MOTHERS' SOCIAL NETWORKS AND SOCIOECONOMIC GRADIENTS OF ISOLATION

By

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

November 2020

COWLES FOUNDATION DISCUSSION PAPER NO. 2261

COWLES FOUNDATION FOR RESEARCH IN ECONOMICS
YALE UNIVERSITY
Box 208281
New Haven, Connecticut 06520-8281

http://cowles.yale.edu/

Mothers' Social Networks and Socioeconomic Gradients of Isolation

Alison Andrew (Institute for Fiscal Studies (IFS), University College London (UCL))
alison_a@ifs.org.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 (Ambedkar University)
monimalika.day@gmail.com

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

Costas Meghir (Yale University)
c.meghir@yale.edu

Angus Phimister (IFS, UCL)
angus_p@ifs.org.uk

Acknowledgements: 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 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 programme (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 the IFS (grant reference ES/M010147/1). We bear sole responsibility for all errors and opinions expressed.

Abstract

Social connections are fundamental to human wellbeing. This paper examines the social networks of young married women in rural Odisha, India.. This is a group, for whom highly-gendered norms around marriage, mobility, and work are likely to shape opportunities to form and maintain meaningful ties with other women. We track the social networks of 2,170 mothers over four years, and find a high degree of isolation. Wealthier women and women more-advantaged castes have smaller social networks than their less-advantaged peers. These gradients are primarily driven by the fact that more-advantaged women are less likely to know other women within their same socioeconomic group than are less-advantaged women are. There exists strong homophily by socioeconomic status that is symmetric across socioeconomic groups. Mediation analysis shows that SES differences in social isolation are strongly associated to caste, ownership of toilets and distance. Further research should investigate the formation and role of female networks.

Highlights

1. We explore a novel primary dataset of social networks between young mothers in rural Odisha, India, and show quantitively a high degree of social isolation.

2. Mothers of higher socio-economic status (SES) - in terms of caste and wealth – have significantly fewer connections than mothers of lower SES.

3. Caste differences in network size are a result of both village composition and differences in within-caste connectedness.

4. Differences in network size by wealth are driven by differences in within group connectedness.

5. Mediation analysis suggests that toilet ownership and labor force participation are important correlates of network size.

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 connectedness of 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 broader communities (Sanyal, 2009; Prillaman, 2017).[3] Therefore, isolation may be part of a vicious cycle that entrenches the disadvantages that women face in terms of political representation, and voice and involvement in decision-making in their households and communities. Likewise, since social networks are important transmitters of information (Kohler, Behrman and Watkins, 2000, 2007; Kohler and Bühler, 2001; Behrman, Kohler and Watkins, 2002; Beaman *et al.*, 2018; BenYishay and Mobarak, 2019) isolation may limit women’s knowledge, particularly of heavily gendered subjects, such as sexual and reproductive health or child development, that are not typically discussed within married couples or within male social networks (Mason and Smith, 2000). And finally, networks can be an important means of access to capital, markets and insurance (e.g. Barnhardt et al., 2017; Fafchamps & Lund, 2003; Feigenberg et al., 2013; Field et al., 2016) 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 contexts with strict gender norms. How connected, or isolated, are women on average and how does this

[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; Sawyer, Ayers and Smith, 2010)
[2] For example, Ambrus el al (2018), Ambrus et al (2020) and Attanasio and Krutikova (2020) analyse the role of networks in providing insurance.
[3] A recent review by Dìaz-Martin *et al.*, (2020) found positive effects of women’s decision making in roughly half the evaluations of women’s groups they studied.

change over time? What is the depth of the social connections that women do have? How does connectedness or isolation vary by women's socio-economic status? What drives socio-economic gradients in isolation?

In this paper, we answer these questions by documenting the social ties of 2170 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, 2019). 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 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 0 and 30 months in each village whether, and how well, they knew each of the other 11 in the village. On average, we asked one third (quasi-randomly selected) 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 did not report knowing 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 of 3.2. The panel nature of the data allows us to document high persistence of this isolation over the four years that our annual surveys cover. The high level of isolation we document is consistent with qualitative (Sanyal, 2009; Crivello *et al.*, 2018) and quantitative (Kandpal and Baylis, 2019; Anukriti *et al.*, 2020) evidence from rural India.

We next describe the socioeconomic gradient of isolation and examine its correlates. We might expect social and economic characteristics, like 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 are less valuable sources of information (Magnan *et al.*, 2015). More advantaged households may also be able to "afford" to adhere to more restrictive gendered roles for women. 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). This may lead to women in more advantaged households to face 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 like 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. 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 socio-economic 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. To the best of our knowledge, this is the first quantitative study to examine the socio-economic gradient of women's isolation.

We next analyze the drivers of the SES gradients we observe. We develop a decomposition method and show that the gradients are composed of three parts: first, differences by SES in women's propensities to have social connections *within* their SES group; second, SES-differences in women's propensities to have ties *across* SES groups (i.e. differences in the degree of homophily); and third, the SES-composition of villages coupled with the initial degree of homophily. Our data suggest that the first 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.

Finally, we examine the 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 mediate 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, 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.

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. Odisha is poorer and more rural than India 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). 16.5% of the population belong to a scheduled caste (SC) which is greater than the national proportion (16.2%) while, at 22.5%, the proportion belonging to a scheduled tribe (ST) is far greater than the national average (8.2%).[4]

This study uses data gathered as part of a randomized control trial (RCT) of an early childhood intervention in 192 villages across these three blocks (districts) of Odisha: Balangir (in Balangir district), Soro (Balasore), and Salepur (Cuttack) (Figure A1 in appendix A). The study sample consists of a panel of 2,170 mothers with infants between the ages of 2 and 20 months at the time of the first survey (wave 1) in November 2015, with an average of 11 study participants per village.[5]

[4] 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 (Chakravarti, 2018; Rao, 2009; Kapadia, 1999).

[5] 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

We identified the participants who met the study's eligibility criteria, which were based on the requirements of the early childhood intervention, before wave 1 fieldwork; see (Attanasio *et al.*, 2016) for full details of eligibility criteria. Pregnant women and mothers with children under the age of 2 were identified through a census of each village in the summer of 2015. The sample was split into two groups: target children, who met the age criteria for the intervention (between 7 and 14 months at the start of the intervention), and spillover children, who were aged just above or just below the age criteria. In villages where there were less than 8 eligible target children (roughly 38% of villages), all were selected. Villages with more than 8 eligible target children had a median of 15 eligible children. In these villages, one child was selected at random, and that child's 7 nearest neighbors were then targeted for enrollment. All surplus children (children in the eligible age range who lived further than first eight children form 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. 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 4 spillover children per village from the ordered list were enrolled in the sample[6]. This generated an overall sample for wave 1 (target and spillover children combined) of 2,170 children with ages between 2 and 20 months, and between 10-12 households in each village (1,401 target, 769 spillover). Since households of target and spillover children are observationally equivalent on key margins, we make no distinction between the two groups and use only the mother-, or primary caregiver-level surveys.

Mothers were re-interviewed in three further surveys in November 2016 (wave 2), November 2017 (wave 3) and March 2019 (wave 4). 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 and an average household per capita income per day

the networks module; where this was not the case, we replaced her with the primary caregiver.

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

of $0.84 (2019 USD); 93% live below the US$1.90 per day international poverty line. Around 65% of 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 2 years 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 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
Distance from other sample mothers within village (m)	237.4	213.38	2,168

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.[7] 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 if they could borrow food from this person.[8] 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.[9]

It is important to consider the implications of our sampling strategy on our network data. Our social networks data are incomplete in two senses. First, from our census data, we estimate that our study captures around 1 in 3 mothers with children between the ages of 0 and 30 months in each village. As elaborated in section 3.1, we 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. Given the location-based nature of our sampling, our mothers are on average 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. Second, our network is incomplete in the sense that we only analyze connections to other mothers of young children. While this is a subset of the overall social networks of these mothers, women of similar ages and circumstances represent an important group, which, in many contexts, is a primary 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.

A final implication of our sampling strategy, in which spillover children were selected from narrower age

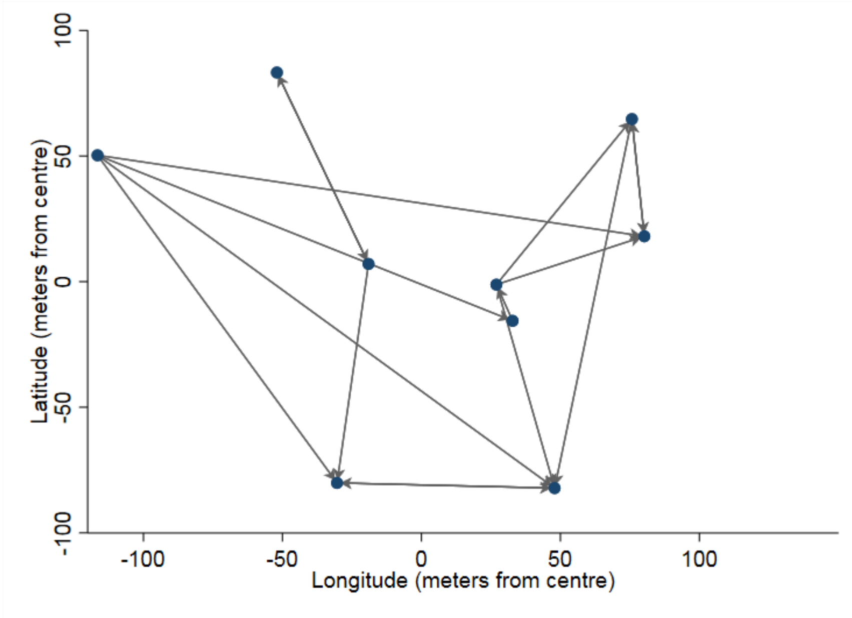
[7] In wave 1, this list was populated with the 12 mothers targeted for inclusion in the study on the basis of the census. Since sometimes not all these mothers were actually enrolled (due to refusals, incorrect information about the children's ages having been recorded during the census, or the interviewers being unable to relocate the house), in some cases in wave 1, participants were asked about other mothers in their village who were not subsequently enrolled in the study. In these cases, we include social connections with mothers who were never in fact enrolled in the total number of social connections mothers have, but not in the dyad-level analysis since for the dyad-level analysis we require characteristics of the mothers which are not available for mothers who did not end up participating. In waves 2-4 the list of actual study participants in the previous wave was used.
[8] For full module see appendix B.
[9] We primarily used GPS measurements taken at the census carried out at the start of the study. However, in cases where these co-ordinates suggested that a respondent lived move than one kilometer from their nearest neighbour, we manually compared these measures to those taken at later rounds and took the measure that appeared most reasonable.

bins than target children, is that mothers of spillover children live slightly further away from the central mother in the sample than the mothers of other target children. However, since the magnitude of this difference is small (216m vs. 276m), and since spillover mothers are similar in all other respects to mothers of target children (Table A1 in Appendix A), we do not expect this to have a substantial impact on our results.

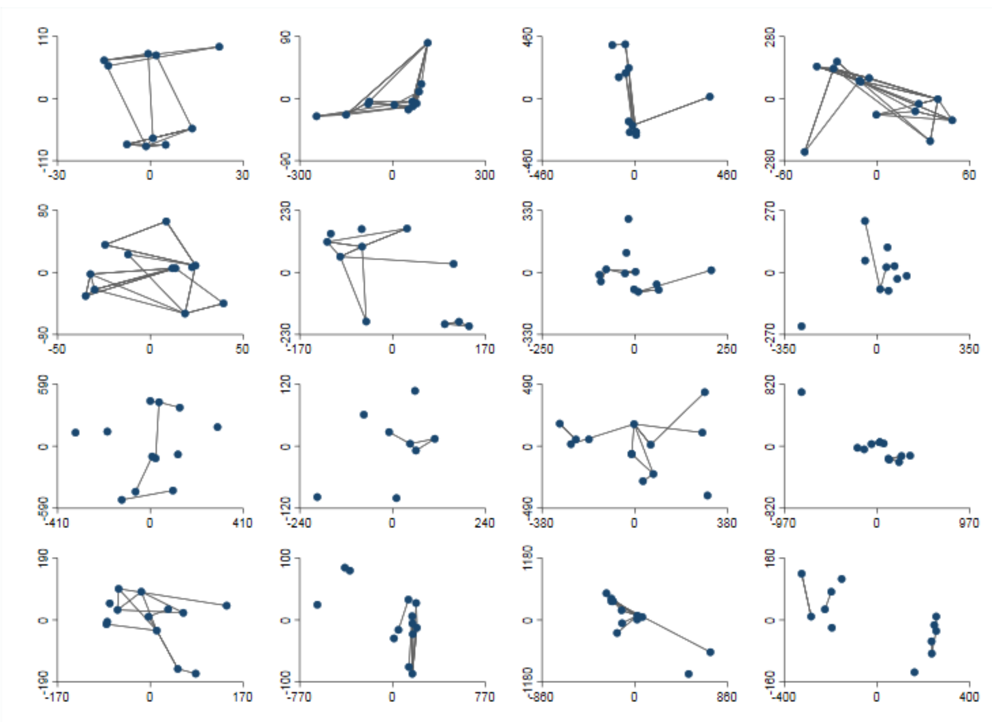
Figure 2 shows examples of village networks in wave 1. Figure 2a shows an example sample village where each dot represents a respondent plotted, on the basis of their geographic position in the village, on a Cartesian coordinate system with the village center at (0,0), and each arrow represents a connection 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 5 target children and 4 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. Figure 2a 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 likely that some respondents knew who the other mother was or had a brief acquaintance with her but that the connection wasn't reciprocated. This might mean that 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 2b shows 16 other randomly-selected villages displayed in the same way, where lines between respondents indicate any connection between the two. Figure 2b 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 2: 17 Randomly-Selected Network Graphs from Wave 1



(a) Directed Network graph



(b) 16 Randomly-Selected Undirected Network Graphs

Notes; Data from all villages in wave 1 of data collection.

3 Isolation and its Socioeconomic Gradient

3.1 Isolation

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 3 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 village, in the control group the average number of connections reported is 1.21 the median is 1, and mode is 0 (reported by 39% of sampled women). This number increases over time but remains relatively small, with a mean network size of 1.99 by Wave 4 (Table 2).[10]

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 kids of a similar age that respondents know in the *whole* village we perform an out-of-sample prediction exercise. For the sample 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).[11] 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 similar-age kids but whom we did not ask the respondent about. We then sum these probabilities to get 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 have an average of 3.2 connections to other mothers of similarly aged 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 median of 4 (Figure A2 in Appendix A). Considering the proximity of these households and the small communities in which they reside, these are a strikingly small peer groups.

[10] As discussed in footnote 4, in waves 2-4 we asked about a different set of mothers. This may explain the reduction in total connections from wave 1 to 2.
[11] Since we didn't collect socio-economic characteristics of non-sample mothers, we were unable to include socio-economic characteristics as predictors in this exercise.

Figure 3: Distribution of Connections in Wave 1

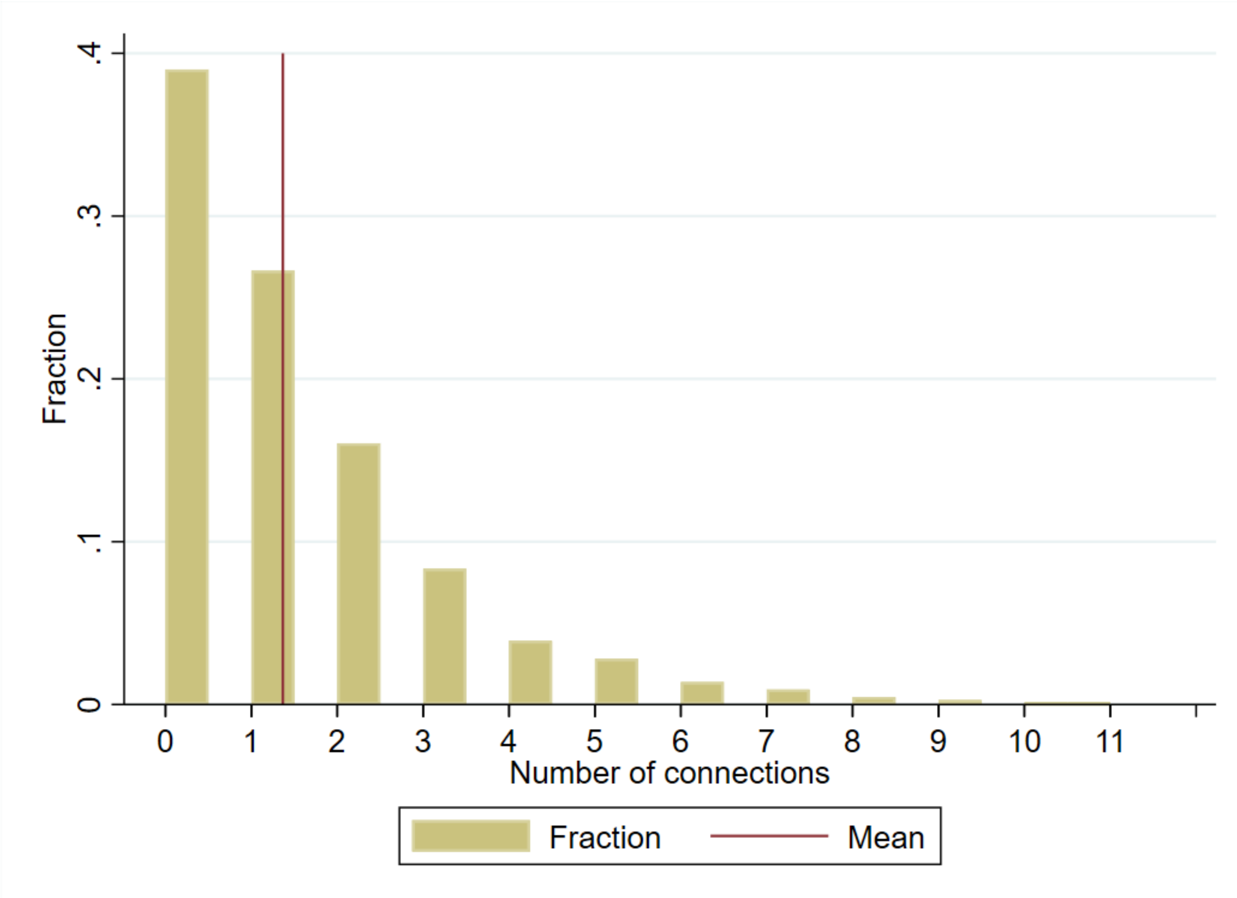


Table 2: Network Size by Wave in the Control Group

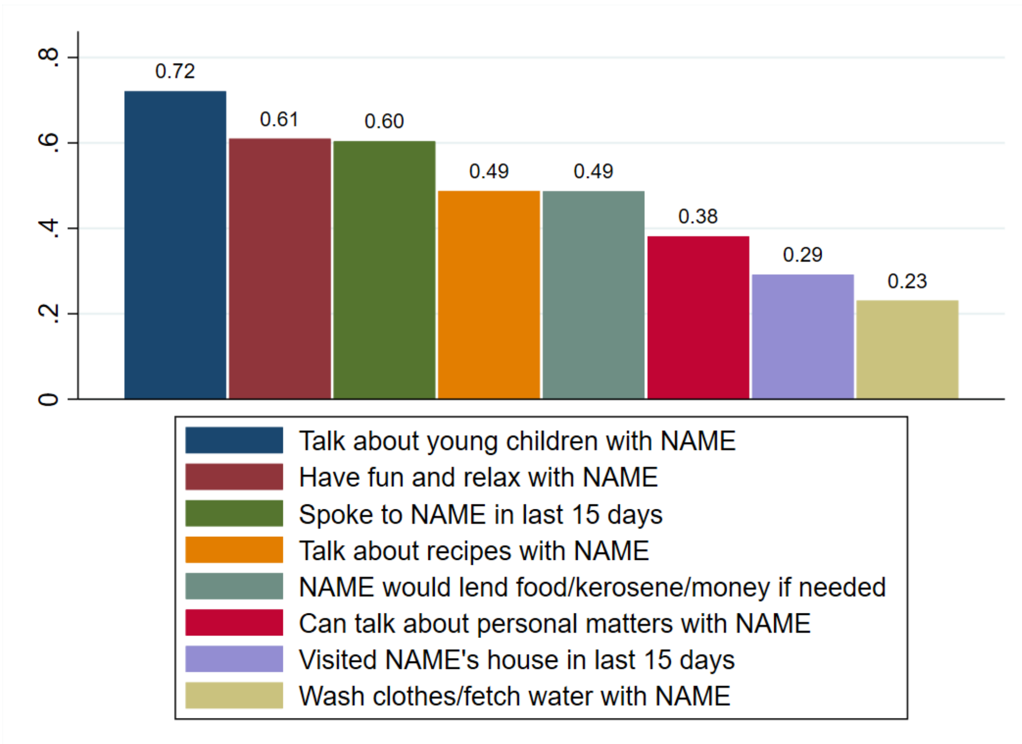
	Mean	SD
Wave 1	1.21	1.54
Wave 2	0.93	1.26
Wave 3	1.63	1.70
Wave 4	1.99	1.95

3.2 Strength of Connections

Figure 4 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 (72%) shared activity was talking about young children. 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, respondents reported being able to talk about personal matters.

Figure 4: Strength of Social Ties (Wave 1)



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 doesn't 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 4). 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,…8},* listed in Figure 4 regarding other mother *j* as the following function of the underlying connectedness of mother *i* to mother *j*, θ_{ij}:

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

Conditional on a connection existing between i and j at all, we assume that the 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 we can also define a level for this connectedness index for connections that don't 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 are 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.

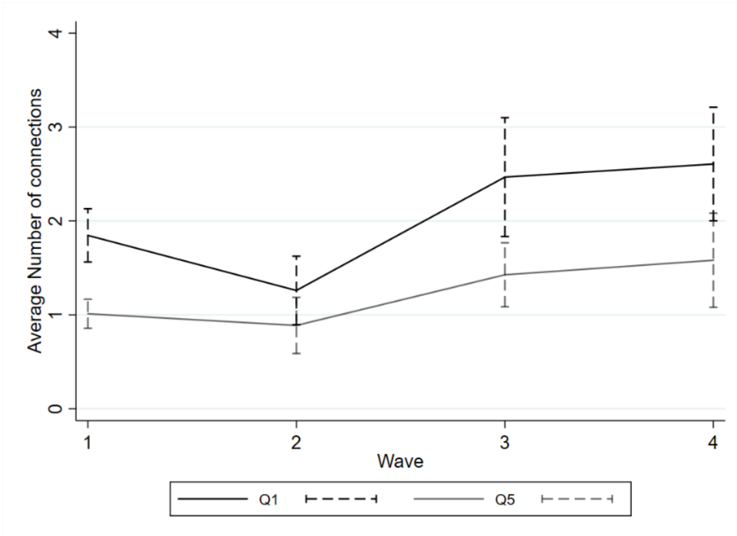
3.3 Socio-Economic Gradient of Connections

We next consider how the size of mothers' networks vary by socioeconomic status (SES), specifically by wealth, and caste and tribe.[12] Figures 5a and 5b 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 (ST/SC/OBC) report *more* connections than their wealthier peers and peers from more advantaged castes or tribes (non-ST/SC/OBC). At wave 1 this amounted to an average of 0.90 fewer connections for mothers in the highest wealth quintile

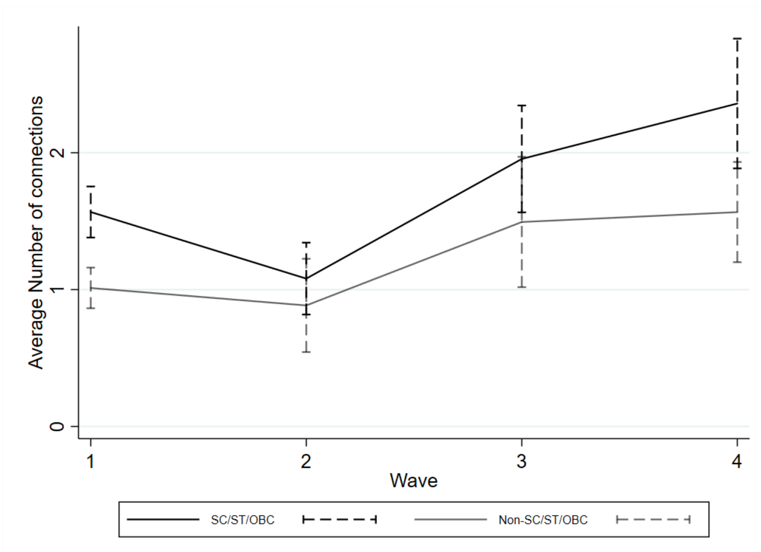
[12] An individual's wealth score is calculated using a principle component analysis of assets in wave 1. Across all groups wealth is low, with an average per capita daily income of $0.55 USD in the lowest wealth quintile compared to $1.39 USD in the highest.

relative to the lowest, and of 0.56 fewer connections for non-ST/SC/OBC women relative to ST/SC/OBC women. Given the median network size in wave 1 is one, these differences are substantial.

Figure 5: Socioeconomic gradient over time for controls



(a) **Network size and Wealth Quintile**



(b) Network size and Caste

Note: Averages include the whole sample in wave 1, and only control villages thereafter.

Both figures show an increase in network size over time for all groups, 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. In columns 1-4 the outcome variable is total outward network size, and in columns 5-8 we weight each connection by their estimated "connectedness", the index between 0 and 1 defined in the previous section. 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. Column 3 shows that each dimension (caste/tribe and wealth) is statistically significant even when both are included in the regression suggesting that both are important predictors of network size. Column 4 shows that this effect of caste/tribe persists even when we control for other covariates. Conditional on other covariates, the wealth index alone 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 village center.

Mothers' ages are a strong positive predictor of network size - likely proxying for how long mothers have been in their current villages of residence - mothers who have been around longer have had more opportunities to expand their networks. Interestingly network size is also strongly predicted by labor force participation, indicative of working mothers being more mobile around their villages. Likewise toilet ownership, even conditional on wealth, is associated with 0.53 fewer connections, likely due to women who own toilets not travelling with other mothers in their villages to defecate.

Moving to columns 5 through 8 we see that these associations persist once we weight the number of connections by how well mothers know each. Wealth conditional on other covariates is significantly negatively correlated with having a lower, weighted, number of connections suggesting that after conditioning on other factors, higher wealth may be particularly associated with knowing connections less well.

Table 3. Correlates of outward network size at wave one

	Number of Outward Connections			
Wealth index	-0.308***		-0.259***	-0.0654
	(0.0546)		(0.053)	(0.0531)
Scheduled caste or tribe + OBC		0.555***	0.453***	0.359***
		(0.103)	(0.099)	(0.0912)
Distance from sample center (km)				-2.169***
				(0.249)
HH owns a toilet				-0.533***
				(0.0972)
Mother's age (years)				0.0371***
				(0.00968)
Mother in labor force				0.735***
				(0.215)
Constant	1.354***	1.012***	1.075***	0.902***
	(0.0725)	(0.0752)	(0.790)	(0.250)
Observations	2153	2144	2144	2139
Adjusted R-squared	0.028	0.026	0.045	0.148
	Number of Outward Connections weighted by Connectedness			
Wealth index	-0.204***		-0.177***	-0.0676**
	(0.0309)		(0.0307)	(0.0303)
Scheduled caste or tribe + OBC		0.316***	0.251***	0.199***
		(0.0615)	(0.0606)	(0.0545)
Distance from sample center (km)				-1.180***
				(0.153)
HH owns a toilet				-0.297***
				(0.0555)
Mother's age (years)				0.0201***
				(0.00571)
Mother in labor force				0.520***
				(0.145)
Constant	0.664***	0.471***	0.509***	0.414***
	(0.0412)	(0.0449)	(0.0476)	(0.152)
Observations	2153	2160	2144	2139
Adjusted R-squared	0.035	0.024	0.051	0.148

3.4 Who knows whom? Decomposing SES gradients

There are several potential drivers of these observed gradients for caste and wealth. High-SES women with young children might be less likely than low-SES women to know other women *within* their own SES group or might be less likely to know women from *outside* their SES group. For the negative caste/tribe gradient, it may also be the case that the caste/tribe composition of villages is such that non-SC/ST/OBC caste women systematically live in villages where they are in the minority.[13] Under homophily, this would lead to an aggregate difference in the total number of connections even if non-SC/ST/OBC and non- SC/ST/OBC women were as likely as each other to know women of their own and outside their caste/tribe groups. In this section, we decompose the average differences we see in reported network size into the portions driven by these different components.

Consider that there are higher and lower-SES mothers whom we denote, respectively, H and L. Let $\bar{T}_H$ be the sample average of the total number of other mothers that the high-SES sample mothers know in each village. Mechanically $\bar{T}_H$, is the weighted sum of the sample averages of the total number of other high-SES ($\bar{n}_{HH}$) and of low-SES sample mothers ($\bar{n}_{HL}$) living in villages where high-SES mothers live, each weighted by the in-sample probability that a high-SES mother reports knowing, respectively, another high-SES sample mother, $\hat{p}_{HH}$, or a low-SES sample mother, $\hat{p}_{HL}$:

$$\bar{T}_H = \hat{p}_{HH} * \bar{n}_{HH} + \hat{p}_{HL} * \bar{n}_{HL} \qquad (1)$$

Correspondingly, the sample average of the total number of other sample mothers that low-SES sample mothers report knowing, $\bar{T}_L$, is a function of the number of other low-SES ($\bar{n}_{LL}$)and of high-SES ($\bar{n}_{HL}$) sample mothers living in villages where low-SES women live, and of the in-sample probability of low-SES mothers reporting knowing these other low-SES mothers ($\hat{p}_{LL}$) and high-SES mothers ($\hat{p}_{LH}$):

$$\bar{T}_L = \hat{p}_{LL} * \bar{n}_{LL} + \hat{p}_{LH} * \bar{n}_{LH} \qquad (2)$$

[13] Note, that since we have defined high wealth and low wealth by the household being above or below the median value of an asset index, a similar village composition explanation is not relevant to the wealth gradient; if all sample villages contain the same number of sample women then, mechanically, high- and low-wealth women will live in villages with equal numbers of other mothers of their own and of the opposite wealth group. 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.

By taking the difference between equations (1) and (2), and by rearranging terms, we can decompose the *difference* in the number of connections that low- and high-SES mothers report:[14]

$$\bar{T}_L - \bar{T}_H = \hat{p}_{LL}(\bar{n}_{LL} - \bar{n}_{HH}) + \hat{p}_{LH}\,(\bar{n}_{LH} - \bar{n}_{HL})$$

$$+(\bar{n}_{HH})(\hat{p}_{LL} - \hat{\mathrm{p}}_{\mathrm{HH}})$$

$$+(\bar{n}_{HL})(\hat{p}_{LH} - \hat{p}_{HL}) \qquad (3)$$

The first line is the component of the SES-gradient that can be attributed to village composition *if* the degree of homophily were symmetric between the two groups, i.e., if $\hat{p}_{LL} = \hat{p}_{HH}$ and $\hat{p}_{LH} = \hat{p}_{HL}$. If this were the case, then equation (3) reduces to this first line and all the observed difference in network size must arise from differences in the composition of villages, i.e., under homophily, that L type women on average live in villages with a higher proportion of other L type women than H type do with other H types.

The second line is the component of the gradient that can be attributed to differences in the within-group connectedness of low- and high-SES sample mothers. For example, if village composition were identically symmetric – i.e., $\bar{n}_{HH} = \bar{n}_{LL}$ and $\bar{n}_{HL} = \bar{n}_{LH}$ – and the probability of knowing mothers from the opposite group were the same for low-SES as for high-SES mothers – i.e. $\hat{p}_{LH} = \hat{p}_{HL}$ – then differences in observed network size must come from differences between the probability of members of each group having connections within their group, i.e. differences between $\hat{p}_{LL}$ and $\hat{p}_{HH}$.

The third line is the component of the gradient due to across-group connectedness of low- and high-SES sample mothers. Namely, this is the component driven by differences in the rate at which low-SES sample mothers report knowing high-SES sample mothers and vice-versa.

With our detailed dyad-level data, we can estimate every term in equation 3 and thus we can provide a complete decomposition of the SES gradient into components driven by: (i) village composition, (ii) within-group connectedness, and (iii) across-group connectedness. Below, we discuss each in turn for all villages in wave 1 for those dyads for which we have complete information.

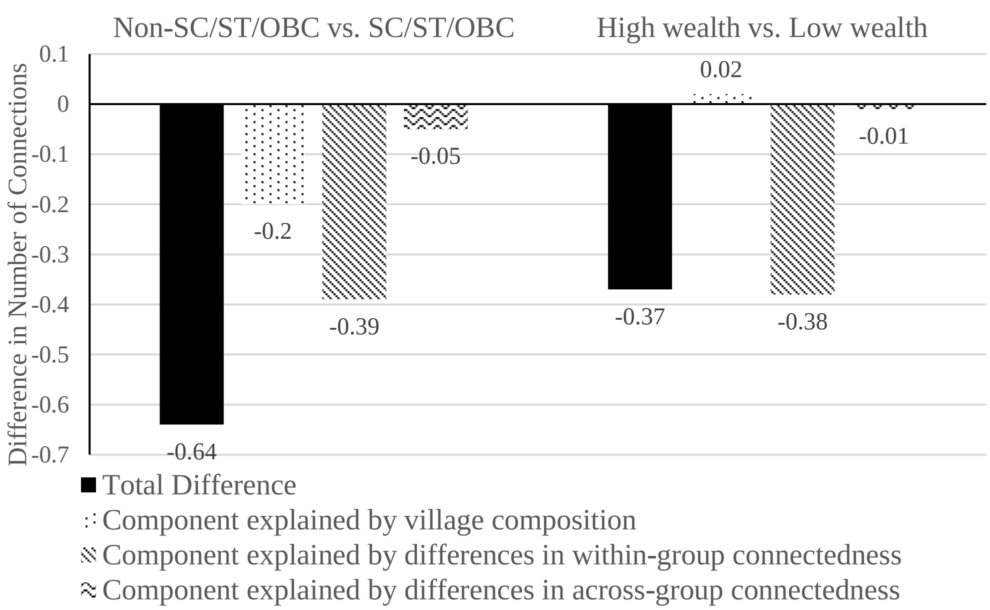
[14] The logic of this decomposition is similar to the Oaxaca-Blinder decomposition (Blinder, 1973; Oaxaca, 1973). For derivation see Appendix C.

(i) Village composition

We find 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. When we evaluate the first line of equation 3, we find that this village composition effect, under symmetric homophily, would predict that non-SC/ST/OBC women have 0.2 fewer connections than SC/ST/OBC women. In other words, village composition can explain 36% of the actual caste/tribe gradient of -0.55 connections (Figure 6).

As noted earlier, since our wealth grouping is simply defined as being above or below the medians on an asset index, the only reason why we would see village composition playing a role for wealth would be due to 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 6: Decomposition of SES gradients in network size



(ii) Within-group connectedness

The top two bars on Figure 7a plot the in-sample probability, and corresponding 95% confidence interval, that non-SC/ST/OBC mothers report knowing the other non-SC/ST/OBC mothers in their villages and that SC/ST/OBC sample mothers report knowing the other SC/ST/OBC sample mothers in their villages. 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%). The difference in the within-group probability of connections for non-SC/ST/OBC versus SC/ST/OBC can thus account for a difference of -0.39 in their number of connections (following line two of equation 3), or 61% of the overall observed difference (Figure 6).

The top two bars of Figure 7b plot the probabilities that high- versus low-wealth sample mothers report knowing the other sample mothers in their village of their same wealth group. 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%). Overall, this difference can account for a -0.38 difference in the total connections of high-wealth and low-wealth mothers (Figure 6), completely explaining the SES gradient for wealth.

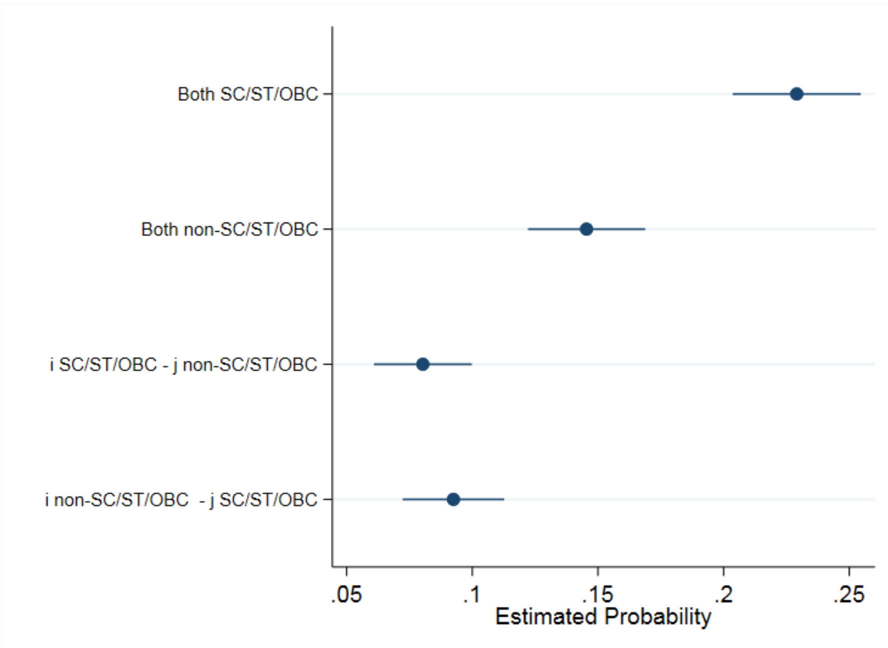
(iii) Across-group connectedness

The bottom two bars of Figures 7a and 7b plot the across-group connectedness of sample mothers by both wealth and caste/tribe. The probability of across-group connections is substantially lower than the probability of within-group connections. This is true along both the caste and the wealth dimensions, and for both higher- and lower-SES mothers. Our social networks thus exhibit substantial homophily.

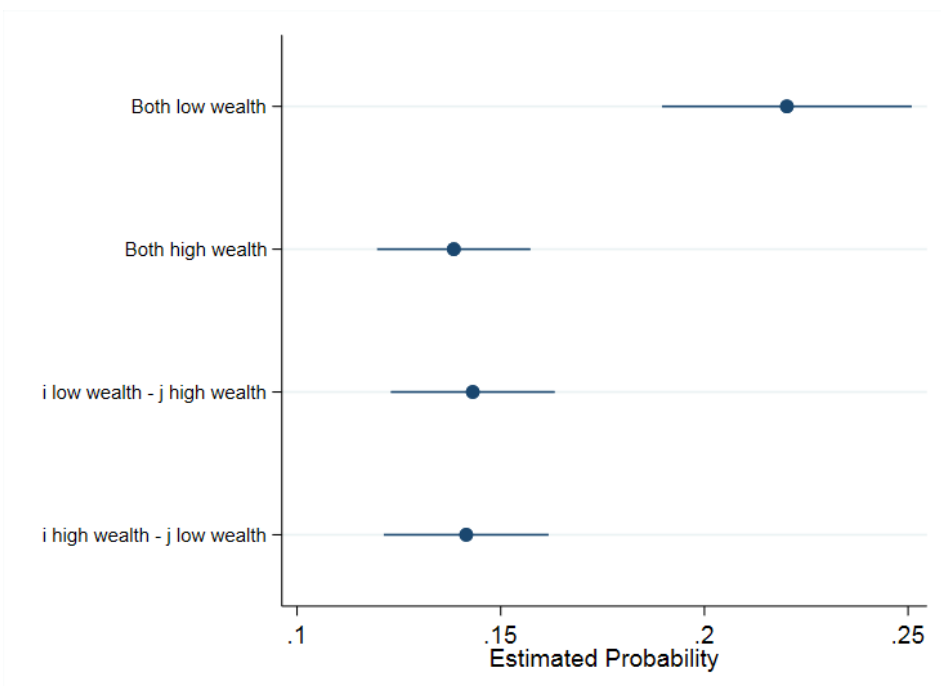
For neither caste/tribe nor wealth do we see differences in the probability of across-group connections by the mothers' SES. In other words, high-SES mothers are as likely to report knowing a randomly-chosen lower-SES mother in their village than vice versa. This implies that differences in the probabilities of across-group connections contribute little to the overall SES gradients (Figure 6).

Figure 7: Dyad Level Probabilities in wave 1

(a) Caste/Tribe



(b) Wealth



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 two-thirds of the caste/tribe gradient. However, this decomposition does not tell us *why* lower-SES women have higher within-group connectedness than higher-SES women. One explanation is that higher-SES women face more restrictions in interacting with peers, even peers of the same wealth and cast/tribe groups. This could stem from women in higher-SES households facing greater mobility restrictions, especially if it is less necessary for women to leave the household frequently for work or for using the toilet. Albeit consistent with the prescriptions of the caste system which discourages interactions between higher and lower castes and promotes within caste interactions, our analysis so far does not tell us much about *why* our networks exhibit homophily; could it be explained by mothers of the same group living closer together, having similar habits in terms of leaving the household, or does it remain unexplained by the characteristics we observe?

To probe the drivers of within- and across-group connectedness by caste/tribe and wealth we assess whether other observed characteristics of respondents and the asked-about mother can mediate the observed SES gradients using a dyad-level mediation analysis. We first regress, by OLS[15], 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 probability of a high-SES mother having a randomly-chosen within-group connection and the same

[15] The benefit of using OLS over the probit estimator in this exercise 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 for a low-SES mother (a.k.a. $\widehat{p^{HH}} - \widehat{p^{LL}}$). $\hat{\beta}^{LH}$ ($\hat{\beta}^{HL}$) is the difference between the probability of a high-SES (low-SES) mother having a randomly-chosen across-group connection than 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 $\hat{\beta}^{HH}$is indicative of the degree to which low-SES women have within-group connections at a higher rate than high-SES women.

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

$$Y_{ijv} = \beta^0 + \beta^{HH} C_{ijv}^{HH} + \beta^{LH} C_{ijv}^{LH} + \beta^{HL} C_{ijv}^{HL} + \alpha^1 X_{iv} + \alpha^2 X_{jv} + \alpha^3 X_{iv} \cdot X_{jv} + \epsilon_{ijv} \quad (4)$$

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. This analysis is descriptive and is not necessarily causal: 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.

Figure 8a shows how different characteristics mediate the gaps in probabilities of different groups reporting connections relative to the probability of the "ST/SC/OBC to ST/SC/OBC" connection. The figure starts with the caste/tribe-only model, sequentially adding wealth, age, household toilet ownership, maternal labor force participation, and then 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.

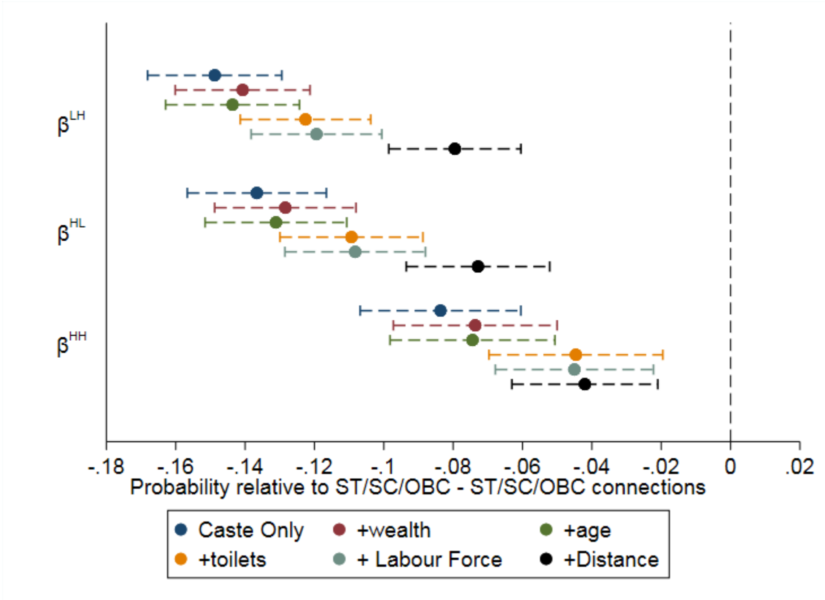
Controlling for household toilet ownership reduces the difference in within-group connectedness by caste/tribe by roughly 3 p.p.. It also reduces the difference between the probability between across-group existing and within-group connections existing by a similar magnitude. Non-SC/ST/OBC households 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 reduces the difference in probability of across-caste/tribe vs. within-group connections by 2-3 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), to avoid close interactions. 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.

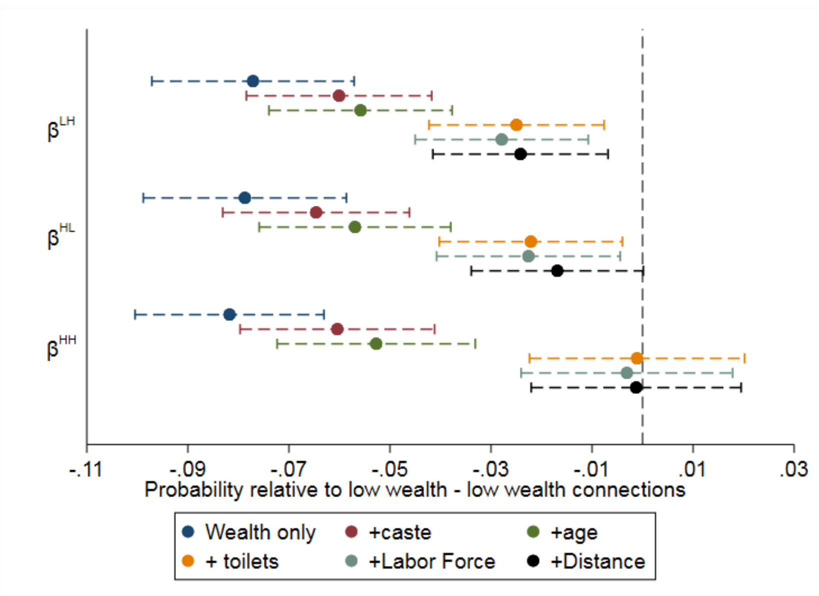
Figure 8b shows the same mediation analysis for the wealth gradient, showing probabilities relative to low-wealth-to-low-wealth connections. Controlling for caste/tribe and age reduce by roughly 2 p.p. the wealth-difference both in the within- and across-group connectedness. Labor force participation 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 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: Figures a and b plot the coefficients $\widehat{\beta^{HH}}, \widehat{\beta^{LH}}$ and $\widehat{\beta^{HL}}$ from equation (4) as controls are sequentially added to the model. Wealth is a binary indicator if a household has a Principal Component Analysis (PCA) asset score above the village median. Caste is a binary indicator equal to one if a household head is SC/ST/OBC. Age is mothers' age in years. Toilets is binary indicator of household toilet ownership. Labor force is binary indicator of mothers' labor force participation. Distance is distance to other mother in meters (included quadratically).

4 Discussion and Conclusions

In this paper, we provide novel quantitative evidence on the degree of isolation for young mothers in rural India. We demonstrate that mothers are, on average, extremely isolated. This is worrying given existing evidence, from various contexts, that social isolation is associated with poor women's mental and physical health (Berkman et al., 2000; Cacioppo & Hawkley, 2003; De Silva et al., 2007; Fowler & Christakis, 2009; Kohler et al., 2007; Sawyer et al., 2010; Smith & Postmes, 2011) and with women more likely to be victims of domestic violence (Lanier and Maume, 2009; Choi, Cheung and Cheung, 2012). Adverse effects of social isolation on mothers may have knock-on impacts on their children (Sawyer, Ayers and Smith, 2010; Kingston and Tough, 2014; Bennett *et al.*, 2016). 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 restrictions 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.

We find significant heterogeneity in the degree of this isolation and, in particular, we demonstrate large negative SES gradients, where higher-SES mothers report significantly fewer connections than their lower-SES peers. This gap is persistent, remaining large and significant over a period of four years. We decompose these gradients for caste/tribe and wealth into three components explained by: village-level composition and homophily; differences in between-group connectedness; and differences in within-group connectedness. We find that around a third of the gap for caste/tribe is associated with village-level composition and homophily. The majority of both gradients is associated with differences in within-group connectivity, with lower-SES dyads being significantly more likely to be connected than higher-SES dyads.

Our mediation analysis suggests that higher rates of toilet ownership amongst higher-SES households may be an important explanation of the SES gradients, both by wealth and by caste/tribe, and of homophily by SES. Toilet ownership likely further limits the opportunities for young mothers in this context to have to leave their home, and thus restricts their opportunities to form ties with peers. Their higher toilet-ownership rates appear key in associations such that higher-SES women have fewer connections with other women of their own SES group than do lower-SES women. They also appear important in associations for social ties across SES groups. Our analysis suggests that female labor force participation might also be important in causing these SES gradient associations.

We cannot definitively disentangle the causes of the SES gradients we observe. Networks are formed endogenously, in part to serve individuals' and households' economic and social interests. It may be that lower-SES women have more to gain from social connections if, for example, they are more actively involved in agricultural or other economic production and social connections are important sources of information, credit, or business (Banerjee *et al.*, 2013; Munshi and Rosenzweig, 2016). However, the negative SES gradients we observe and especially the fact that they are partially mediated by actions households are likely to take as they grow wealthier are also consistent with a more troubling picture. It is well-documented that women living secluded lives focused on the home can bring social status to households and that this can lead to more advantaged households, for whom women leaving the house is less of an economic or practical necessity, to place greater restrictions on women's mobility and work (Boserup, 1970; Miller, 1982; Chen, 1983, 1995; Klasen and Pieters, 2015).This phenomenon may in part drive the negative SES gradients we observe in social connectedness if greater restrictions constrain women's ability to gain social connections, either made incidentally, through working and spending time outside the house, or purposefully to serve their economic and social interests.

Given the very high degree of social isolation among young married women in rural India and given that our analysis suggests that increasing wealth alone may not improve the situation 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 like 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 how changing wealth and amenities in a village can impact causally network formation and social isolation. 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 important 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, Yuan Gao Wayne, Milan Pau (2020). "Informal Risk Sharing with Local Information." Available at SSRN: https://ssrn.com/abstract=3220524 or http://dx.doi.org/10.2139/ssrn.3220524.

Ambrus Attila, Mobius Markus, Szeidl Adam (2014). "Consumption Risk-Sharing in Social Networks." American Economic Review, 104(1), 149–182.

Anukriti, S. et al. (2020) Curse of the Mummy-ji: The Influence of Mothers-in-Law on Women in India, Boston University - Department of Economics - The Institute for Economic Development Working Papers Series. dp-337. Boston University - Department of Economics. Available at: https://ideas.repec.org/p/bos/iedwpr/dp-337.html (Accessed: 26 August 2020).

Attanasio, O. et al. (2016) Early Childhood Development for the Poor: Impacting at Scale, Odisha, India. Baseline Report, pp. 1–107.

Attanasio, O. and Krutikova, S. (2020) Consumption Insurance in Networks with Asymmetric Information. Available at: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3615487 (Accessed: 5 October 2020).

Banerjee, A. et al. (2013) 'The Diffusion of Microfinance', Science, 341(6144), pp. 1236498–1236498. doi: 10.1126/science.1236498.

Barnhardt, S., Field, E. and Pande, R. (2017) 'Moving to opportunity or isolation? network effects of a randomized housing lottery in urban india', American Economic Journal: Applied Economics, 9(1), pp. 1–32.

Beaman, L. et al. (2018) Can Network Theory-based Targeting Increase Technology Adoption? Working Paper 24912. National Bureau of Economic Research. doi: 10.3386/w24912.

Behrman, J. R., Kohler, H.-P. and Watkins, S. C. (2002) 'Social networks and changes in contraceptive use over time: Evidence from a longitudinal study in rural Kenya', Demography, 39(4), pp. 713–738.

Bennett, I. M. et al. (2016) 'Maternal mental health, and child growth and development, in four low-income and middle-income countries', J Epidemiol Community Health, 70(2), pp. 168–173.

BenYishay, A. and Mobarak, A. M. (2019) 'Social Learning and Incentives for Experimentation and Communication', The Review of Economic Studies, 86(3), pp. 976–1009. doi: 10.1093/restud/rdy039.

Berkman, L. F. et al. (2000) 'From social integration to health: Durkheim in the new millennium', Social Science & Medicine (1982), 51(6), pp. 843–857. doi: 10.1016/s0277-9536(00)00065-4.

Blinder, A. S. (1973) 'Wage Discrimination: Reduced Form and Structural Estimates', The Journal of Human Resources, 8(4), pp. 436–455. doi: 10.2307/144855.

Boserup, E. (1970) Woman's Role in Economic Development. Earthscan.

Cacioppo, J. T. and Hawkley, L. C. (2003) 'Social isolation and health, with an emphasis on underlying mechanisms', Perspectives in Biology and Medicine, 46(3 Suppl), pp. S39-52.

Chatterjee, E., Desai, S. and Vanneman, R. (2018) 'Indian paradox: Rising education, declining womens' employment', Demographic Research, 38, pp. 855–878. doi: 10.4054/DemRes.2018.38.31.

Chen, M. (1983) A quiet revolution : women in transition in rural Bangladesh. Cambridge, MA : Schenkman Pub. Co. Available at: https://trove.nla.gov.au/work/21577368 (Accessed: 6 October 2020).

Chen, M. (1995) A Matter of Survival: Women's Right to Employment in India and Bangladesh, Women, Culture, and Development. Oxford University Press. Available at: https://oxford.universitypressscholarship.com/view/10.1093/0198289642.001.0001/acprof-9780198289647-chapter-2 (Accessed: 6 October 2020).

Choi, S. Y. P., Cheung, Y. W. and Cheung, A. K. L. (2012) 'Social Isolation and Spousal Violence: Comparing Female Marriage Migrants With Local Women', Journal of Marriage and Family, 74(3), pp. 444–461. doi: 10.1111/j.1741-3737.2012.00963.x.

Crivello, G. et al. (2018) Marital and fertility decision-making: The lived experiences of adolescents and young married couples in Andhra Pradesh and Telangana, India. Young Lives. Available at: https://ora.ox.ac.uk/objects/uuid:315e83f5-753f-42c8-afec-237bd0e96f64 (Accessed: 6 February 2020).

Curtis, V. (2019) 'Explaining the outcomes of the "Clean India" campaign: institutional behaviour and sanitation transformation in India', BMJ Global Health, 4(5). doi: 10.1136/bmjgh-2019-001892.

De Silva, M. J. et al. (2007) 'Social capital and mental health: A comparative analysis of four low income countries', Social Science and Medicine, 64(1), pp. 5–20. doi: 10.1016/j.socscimed.2006.08.044.

Dìaz-Martin, L. et al. (2020) 'Greater than the Sum of the Parts? Evidence on Mechanisms Operating in Women's Groups'. Available at: https://faculty.wcas.northwestern.edu/~sjv340/womens_groups.pdf.

Eswaran, M., Ramaswami, B. and Wadhwa, W. (2013) 'Status, Caste, and the Time Allocation of Women in Rural India', Economic Development and Cultural Change, 61(2), pp. 311–333. doi: 10.1086/668282.

Fafchamps, M. and Lund, S. (2003) 'Risk-sharing networks in rural Philippines', Journal of Development Economics, 71(2), pp. 261–287. doi: 10.1016/S0304-3878(03)00029-4.

Feigenberg, B., Field, E. and Pande, R. (2013) 'The Economic Returns to Social Interaction: Experimental Evidence from Microfinance', The Review of Economic Studies, 80(4), pp. 1459–1483. doi: 10.1093/restud/rdt016.

Field, E. et al. (2016) 'Friendship at Work: Can Peer Effects Catalyze Female Entrepreneurship?', American Economic Journal: Economic Policy, 8(2), pp. 125–153. doi: 10.1257/pol.20140215.

Field, E. et al. (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. doi: 10.2139/ssrn.3456234.

Fowler, J. H. and Christakis, N. A. (2009) 'Dynamic spread of happiness in a large social network: Longitudinal analysis over 20 years in the Framingham Heart Study', BMJ (Online), 338(7685), pp. 23–26. doi: 10.1136/bmj.a2338.

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

Jayachandran, S. (2019) 'Social Norms as a Barrier to Women's Employment in Developing Countries'.

Kabeer, N. (1994) Reversed Realities: Gender Hierarchies in Development Thought. Verso.

Kandpal, E. and Baylis, K. (2019) ‘The social lives of married women: Peer effects in female autonomy and investments in children’, Journal of Development Economics. doi: 10.1016/j.jdeveco.2019.05.004.

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

Kingston, D. and Tough, S. (2014) ‘Prenatal and postnatal maternal mental health and school-age child development: a systematic review’, Maternal and child health journal, 18(7), pp. 1728–1741.

Klasen, S. and Pieters, J. (2015) ‘What Explains the Stagnation of Female Labor Force Participation in Urban India?’, World Bank Economic Review, 29(3), pp. 449–478. doi: 10.1093/wber/lhv003.

Kohler, H. P. and Bühler, C. (2001) ‘Social Networks and Fertility’, in Smelser, N. J. and Baltes, P. B. (eds) International Encyclopedia of the Social & Behavioral Sciences. Oxford: Pergamon, pp. 14380–14384. doi: 10.1016/B0-08-043076-7/02176-8.

Kohler, H.-P., Behrman, J. R. and Watkins, S. C. (2000) ‘Empirical Assessments of Social Networks, Fertility and Family Planning Programs: Nonlinearities and their Implications’, Demographic Research, 3. Available at: https://www.jstor.org/stable/26348012 (Accessed: 5 October 2020).

Kohler, H.-P., Behrman, J. R. and Watkins, S. C. (2007) ‘Social networks and HIV/AIDS risk perceptions’, Demography, 44(1), pp. 1–33.

Lanier, C. and Maume, M. O. (2009) ‘Intimate partner violence and social isolation across the rural/urban divide’, Violence Against Women, 15(11), pp. 1311–1330. doi: 10.1177/1077801209346711.

Magnan, N. et al. (2015) ‘Leveling with friends: Social networks and Indian farmers’ demand for a technology with heterogeneous benefits’, Journal of Development Economics, 116, pp. 223–251. doi: 10.1016/j.jdeveco.2015.05.003.

Mason, K. O. and Smith, H. L. (2000) ‘Husbands’ versus wives’ fertility goals and use of contraception: The influence of gender context in five Asian countries’, Demography, 37(3), pp. 299–311. doi: 10.2307/2648043.

Mehrotra, S. and Parida, J. K. (2017) 'Why is the Labour Force Participation of Women Declining in India?', World Development, 98(C), pp. 360–380.

Miller, B. D. (1982) 'Female Labor Participation and Female Seclusion in Rural India: A Regional View', Economic Development and Cultural Change, 30(4), pp. 777–794.

Munshi, K. and Rosenzweig, M. (2016) 'Networks and Misallocation: Insurance, Migration, and the Rural-Urban Wage Gap', American Economic Review, 106(1), pp. 46–98. doi: 10.1257/aer.20131365.

Nussbaum, M. C. (2011) Creating capabilities. Harvard University Press.

Oaxaca, R. (1973) 'Male-Female Wage Differentials in Urban Labor Markets', International Economic Review, 14(3), pp. 693–709. doi: 10.2307/2525981.

Patil, K. (2019) 'A Social Network Analysis of Open Defecation Practices in India', p. 48.

Prillaman, S. A. (2017) 'Strength in Numbers: How Women's Groups Close India's Political Gender Gap', p. 71.

Richardson, R. A., Barbour, N. E. and Bubenzer, D. L. (1995) 'Peer Relationships as a Source of Support for Adolescent Mothers', Journal of Adolescent Research, 10(2), pp. 278–290. doi: 10.1177/0743554895102005.

Rowlands, J. (1997) Questioning Empowerment: Working with Women in Honduras. Oxfam.

Sanyal, P. (2009) 'From Credit to Collective Action: The Role of Microfinance in Promoting Women's Social Capital and Normative Influence', American Sociological Review, 74(4), pp. 529–550. doi: 10.1177/000312240907400402.

Sawyer, A., Ayers, S. and Smith, H. (2010) 'Pre- and postnatal psychological wellbeing in Africa: A systematic review', Journal of Affective Disorders, 123(1–3), pp. 17–29. doi: 10.1016/j.jad.2009.06.027.

Smith, L. G. E. and Postmes, T. (2011) 'The power of talk: Developing discriminatory group norms through discussion', British Journal of Social Psychology, 50(2), pp. 193–215. doi: 10.1348/014466610X504805.

Townsend, R. (1994) 'Risk and Insurance in Village India', Econometrica, 62(3), pp. 539–91.

Appendix A: Additional Tables and Figures

Figure A1: Study Area

(a) Odisha within India

(b) Study districts within Odisha

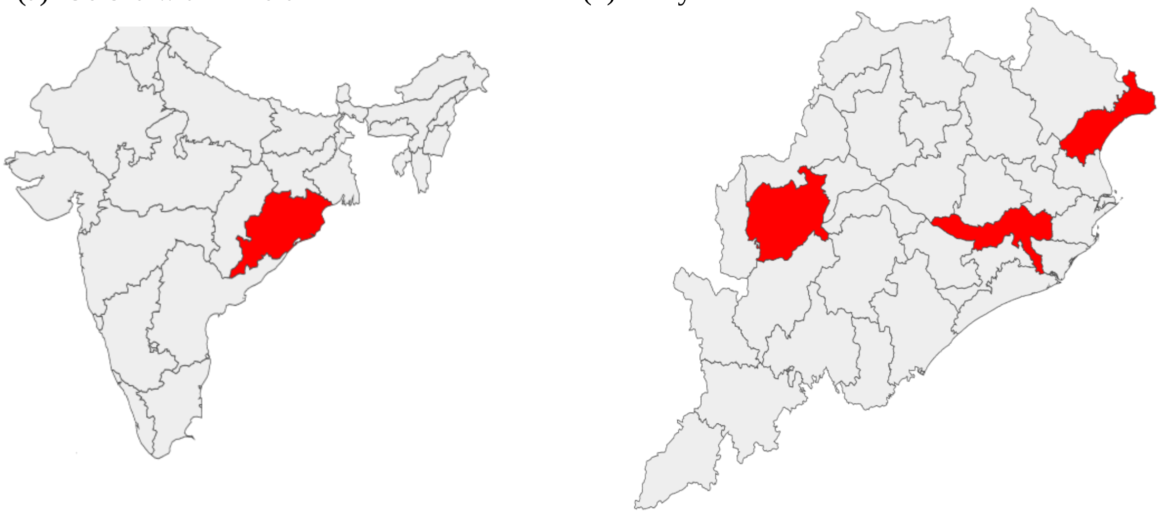
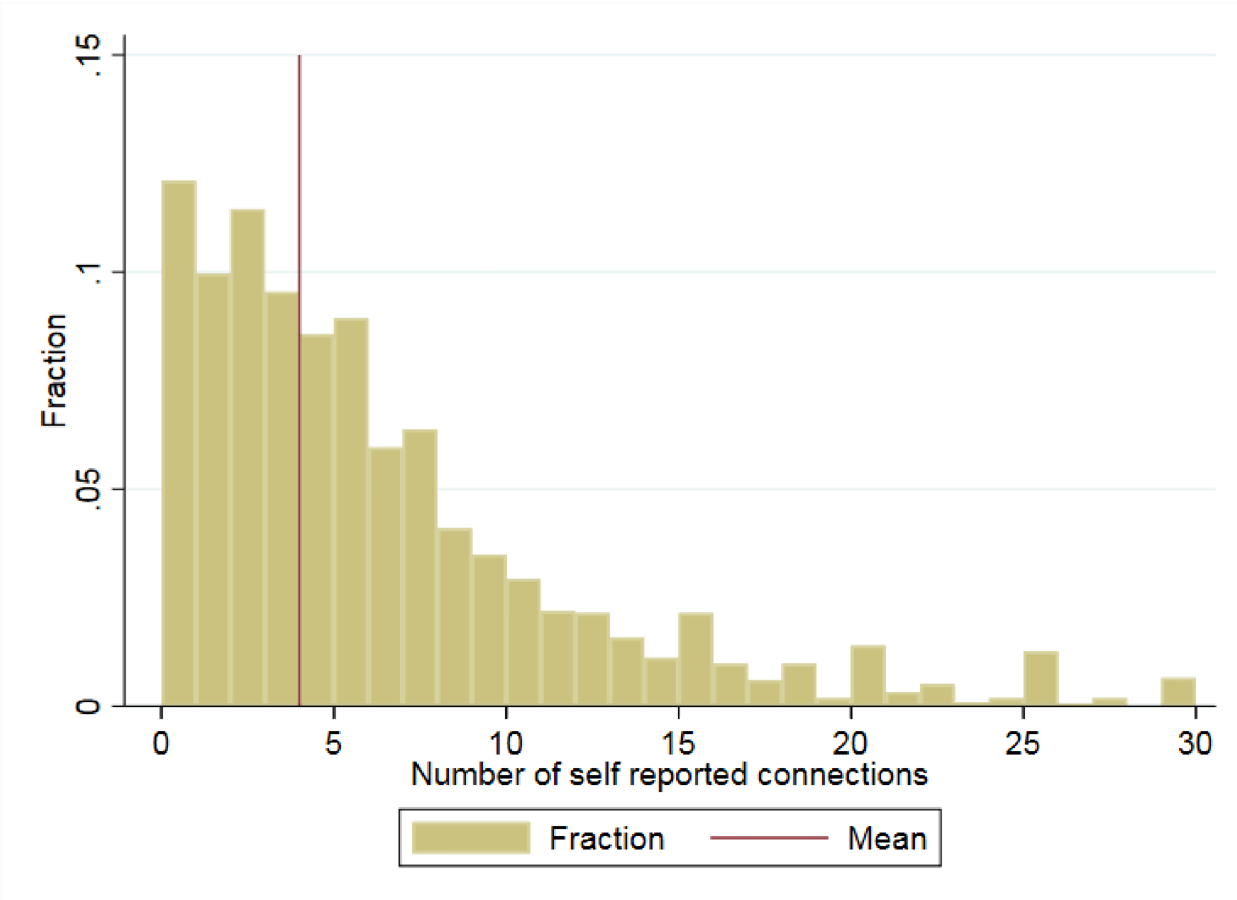


Table A1: Spillover vs Target Mothers

	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 (0.5)	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

Means (sds) for selected characteristics of target and spillover mothers. P-value is for the t-test of means equality.

Figure A2: Distributions of Self Reported Connections



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

We collect data on network graphs for a (quasi-random) sample of the village network of mothers with young children. However, in order to assess eligibility for the study we collected village-level censuses of all mothers with children under the age of two years before the study began (August 2015). In this data we collect information on GPS location, caste, gender and the age of 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 $y_{ikv} = 1$ 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

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

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

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