Distribution of caste and tribal status among children of different religions.
Table A1 shows the distribution of caste and tribal status by religion among rural children whose heights were measured in the NFHS-3. The religion of the child's household head is used as the measure of religion.

CDFs of height by population group.
The summary statistics table (Table 1) in the main text described the 25th, 50th, and 75th percentiles of height-for-age by population group. Figure A1 plots the empirical CDFs of child height-for-age for the four population groups in rural India that we study. Note that the distribution for general caste children stochastically dominates (is everywhere to the right of) the distribution for OBC children, which stochastically dominates the distributions for SC and ST children.

Floor type predicts child height in rural India.
In the reweighting decomposition described in Section 4 of the main text, we use 4 categories of floor type as a measure of child SES. Table A2 uses OLS regression to show that floor type predicts child height in rural India. This is true whether or not we additionally control for mother's education, the second variable used in the reweighting decomposition.

Using father's education in the reweighting decomposition produces similar results as using mother's education.
In the main text, we perform reweighting decompositions that ask the question: what would be the average heights of rural ST, SC, and OBC children be if they had the same distribution of SES (as measured by flooring material and mother's education) as general caste children? Here, we perform the same analysis, but use father's education instead of mother's education. The results are similar: reweighting over the intersection of father's education and floor type entirely explains the ST–general height gap, but does not entirely explain the SC–general gap, nor the OBC–general gap. Figure A2 reproduces Figure 3 from the main text using father's education instead of mother's education.

Regression analysis: Robustness to omitting OBC children.

The caste composition of the OBC category varies at the state-level throughout India, as does the everyday experience of being someone from an OBC caste. In some contexts, some OBC households may not be considered socially lower ranking than some households classified here as general. However, SC households are unambiguously considered socially lower ranking than OBC and general caste households. To verify that no special property of OBC children or households is responsible for the pattern of our results, Table A3 replicates our main results excluding all OBC children from the sample. The same results are found: SES accounts for all of the ST height gap, but not the SC height gap, which is explained by SES variables and local caste rank.

OLS regression analysis, separately by boys and girls.

Many demographic process in India differ by sex; gender attitudes and practices have important consequences for health, nutrition, and mortality. Our demographic controls in the regression results include a control for child sex. Table A4 verifies that the general pattern of our results is robust to splitting the sample by sex. For both girls and boys, there are large height gaps between SC, ST, and OBC children, and general caste children. For ST children, the gap can be explained by SES controls.

Regression analysis: Robustness to controlling for local open defecation and other controls.

Recent research has demonstrated the relevance of the local gastro-intestinal disease environment to child height in India (for example, Hathi et al., 2017). Table A5 reproduces the main results (from Table 3) and adds a control for local open defecation. That is, we control for the fraction of households in a child's PSU that report practicing open defecation rather than toilet or latrine use. Consistent with the prior literature, we find that local open defecation predicts child height. However, the inclusion of this control variable does not change our main results: SES differences still explain height gaps between general caste and ST children, and a combination of SES differences and local caste rank still explain height gaps between general caste and SC children, and between general caste and OBC children.

Table A6 reproduces Table A5, and uses additional controls in columns 6 and 7: the number of siblings a child has at the time of the survey (no siblings is omitted category), the religion of the household (Hindu is the omitted category), and whether the households owns a mattress, a pressure cooker, a chair, a bed, a table, or an electric fan. These additional controls are added as dummy variables. The results in Table A6 show that the addition of these new controls does not change the interpretation of the main results.

SC respondents are more likely to report open defecation near their houses.
Table A5 controls for PSU-level open defecation, but does not control for individual disease burdens. Section 6 of the main text discusses the possibility that differences in sanitation externalities might partially explain why, conditional on SES, lower caste children are shorter than general caste children in localities where they are outranked, but not in homogeneous villages. Table A7 finds suggestive evidence that in mixed villages, SC disease burdens may be higher than SC disease burdens in homogeneous villages. Table A7 uses data from a sanitation survey collected in five states in north India in 2013-14 (see Coffey et al. (2014) for details of the survey) to show that SC respondents are more likely to report seeing someone defecate in the open near their houses than general caste respondents. This result is robust to controls for whether the respondent defecates in the open, for estimated village level open defecation, and for the respondent's sex and age. We note, however, that this is not true for OBC respondents.

The heights of Scheduled Caste children are similar in large and small villages.
Section 6 of the main text discusses the possibility that the fraction of a child's village that outranks her household, a key variable in our analysis, may be correlated with village size. If lower caste children living in smaller villages are shorter than lower caste children living in larger villages, this would suggest that village size could be an important omitted variable in our analysis. Unfortunately, the NFHS-3 does not report village size. In order to investigate the correlation between village size and child height, we turn to the India Human Development Survey (IHDS), 2005 (Desai et al., 2005). The IHDS is a smaller, but nationally representative, sample of children for whom height, social category, and village size were collected. Figure A3 shows that at every age at which rural children were measured (ages zero to five and eight to eleven), Scheduled Caste children living in smaller villages (less than 500 households) are as tall as children living in larger villages (500+ households). Further, Table A8 shows the same conclusion using an OLS regression: it uses the same IHDS data on rural children ages zero to five and eight to eleven to show that, controlling for state fixed effects, household consumption per capita, and the education level of the most educated adult in the household, living in a large village does not predict height for rural SC children.

The fraction of households that are Scheduled Caste is similar in NFHS-3 PSUs and in 2001 Census villages.
Section 6 of the main text discusses the fact that the NFHS data allow us to observe local caste rank at the primary sampling unit (PSU) level, rather than at the village level. Figure A4 provides suggestive evidence that if we could observe a child's local caste rank at the village level, we would find similar results. It plots the cumulative distribution of the fraction of households in the NFHS PSUs that are Scheduled Caste, as well as the cumulative

distribution of the fraction of households in 2001 Census villages that are Scheduled Caste. The two cumulative distributions are very similar.

Figure A1: Empirical CDF of height-for-age, by population group
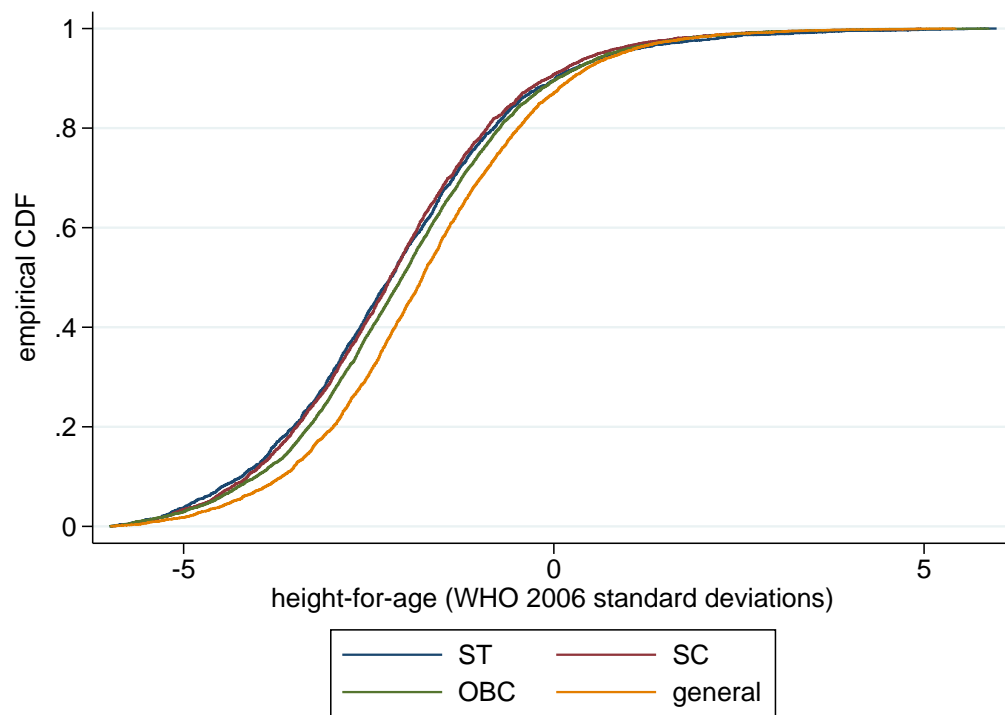
Observations are rural children whose heights were measured by the NFHS-3.

Figure A2: Decomposition results: Fraction of the height gap between general case children and ST, SC, and OBC children unexplained by socioeconomic variables
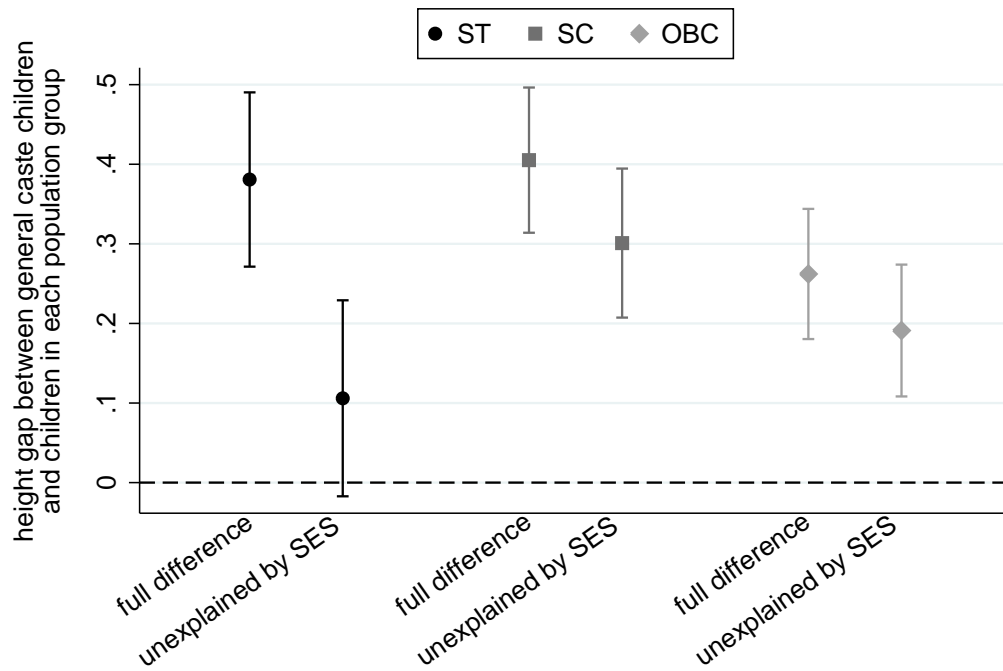
Estimates represent reweighting decomposition results that describe height gaps between general caste children and children in other population groups. Confidence intervals represent 95% confidence, computed using clustered standard error to reflect the survey design of the DHS. The socioeconomic variables used in the decomposition are floor type (4 categories) and education level of her father (6 categories), for a combined set of 24 bins. These variables and the reweighting decomposition method are described in section 3. The figure in the main text uses mother's education instead of father's education.

Figure A3: Heights of Scheduled Caste children are similar in large and small villages
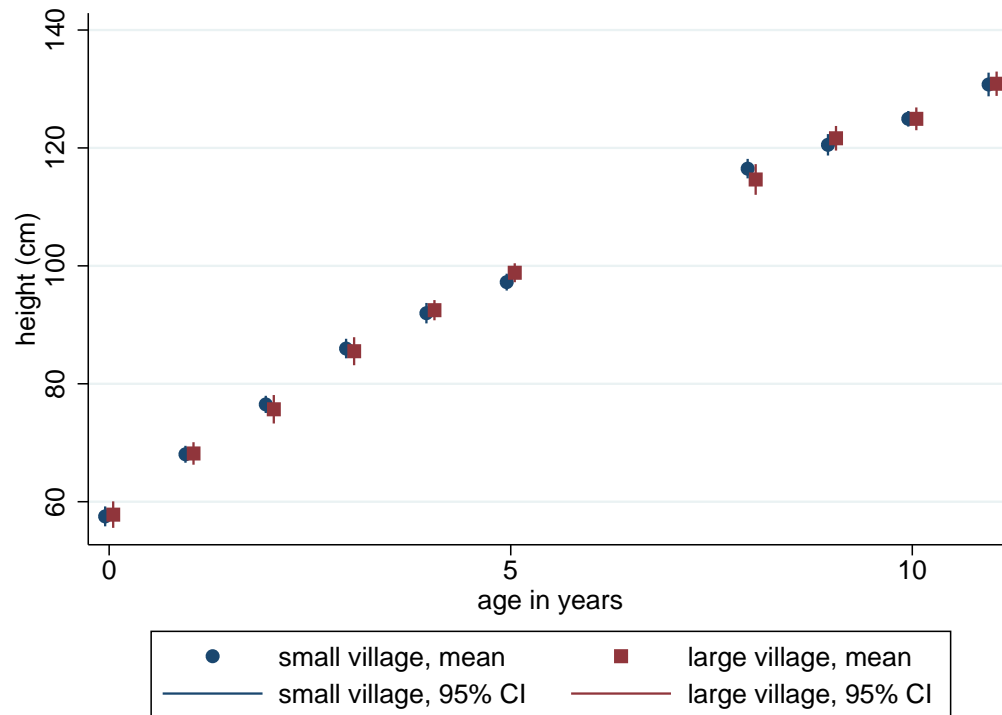
Observations are rural SC children whose heights were measured by the India Human Development Survey, 2005. The graph shows average height and 95% confidence intervals for children of different ages. $n = 5,445$. Large villages have 500 or more households; small villages have less than 500 households.

Figure A4: The fraction of households that are Scheduled Caste is similar in NHFS-3 PSUs and in 2001 Census villages
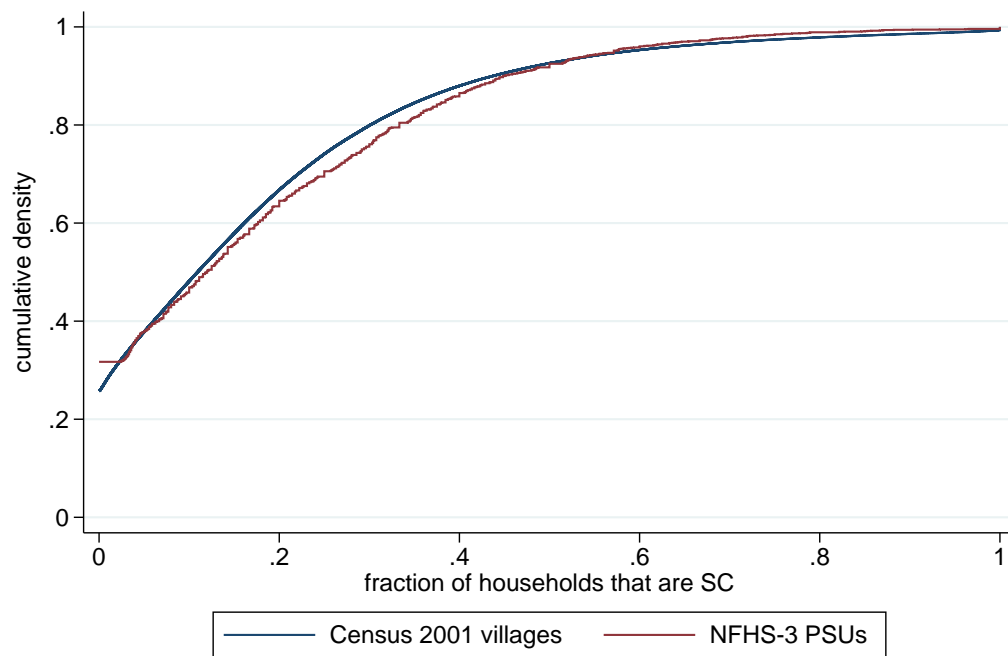
Observations are rural primary sampling units in the NFHS data and villages in the 2001 Census data (Office of Registrar General & Census Commissioner, 2001). Cumulative distributions of fraction of the PSU or village that are SC are shown.

Table A1: Distribution of caste and tribal status among children of different religions

	religion of the child's household head		
	Hindu	Muslim	Other religion
ST	0.12	0.01	0.34
SC	0.24	0.02	0.28
OBC	0.45	0.32	0.10
General	0.18	0.58	0.23
social group data missing	0.01	0.17	0.05
n	18,485	3,587	3,993

This table shows the caste and tribal status of children by the religion of their household head.
Observations are rural children whose heights were measured by the NFHS-3.

Table A2: Floor type predicts child height-for-age

	(1)	(2)
rudimentary floor	0.307***	0.176***
	(0.0504)	(0.0502)
finished floor	0.569***	0.346***
	(0.0249)	(0.0263)
other floor type	0.461***	0.327***
	(0.0382)	(0.0382)
mother did not complete primary		0.226***
		(0.0398)
mother completed primary		0.281***
		(0.0414)
mother did not complete secondary		0.490***
		(0.0260)
mother completed secondary		0.815***
		(0.0662)
mother has higher education		1.094***
		(0.0704)
constant	-2.158***	-2.290***
	(0.0129)	(0.0145)
n	24840	24840
R^2	0.023	0.046

Observations are rural children whose heights were measured by the NFHS-3. The omitted category of floor type is a dirt floor. The omitted category for mother's education is no education.

Table A3: Regression robustness: Similar results omitting OBC children

dependent variable:	(1)	(2)	(3)	(4)
		height-for-age z-score		
ST	-0.381***	-0.0979		
	(0.0559)	(0.0636)		
SC	-0.405***	-0.193***	-0.204***	-0.164
	(0.0465)	(0.0515)	(0.0526)	(0.207)
own SES bins		✓	✓	✓
local caste rank bins				✓
n	16,227	16,227	11,082	11,082
R^2	0.014	0.190	0.226	0.339

† two-sided $p < 0.10$; * $p < 0.05$; ** $p < 0.01$. Standard errors are clustered by survey PSU. These regressions fully exclude OBC children, but are otherwise identical to the same numbered columns in the main regression results in Table 3 of the main text.

Table A4: OLS regression results, by sex

	girls before SES	girls after SES	boys before SES	boys after SES
ST	-0.30***	0.03	-0.46***	-0.15
	(0.07)	(0.10)	(0.07)	(0.09)
	$n = 5,572$		$n = 5,925$	
SC	-0.36***	-0.17*	-0.44***	-0.24**
	(0.06)	(0.07)	(0.06)	(0.07)
	$n = 5,262$		$n = 5,822$	
OBC	-0.23***	-0.11	-0.29***	-0.13*
	(0.05)	(0.07)	(0.05)	(0.06)
	$n = 7,118$		$n = 7,849$	

Each estimate is from a separate comparison of children in the disadvantaged population group (ST, SC, or OBC) with general caste children. The standard errors are clustered by survey PSU. SES controls include an indicator for the household having electricity, a phone, a radio, a TV, a refrigerator, a bicycle, a motorcycle, a car, land, and whether the household uses a toilet or latrine. These are the same SES controls used in Table 3 of the main text. Two-sided p-values: * $p < 0.05$, ** $p < 0.01$, ***$p < 0.001$.

Table A5: Main results (Table 3), adding a control for local open defecation

dependent variable:	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	height for age z-score (general caste children are omitted category)						
ST	-0.381***	-0.0560					
	(0.0559)	(0.0648)					
SC	-0.405***	-0.147**	-0.148**	0.0636	-0.125+	-0.148**	0.0636
	(0.0465)	(0.0544)	(0.0551)	(0.119)	(0.0643)	(0.0551)	(0.119)
OBC	-0.262***	-0.112*	-0.122*	0.00217	-0.129*	-0.122*	0.00217
	(0.0417)	(0.0482)	(0.0493)	(0.0783)	(0.0595)	(0.0493)	(0.0783)
local open defecation		-0.213*	-0.235*	-0.231+	-0.237+	-0.235*	-0.231+
		(0.0949)	(0.103)	(0.134)	(0.138)	(0.103)	(0.134)
n	24840	23111	18141	18141	18141	18148	18141
R^2	0.008	0.201	0.222	0.364	0.344	0.222	0.364
own SES bins		✓				✓	
own SES × caste rank bins			✓	✓			✓
own SES × SES rank bins					✓		
demographic & neighborhood						✓	✓
composition controls						✓	✓
STs in the sample	yes	yes	no	no	no	no	no

OLS regressions, weighted using NFHS sample weights. Observations are children under five in the NFHS-3. Two-sided *p*-values: † $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, ***$p < 0.001$. Standard errors clustered by primary sampling unit (PSU) in parentheses. ST = Scheduled Tribe, SC = Scheduled Caste, OBC = Other Backwards Class. The construction of controls for "own SES bins," "own SES × caste rank bins," and "own SES × SES rank bins" are discussed in Section 4 of the main text. "demographic controls" include child birth order, child sex, and whether the child lives in a joint family with his/her grandparents. "neighborhood controls" include the fraction of households in a child's PSU with electricity, the fraction of births (last births to the mother) in the child's PSU that got prenatal care, the fraction of SC households, the fraction of OBC households, and the fraction of general caste households in a child's PSU. "local open defecation" is the fraction of households sampled in a child's primary sampling unit that defecate in the open.

Table A6: Main results (Table 3), with additional controls

dependent variable:	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	height for age z-score (general caste children are omitted category)						
ST	-0.381***	-0.0705					
	(0.0559)	(0.0653)					
SC	-0.405***	-0.153**	-0.154**	0.0383	-0.132*	-0.219***	-0.0345
	(0.0465)	(0.0545)	(0.0551)	(0.118)	(0.0646)	(0.0618)	(0.161)
OBC	-0.262***	-0.123*	-0.134**	-0.0200	-0.142*	-0.152**	-0.0171
	(0.0417)	(0.0485)	(0.0494)	(0.0774)	(0.0598)	(0.0588)	(0.104)
local open defecation						-0.168	-0.122
						(0.109)	(0.138)
one sibling						-0.322***	-0.327***
						(0.0639)	(0.0823)
two siblings						-0.562***	-0.602***
						(0.0861)	(0.110)
three siblings						-0.812***	-0.853***
						(0.116)	(0.145)
four or more siblings						-0.998***	-1.095***
						(0.168)	(0.205)
Muslim						-0.0284	-0.0371
						(0.0664)	(0.0798)
neither Hindu nor Muslim						0.0897	0.281*
						(0.0859)	(0.125)
n	24840	23111	18141	18141	18141	18089	18082
R^2	0.008	0.201	0.222	0.364	0.344	0.235	0.375
own SES bins		✓	✓			✓	
own SES × caste rank bins				✓			✓
own SES × SES rank bins					✓		
additional controls						✓	✓
demographic & neighborhood						✓	✓
composition controls						✓	✓
STs in the sample	yes	yes	no	no	no	no	no

The same note as for table A5 applies here. This table uses additional controls in columns 6 and 7: the number of siblings a child has at the time of the survey (no siblings is omitted category), the religion of the household (Hindu is the omitted category), and whether the households owns a mattress, a pressure cooker, a chair, a bed, a table, or an electric fan. These additional controls are added as dummy variables.

Table A7: SC villagers are more likely to see others defecate in the open near their home

	(1)	(2)
dependent variable: has seen OD near home		
SC household	0.0588*	0.0412†
	(0.0229)	(0.0235)
OBC household	0.0123	-0.00384
	(0.0180)	(0.0182)
village OD		0.284**
		(0.0332)
respondent ODs		-0.0222
		(0.0197)
respondent is female		0.00261
		(0.0161)
respondent's age		0.000780
		(0.000612)
n (rural households)	2,947	2,914

Two-sided p-values: $\dagger\ p < 0.10$, * $p < 0.05$, ** $p < 0.01$, ***$p < 0.001$. OD = open defecation; village OD is the fraction 0-1 of households in the village that defecate in the open. Data are from the SQUAT survey; see Coffey et al. (2014) for more information about the survey.

Table A8: Heights of Scheduled Caste children are similar in large and small villages: OLS results

dependent variable: height in cm	
large village (> 500 HH)	-0.05
	(0.31)
age 1	9.73***
	(0.94)
age 2	18.84***
	(0.62)
age 3	27.92***
	(0.57)
age 4	33.84***
	(0.66)
age 5	40.81***
	(0.69)
age 8	58.51***
	(0.43)
age 9	64.17***
	(0.42)
age 10	67.91***
	(0.70)
age 11	72.80***
	(0.73)
state fixed-effects	✓
fixed effects for years of education of the most educated adult in HH	✓
consumption per capita	0.00139***
	(0.000334)
constant	47.96***
	(3.14)
n	5434
R^2	0.81

Two-sided p-values: † $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, ***$p < 0.001$. Observations are rural SC children whose heights were measured by the India Human Development Survey, 2005. The table shows coefficients from an OLS regression of height in centimeters on village size, age in years, and controls. Large villages have 500 or more households; small villages have less than 500 households.