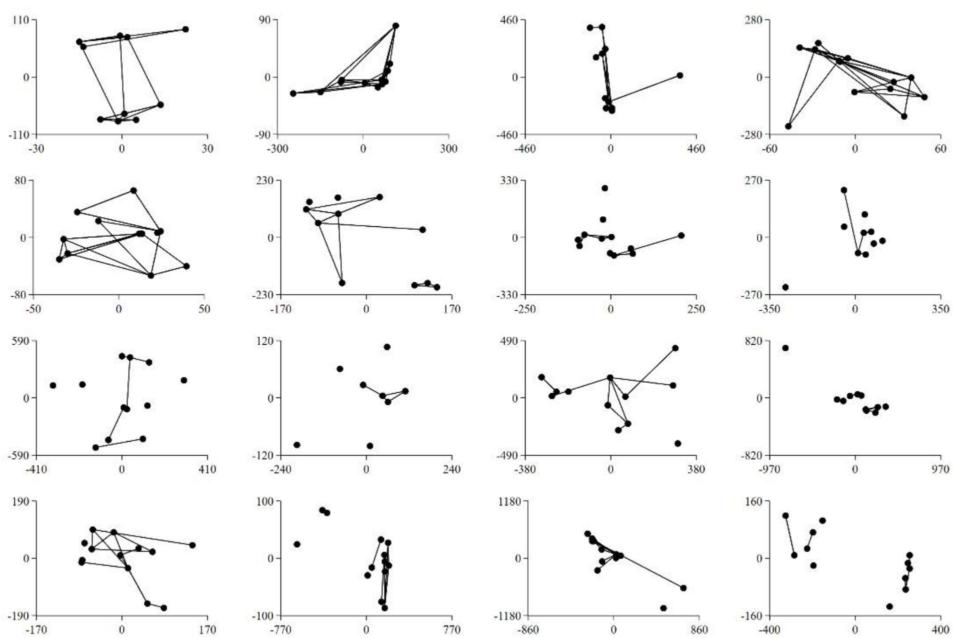
Figure 1b shows 16 other randomly selected villages displayed in the same way, where lines between respondents indicate any connection between the two. Figure 1b shows that there is considerable heterogeneity in the geographical spread of the sample in different villages, with many containing small sub-hamlets where a few households live outside of the main village.

Figure 1: 17 Randomly Selected Network Graphs from Wave 1

(a) Directed Network Graph



(b) 16 Randomly Selected Undirected Network Graphs



Notes: Panel A plots a directed network graph for a single randomly-chosen village. Panel B plots an undirected network from for random selection of villages in wave 1 of data collection. Mothers are positioned according to their geographic position in the village with the x- and y-axis showing meters from the village center along the longitude and latitude dimension respectively.

3 Isolation and its Socioeconomic Gradient

3.1 Isolation

Number of connections reported: We examine outward connections (that the mother identifies between herself and other mothers in the village sample) and inward connections (where other mothers have reported a connection with a particular sample mother within the village). Figure 2 shows the distribution of outward connections for all respondents in wave 1. The first feature of social networks in this sample is their sparsity. Out of an average of 11 possible connections within a village, in the control group the average number of connections reported is 1.21, the median is 1 and the mode is 0 (reported by 39% of sampled women). This number increases over time but remains relatively small, with a mean network size of 2.07 by wave 4 (Table 2).[15] We also show in Table 2 that these patterns over time in network size are almost identical when we consider the balanced panel who appear in all four waves, reassuring us that patterns are not being driven by differential attrition.

Estimating total within-village peer networks: A limitation of this exercise is that our data only contain connections between the mothers selected to be a part of the study. To estimate the average number of other mothers with children of a similar age that respondents know in the *whole* village, we perform an out-of-sample prediction exercise. For the sample for whom we have detailed network information, we estimate the probability that a connection exists between any two mothers (allowing the probability to vary with the children's ages, the mothers' ages, the mothers' castes and the mothers' geographic proximity to one another).[16] We then use these probabilities to predict the likelihood that our respondents know each of the other mothers in the village identified in the census with similarly aged children but whom we did not ask the respondent about. We then sum these probabilities to obtain an estimate of the total number of connections that mothers have, including those we did not directly enquire about. See Appendix C for details of this method. We estimate that each mother has an average of 3.2 connections to other mothers of similarly aged children in the village.

As an alternative to this extrapolation exercise, in wave 1 we additionally ask respondents how many other mothers they know with children between 0 and 24 months inside the village. While the networks literature reports asking about a random set of peers to be the preferred elicitation method when cost is not a

---

[15] As discussed in footnote 7, in waves 2–4 we asked about a different set of mothers. This may explain the reduction in total connections from wave 1 to wave 2.

determining factor (Griffith, 2022), the results show peer groups with a median size of 4 (see Figure A2 in Appendix A). Considering the proximity of these households and the small communities in which they reside, these are strikingly small peer groups.

Out-of-village and online connections: In the main survey waves, we only have information about social networks with other young mothers within the village. It could, therefore, be the case that although these young women are isolated from others in their situation in their immediate vicinity, they have other robust social networks that we are missing. In wave 1, we have some information about women's contact with their natal families. We plot this data in Figure A3a. The data confirms that patrilocality is indeed the overwhelming practice; While 94% of respondents whose in-laws are still alive live in the same village as them, only 9% of respondents live in the same village as their mother.[17] We further see that women's connections with their natal family are limited. For instance, the median woman living in a different village from her mother had seen her mother only 4 times over the past 12 months. We see similar patterns for seeing siblings. This data suggests that strong in-person connections with women's natal family are unlikely to be compensating for a sparse social network within their marital village.

This still leaves uncertain whether young women have other connections, and in particular whether they have connections (either to their natal family or to other people) over mobile or online that provide important sources of social support. To ascertain whether this holds true, we fielded some additional questions on a phone survey in 2022 which asked about the extent of women's networks outside the group of other young mothers in their village, which we plot in Figure A3b. The first thing to note is that even after having lived in their husbands' villages for upwards of 7 years, 96% of mothers say that most of her social interactions are with members of her own household. Mobile phones or social media do not seem reduce this degree of isolation. 86% of mothers say they did not interact with anyone on mobile or online (including through text messaging, WhatsApp, Facebook, etc.) with someone outside her immediate household in the past week. Those that did reported on average two such interactions in the last week. We conclude that women of similar ages and circumstances represent the key social network of these women, in line with many other contexts where peers have been shown to be the primary source of advice and support (Richardson, Barbour, and Bubenzer 1995).

[17] We note that the wealth and caste gradients we observe in Section 3.3 remain unchanged when we control for whether a woman lives in the same village as her mother and when we condition on those women who moved.

Figure 2: Distribution of Connections in Wave 1

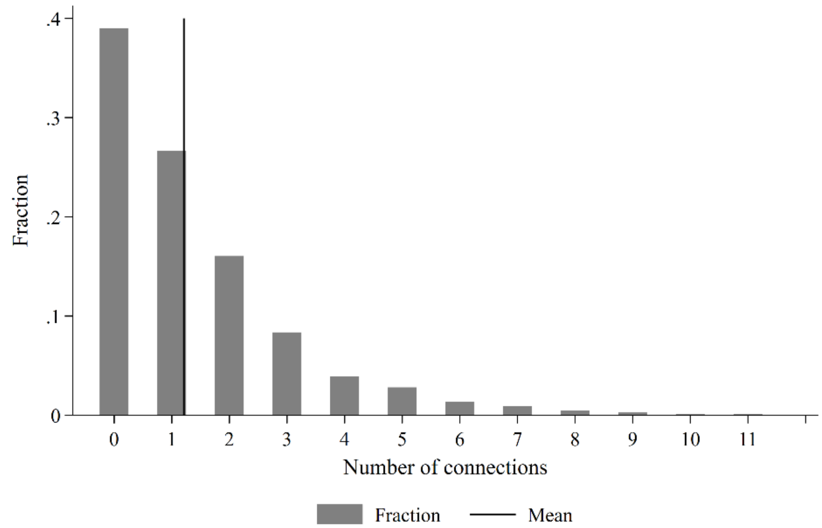


Table 2: Network Size by Wave in the Control Group

	All observations			*Balanced Panel*		
	Mean	SD	N	Mean	SD	N
Wave 1	1.21	1.54	530	1.28	1.57	455
Wave 2	1.01	1.28	487	1.02	1.26	455
Wave 3	1.79	1.70	486	1.82	1.71	455
Wave 4	2.07	1.95	490	2.09	1.91	455

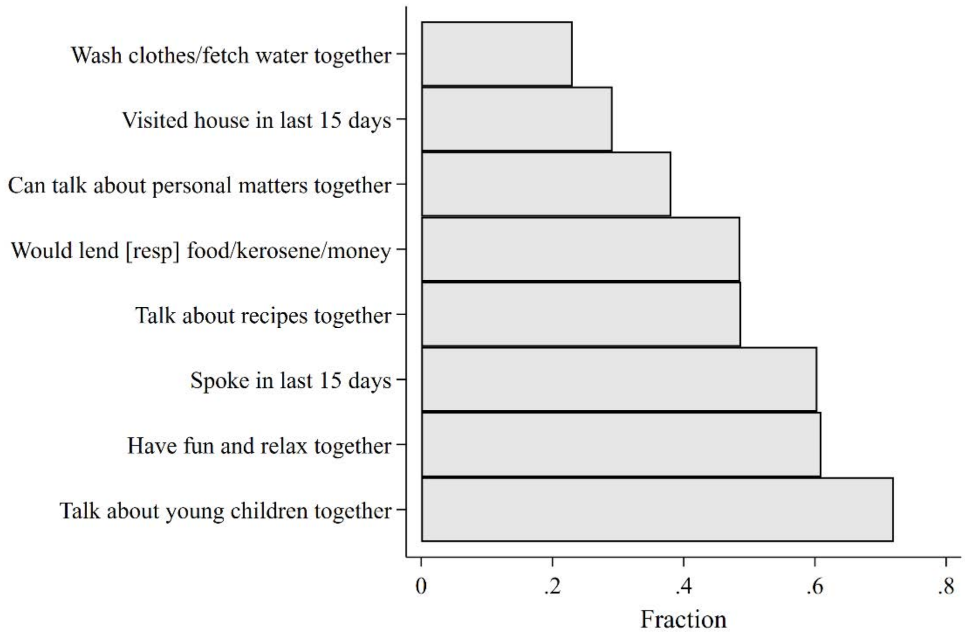
Note: Mean and standard deviation of the number of outward connections by wave in the control group. Left-hand panel contains all observations while right-hand side shows is a balanced panel of mothers who appear in every wave.

3.2 Strength of Connections

Figure 3 shows the strength of social ties that women in our sample report in wave 1. It displays the proportion of connections for which respondents report doing a certain activity together or being able to draw on the connection for support.

Of those we asked about, the most common shared activity was talking about young children (72%). This suggests that motherhood is a defining identity in structuring young women's relationships in this context. 60% of respondents reported having spoken to a given connection in the last 15 days. Only 29% had visited the connection's house during the same period. Given that the sample villages are small and respondents live close together, this suggests that women have relatively infrequent contact, and even less frequent private contact, even with the connections that they do have. Only for 38% of connections did respondents report being able to talk about personal matters.

Figure 3: Strength of Social Ties (Wave 1)



Notes: Bar chart shows the fraction of times that (conditional on a connection existing) the respondent reports that the connection meets each of eight criteria.

For some analysis, it is useful to summarize all information about how well members of such connections know each other into a single 'connectedness' index defined between each mother and every other mother in the sample in her village. This index takes on a value between 0 (indicating the respondent does not know that mother at all) and 1 (indicating that the respondent knows that mother and answered 'yes' to every one of the indicators listed in Figure 3). We create this indicator through a latent factor model. We model respondent *i*'s response (Z_{ijk}) to each of the eight indicators, $k = \{1,\ldots,8\}$, listed in Figure 3 regarding other mother *j* as the following function of the underlying connectedness of mother *i* to mother *j*, θ_{ij}:

$$Z_{ijk} = \frac{\exp(a_k\theta_{ij} + b_k)}{1 + \exp(a_k\theta_{ij} + b_k)}$$

Conditional on a connection existing between i and j at all, we assume that θ_{ij} is distributed normally with mean 0 and variance 1. This is a standard two-parameter item response theory model. We estimate the parameters, $\{a_k, b_k\}$, through maximum likelihood. We then predict values of θ_{ij} for each i to j connection by taking the mean of the posterior distribution of θ_{ij} conditional on Z_{ijk} and the estimated parameters. So that we can also define a level for this connectedness index for connections that do not exist, we assume that a connection not existing is the same as a connection where none of the indicators about the strength of the connection is nonzero. Finally, we rescale the connectedness index to lie on the [0,1] interval, where it takes the value 0 when $Z_{ijk} = 0$ for all k, and the value 1 when $Z_{ijk} = 1$ for all k.

3.3 Heterogeneity and the Socioeconomic Gradient of Connections

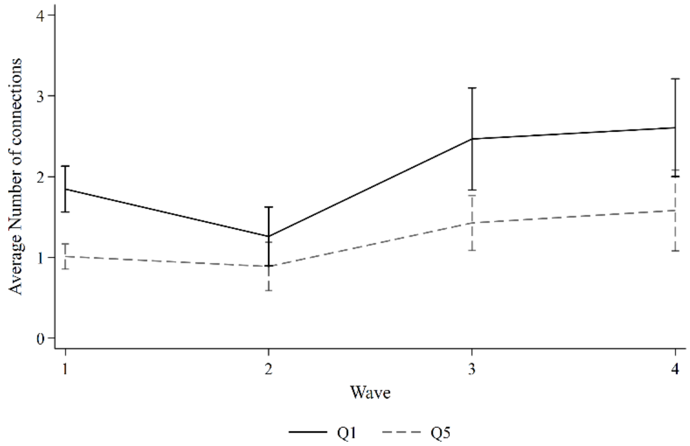
We next explore heterogeneity in mothers' networks. As we saw in Figure 2, there is considerable variation in the number of connections reported. The first thing we check is whether this variation exists *within* or *between* villages. In other words, is it just that some villages are much more densely connected than others or is it the case that even within villages there is heterogeneity in mothers' connectedness? A simple decomposition of variance suggests that 59.7% of the overall variance in reported connections comes from *within* village variation. If we condition on mother's age, age of the eldest child, distance from the village center and the number of kids, 63.5% of the residualized variance comes from within-village variation.

We next consider how the size of mothers' networks varies by socioeconomic status (SES), specifically by wealth, and caste and tribe.[18] Figures 4a and 4b plot, respectively, the average number of outward connections by wealth and by caste and tribe across the four survey waves for the controls. Across both dimensions of socioeconomic status, there are large and persistent negative gradients in network size. Namely, poorer mothers and mothers from more disadvantaged castes and tribes (SC/ST/OBC) report *more* connections than their wealthier peers and peers from more advantaged castes or tribes (non-SC/ST/OBC). At wave 1 this amounted to an average of 0.90 fewer connections for mothers in the highest wealth quintile relative to the lowest and of 0.56 fewer connections for non-SC/ST/OBC women relative to SC/ST/OBC women. Given the median network size in wave 1 is 1, these differences are substantial.

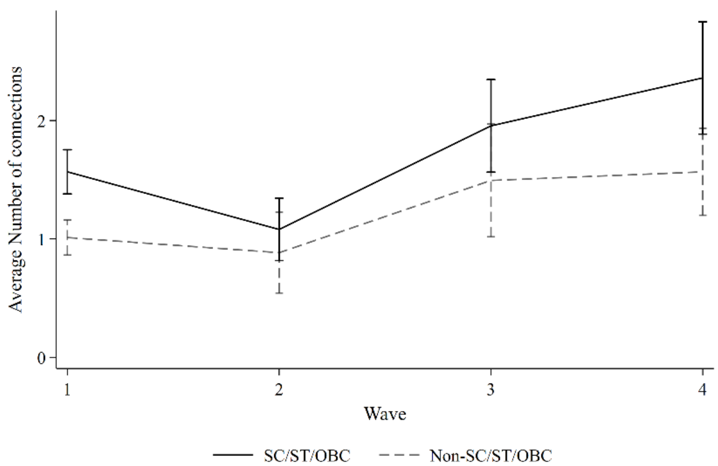
[18] An individual's wealth score is calculated using a principal component analysis of assets in wave 1. The principal component is calculated across, not within, villages. Across all groups, wealth is low, with an average per-capita daily income of $0.55 in the lowest wealth quintile compared with $1.39 in the highest (2019 USD). While wealth and caste are significantly correlated (correlation coefficient of 0.24), there is a lot of variation in wealth within caste and we therefore consider this to be a conceptually separate dimension of SES that is important to analyze.

Figure 4: Socioeconomic Gradient over Time for Controls

(a) Network Size and Wealth Quintile



(b) Network Size and Caste or Tribe



Note: Averages include the whole sample in wave 1, and only control villages thereafter. Dashed lines indicate 95% confidence intervals

Both panels of Figure 4 show an increase in network size over time for each group, yet both the caste/tribe and wealth gradients persist, and arguably increase, between waves 2 and 4 and persist thereafter. This suggests that the determinants of these gradients are pertinent throughout the period in which mothers have young children.

We run a regression analysis of total network size at baseline against a series of covariates to estimate the conditional correlation between certain key characteristics and network size (Table 3). In Panel A the outcome variable is total outward network size, and in Panel B we weight each connection by its estimated ‘connectedness’, the index between 0 and 1 defined in the previous subsection. This weighted measure thus combines both the number of connections and how well each connection is known.

Columns 1 and 2 show us again what we saw in the above figures: dominant caste and wealthier women have fewer connections. While caste is a binary indicator, our wealth index is continuous and Figure 5a shows a binscatter plot of the number of connections by the wealth indicator. We see evidence that average number of connections is monotonically decreasing in wealth across the whole of the wealth distribution. Column 3 shows that women who had their first child longer ago (proxying for the length of time in the village) also have larger networks which is what we would expect. Figure 5b plots a binscatter plot of network size on time since the birth of the eldest child and shows that while the mean appears to be increasing throughout the distribution of child age, the rate of change decreases suggesting that network size might plateau. Column 4 shows that each of these dimensions (caste/tribe, wealth and age of the eldest child) is statistically significant even when both are included in the regression, suggesting that all are important predictors of network size. Column 5 shows that this effect of caste/tribe and age of the eldest child persists even when we control for covariates mothers’ age, number of children and distance from the village center. These are important to control for since all these might systematically differ by SES. Each of the three dimensions remains highly statistically significant and similar in size after controlling for these covariates. Over and above age of the eldest child, mothers’ is not predictive of network size. Interestingly, network size is also strongly predicted by labor force participation, indicative of working mothers being more mobile around their villages.

Finally, in Column 6, we add further covariates that we might imagine are co-determined with networks including toilet ownership and labor force participation. Toilet ownership, even conditional on wealth, is associated with 0.52 fewer connections, likely due to women who own toilets not travelling with other mothers in their villages to defecate. Conditional on other covariates, the wealth index is not statistically significant, which could indicate that the effect of wealth is operating through these other characteristics, such as toilet ownership, labor force participation and distance from the village center.

Moving to Panel B, we see that these associations persist once we weight the number of connections by how well mothers know each other. Wealth conditional on other covariates is significantly negatively correlated with having a higher weighted number of connections, suggesting that after conditioning on other factors, higher wealth may be particularly associated with knowing connections less well. Finally, in Appendix Table A5, we show that we also see the same large and highly statistically significant socioeconomic gradients in respondents’ self-reports of their overall number of connections.

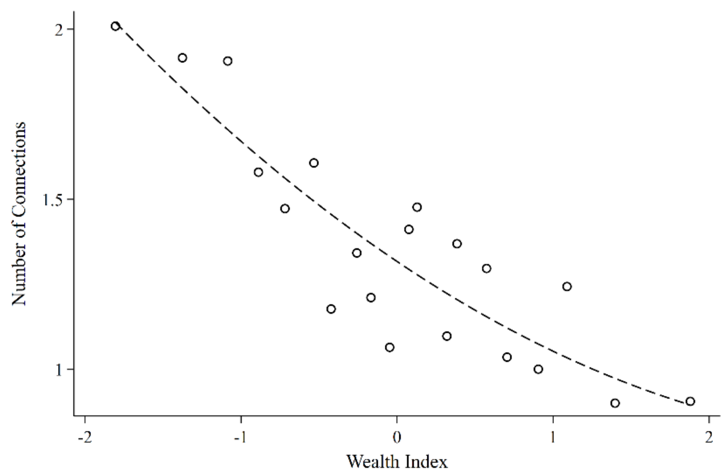
Table 3. Correlates of Outward Network Size at Wave 1

	Panel A: Number of Outward Connections					
	(1)	(2)	(3)	(4)	(5)	(6)
Wealth index	-0.308***			-0.198***	-0.180***	-0.0341
	(0.0546)			(0.0499)	(0.0482)	(0.0536)
Scheduled caste or tribe + OBC		0.555***		0.445***	0.430***	0.344***
		(0.103)		(0.0984)	(0.0942)	(0.0904)
Age of eldest child (years)			0.0891***	0.0748***	0.0568***	0.0527***
			(0.0133)	(0.0124)	(0.0198)	(0.0193)
Mother's age (years)					0.00544	0.00669
					(0.0125)	(0.0125)
Dist. from sample center (km)					-2.269***	-2.177***
					(0.256)	(0.244)
Number of children					0.0639	0.0521
					(0.0685)	(0.0693)
Household owns toilet						-0.521***
						(0.0973)
Mother in labor force						0.682***
						(0.210)
Constant	1.354***	1.012***	1.040***	0.817***	1.169***	1.405***
	(0.0725)	(0.0752)	(0.0678)	(0.0796)	(0.287)	(0.294)
Observations	2153	2144	2153	2144	2139	2139
	Panel B: Number of Outward Connections weighted by Connectedness					
	(1)	(2)	(3)	(4)	(5)	(6)
Wealth index	-0.204***			-0.140***	-0.136***	-0.0530*
	(0.0309)			(0.0277)	(0.0280)	(0.0308)
Scheduled caste or tribe + OBC		0.321***		0.241***	0.240***	0.193***
		(0.0613)		(0.0573)	(0.0576)	(0.0540)
Age of eldest child (years)			0.0469***	0.0310***	0.0218*	0.0186
			(0.00768)	(0.00908)	(0.0122)	(0.0118)
Mother's age (years)					0.00658	0.00716
					(0.00807)	(0.00801)
Dist. from sample center (km)					-1.243***	-1.182***
					(0.160)	(0.151)
Number of children					0.0465	0.0405
					(0.0431)	(0.0428)
Household owns toilet						-0.291***
						(0.0560)
Mother in labor force						0.499***
						(0.143)
Constant	0.664***	0.466***	0.498***	0.524***	0.481***	0.609***
	(0.0412)	(0.0451)	(0.0398)	(0.189)	(0.182)	(0.182)
Observations	2153	2144	2153	2139	2139	2139

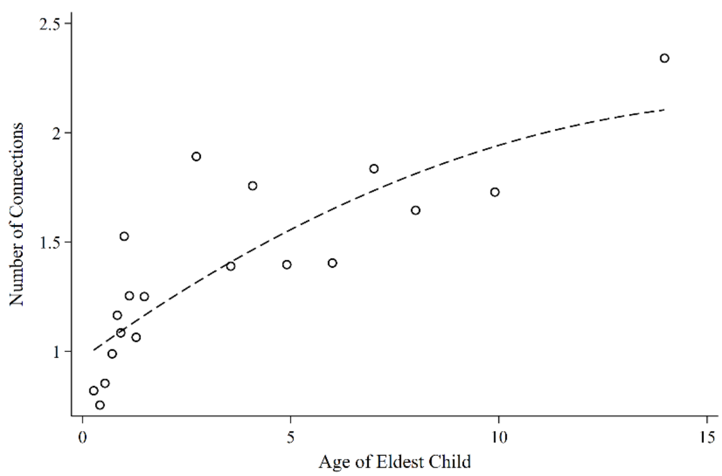
Notes: Table shows regression coefficients and standard errors from regressing the number of outward connections an individual has (Panel A) and that number weighted by how well they know each connection (Panel B) on wealth, caste/tribe, toilet ownership, age and labor force participation. * p<0.1, ** p<0.05, *** p<0.001.

Figure 5: Relationship between Number of Outward Connections, Household Wealth and Age of the Eldest Child.

(a) Wealth Index



(b) Age of Eldest Child



Note: Figures plot binscatter plots between the number of connections reported and (a) the household's wealth index and (b) the age of the eldest child. In particular, we use 20 quantiles for both variables and within each quantile, we plot the average number of outward connections. We also add a quadratic trend line.

3.4 Who Knows Whom? Decomposing SES Gradients

Mechanically, the SES gradients we observe can be decomposed into three components:[19]

i. *Differences in within-group connectedness between SES groups.* Such differences could drive our SES gradients if high-SES women with young children were less likely than low-SES women to know other women *within* their own SES group. We define this component of the gradient as that that can be attributed to differences in the within-group connectedness of low- and high-SES sample mothers; It is the portion of the gradient that would remain if village composition was symmetric (i.e. both SES groups lived in villages with the same number of potential within- and across-group connections) and if across-group connection rates were the same across groups.

ii. *Differences in across-group connectedness between SES groups.* Such differences could drive our SES gradients if high-SES women were less likely to know women from outside their SES group (low-SES women) than low-SES women were. This is the component of the gradient driven by differences in the rate at which low-SES sample mothers report knowing high-SES sample mothers and vice versa; It is the portion that would remain if village composition was symmetric and if within-group connection rates were the same across groups.

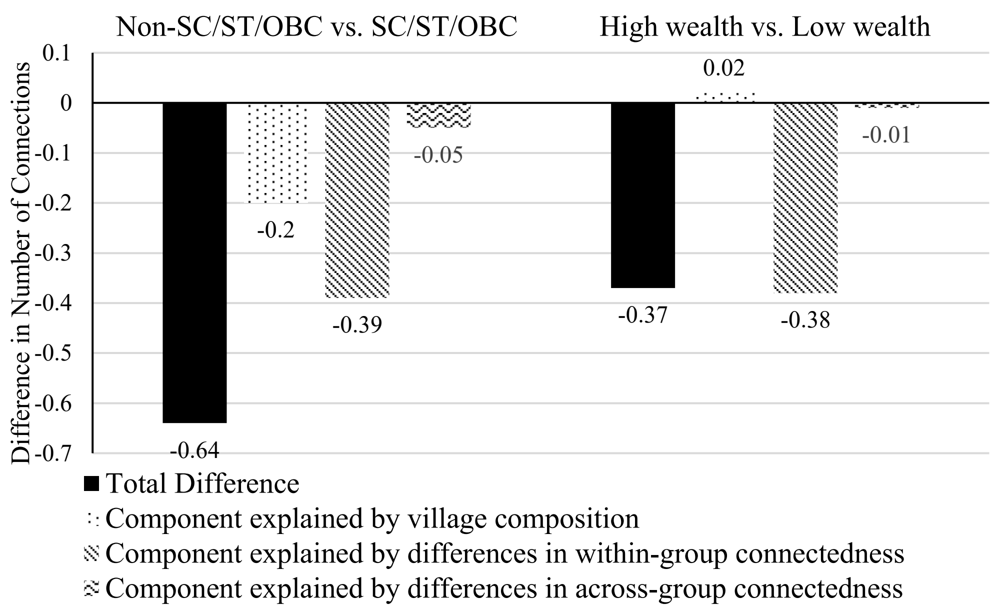
iii. *Village composition effects.* A final factor is that even with identical within- and across-group rates of connections, a negative gradient could result if the composition of the village is such that women of higher SES systematically live in villages where they are in the minority, and those of lower SES in villages where they are in the majority. Under homophily (a higher rate of within vs. across group connections), this would lead to an aggregate difference in the total number of connections even if high SES and low SES women were as likely as each other to know women of their own and outside their groups. We defined these village composition effects as the portion of the gradient that would remain if the between- and across-group connection rates were the same across SES groups.

[19] Denote higher- and lower-SES mothers as, respectively, H and L. Let $\bar{T}_H$ be the sample average of the total number of other sample mothers that the high-SES sample mothers know in each village. Mechanically, we have that $\bar{T}_H = \hat{p}_{HH} * \bar{n}_{HH} + \hat{p}_{HL} * \bar{n}_{HL}$ where $\bar{n}_{HH}$ and $\bar{n}_{HL}$ are sample averages of the total number of other high-SES sample mothers and of low-SES sample mothers living in villages where high-SES mothers live, and $\hat{p}_{HH}$ and , $\hat{p}_{HL}$ are the in-sample probabilities that a high-SES mother reports knowing, respectively, another high-SES sample mother, or a low-SES sample mother. Correspondingly, we have: $\bar{T}_L = \hat{p}_{LL} * \bar{n}_{LL} + \hat{p}_{LH} * \bar{n}_{LH}$ with analogous definitions. Taking the difference and by rearranging terms, we can decompose the *difference* in the number of connections that low- and high-SES mothers report into:

$$\bar{T}_L - \bar{T}_H = \underset{(i) Within-group}{\bar{n}_{HH}(\hat{p}_{LL} - \hat{p}_{HH})} + \underset{(ii) Across-group}{\bar{n}_{HL}(\hat{p}_{LH} - \hat{p}_{HL})} + \underset{(iii) Village\ Composition\ Effects}{\hat{p}_{LL}(\bar{n}_{LL} - \bar{n}_{HH}) + \hat{p}_{LH}(\bar{n}_{LH} - \bar{n}_{HL})}$$

With our detailed dyad-level data, we can provide an exact decomposition of the SES gradient into components driven by the three components. In this exercise, we again define high wealth and low wealth by the household being above or below the sample median value of a wealth index. This makes village composition effects (effect iii) irrelevant for explaining wealth gradients since, by definition, across the sample high and low wealth women live in villages with the same number of within- and across-group potential connections.[20]

Figure 6: Decomposition of SES Gradients in Network Size



We show results from this decomposition exercise in Figure 6. Overall, we see that the largest share of both the wealth and the caste gradient comes from differences in *within-group* connectedness. Differences in the rate at which high- and low-SES women report knowing the other women in their village from their same SES group accounts for 61% of the caste gradient and 100% of the wealth gradient. By contrast, only 7% of the caste gradient is explained by differences in *across*-group connection rates while this component plays no role for wealth.

These patterns can be seen in Figures 7a and 7b, which plot the dyadic probabilities of connections within

[20] In practice, villages do not contain the identical number of other sample women and so this is true for the proportions but not the numbers of women in their own and of the opposite wealth group. This is why we estimate a non-zero (but very small) village composition effect in Figure 6.

and across SES groups. Starting with *within*-group connections, we see that SC/ST/OBC sample mothers are substantially more likely to report knowing a randomly chosen other sample mother from their village from their broadly-defined caste/tribe group (around 23%) than non-SC/ST/OBC mothers are (around 15%; Figure 7a). Low-wealth mothers are substantially more likely to report knowing a randomly chosen mother in their same wealth group than are high-wealth mothers (22% versus 14%; Figure 7b).

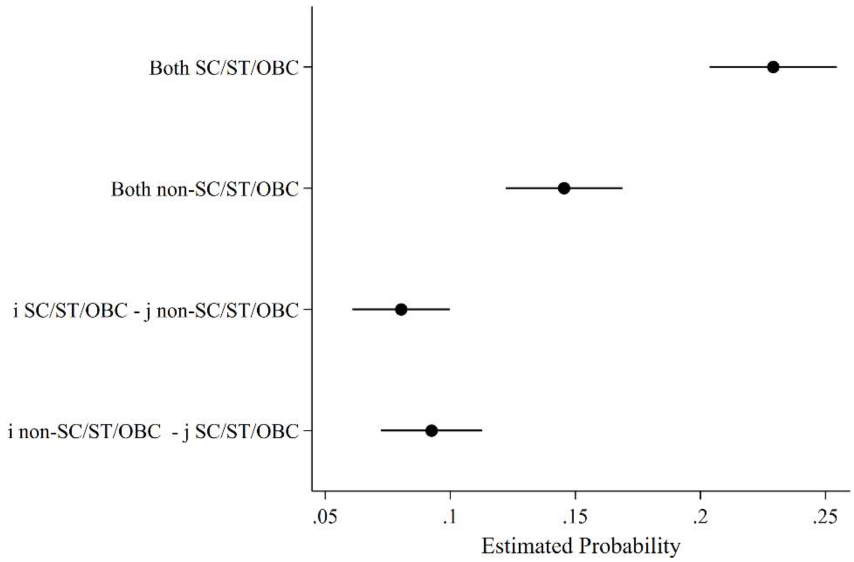
Moving to *across*-group connections, the bottom two bars of Figures 7a and 7b plot the across-group connectedness of sample mothers by caste/tribe and wealth respectively. The probability of across-group connections is substantially lower than the probability of within-group connections. This is true along both the caste/tribe and the wealth dimensions, and for both higher- and lower-SES mothers. Our social networks thus exhibit substantial homophily. For neither caste/tribe nor wealth do we see differences in the probability of across-group connections by mothers' SES. In other words, high-SES mothers are as likely to report knowing a randomly chosen lower-SES mother in their village as low-SES mothers are to report knowing a high-SES mother. This is why across-group differences contribute little to the overall SES gradients (Figure 5).

The remaining component is what can be explained by village composition given identical within- and across-group connection rates. As noted earlier, since our wealth grouping is simply defined as being above or below the median on an wealth index, this is not relevant for the wealth gradient.[21] Village composition can, though, explain 31% of the caste gradient. Simple descriptive statistics tell us that SC/ST/OBC sample mothers, on average, live in villages with 6.0 other SC/ST/OBC sample mothers and 2.3 non-SC/ST/OBC sample mothers. This contrasts to non-SC/ST/OBC sample mothers who, on average, live in villages with 4.7 other non-SC/ST/OBC sample mothers and 3.5 SC/ST/OBC sample mothers. Even with identical probabilities of forming connections within and across groups, the fact that mothers from more advantaged caste/tribe groups systematically live in villages with fewer other mothers from their own caste/tribe group could contribute to the SES caste gradient we observe under homophily.

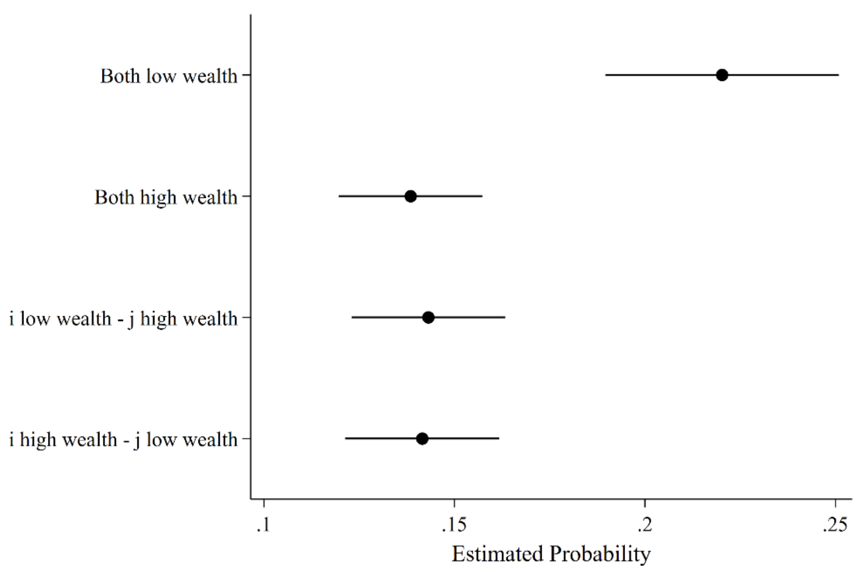
[21] The only reason we would see village composition playing a role for wealth would be that differences in village sizes and/or differential non-response to the network questions. Reassuringly, then, our decomposition finds that village composition would predict a difference between the number of connections of high- and low-wealth women of just 0.02.

Figure 7: Dyad-Level Probabilities in Wave 1

(a) Caste/Tribe



(b) Wealth



Notes: Figures plot the probability that a respondent *i* reports knowing another respondent *j* depending on *i* and *j*'s caste/tribe and wealth alongside 95% confidence intervals.

3.5 Mediating the Gap in Connectedness

Our decomposition exercise shows that differences in within-group connectedness can explain the majority of the negative SES gradients in connectedness by caste/tribe and by wealth; a lower within-group connectedness can explain the entirety of the wealth gradient and three fifths of the caste/tribe gradient. In this section, we look at what mother characteristics correlate with the observed gradient, to get a better idea of *why* lower-SES women have higher within-group connectedness than higher-SES women. One explanation could for example be that higher-SES women face more restrictions in interacting with peers, even peers of the same wealth and caste/tribe groups. These could stem from women in higher-SES households facing greater mobility restrictions, especially if it is less necessary for these women to leave the household frequently for work or for using the toilet.

To probe the relevance of such drivers of within- and across-group connectedness by caste/tribe and wealth, we conduct a mediation analysis, which helps us to assess whether other observed characteristics of respondents and the asked-about mother can mediate the observed SES gradients using a dyad-level analysis. The approach we take, while routinely used should be seen as descriptive in nature: control variables that can explain a portion of the SES gap do not necessarily themselves 'cause' social connections; they may simply be correlated to underlying causes of connections (Rosenbaum, 1984). A causal mediation analysis would require an instrumental variable approach to control for the potential endogeneity of the mediators[22]. However, no credible instruments exist in our context, and we offer this mediation analysis as tentative evidence to stimulate further research on the topic.

We first regress, by ordinary least squares,[23] a binary indicator of whether or not a connection between mother *i* and mother *j* in village *v* exists (Y_{ijv}) on indicators of whether this is a low-to-high-SES connection (C_{ijv}^{LH}), a high-to-low-SES connection (C_{ijv}^{HL}) or a high-to-high-SES connection (C_{ijv}^{HH}) based on i's and j's caste or wealth group, with the omitted group being low-to-low-SES connections:

$$Y_{ijv} = \beta_0 + \beta^{HH} C_{ijv}^{HH} + \beta^{LH} C_{ijv}^{LH} + \beta^{HL} C_{ijv}^{HL} + \epsilon_{ijv}$$

We allow the error term, ϵ_{ijv}, to be arbitrarily correlated within the same village but assume independence across villages. These estimates are equivalent to those in Section 3.4. $\hat{\beta}^{HH}$ is the difference in the

22 See for example the analysis of the role of parental investments as a mediator for an early childhood intervention in Colombia (Attanasio et al., 2020).

23 The benefit of using OLS over the probit estimator in this exercise is that we can use simple linear combinations of the β parameters to exactly recover the estimated probability of two individuals being connected, and do not have to make assumptions about the distribution of ϵ_{ijv}. Repeating the analysis with probit yields almost identical results (available upon request).

probability of a high-SES mother having a randomly chosen within-group connection and the same probability for a low-SES mother (i.e. $\hat{p}^{HH} - \hat{p}^{LL}$). $\hat{\beta}^{LH}$ ($\hat{\beta}^{HL}$) is the difference between the probability of a low-SES (high-SES) mother having a randomly chosen across-group connection and the probability that a low-SES mother has a randomly chosen within-group connection, and thus is equal to $\hat{p}^{LH} - \hat{p}^{LL}$ ($\hat{p}^{HL} - \hat{p}^{LL}$). The magnitudes of $\hat{\beta}^{LH}$ and $\hat{\beta}^{HL}$ are indicative of the degree of homophily while the magnitude of $\hat{\beta}^{HH}$ is indicative of the degree to which low-SES women have within-group connections at a different rate from high-SES women.

We sequentially add other characteristics of mother *i* (X_{iv}), mother *j* (X_{jv}) and their interactions ($X_{iv} * X_{jv}$):

$$Y_{ijv} = \beta_0 + \beta^{HH} C_{ijv}^{HH} + \beta^{LH} C_{ijv}^{LH} + \beta^{HL} C_{ijv}^{HL} + \alpha_1 X_{iv} + \alpha_2 X_{jv} + \alpha_3 X_{iv} * X_{jv} + \epsilon_{ijv}$$

We observe how the unexplained differences in the probability of a connection existing (β^{HH}, β^{LH} and β^{HL}) change as a result of adding these controls. This provides an indication of whether these observed characteristics can 'explain' SES differences we see by caste/tribe and by wealth in the probability of having connections. Figure 8a shows how different characteristics mediate the gaps in probabilities of different groups reporting connections relative to the probability of the 'SC/ST/OBC to SC/ST/OBC' connection. The figure starts with the caste/tribe-only model, sequentially adding wealth, age, household toilet ownership, maternal labor force participation, and finally the distance between respondents in the same village (quadratically). While independently important predictors of connectedness, controlling for wealth and age does not substantially alter the gap between the within-group connectedness SC/ST/OBC mothers and that of non-SC/ST/OBC mothers, or the degree of homophily exhibited.

Controlling for household toilet ownership reduces the difference in within-group connectedness by caste/tribe by roughly 3 percentage points (p.p.). It also reduces the difference between the probability of cross-group connections and within-group connections existing by a similar magnitude. Non-SC/ST/OBC households are more likely to own a toilet in our sample (64% vs 36% for SC/ST/OBC) and thus are less likely to defecate in the open, something that women often do in a group (Patil 2019). This analysis suggests that this might be an important feature in explaining why women from non-SC/ST/OBC households have fewer within-group connections, and why they both know fewer and are known by fewer SC/ST/OBC women. Labor force participation, while having little association above and beyond toilet ownership, if included separately is associated with a similar percentage of both within- and across-group connectivity. Labor force participation amongst sample women is rare, but marginally more common amongst SC/ST/OBC women (6.2% vs 6.0%). Taken together, these results suggest that the lower mobility of non-

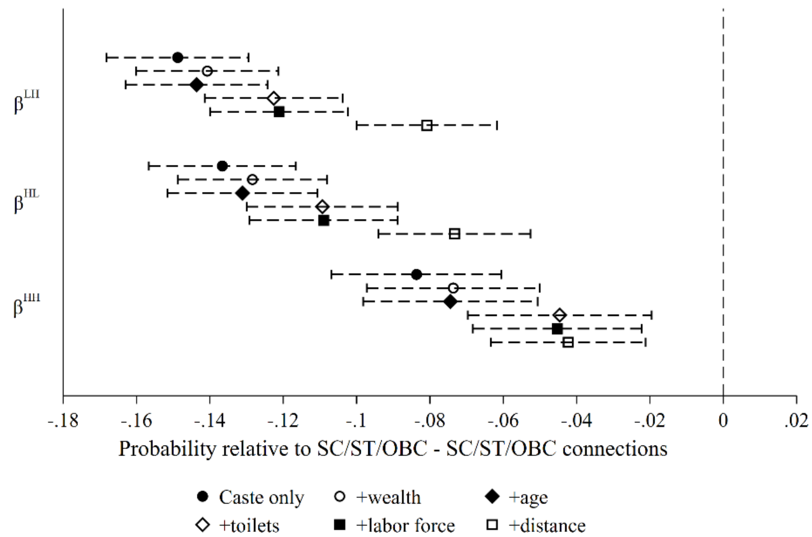
SC/ST/OBC women is associated with their smaller social networks.

Controlling for distance between respondents reduces the difference in probability of a cross-caste/tribe versus within-group connections by around 4 p.p., suggesting it could be an important driver of caste/tribe-based homophily. However, distance is associated with none of the difference in within-group connectedness by caste/tribe conditional on all other covariates. Villages in our sample are segregated by caste and tribe, with the average distance between mothers of different groups being 339m relative to only 244m for mothers of the same groups, in line with the general practice of families from different castes and tribes residing in different parts of the village (or even different villages). This framework does not allow us to determine the causal relationship between distance and network size; villages could be segregated because households do not want to form ties across caste/tribe lines, and segregated villages could simultaneously limit the opportunities for individual connections to be made.

Figure 8b shows the same mediation analysis for the wealth gradient, plotting probabilities relative to low-wealth-to-low-wealth connections. Controlling for caste/tribe and age reduces by roughly 2 p.p. the wealth difference both in within- and across-group connectedness. Labor force participation and toilet ownership, as with the caste/tribe gradient, can also explain some of the wealth difference. This lends weight to the argument that mobility plays a role in the size of one’s network. Indeed, once toilet ownership is controlled for, there is no remaining within-group difference in the probability of connections between low- and high-wealth mothers. Distance is associated with little of the wealth gradient, likely due to a lower degree of segregation (233m for within-wealth connections vs 294m for across-wealth connections).

Figure 8: Mediation Analysis of SES Differences in Connection Probability in Wave 1

(a) Caste/Tribe



(b) Wealth

Note: Panels a and b plot the coefficients $\hat{\beta}^{LH}$, $\hat{\beta}^{HL}$ and $\hat{\beta}^{HH}$ from equation 4 as controls are sequentially added to the model. Wealth is a binary indicator equal to 1 if a household has a principal component analysis (PCA) asset score above the village median. Caste is a binary indicator equal to 1 if a household head is SC/ST/OBC. Age is mother's age in years. Toilets is a binary indicator of household toilet ownership. Labor force is a binary indicator of mother's labor force participation. Distance is distance to other mother in meters (included quadratically).

4 Discussion and conclusion

The extreme isolation of young mothers in rural India that we document is worrying given existing evidence, from various contexts, that social isolation is associated with poor mental and physical health for women (Berkman et al., 2000; Cacioppo and Hawkley, 2003; De Silva et al., 2007; Kohler, Behrman and Watkins, 2007; Fowler and Christakis, 2009; Sawyer, Ayers and Smith, 2010; Smith and Postmes, 2011) and with women more likely to be victims of domestic violence (Choi, Cheung, and Cheung 2012; Lanier and Maume 2009). Adverse effects of social isolation on mothers may have knock-on impacts on their children (Bennett et al. 2016; Kingston and Tough 2014; Sawyer, Ayers, and Smith 2010). Much of the existing evidence on the effects of social isolation comes from high-income countries where the reasons for and consequences of isolation probably are substantially different from the context we study due to, for example, fewer restrictions on women's mobility, higher incomes and higher rates of women working outside of the home, and less restriction due to social structures such as the caste system. More evidence on the correlates of isolation for young women in contexts with highly restrictive gender norms and in high-poverty settings is useful to understand the costs borne by women and communities as a result of female isolation.

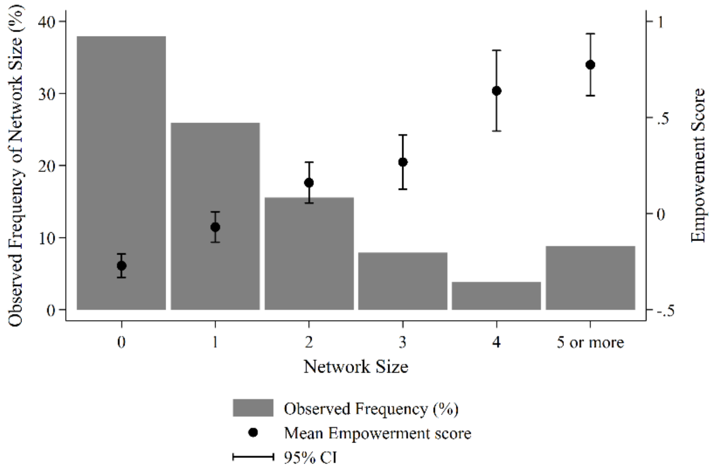
Within our own data, we can examine correlations between social networks and broader indicators of women's wellbeing. Figure 9a plots these associations for both an indicator of women's empowerment, as measured by an index summarizing 10 questions covering household decision making and independence.[24] Figure 9b plots the same for symptoms of depression, as measured by the 10-item versions of the Center for Epidemiologic Studies depression scale (CES-D) (Andresen et al., 1994). Figure 9a shows clearly that larger network sizes are associated with women being more empowered. The empowerment index is scaled to have a zero mean and unit standard deviation in our sample. We find that while women who report having zero connections have an average empowerment score of –0.27, those who report having 5 or more connections have an average empowerment score of 0.77. Appendix Table A7, which reports the same results in a regression framework, confirms this very substantial correlation and shows that it remains similar in magnitude and significance once controlling for covariates. Figure 9b below shows the correlation between network size and symptoms of depression, which has a mean of 3.6 and a standard deviation of

[24] 5 questions relate to who (within the family) decides what to do in different scenarios, e.g. "Who decides how much money is spent on food?". 5 relate to independence, e.g. "Do you own any asset of value you could sell to make a needed payment?" We combine these using a 2-parameter IRT factor model to create a summary score with zero mean and unit variance. The index is coded such that higher values represent more empowerment.

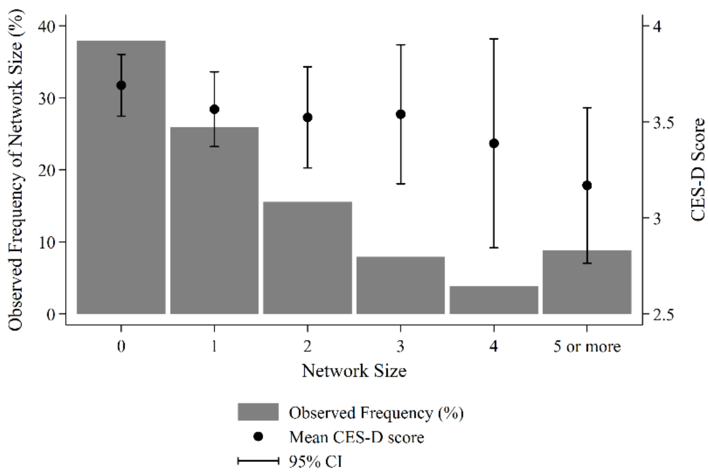
2.4 in our sample. This correlation is substantially weaker than that with empowerment although there is nevertheless a statistically significant negative relationship whereby women with larger networks report fewer symptoms of depression (also see Table A7).

Figure 9: Correlations between Network Size, Empowerment and Symptoms of Depression

(a) Empowerment



(b) Symptoms of Depression



Given the very high degree of social isolation among young married women in rural India it is important to understand more about the impact of governmental policies and large-scale programs on connectedness. Recent work has shown that women's educational programs can be successful at expanding women's social networks (Kandpal and Baylis 2019). Conversely, relocation programs for slum dwellers can shrink networks (Barnhardt, Field, and Pande 2017). However, little evidence exists about the impact of national programs, including employment programs such as the National Rural Employment Guarantee Act in India, that may indirectly expand or contract women's social networks. The Indian government has recently made huge investments in expanding access to private toilets through the Swachh Bharat Mission (Curtis 2019) and our results suggest that evaluations of this effort may want to consider the policy's unintended impacts on female isolation.

We need to better understand the nature of the relationship between social isolation on the one hand and wealth and amenities, since here we cannot establish causality. Further research should study how women's networks relate to those of men and how important each of these networks is for information dissemination, insurance and other economic and social activities. With that understanding we may start to see how economic growth may affect social networks, which can be crucial for individual wellbeing.

The analysis we have presented in this paper is descriptive and thus we do not draw firm causal conclusions about the causes and the consequences of women's isolation, which can include negative impacts on their wellbeing and the development of their children, thereby deepening the intergenerational transmission of poverty and inequalities. However, we consider the extent of isolation we document, and its association with socioeconomic status, to be a cause for concern, and a motivator for future research on this topic.

References

Ambrus, Attila, Wayne Gao, and Pau Milán. 2020. "Informal risk sharing with local information." The Review of Economic Studies 89, no. 5: 2329-2380.

Ambrus, Attila, Markus Mobius, and Adam Szeidl. 2014. “Consumption Risk-Sharing in Social Networks.” American Economic Review 104 (1): 149–82. https://doi.org/10.1257/aer.104.1.149.

Andresen EM, Malmgren JA, Carter WB, & Patrick DL (1994). Screening for depression in well older adults: evaluation of a short form of the CES-D (Center for Epidemiologic Studies Depression Scale). American journal of preventive medicine, 10(2), 77–84.

Anukriti, S., Catalina Herrera-Almanza, and Praveen Kumar Pathak Mahesh Karra. 2020. “Curse of the Mummy-Ji: The Influence of Mothers-in-Law on Women in India.” American Journal of Agricultural Economics. Volume 102, Issue 5, October 2020, Pages 1328-1351. https://ideas.repec.org/p/bos/iedwpr/dp-337.html.

Anukriti, S. Catalina Herrera-Almanza, and Praveen Kumar Pathak Mahesh Karra. 2022. " Bring A Friend: Leveraging Financial and Peer Support to Improve Women’s Reproductive Agency in India". Boston University Working Paper.

Attanasio, Orazio, Sarah Cattan, Emla Fitzsimons, Costas Meghir, and Marta Rubio-Codina. 2020. Estimating the Production Function for Human Capital: Results from a Randomized Controlled Trial in Colombia, American Economic Review, Volume 110(1) pages 48-85.

Attanasio, Orazio, and Sonya Krutikova. 2020. "Consumption Insurance in Networks with Asymmetric Information: Evidence from Tanzania." Journal of the European Economic Association 18, no. 4.

Attanasio, Orazio, Costas Meghir, Britta Augsburg, Jere Behrman, Bet Caeyers, Monimalika Day, Sally Grantham-McGregor, Smitri Pahwa, and Marta Rubio-Codina. 2016. “Early Childhood Development for the Poor: Impacting at Scale, Odisha, India.” Baseline Report.

Banerjee, A., A. G. Chandrasekhar, E. Duflo, and Matthew O Jackson. 2013. “The Diffusion of Microfinance.” Science 341 (6144): 1236498–1236498. https://doi.org/10.1126/science.1236498.

Barnhardt, Sharon, Erica Field, and Rohini Pande. 2017. “Moving to Opportunity or Isolation? Network Effects of a Randomized Housing Lottery in Urban India.” American Economic Journal: Applied Economics 9 (1): 1–32.

Beaman, Lori, Ariel BenYishay, Jeremy Magruder, and Ahmed Mushfiq Mobarak. 2021. "Can network theory-based targeting increase technology adoption?." American Economic Review. 111(6), 1918-1943.

Behrman, Jere R, Hans-Peter Kohler, and Susan Cotts Watkins. 2002. "Social Networks and Changes in Contraceptive Use over Time: Evidence from a Longitudinal Study in Rural Kenya." Demography 39 (4): 713–38.

Bennett, Ian M, Whitney Schott, Sofya Krutikova, and Jere R Behrman. 2016. "Maternal Mental Health, and Child Growth and Development, in Four Low-Income and Middle-Income Countries." J Epidemiol Community Health 70 (2): 168–73.

BenYishay, Ariel, and A. Mushfiq Mobarak. 2019. "Social Learning and Incentives for Experimentation and Communication." The Review of Economic Studies 86 (3): 976–1009. https://doi.org/10.1093/restud/rdy039.

Berkman, L. F., T. Glass, I. Brissette, and T. E. Seeman. 2000. "From Social Integration to Health: Durkheim in the New Millennium." Social Science & Medicine (1982) 51 (6): 843–57. https://doi.org/10.1016/s0277-9536(00)00065-4.

Blinder, Alan S. 1973. "Wage Discrimination: Reduced Form and Structural Estimates." The Journal of Human Resources 8 (4): 436–55. https://doi.org/10.2307/144855.

Boserup, Ester. 1970. Woman's Role in Economic Development. Earthscan.

Breza, E., A.G. Chandrasekhar, T.H. McCormick, M. Pan (2020). Using Aggregated Relational Data to Feasibly Identify Network Structure without Network Data. American Economic Review, 110(8): 2454–2484

Byres, T. J., Karin Kapadia, and Jens Lerche. 1999. Rural Labour Relations in India. Routledge.

Cacioppo, John T., and Louise C. Hawkley. 2003. "Social Isolation and Health, with an Emphasis on Underlying Mechanisms." Perspectives in Biology and Medicine 46 (3 Suppl): S39-52.

Chakravarti, Uma, and Maithreyi Krishnaraj. 2018. Gendering Caste: Through a Feminist Lens. First Edition. Mathura Road, New Delhi. https://doi.org/10.4135/9789353287818.

Chatterjee, Esha, Sonalde Desai, and Reeve Vanneman. 2018. "Indian Paradox: Rising Education, Declining Womens' Employment." Demographic Research 38 (March): 855–78.

https://doi.org/10.4054/DemRes.2018.38.31.

Chen, Martha. 1983. A Quiet Revolution : Women in Transition in Rural Bangladesh. Cambridge, MA : Schenkman Pub. Co. https://trove.nla.gov.au/work/21577368.

———. 1995. A Matter of Survival: Women's Right to Employment in India and Bangladesh. Women, Culture, and Development. Oxford University Press. https://oxford.universitypressscholarship.com/view/10.1093/0198289642.001.0001/acprof-9780198289647-chapter-2.

Choi, Susanne Y. P., Y. W. Cheung, and Adam K. L. Cheung. 2012. "Social Isolation and Spousal Violence: Comparing Female Marriage Migrants With Local Women." Journal of Marriage and Family 74 (3): 444–61. https://doi.org/10.1111/j.1741-3737.2012.00963.x.

Crivello, G., J. Roest, U. Vennam, R. Singh, and F. Winter. 2018. Marital and Fertility Decision-Making: The Lived Experiences of Adolescents and Young Married Couples in Andhra Pradesh and Telangana, India. Young Lives. https://ora.ox.ac.uk/objects/uuid:315e83f5-753f-42c8-afec-237bd0e96f64.

Curtis, Val. 2019. "Explaining the Outcomes of the 'Clean India' Campaign: Institutional Behaviour and Sanitation Transformation in India." BMJ Global Health 4 (5). https://doi.org/10.1136/bmjgh-2019-001892.

De Silva, Mary J., Sharon R. Huttly, Trudy Harpham, and Michael G. Kenward. 2007. "Social Capital and Mental Health: A Comparative Analysis of Four Low Income Countries." Social Science and Medicine 64 (1): 5–20. https://doi.org/10.1016/j.socscimed.2006.08.044.

Díaz-Martin, Lucía, Akshara Gopalan, Eleonora Guarnieri, and Seema Jayachandran. 2023. "Greater than the sum of the parts? Evidence on mechanisms operating in women's groups." The World Bank Research Observer. *38*(1), pp.1-35.

Eswaran, Mukesh, Bharat Ramaswami, and Wilima Wadhwa. 2013. "Status, Caste, and the Time Allocation of Women in Rural India." Economic Development and Cultural Change 61 (2): 311–33. https://doi.org/10.1086/668282.

Fafchamps, Marcel, and Susan Lund. 2003. "Risk-Sharing Networks in Rural Philippines." Journal of Development Economics 71 (2): 261–87. https://doi.org/10.1016/S0304-3878(03)00029-4.

Feigenberg, Benjamin, E. Field, and Rohini Pande. 2013. "The Economic Returns to Social Interaction:

Experimental Evidence from Microfinance." The Review of Economic Studies 80 (4): 1459–83. https://doi.org/10.1093/restud/rdt016.

Field, Erica, Seema Jayachandran, Rohini Pande, and Natalia Rigol. 2016. "Friendship at Work: Can Peer Effects Catalyze Female Entrepreneurship?" American Economic Journal: Economic Policy 8 (2): 125–53. https://doi.org/10.1257/pol.20140215.

Field, Erica, Rohini Pande, Natalia Rigol, Simone Schaner, and Charity Moore. 2019. "On Her Own Account: How Strengthening Women's Financial Control Impacts Labor Supply and Gender Norms." SSRN Scholarly Paper ID 3456234. Rochester, NY: Social Science Research Network. https://doi.org/10.2139/ssrn.3456234.

Fowler, James H., and Nicholas A. Christakis. 2009. "Dynamic Spread of Happiness in a Large Social Network: Longitudinal Analysis over 20 Years in the Framingham Heart Study." BMJ (Online) 338 (7685): 23–26. https://doi.org/10.1136/bmj.a2338.

Griffith, A. 2022. Name Your Friends, but Only Five? The Importance of Censoring in Peer Effects Estimates Using Social Network Data. Journal of Labor Economics. Volume 40, Number 4

Government of Odisha. 2018. "Odisha Economic Survey 2018-2019." Report. Government of Odisha. https://pc.odisha.gov.in/Download/Economic_Survey_2018-19.pdf.

Jayachandran, Seema, 2021. Social norms as a barrier to women's employment in developing countries. IMF Economic Review, 69(3), pp.576-595.

Kabeer, Naila. 1994. Reversed Realities: Gender Hierarchies in Development Thought. Verso.

Kandpal, Eeshani, and Kathy Baylis. 2019. "The Social Lives of Married Women: Peer Effects in Female Autonomy and Investments in Children." Journal of Development Economics. https://doi.org/10.1016/j.jdeveco.2019.05.004.

Kapadia, Karin. 1995. Siva And Her Sisters: Gender, Caste, And Class In Rural South India. Routledge.

Kingston, Dawn, and Suzanne Tough. 2014. "Prenatal and Postnatal Maternal Mental Health and School-Age Child Development: A Systematic Review." Maternal and Child Health Journal 18 (7): 1728–41.

Klasen, Stephan, and Janneke Pieters. 2015. "What Explains the Stagnation of Female Labor Force Participation in Urban India?" World Bank Economic Review 29 (3): 449–78.

https://doi.org/10.1093/wber/lhv003.

Kohler, H. P., and C. Bühler. 2001. "Social Networks and Fertility." In International Encyclopedia of the Social & Behavioral Sciences, edited by Neil J. Smelser and Paul B. Baltes, 14380–84. Oxford: Pergamon. https://doi.org/10.1016/B0-08-043076-7/02176-8.

Kohler, Hans-Peter, Jere R Behrman, and Susan C Watkins. 2007. "Social Networks and HIV/AIDS Risk Perceptions." Demography 44 (1): 1–33.

Kohler, Hans-Peter, Jere R. Behrman, and Susan Cotts Watkins. 2000. "Empirical Assessments of Social Networks, Fertility and Family Planning Programs: Nonlinearities and Their Implications." Demographic Research 3. https://www.jstor.org/stable/26348012.

Lanier, Christina, and Michael O. Maume. 2009. "Intimate Partner Violence and Social Isolation across the Rural/Urban Divide." Violence Against Women 15 (11): 1311–30. https://doi.org/10.1177/1077801209346711.

Magnan, Nicholas, David J. Spielman, Travis J. Lybbert, and Kajal Gulati. 2015. "Leveling with Friends: Social Networks and Indian Farmers' Demand for a Technology with Heterogeneous Benefits." Journal of Development Economics 116 (September): 223–51. https://doi.org/10.1016/j.jdeveco.2015.05.003.

Mason, Karen Oppenheim, and Herbert L. Smith. 2000. "Husbands' versus Wives' Fertility Goals and Use of Contraception: The Influence of Gender Context in Five Asian Countries." Demography 37 (3): 299–311. https://doi.org/10.2307/2648043.

McCormick, T.H., M.J. Salganik, T. Zheng. 2012. How many people do you know?: Efficiently estimating personal network size. J Am Stat Assoc. 2010 Mar 1; 105(489): 59–70.

Mehrotra, Santosh, and Jajati K. Parida. 2017. "Why Is the Labour Force Participation of Women Declining in India?" World Development 98 (C): 360–80.

Miller, Barbara D. 1982. "Female Labor Participation and Female Seclusion in Rural India: A Regional View." Economic Development and Cultural Change 30 (4): 777–94.

Munshi, Kaivan, and Mark Rosenzweig. 2016. "Networks and Misallocation: Insurance, Migration, and the Rural-Urban Wage Gap." American Economic Review 106 (1): 46–98. https://doi.org/10.1257/aer.20131365.

Nussbaum, Martha C. 2011. Creating Capabilities. Harvard University Press.

Oaxaca, Ronald. 1973. “Male-Female Wage Differentials in Urban Labor Markets.” International Economic Review 14 (3): 693–709. https://doi.org/10.2307/2525981.

Patil, Kajal. 2019. “A Social Network Analysis of Open Defecation Practices in India,” 48.

Prillaman, Soledad Artiz., 2023. Strength in numbers: how women's groups close India's political gender gap. American Journal of Political Science, 67(2), pp.390-410.

Richardson, Rhonda A., Nancy E. Barbour, and Donald L. Bubenzer. 1995. “Peer Relationships as a Source of Support for Adolescent Mothers.” Journal of Adolescent Research 10 (2): 278–90. https://doi.org/10.1177/0743554895102005.

Rosenbaum P. R. (1984). The consequences of adjustment for a concomitant variable that has been affected by the treatment. Journal of the Royal Statistical Society A, 147(5), 656–666. https://doi.org/10.2307/2981697

Rowlands, Jo. 1997. Questioning Empowerment: Working with Women in Honduras. Oxfam.

Sanyal, Paromita. 2009. “From Credit to Collective Action: The Role of Microfinance in Promoting Women’s Social Capital and Normative Influence.” American Sociological Review 74 (4): 529–50. https://doi.org/10.1177/000312240907400402.

Sawyer, Alexandra, Susan Ayers, and Helen Smith. 2010. “Pre- and Postnatal Psychological Wellbeing in Africa: A Systematic Review.” Journal of Affective Disorders 123 (1–3): 17–29. https://doi.org/10.1016/j.jad.2009.06.027.

Smith, Laura G. E., and Tom Postmes. 2011. “The Power of Talk: Developing Discriminatory Group Norms through Discussion.” British Journal of Social Psychology 50 (2): 193–215. https://doi.org/10.1348/014466610X504805.

Townsend, Robert. 1994. “Risk and Insurance in Village India.” Econometrica 62 (3): 539–91.

Mothers' Social Networks and Socioeconomic Gradients of Isolation

Appendices

Alison Andrew (Oxford University, Institute for Fiscal Studies (IFS))
alison.andrew@economics.ox.ac.uk

Orazio Attanasio (Yale University, IFS, UCL)
orazio.attanasio@yale.edu

Britta Augsburg (IFS)
britta_a@ifs.org.uk

Jere Behrman (University of Pennsylvania)
jbehrman@sas.upenn.edu

Monimalika Day (Dr. B. R. Ambedkar University Delhi)
monimalika@aud.ac.in

Pamela Jervis (University of Chile)
pjervisr@uchile.cl

Costas Meghir (Yale University, IFS, NBER, CEPR, IZA)
c.meghir@yale.edu

Angus Phimister (UCL)
angusdav@outlook.com

Appendix A: Additional Table and Figures

Appendix Table A1: Representativeness of Selected Sample relative to Census Sample

	Panel A: Target Child Sample, No Fixed Effects				
	Dependent Variable: Targeted for Inclusion in "Target Child" Sample				
	(1)	(2)	(3)	(4)	(5)
Dominant caste	0.0298				0.0338
	(0.0338)				(0.0334)
Number of HH members		-0.00731			-0.00639
		(0.00497)			(0.00482)
Main respondent has mobile			-0.0326		-0.0313
			(0.0272)		(0.0269)
Live in multistory building				0.0317	0.0144
				(0.0493)	(0.0484)
Constant	0.527***	0.570***	0.553***	0.500***	0.570***
	(0.0239)	(0.0335)	(0.0287)	(0.0541)	(0.0579)
Observations	2695	2735	2735	2735	2695
	Panel B: Target Child Sample, Village Fixed Effects				
	Dependent Variable: Targeted for Inclusion in "Target Child" Sample				
Dominant caste	0.0201				0.0204
	(0.0346)				(0.0346)
Number of HH members		0.00333			0.00258
		(0.00425)			(0.00438)
Main respondent has mobile			0.00496		0.00122
			(0.0256)		(0.0259)
Live in multistory building				-0.0337	-0.0344
				(0.0398)	(0.0402)
Observations	2695	2735	2735	2735	2695
	Panel C: Spillover Child Sample, No Fixed Effects				
	Dependent Variable: Targeted for Inclusion in "Spillover Child" Sample				
Dominant caste	0.0209				0.0216
	(0.0284)				(0.0283)
Number of HH members		-0.00309			-0.00304
		(0.00446)			(0.00446)
Main respondent has mobile			-0.00125		-0.00234
			(0.0244)		(0.0244)
Live in multistory building				0.0137	0.00859
				(0.0497)	(0.0493)
Constant	0.351***	0.371***	0.356***	0.343***	0.359***
	(0.0158)	(0.0258)	(0.0224)	(0.0491)	(0.0568)
Observations	2123	2127	2127	2127	2123
	Panel D: Spillover Child Sample, Village Fixed Effects				
	Dependent Variable: Targeted for Inclusion in "Spillover Child" Sample				
Dominant caste	0.0162				0.0128
	(0.0323)				(0.0321)
Number of HH members		0.00296			0.00233
		(0.00473)			(0.00480)
Main respondent has mobile			0.0319		0.0288
			(0.0246)		(0.0249)
Live in multistory building				-0.0162	-0.00927
				(0.0505)	(0.0510)
Observations	2123	2127	2127	2127	2123

Notes: Table shows regression estimates and clustered standard errors (in parentheses) for regression of a binary indicator of whether the mother-child pair was selected for inclusion as a target child (Panels A and B) or spillover child (C and D) in the sample on caste, household size and indicators of wealth. The sample is all mother-child pairs recorded in the census where the child was in the eligible age range. Panels B and D included village-level fixed effects.

Table A2: Spillover versus Target Mothers

	Target Mothers	Spillover Mothers	p-value
Male child	0.51 (0.50)	0.50 (0.50)	0.757
Age (years)	25.38 (4.37)	25.34 (4.42)	0.838
Age of child (months)	11.09 (2.70)	10.11 (6.41)	0.000
Years of education	7.34 (3.49)	7.46 (3.53)	0.428
Toilet ownership	0.47 (0.50)	0.47 (0.50)	0.932
Wealth index	–0.02 (0.92)	0.03 (0.92)	0.242
Raven progressive matrix IRT score	0.00 (0.86)	0.01 (0.84)	0.844
Labor force participation	0.06 (0.24)	0.06 (0.24)	0.845
SC/ST/OBC	0.62 (0.49)	0.61 (0.49)	0.592

Notes: Means (SDs) for selected characteristics of target and spillover mothers. p-value is for the t-test of means equality.

Table A3: Network Size by Wave in the Control Group (Small villages only)

	Mean	SD
Wave 1	1.81	1.86
Wave 2	1.43	1.51
Wave 3	2.60	1.88
Wave 4	2.94	2.00

Notes: Mean and standard deviation of the number of outward connections by wave in the control group for villages where there were 8 or fewer eligible children recorded in the census and thus where all eligible children were targeted for the sample.

Table A4: Correlates of Outward Network Size at Wave 1 (Small Villages Only)

	Panel A: Number of Outward Connections					
	(1)	(2)	(3)	(4)	(5)	(6)
Wealth index	-0.306**			-0.118	-0.0858	0.0515
	(0.124)			(0.109)	(0.106)	(0.117)
Scheduled caste or tribe + OBC		0.840***		0.752***	0.758***	0.658***
		(0.178)		(0.180)	(0.176)	(0.193)
Age of eldest child (years)			0.0878***	0.0763***	0.0931**	0.0795**
			(0.0246)	(0.0222)	(0.0385)	(0.0378)
Mother's age (years)					0.00549	0.0133
					(0.0241)	(0.0241)
Dist. from sample center (km)					-1.864***	-1.741***
					(0.560)	(0.524)
Number of children					-0.111	-0.105
					(0.0981)	(0.101)
Household owns toilet						-0.528**
						(0.205)
Mother in labor force						0.435
						(0.354)
Constant	1.748***	1.239***	1.468***	0.989***	1.352**	1.435**
	(0.138)	(0.141)	(0.146)	(0.149)	(0.573)	(0.578)
Observations	609	607	609	607	606	606
	Panel B: Number of Outward Connections weighted by Connectedness					
	(1)	(2)	(3)	(4)	(5)	(6)
Wealth index	-0.246***			-0.140**	-0.124*	-0.0351
	(0.0727)			(0.0668)	(0.0668)	(0.0712)
Scheduled caste or tribe + OBC		0.542***		0.454***	0.457***	0.393***
		(0.109)		(0.120)	(0.118)	(0.122)
Age of eldest child (years)			0.0511***	0.0395***	0.0486**	0.0393*
			(0.0147)	(0.0130)	(0.0223)	(0.0223)
Mother's age (years)					0.00526	0.0104
					(0.0128)	(0.0129)
Dist. from sample center (km)					-0.973***	-0.889***
					(0.355)	(0.326)
Number of children					-0.0686	-0.0659
					(0.0584)	(0.0617)
Household owns toilet						-0.339***
						(0.115)
Mother in labor force						0.352
						(0.252)
Constant	0.875***	0.555***	0.723***	0.440***	0.587*	0.636**
	(0.0798)	(0.0790)	(0.0849)	(0.0878)	(0.299)	(0.289)
Observations	609	607	609	607	606	606

Notes: Table replicates Table 3 but including only observations from villages with eight or fewer eligible target children identified in the census, i.e. where all children in the eligible age range were targeted for inclusion. Table shows regression coefficients and standard errors from regressing the number of outward connections an individual has (Panel A) and that number weighted by how well they know each connection (Panel B) on wealth, caste/tribe, toilet ownership, age and labor force participation. * p<0.1, ** p<0.05, *** p<0.001.

Table A5: Correlates of Network Size at Wave 1 (Self-Reported Total Connections)

	Number of Outward Connections					
	(1)	(2)	(3)	(4)	(5)	(6)
Wealth index	-1.043***			-0.573***	-0.479**	0.0375
	(0.204)			(0.198)	(0.199)	(0.216)
Scheduled caste or tribe + OBC		2.589***		2.235***	2.185***	1.868***
		(0.384)		(0.378)	(0.365)	(0.340)
Age of eldest child (years)			0.329***	0.280***	0.125	0.119
			(0.0447)	(0.0451)	(0.0840)	(0.0839)
Mother's age (years)					-0.00369	0.00193
					(0.0493)	(0.0497)
Dist. from sample center (km)					-5.921***	-5.694***
					(0.931)	(0.880)
Number of children					0.766**	0.718**
					(0.350)	(0.344)
Household owns toilet						-1.884***
						(0.367)
Mother in labor force						1.351**
						(0.639)
Constant	6.336***	4.730***	5.174***	3.959***	4.693***	5.596***
	(0.269)	(0.201)	(0.280)	(0.245)	(1.095)	(1.123)
Observations	2170	2161	2161	2152	2145	2145

Notes: Table replicates Table 3 but uses total self-reported connections (as opposed to total connections from the set of nearby mothers asked about) Table shows regression coefficients and standard errors from regressing the number of outward connections an individual has on wealth, caste/tribe, toilet ownership, age and labor force participation. * p<0.1, ** p<0.05, *** p<0.001.

Table A6: Correlates of Women's Mobile Phone Use

	Number of times have had a mobile or online interactions (including through text messaging, whatsapp, facebook, etc) with someone outside your immediate household in the past week.				
	(1)	(2)	(3)	(4)	(5)
Wealth index	0.0839			0.0699	0.0794
	(0.0587)			(0.0642)	(0.0658)
Scheduled caste or tribe + OBC		-0.0243		0.00735	-0.000904
		(0.0889)		(0.0979)	(0.0975)
Age of eldest child (years)			-0.0267***	-0.0231***	-0.0276
			(0.00710)	(0.00774)	(0.0187)
Mother's age (years)					0.00121
					(0.0102)
Dist. from sample center (km)					0.00463
					(0.139)
Number of children					0.0257
					(0.0825)
Constant	0.316***	0.328***	0.405***	0.391***	0.329
	(0.0459)	(0.0608)	(0.0643)	(0.0824)	(0.251)
Observations	366	365	365	364	362

Notes: Table shows regression coefficients and standard errors from regressing the number of times women report interactions on a mobile phone/online on wealth, caste/tribe, age. * p<0.1, ** p<0.05, *** p<0.001.

Table A7: Correlation between Network Size, Empowerment and Symptoms of Depression

	Empowerment		Symptoms of Depression (CES-D score)	
	(1)	(2)	(3)	(4)
Number of connections	0.183***	0.127***	-0.0776**	-0.101***
	(0.0132)	(0.0152)	(0.0331)	(0.0349)
Birth order		0.0826***		0.198***
		(0.0257)		(0.0660)
Distance from Village center		-0.138		0.674**
		(0.0979)		(0.310)
Wealth index		-0.173***		-0.307***
		(0.0260)		(0.0678)
Scheduled caste or tribe + OBC		0.0906*		-0.176
		(0.0463)		(0.115)
HH owns a toilet		-0.177***		0.0704
		(0.0474)		(0.129)
Mother age (years)		0.0208***		0.0286**
		(0.00461)		(0.0131)
Works for pay (baseline)		0.210**		-0.00846
		(0.0918)		(0.236)
Observations	2137	2137	2137	2137
Adjusted R^2	0.094	0.180	0.002	0.032
Mean of Dep. Var.	0.00	0.00	3.58	3.58
SD of Dep. Var.	1.00	1.00	2.41	2.41

Notes: Table correlates measures of mothers' empowerment (columns 1 and 2) and mothers' symptoms of depression as measured by the CES-D score (columns 3 and 4) with the number of outward connections. We report correlations both without (columns 1 and 3) and with (columns 2 and 4) controlling for covariates. * p<0.1, ** p<0.05, *** p<0.001.

Figure A1: Study districts within Odisha

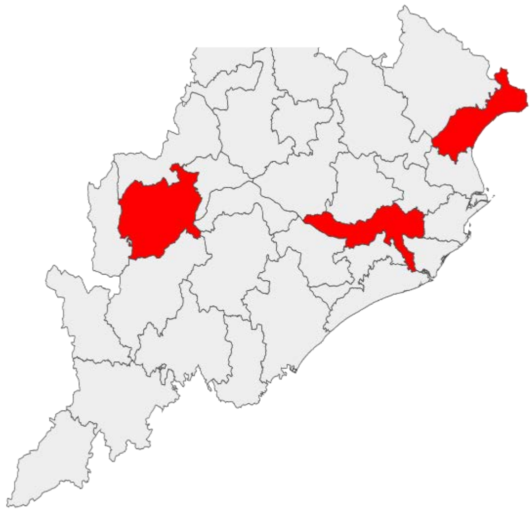


Figure A2: Distribution of Self-Reported Connections

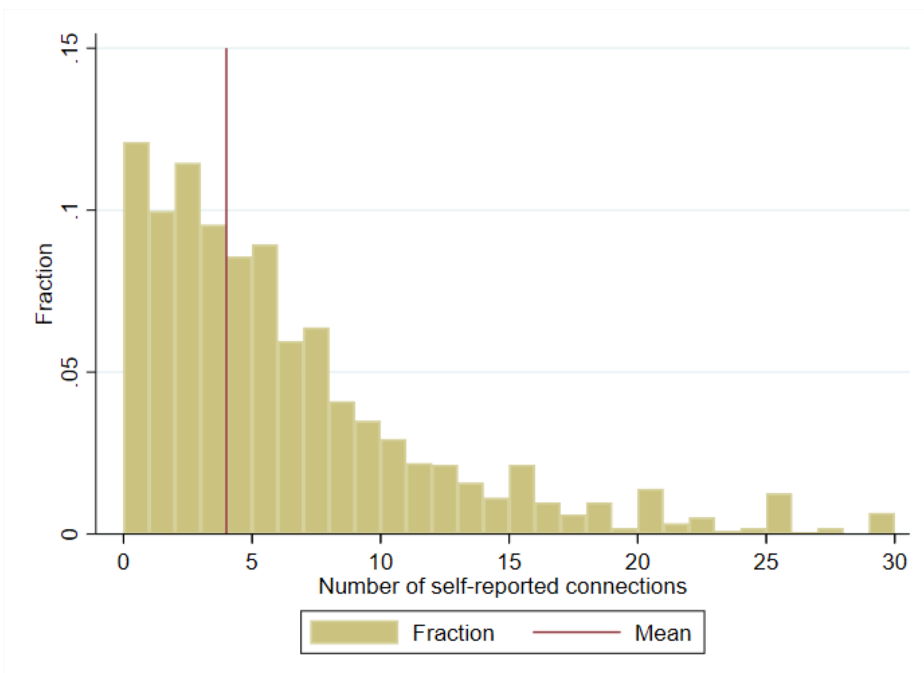
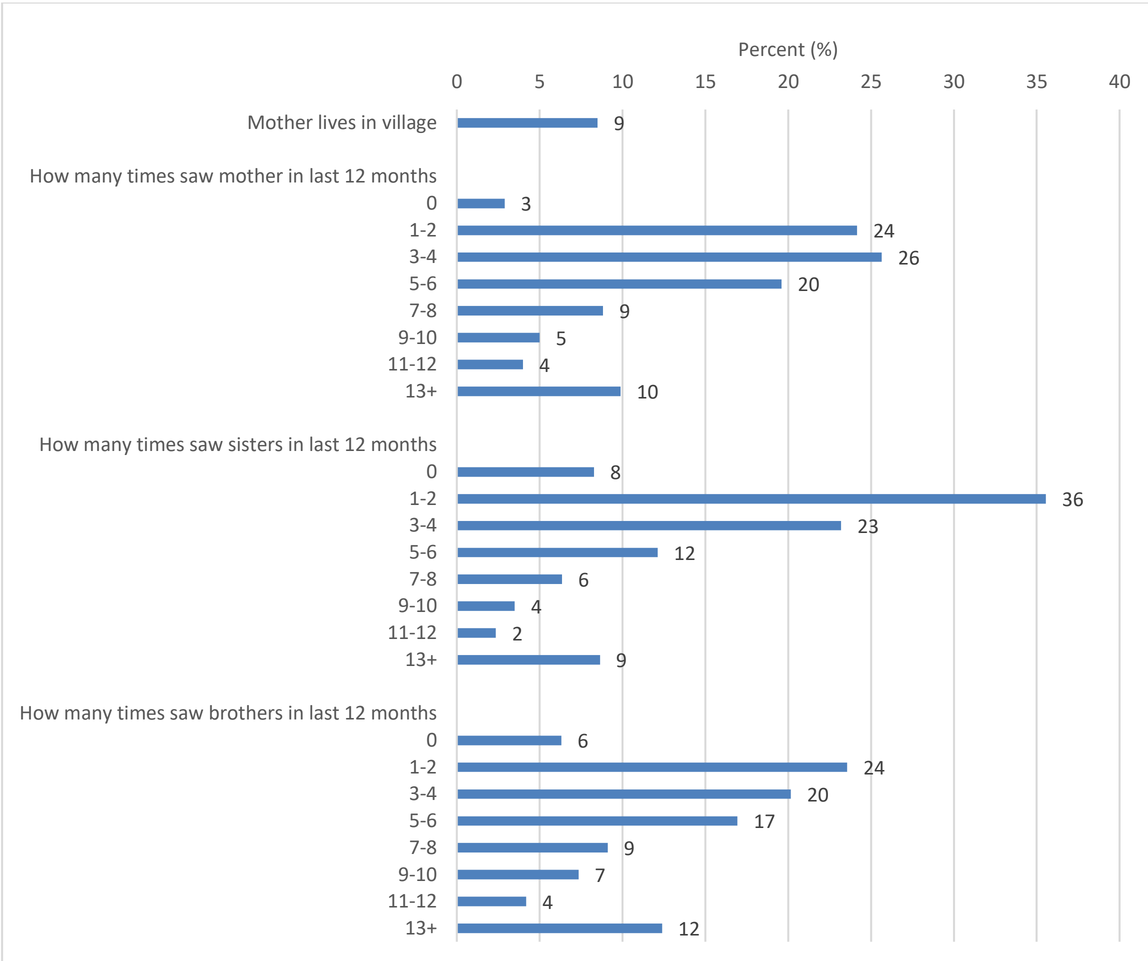
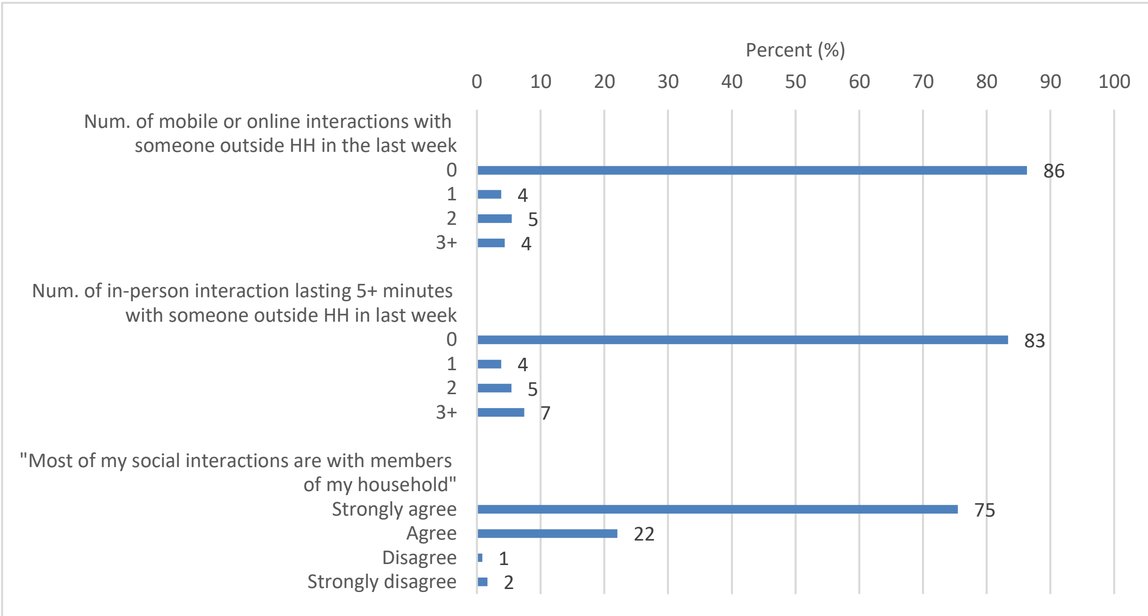


Figure A3: Social Connections with those other than other young mothers from the village

(a) Connections with Natal Family (measured at wave 1)



(b) Virtual and in-person connections outside the household (additional data, after wave 4)



Appendix B: Intensity of Relationship Questions

1. How long have you known [Name]?
2. How many years/months/days ago was the last time you spoke to [Name]?
3. How many times have you visited [Name]'s house in the past 15 days?
4. Do you talk about recipes with [Name]?
5. Do you wash clothes or fetch water with [Name]?
6. Do you talk about your young children (for example their health, nutrition, parenting techniques or play) with [Name]?
7. If you wanted to talk to someone about something personal or private (for instance, if you had something on your mind that was worrying you or making you feel upset) would you talk to [Name]?
8. Would [Name] lend you food, kerosene or money if you needed it?
9. Do you often have fun and relax with [Name]?

Appendix C. Estimating Out-of-Sample Connections

A limitation of this exercise is that our data only contain connections between the mothers selected to be a part of the study. To estimate the average number of other mothers with children of a similar age that respondents know in the *whole* village, we perform an out-of-sample prediction exercise. For the sample for whom we have detailed network information, we estimate the probability that a connection exists between any two mothers (allowing the probability to vary with the children's ages, the mothers' ages, the mothers' castes and the mothers' geographic proximity to one another).[1] We then use these probabilities to predict the likelihood that our respondents know each of the other mothers in the village identified in the census with similarly-aged children but whom we did not ask the respondent about. We then sum these probabilities to obtain an estimate of the total number of connections that mothers have, including those we did not directly enquire about.

In particular, we collected village-level censuses of all mothers with children under the age of 2 years before the study began (August 2015). In these data we collected information on GPS location, caste, and the gender and age of the child. Assuming that the relationships we observe in the village hold for non-sampled mothers, we can use these data to estimate the total size of mothers' networks.

We proceed in two steps: (i) estimate a probit model of the number of connections using the characteristics observed in the census data and (ii) extrapolate from this for unknown connections, calculating the expected number of connections. Consider a village with N eligible mothers. Of those, $l \in L$ are in the sample and $k \in K$ are not. In step (i) we estimate a model of the following form for all mothers *l*, where $y_{ijv} = 1$ if mother *i* reports knowing mother *j*.

$$y_{ijv}^* = \alpha_0 + \beta_1 \boldsymbol{X}_i + \beta_2 \boldsymbol{X}_j + \beta_3 \boldsymbol{X}_i * \boldsymbol{X}_j + dist_{ij} + \gamma_v \varepsilon_{ij}$$

$$y_{ijv} = \mathbf{1}\left[y_{ij}^* \geq 0\right] \text{ and } \varepsilon_{ij} \sim N(0,1)$$

where **X** contains age of mother, age of child and whether the mother was high or low caste, and the variable $dist_{ij}$ is the distance in meters between mother *i* and mother *j*. In step (ii) we use the parameter estimates from the above equation to estimate the probability of mother *i* knowing any out-of-sample mother *k* as

$$\Pr(y_{ikv} = 1 | i, k, v) = \Phi\left(\alpha_0 + \beta_1 \boldsymbol{X}_i + \beta_2 \boldsymbol{X}_j + \beta_3 \boldsymbol{X}_i * \boldsymbol{X}_j + dist_{ik} + \gamma_v \right)$$

The total expected number of connections for mother *i* is then given by

$$\sum_j y_{ijv} + \sum_k \Pr(y_{ikv} = 1 | i, k, v)$$

We estimate that each mother has an average of 3.2 connections to other mothers of similarly-aged

[1] Since we did not have the same socioeconomic characteristics of non-sample mothers, we were unable to include socioeconomic characteristics as predictors in this exercise.

children in the village. In wave 1 we additionally ask respondents how many other mothers they know with children between 0 and 24 months inside the village. The results show peer groups with a mean of 6.3 and median size of 4 (see Figure A2 in Appendix A).