Parental Death and Schooling: Gendered Spheres of Production and Parental Preferences*

Rodrigo Guerrero†

November 24, 2023

Click here for the latest version.

Abstract

One out of every ten people in India experiences the death of a parent by the age of 18. This paper estimates the impact of parental death on educational choices in India and leverages the death shock to offer insights into the role of household resources and parental preferences in schooling decisions. Exploiting variation in the timing of parental loss among nearly 10,000 children, I find that both paternal and maternal death lead to a 7-percentage point decline in school enrollment. Time-use data sheds further light on this finding by showing that the death of either parent induces sons to enter the labor force and daughters to take on domestic responsibilities. Motivated by these facts, this paper examines three channels through which parental death might affect educational choices: household resources, parental inputs into education production, and parental preferences. To disentangle and quantify these mechanisms, I develop a structural model of household consumption and time allocation. Households in the model produce a domestic good and parents bargain over the intrahousehold allocation of resources in a collective framework. The model is estimated using shifts in time allocation and itemized expenditure in response to the death shock. The estimates show that compared to fathers, mothers have a stronger preference for schooling and a lower bargaining weight in intrahousehold decisions. Paternal death affects schooling by reducing household income, while maternal death affects schooling by changing household preferences and removing her contribution to home production. Counterfactual simulations show that the effectiveness of interventions to support orphans' education, such as pensions and conditional cash transfers, depends on the gender of the child and the deceased parent.

JEL Codes: D1, I25, J1, J22, J24, O15

*I am very grateful to Mark Rosenzweig, Kaivan Munshi, Costas Meghir, and Rohini Pande for their invaluable guidance and advice. I would like to thank Joseph Altonji, Orazio Attanasio, Lauren Bergquist, Rossella Calvi, John Hamman, Mushfiq Mobarak, and Cormac O'Dea for helpful comments and insightful conversations.

†Yale University – Department of Economics. Contact: rodrigo.guerrerocastaneda@yale.edu

1 Introduction

In low-income countries, a substantial fraction of the population experiences the death of a parent before reaching adulthood. How does the loss of a parent impact children's educational choices? And what do these effects imply about the distinct roles of mothers and fathers in educational decisions? This paper analyzes the consequences of early parental death on schooling in India, where one in ten individuals faces such a loss before turning 18 (NFHS, 2015). By examining the mechanisms through which parental death affects school enrollment, this study sheds light on how educational choices are shaped by household resources, children's opportunity cost of time, parental preferences, and intrahousehold bargaining.

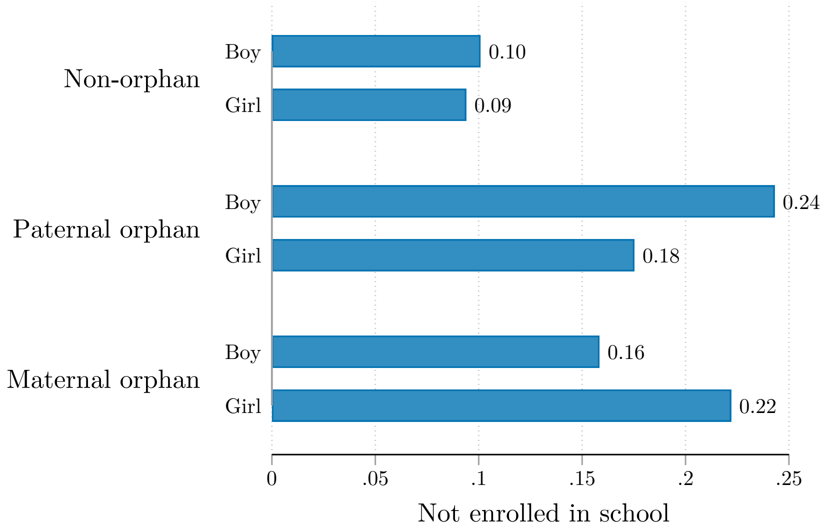
In India, the schooling trajectory of orphans and non-orphans begins to diverge around age 12. Adolescents who have lost a parent are, on average, 10 percentage points less likely to be enrolled in school. Decomposing this average difference by gender reveals a stark interaction between the gender of the child and that of the deceased parent. While there is gender parity in school enrollment among non-orphans, the experience of orphans is different, as depicted in Figure 1. Among paternal orphans, non-enrollment rates are higher for boys than for girls, and this gender gap is reversed among maternal orphans. Several theories could explain this pattern. It may be that mothers prioritize the education of daughters while fathers favor sons. Alternatively, maternal inputs may be more important for the education of daughters, and paternal inputs for the education of sons. Another potential explanation is that parental loss has different implications for the opportunity cost of schooling for sons versus daughters. Distinguishing between these explanations offers valuable insights into the determinants of schooling and the effectiveness of policies to keep children in school.

The first step in understanding how parental loss shapes educational outcomes is to establish a causal link. Several earlier estimates rely on cross-sectional comparisons between orphans and non-orphans.[1] However, it is difficult to convincingly establish a causal relationship between parental death and education with cross-sectional data due to unobserved differences between orphans and non-orphans. To address this concern, some studies use longitudinal data and control for child and time fixed effects.[2] These estimates rely on the assumption that the educational trajectory of orphans and non-orphans would evolve in parallel in the absence of the shock. As discussed in Section 2, this assumption is likely to be violated given that the sample of orphans is highly selected. Thus, the suitability of non-orphans as a control group, either in levels or in differences, is questionable.

[1]See Bicego et al. (2003); Guarcello et al. (2004); Case et al. (2004); Gertler et al. (2004a); Kobiané et al. (2005); Ardington and Leibbrandt (2010); Chuong and Operario (2012).

[2]See Ainsworth et al. (2005); Beegle et al. (2006); Case and Ardington (2006); Evans and Miguel (2007); Himaz (2013); Cas et al. (2014); Senne (2014). Studies that use administrative data from high-income countries leverage cross-sibling variation (Chen et al., 2009; Kalil et al., 2016; Kailaheimo-Lönnqvist and Erola, 2020; Barclay and Hällsten, 2022; Dribe et al., 2022; Liu et al., 2022) and cause of death data (Adda et al., 2011; Gimenez et al., 2013; Kailaheimo-Lönnqvist and Kotimäki, 2020; Goldstein, 2022; Dupraz and Ferrara, 2023).

Figure 1: Non-enrollment by gender of the child and vital status of each parent



Note: Sample consists of children from ages 12 to 18 in the CPHS data.

This paper addresses the endogeneity concerns mentioned above by proposing a staggered difference-in-differences estimator that relies on variation in the timing of unexpected parental deaths among orphans only. Specifically, identification relies on comparing the outcomes of a child before and after an unexpected parental loss, relative to children of the same age observed at the same time, but who will experience the shock in a later wave. This strategy estimates the average treatment effect on the treated under the assumptions of no anticipation and parallel trends among those experiencing a loss between ages 12 to 18. The data requirements of this estimator are high. In particular, it requires longitudinal data from a large sample of orphans. This study uses India's Consumer Pyramids Household Survey (CPHS), a household panel survey where I observe nearly 10,000 children before and after the death of a parent.

I estimate that among children aged 12 to 18, both paternal and maternal death cause a 7-percentage point drop in school enrollment. The loss of a father has a stronger effect on boys (11 percentage points) than on girls (2 percentage points). Conversely, the death of a mother has a similar effect on boys (5 percentage points) and girls (7 percentage points). Importantly, I find that the schooling effects of parental death are severely overestimated if non-orphans are used as a control group. The analysis of pre-trends shows that the school enrollment gap between orphans and non-orphans widens as children age. Thus, the assumption of parallel trends between orphans and non-orphans is unlikely to hold.

Next, I disentangle and quantify the mechanisms underlying the effects of parental death on schooling.[3] Specifically, parental death may affect educational choices through three distinct chan-

[3]Two studies have explored the mechanisms through which parental death affects schooling. Using data from

nels: (1) changes in household resources, (2) loss of parental inputs into education production, and (3) shifts in preferences due to the transition from two-parent to single-parent households. This analysis is made possible by combining time-use data with a structural model of household consumption and time allocation.

The first channel analyzed is the change in household resources. Given the strict gendered division of labor in India, the resource mechanism delivers tight predictions. Paternal death reduces household income and therefore should mostly affect sons, while maternal death constitutes a loss in home production and therefore should mostly affect daughters. Analysis of time-use data confirms that when a father dies, sons drop out of school to enter the labor force, and when a mother dies, daughters drop out of school to manage domestic responsibilities. However, I also find that maternal loss drives sons into the labor force, despite the fact that the shock does not affect household income or the probability of hiring domestic help. This result indicates that the resource mechanism cannot fully explain the observed effects and provides motivation to explore the other two channels.

The second channel concerns the loss of parental inputs into human capital production. If these inputs substitute or complement schooling in the production of human capital, then such a loss may affect the optimal choice of schooling.[4] However, while this mechanism may play a critical role for children who lose a parent during early childhood, it does not seem to be the case for children aged 12 through 18. Time-use data shows that individuals with children in this age range spend, on average, less than 10 minutes a day on childcare and instruction. Nonetheless, time-use data may not capture other forms of parental contributions to human capital production, such as guidance and passive supervision. Moreover, a parent's passing can directly affect a child's academic performance due to the psychological costs associated with the shock. To estimate this direct effect on academic outcomes, I estimate an education production function using data on test scores from the Young Lives - India panel. The estimates show that conditional on school enrollment, parental death does not have a significant impact on academic performance.

The third channel recognizes that children's educational choices are largely shaped by the preferences of their parents.[5] Thus, the loss of either parent may significantly alter these choices, particularly when mothers and fathers have different preferences and bargain over the intrahousehold allocation of resources. For example, the finding that a mother's passing induces sons to drop out of school and enter the labor force may reflect the fact that mothers care more about schooling

Indonesia and Mexico, Gertler et al. (2004b) find that controlling for changes in household per-capita consumption does not reduce the negative effects of parental death, suggesting that the income mechanism cannot fully account for the observed effects. Chen (2012) investigates the causes of the gender gap in schooling in Indonesia by comparing households with varying degrees of paternal involvement (both parents present, migrant fathers, sick fathers, deceased fathers). She finds that household production and preferences play a key role in explaining this gap.

[4]The early child development literature has stressed the importance of parental inputs at an early age. See, for instance, Del Boca et al. (2014, 2017); Attanasio et al. (2022); Wong (2023).

[5]A large literature has studied the extent to which parental preferences and intrahousehold bargaining affect children's wellbeing. See, for example, Thomas (1990); Lundberg et al. (1997); Phipps and Burton (1998); Glick and Sahn (2000); Blundell et al. (2005); Emerson and Souza (2007); Cherchye et al. (2012); Ringdal and Sjursen (2021).

than fathers. Importantly, even if fathers and mothers had identical preferences, the death of either parent would alter the tradeoff between private consumption and investment in children. In a two-parent household, resources devoted to children benefit both parents, and so the couple gets "two for the price of one" when spending on children (and other public goods). This force shifts the efficient allocation of resources in favor of children in two-parent households (Weiss and Willis, 1985). Although several studies have documented that single-parent households invest less on children's education compared to two-parent households, the tradeoff between private and public goods has been largely overlooked in this literature (Krein and Beller, 1988; Beller and Chung, 1992; Haveman and Wolfe, 1995; Ermisch and Francesconi, 2001; Gennetian, 2005; Greenwood et al., 2017; Mencarini et al., 2019).

To disentangle and quantify the role of parental preferences and resources in the schooling effects of parental death, I estimate a structural model of household consumption and time allocation that incorporates household production, public goods, and parental bargaining. Households in the model consist of some or all of the following key members: a father, a mother, a son and a daughter. On the production side, households combine time and raw materials to produce a domestic good. The model takes into account the stark division of labor by gender in Indian households, whereby only women spend time on domestic production. The preferences of each parent are defined over a vector of privately consumed market goods and a vector of non-market public goods: the home-produced good, children's schooling and the leisure of each member. Parents bargain over the intra-household allocation of resources in a collective household framework and decide how various members allocate their time between school, work and home production. The death of a parent affects the household problem in two ways. First, there is a change in household resources as the time endowment of the deceased parent is constrained to zero and households lose economies of scale. Second, the household's objective function changes as the Pareto weight of the deceased parent is set to zero.

Identification of this model is complicated by the fact that individual consumption and the level of the home-produced good are unobserved. Estimation of the home production function is possible by examining the choice of time and itemized food ingredients as a function of wages and food prices. The key assumption to identify the rest of the model is that individual preferences do not change upon widowhood, and so the choices of widows and widowers are informative of the preferences of mothers and fathers, respectively. This assumption is more flexible than it may initially appear, as parents in the model care about the leisure of their spouse and so widowhood results in a direct utility loss. Borrowing the methodology developed by Browning et al. (2013); Lewbel and Pendakur (2008), intrahousehold bargaining and economies of scale are identified from a system of Engel curves before and after the death shock. Finally, the staggered difference-in-differences estimates from the first part of the paper provide the identifying variation needed to estimate parental preferences via indirect inference.

The estimated model allows us to quantify the relative importance of the mechanisms through which parental death impacts educational choices. It shows that paternal death affects schooling decisions by reducing household income. Conversely, maternal death affects schooling decisions by changing household preferences and removing her contribution to home production. I then evaluate the effectiveness of various policies in alleviating the adverse effects of parental death on schooling. Unconditional cash transfers are an effective, albeit expensive, policy tool to keep male orphans in school. A monthly pension of Rs. 4,000 following the death of either parent is sufficient to keep male orphans in school. In contrast, unconditional transfers prove ineffective at keeping female orphans in school. To eliminate the schooling effects of maternal death on girls, a cash transfer conditional on enrollment of Rs. 3,000 a month is an effective policy tool.

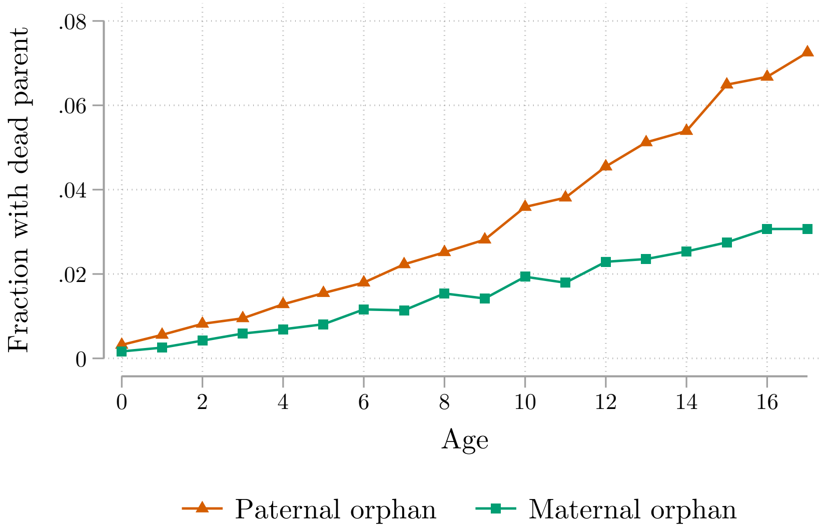
The model also provides broader insights into the role of parental preferences and resources as determinants of schooling. Understanding whether mothers and fathers exhibit different preferences for education has been a long-standing interest in economics (Thomas, 1990; Lundberg et al., 1997; Phipps and Burton, 1998; Glick and Sahn, 2000; Blundell et al., 2005; Cherchye et al., 2012, 2021; Ringdal and Sjursen, 2021). However, it is difficult to identify paternal and maternal preferences separately given that children's choices typically reflect preferences of both parents. This paper tackles this challenge by using the death shock as a source of exogenous variation to estimate a collective household model. The estimated model shows that compared to fathers, mothers have a stronger preference for schooling and a lower bargaining weight in intrahousehold decisions. Specifically, I estimate that the father's bargaining weight is, on average, 50% higher than the mother's. Counterfactual simulations further show that equalizing the bargaining weight of mothers and fathers would lead to increases in school enrollment, especially among girls.

The remainder of the paper proceeds as follows. The next section presents estimates of the causal effects of parental death on school enrollment in India. Section 3 discusses the mechanisms underlying these effects and provides descriptive evidence for their relative importance. Motivated by these findings, Section 4 develops a structural model of household consumption and time allocation whose estimation is presented in Section 5. In Section 6, I use the estimated model to decompose the mechanisms and evaluate various counterfactual policies. The last section offers some concluding remarks.

2 The Causal Effects of Parental Death

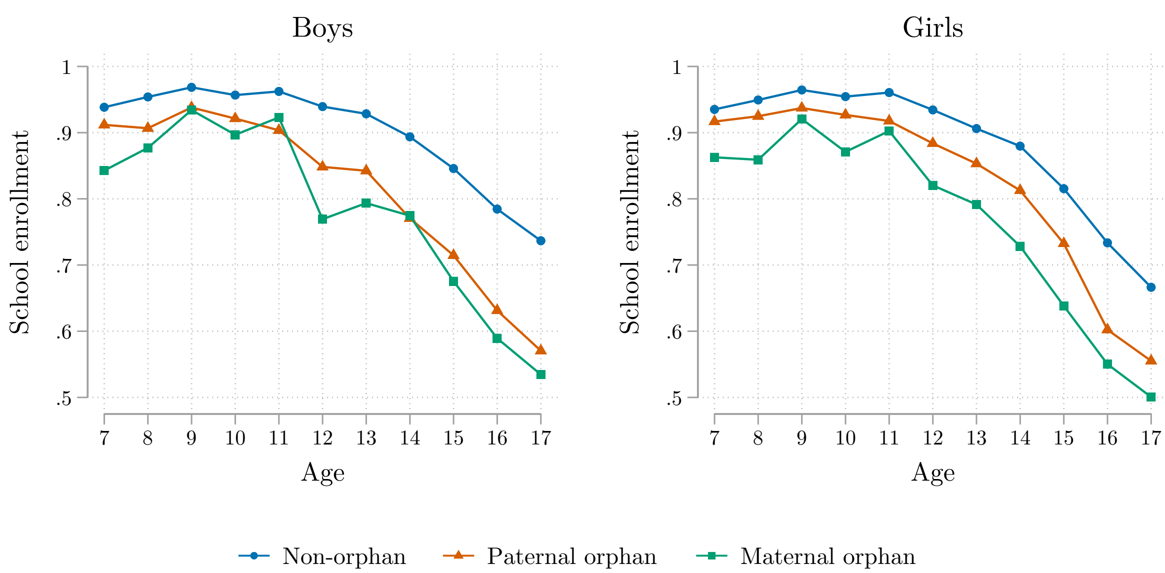
In India, 10% of the population experiences the death of a parent before the age of 18 (NFHS, 2015-2016). Paternal orphans (6.6%) are more common than maternal orphans (2.4%), and double orphans are relatively rare (0.6%). As shown in Figure 2, the death of a parent is more likely to occur during teenage years than during infancy. In fact, nearly half of orphans lose a parent between ages 12 and 18. This is also the age range in which children are more likely to drop out of school. As seen in Figure 3, school enrollment peaks around 95% between ages 9 and 11, after which enrollment drops and the schooling gap between orphans and non-orphans widens. Because parental death is more common among teenagers and schooling is a non-trivial choice for this age group, this study focuses on children aged 12-18.

Figure 2: Proportion of children with deceased parents by age



Note: Sample consists of children from ages 0 to 17 in the NFHS 4 (2015-16).

Figure 3: School enrollment by age and orphan status



Note: Sample consists of children from ages 7 to 17 in the NFHS 4 (2015-16).

The loss of a parent during childhood is often the result of a premature and unexpected death. Between 2014 and 2021, average life expectancy in India was 68 years for men and 71 years for women (United Nations Population Division, 2022). In my sample of orphans, the average age of death is 45 for fathers (25th percentile = 39; 75th percentile = 50) and 39 for mothers (25th percentile = 33; 75th percentile = 44). Self-reported health data indicates that these deaths are largely unexpected. In less than 10% of cases, the deceased parent reported being sick at any point in the year prior to death. Among Indian men who die between ages 30 and 50, the most common causes of death are cardiovascular diseases (24%), digestive diseases (12%), respiratory infections (11%), and transport injuries (10%); for women, these are cardiovascular diseases (21%), neoplasms (19%), respiratory infections (10%), and self-harm and violence (7%) (Global Burden of Disease Collaborative Network, 2019).

2.1 Data

The main data source of this study is India's Consumer Pyramids Household Survey (CPHS), conducted by the Centre for Monitoring the Indian Economy (CMIE). Launched in 2014, the CPHS is a nationally representative[6] longitudinal survey that interviews households every four months. The survey is a panel of residences that does not follow household members when they move. On average, 87.6% of households survive over two waves, 77.5% survive over three waves, and 68.8% survive over four waves. As of June 2023, the CPHS has surveyed over 236,000 households.

The CPHS collects data for each household member on vital status, self-reported health, education, occupation, employment, income, and time use. In addition, it collects household-level data on household income, itemized expenses of food and non-food categories, and assets.

Five features of the survey make the CPHS a particularly attractive source of data for this study. First, the large sample size ensures that I observe a substantial number of households that experience an early parental death. Specifically, I observe nearly 10,000 children before and after the death of a parent. For comparison, this number is 1,800 in the Indian Human Development Survey, another large longitudinal survey of Indian households. Second, the CPHS collects occupation and time use data on all household members.[7] This information provides a window into the shift of roles that takes place after the death of a household member. Third, the availability of itemized expenditure data allows one to examine changes in consumption patterns following the death shock. This is important to understand how death alters intrahousehold bargaining over resources. Fourth,

[6]The CMIE maintains that the survey is nationally representative, yet the CPHS has been criticized for under-representing the poor (Somanchi, 2021). The sampling design is as follows. Within states, districts with similar agroclimatic conditions, urbanization levels, and literacy rates are combined to form 110 homogeneous regions. Each homogeneous region is divided into 5 strata (one rural and four urban). Next, 25-30 villages and 1 town were randomly sampled from the rural and urban strata, respectively. From each sample town, at least 21 census enumeration blocks were randomly sampled. Finally, 16 households are chosen from each sample village and census enumeration block: starting from the main street, every nth household from a street was surveyed, where n was a randomly selected integer between 5 and 15.

[7]Time use data is only asked for household members age 12 and above.

the high frequency of the panel provides a detailed picture of the effects of parental death over time. Finally, with high-frequency data on self-reported health, one can focus on deaths that were largely unexpected.

2.2 The Estimator

What is the impact of parental death on educational choices among children who experience this shock? The estimand of interest is the average treatment effect on the treated (ATT) and the main estimation challenge is that early parental death is not random in the population. This is because the sample of adults with school-aged children who receive a life-threatening shock is not random. For example, individuals living in urban areas are more likely to be involved in a car accident than those living in rural areas. Moreover, the probability of surviving a life-threatening shock is certainly not random, as it depends on wealth, access to healthcare, and health literacy.

In the CPHS, there are 6,591 children who experienced the death of their father at age 18 or below. Hereinafter, these individuals are referred to as paternal orphans. Following an analogous definition, there are 2,557 maternal orphans in the CPHS. Table 1 presents some descriptive statistics at baseline on the samples of paternal orphans, maternal orphans, and non-orphans. Compared to non-orphans, both paternal and maternal orphans are more likely to live in rural ares, they are more likely to live in larger households, and their parents are less educated.

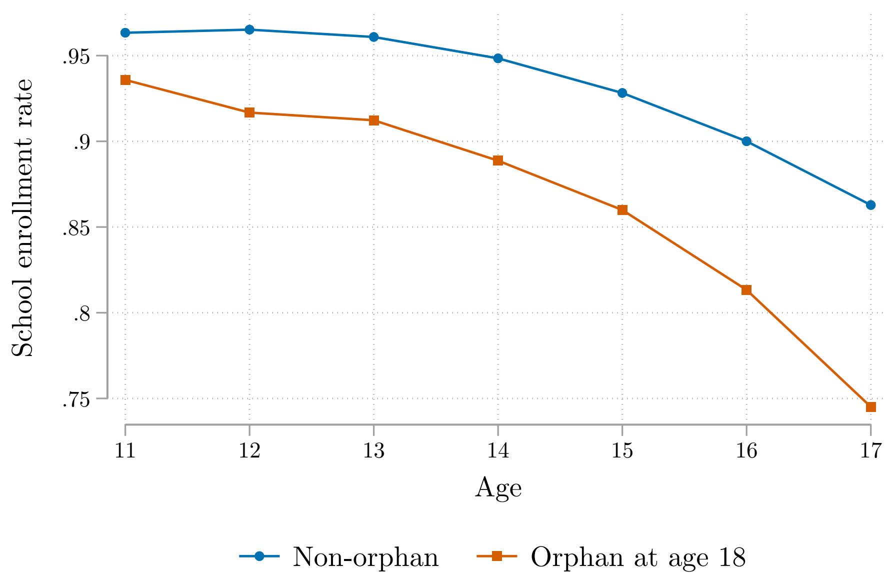
Table 1: Descriptive statistics at baseline

	Paternal orphans mean	Maternal orphans mean	Non-orphans mean
Urban	0.25	0.20	0.27
Upper caste	0.18	0.16	0.19
Household size	5.24	5.40	4.98
Father completed high school	0.14	0.15	0.21
Mother completed high school	0.07	0.07	0.12
Observations	6591	2557	334107

Note: The sample consists of all children (age 0-18) surveyed by the CPHS. The differences in means between paternal orphans and non-orphans are all significant at the 1% level (except for caste). The differences in means between maternal orphans and non-orphans are all significant at the 1% level.

Given the striking differences between orphans and non-orphans at baseline, non-orphans are not an appropriate control group to estimate the impact of parental death. Most studies address this concern by controlling for child and time fixed effects. The identifying assumption is that, in the absence of the shock, the schooling trajectory of orphans and non-orphans would evolve in parallel. However, the analysis of pre-trends reveals that this assumption is likely to be violated. Figure 4 shows school enrollment rate from age 11 to 17 among non-oprhans (blue dots) and among children who lost a father at age 18 (orange squares). The gap in enrollment between non-orphans and orphans widens even years before the shock takes place. This is not surprising given that orphans live in poorer households. While primary school enrollment is close to 100% across the income distribution, poorer children start dropping out school during adolescence.

Figure 4: Non-parallel pre-trends in school enrollment between orphans and non-orphans



Note: The blue dots correspond to children with both parents alive by age 18 and the orange squares correspond to children who lose a father at age 18. Data source: CPHS.

Since the suitability of non-orphans as a control group, either in levels or in differences, is questionable, this study exploits variation in the timing of parental death among orphans only. The identification strategy compares the outcomes of a child before and after parental death, relative to children of the same age observed at the same time, but who will experience the shock in a later wave. Specifically, I estimate the following equation on the sample of orphans,

$$y_{it} = \beta^f d_{it}^f + \beta^m d_{it}^m + \alpha_i + \gamma_t + \sum_{a=12}^{18} \zeta_a a_{it} + \epsilon_{it} \tag{1}$$

where y_{it} is the outcome of child i in period t, and d_{it}^p is an indicator for whether parent $p = \{m, f\}$ is dead.

There are two key identifying assumptions: parallel trends and no anticipation. The first assumption states that in the absence of the shock, the outcomes of orphans would evolve in parallel. To support this assumption, Table A1 in the appendix shows that among orphans, the timing of parental death is uncorrelated with a series of relevant characteristics at baseline. The second assumption states that the treatment effects of parental death do not start to occur before the child experiences the shock. This assumption may be violated in cases where death is the result of a long, debilitating, and costly disease. However, data on self-reported health shows that parental deaths in my samle are largely unexpected. In over 90% of cases, the deceased parent reported being healthy in the last year of his/her life. All of this paper's findings are robust to excluding parental deaths in which the parent reported being sick in the year prior to death.

Several recent studies have shown that in a setting with variation in treatment timing, the coefficients from two-way fixed effects (TWFE) models often do not correspond with an intuitive causal parameter (see Roth et al. (2022) for a survey of this literature). To alleviate this concern, I implement the estimator developed by Callaway and Sant'Anna (2021) and find nearly identical point-estimates. Appendix A describes the details of this estimator.

Table 2 presents the ATT estimates of parental death on school enrollment according to four estimators. Column (1) shows the least-square estimates, controlling for wave fixed effects and age dummies. On average, school enrollment among paternal and maternal orphans is about 12 percentage points lower compared to non-orphans of the same age observed in the same wave.

Table 2: ATT estimates of parental death on schooling

	Full sample		Only orphans	
D.V.: School enrollment	(1)	(2)	(3)	(4)
β^f: Paternal death	-0.13***	-0.12***	-0.07***	-0.07***
	(0.01)	(0.01)	(0.01)	(0.02)
β^m: Maternal death	-0.11***	-0.11***	-0.06***	-0.06*
	(0.01)	(0.01)	(0.01)	(0.03)
p-val: $\beta^f = \beta^m$	0.16	0.52	0.54	
Wave FE	Yes	Yes	Yes	
Child FE	No	Yes	Yes	
Age dummies	Yes	Yes	Yes	
Control mean	0.90	0.90	0.87	0.87
Age range	[12,18]	[12,18]	[12,18]	[12,18]
No. of paternal orphans	4912	4695	4695	4695
No. of maternal orphans	1718	1621	1621	1621
No. of non-orphans	240790	217519	0	0
N	2057252	2033515	74444	2340855
Estimator	FE	FE	FE	CS

Note: The table presents four different estimators of the ATT of parental death on schooling. The sample of columns (1) and (2) include non-orphans, whereas the sample of columns (3) and (4) consists of children who lose a parent by age 18. Standard errors clustered at the household level.

To control for child-specific time averages, column (2) adds child fixed effects. This is the two-way fixed effect estimator from equation (1) estimated on the full sample (including non-orphans). As discussed above, the use of non-orphans as a control group is likely to overestimate the schooling effects of parental death. Thus, column (3) replicates the column (2) specification using only the sample of children who experience a parental death by age 18. The estimates indicate a 7-percentage point drop in school enrollment associated with paternal deaths and a 6-percentage point reduction associated with maternal deaths. Finally, column (4) follows the Callaway and Sant'Anna (2021) estimator, which yields nearly identical point estimates. The remainder of the paper follows the specification in column (3).

3 The Mechanisms

The death of a parent affects educational choices through three mechanisms: household resources, parental inputs into education production, and parental preferences. This section describes these channels and provides descriptive evidence regarding their importance.

3.1 Resources

When a parent dies the household loses a productive household member, which affects the level of household resources. In my sample, 96% of fathers were full-time workers at baseline. Mothers, on the other hand, rarely work for pay in India. Only 12% of mothers in my sample work for pay at baseline. Thus, paternal death constitutes a large negative shock to household income, whereas maternal death does not. Table 3 reports the two-way fixed effects estimates of parental death on monthly household income. It is worth noting that the impact of parental death on household income encompasses two effects: the direct loss resulting from the death of an income-generating member and the subsequent behavioral response of surviving members.

Table 3: Effect of parental death on household income

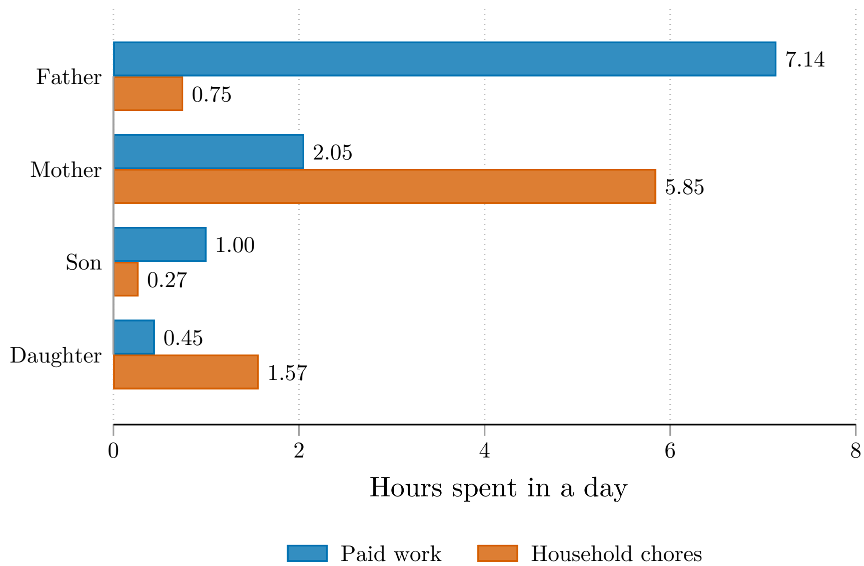
D.V.: Monthly HH income (Rs.)	(1) Total	(2) Wages	(3) Profits	(4) Pensions	(5) Private transfers	(6) Public transfers	(7) Total per capita
β^f: Paternal death	-3018.21***	-2542.32***	-1148.57***	155.69***	356.67***	33.19***	-222.55***
	(291.09)	(216.30)	(245.82)	(38.67)	(86.07)	(10.63)	(61.68)
β^m: Maternal death	-604.85	-300.88	-348.26	62.46	-14.15	-27.87***	140.86
	(414.27)	(344.80)	(323.73)	(49.52)	(62.30)	(10.50)	(87.43)
p-val: $\beta^f = \beta^m$	0.00	0.00	0.01	0.04	0.00	0.00	0.00
Control mean	10815.62	8404.73	1837.45	140.38	141.70	61.11	2127.71
No. of children	7067	7067	7067	7024	7067	7024	7067
N	73865	73865	73865	71674	73865	71674	73865

Note: Sample includes only children that experience parental death by age 18. All regressions include the same controls as those listed in Table 2. Standard errors clustered at the household level.

The death of a father decreases total household income by nearly 30%. This decline is primarily driven by wage income and, to a lesser extent, profits. The drop in wage income and profits is partially attenuated by modest increases in pension income, private transfers, and public transfers. Conversely, the effect of maternal death on household income is not statistically different from zero.

While the loss of a mother does not affect household income, this does not imply that household resources are unaffected by maternal death. In India, as in many other low-income settings, mothers' main contribution to household resources takes the form of domestic production. Figure 5 illustrates marked differences in time allocation across genders. Mothers spend, on average, 6 hours a day performing household chores, whereas fathers spend less than an hour a day in housework. This gender gap is also present in children's time use, with daughters spending six times as much time in housework compared to sons.

Figure 5: Gendered spheres of production in India



Note: Fathers and mothers comprise married individuals aged 30-60. Sons and daughters comprise individuals aged 12-18. Data source: India-TUS data (2019).

Given the strong division of labor across gender lines, the resource mechanism has tight predictions for the effects of parental death: paternal death should induce sons to drop out of school and enter the labor force, whereas maternal death should induce daughters to drop out of school to engage in home production. Table 4 reports the two-way fixed effects estimates of parental death on children's occupation, by the gender of the parent and the child.[8] Appendix C presents additional results on the effects of parental death on household size and composition, as well as on the role of siblings and grandparents in mitigating the schooling effects of parental death.

Consistent with the resource mechanism, paternal loss has a strong effect on the schooling of sons (-11 percentage points), who drop out of school to enter the labor force, as seen in column (2). Interestingly, paternal death also has an effect on daughters. This effect can also be explained by

[8]Table B1 in the appendix reports the estimates based on time-use data, instead of occupation. These estimates are less precise due to a smaller sample size (as time-use data collection in the CPHS commenced only in November 2019) and measurement error in the dependent variable. However, the main findings are qualitatively similar. On average, boys increase their labor supply by 37 minutes per day following the death of their father, while girls increase their time spent on home production by 28 minutes per day following the death of their mother.

Table 4: Effect of parental death on children's occupation

D.V.:	(1) Student	(2) Paid worker	(3) Homemaker
β^{fs}: Paternal death × Boy	-0.11***	0.14***	-0.03***
	(0.01)	(0.01)	(0.01)
β^{fd}: Paternal death × Girl	-0.02*	-0.04***	0.07***
	(0.01)	(0.01)	(0.01)
β^{ms}: Maternal death × Boy	-0.05***	0.05***	-0.00
	(0.02)	(0.01)	(0.01)
β^{md}: Maternal death × Girl	-0.07***	-0.05***	0.12***
	(0.02)	(0.01)	(0.02)
p-val: $\beta^{fs} = \beta^{fd}$	0.00	0.00	0.00
p-val: $\beta^{ms} = \beta^{md}$	0.43	0.00	0.00
p-val: $\beta^{fs} = \beta^{ms}$	0.00	0.00	0.02
p-val: $\beta^{fd} = \beta^{md}$	0.05	0.54	0.03
Control mean: Boys	0.87	0.05	0.06
Control mean: Girls	0.87	0.00	0.11
No. of children	7083	7083	7083
N	74444	74444	74444

Note: Sample includes only children that experience parental death by age 18. All regressions include the same controls as those listed in Table 2. Standard errors clustered at the household level.

changes in household resources once we consider the behavioral response of the surviving mother. The death of a father increases the probability that the widowed mother works for pay by 37 percentage points (see Table B2). Thus, the effect of paternal death on daughters' schooling may be attributed to daughters assuming additional domestic responsibilities as they assist their working mothers.

Turning to the effects of maternal death, daughters experience a 7-percentage point decline in school enrollment and they increase their participation in home production, as predicted by the resource mechanism. Interestingly, sons drop out of school at a similar rate (5 percentage points) to engage in paid work. In fact, the schooling effects of maternal death on sons is not statistically different from the effect on daughters. This result cannot be fully explained by changes in household resources following maternal death. While there is an income effect associated with the loss in home production, household expenditure data suggests that this channel is not quantitatively meaningful. For instance, maternal death does not increase household expenses on domestic help.[9]

In conclusion, household resources could potentially explain most, but not all, of the effects of parental death on school enrollment. Specifically, the large effect of maternal death on sons cannot be attributed to changes in household resources. Thus, the discussion now turns to other mechanisms.

[9] Less than 10% of households in my sample hire domestic help.

3.2 Parenting

Children enroll in school to acquire human capital. However, schooling is not the only input in the production of human capital. Parents may also play an important role educating their children. In fact, the child development literature has found that parental time investments at an early age are a key input in the production of human capital (Del Boca et al., 2014, 2017; Attanasio et al., 2022; Wong, 2023). When a parent dies, the child loses this input and so the optimal choice of schooling may change in response. This behavioral response may either compensate or reinforce the loss of the parental input.

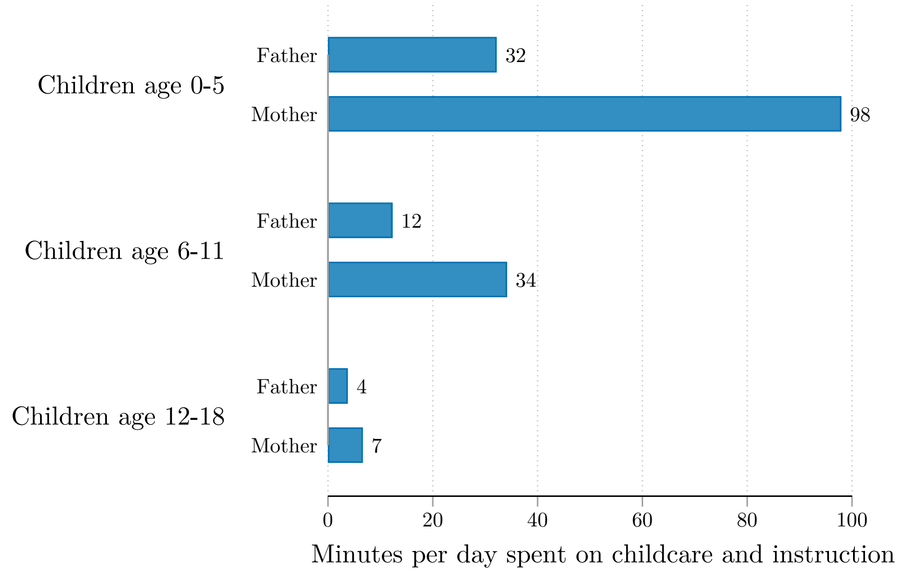
To evaluate the importance of the parenting mechanism, I first examine the amount of time that parents invest in their children. Figure 6 shows the average minutes per day that Indian mothers and fathers spend on "child care and instruction" broken up by the age of their children.[10] I divide parents into three mutually exclusive categories based on the age of their children (parents with children in more than one category are excluded from the sample). The figure shows that parents of children under the age of 6 invest a non-trivial amount of time in their children. On average, mothers spend 100 minutes a day, while fathers spend 30 minutes a day. Nevertheless, parental time inputs look very different among children aged 12-18. On average, parents with children in this age range spend a negligible amount of time on childcare and instruction. Thus, data on parents' time use suggest that the schooling effects of parental death on children age 12-18 are unlikely to be explained by the loss of parental time investments.

While parental time investments are an important aspect of parenting, there are other ways in which parents may contribute to the production of human capital. Parents also provide advice and guidance, and these contributions are unlikely to be captured by a time-use survey. For instance, at dinner, a parent may ask the child how she is doing in school and provide advice. In addition, parenting may take the form of passive supervision. For example, the presence of a mother at home may incentivize children to do their homework, even if the mother does not provide any help or active supervision. Therefore, even if a parent does not actively invest time in the child, his/her death may have a direct impact on the child's human capital. This direct effect of parental death on human capital would also incorporate the psychological toll of losing a parent, which may reduce the productivity of all inputs in the production of human capital.

To examine the direct effect of parental death on human capital, I estimate a human capital production function using test score data from the Young Lives - India panel. Briefly, the production function takes as inputs the child's lagged test score, his/her school enrollment status, and the vital status of his/her parents. I estimate this production function taking into account the endogeneity of inputs and the measurement error associated with test scores. The estimates indicate that parental

[10] Child care and instruction consists of caring for children (including feeding, cleaning, physical care, medical care); instructing, teaching, training, and helping children; talking with and reading to children; playing games and sports with children; minding children (passive care); meetings and arrangements with schools and child care; other activities related to childcare and instruction.

Figure 6: Time spent on childcare and instruction by age of children



Note: Fathers and mothers comprise married individuals aged 30-60. Categories of children composition are mutually exclusive (sample excludes households with children in multiple categories).
Data source: India-TUS data (2019).

death has a large impact on test scores, but that effect only operates through changes in school enrollment. That is, conditional on school enrollment, parental death appears to have no significant impact on academic performance. The full analysis is presented in Appendix D.

3.3 Preferences

To the extent that parental preferences for education influence children's educational choices, the death of either parent will affect these choices. This channel is especially important if mothers and fathers have different preferences. For example, if mothers care more about education than fathers, then we may expect declines in schooling following maternal death. The difference in preference for education between mothers and fathers has been emphasized in academic and policy discussions (Thomas, 1990; Lundberg et al., 1997; Phipps and Burton, 1998; Glick and Sahn, 2000; Blundell et al., 2005; Emerson and Souza, 2007; Cherchye et al., 2012; Ringdal and Sjursen, 2021).

Though it has received considerably less attention, there is an additional channel through which parental preferences affect the relationship between family structure and educational choices. Even if parents had identical preferences, the shift from a two-parent household to a one-parent household alters the tradeoff between investment in children and private consumption. Intuitively, in a two-parent household, every rupee invested in children benefits both parents, and so parents get "two for the price of one" when spending on children. This effect is not present in one-parent households.

Formally, children's education is often modeled as a public good in a two-parent household. That is, children's education is non-excludable and non-rivalrous in the sense that both parents fully accrue the utility gains from any household resources invested in the child. This stands in contrast to private goods such as food and clothing, where parents cannot fully share consumption due to the rival nature of these goods. Thus, when a parent dies, the children's education becomes a private good.[11] To illustrate the consequences of this change, consider a simplified version of the model, in which household income (y) is allocated between private consumption (c_m, c_f) and public consumption (k):

$$\max_{c_m, c_f, k} \mu u_m(c_m, k) + (1-\mu) u_f(c_f, k) \quad \text{s.t.} \quad c_m + c_f + k \leq y. \tag{2}$$

Optimality requires that at an interior solution,

$$\frac{\frac{\partial u_m}{\partial k}}{\frac{\partial u_m}{\partial c_m}} + \frac{\frac{\partial u_f}{\partial k}}{\frac{\partial u_f}{\partial c_f}} = 1, \tag{3}$$

which states that the sum of the mother's and the father's marginal rate of substitution must equal the price of public good. This is the Samuelson condition for optimal provision of public goods.

Now, if the father dies ($\mu = 1$), the optimality condition becomes

$$\frac{\frac{\partial u_m}{\partial k}}{\frac{\partial u_m}{\partial c_m}} = 1. \tag{4}$$

Comparing equations (3) and (4) offers insights into the effect of parental death on the allocation of private and public goods. At the couple's optimal allocation, the mother's marginal rate of substitution is less than one, and so it is lower than her marginal rate of substitution at the widow's optimal bundle. Thus, given enough income to attain the level of utility she experienced while married, a widow spends less on the public good compared to the couple.

Thus, the preference mechanism first predicts a decline in school enrollment for both sons and daughters following the death of either parent. This is because the shock alters the tradeoff between private goods and children's education. However, if mothers and fathers have different preferences for education, then the implications for school enrollment are ambiguous. In an extreme case, parental death may increase school enrollment if the surviving parent places a considerably higher weight on education compared to the deceased parent.

To sum up, the descriptive evidence presented in this section suggests an important role for resources and preferences in explaining the schooling effects of parental death. The next section develops a model of household consumption and time allocation that allows for the presence of

[11] The idea that children are viewed as a public good in a two-parent household and as a private good in a one-parent household was first formalized by Weiss and Willis (1985) in the context of divorce.

these two mechanisms.

4 Model

To formalize and quantify the mechanisms underlying the impact of parental death on schooling, I develop a household model in which parents bargain over consumption and time allocation choices.

Household structure

A household of size N consists of some or all of the following key members: a mother (m), a father (f), a son (s) aged 12-18, and a daughter (d) aged 12-18. In the data, these key household members often live with other individuals. For instance, 60% of households in the sample have children or children-in-law over the age of 18, 30% have children under the age of 12, and 15% have grandparents. In the estimation, I allow for the presence of these other household members in the sense that they contribute to household resources and consume a fraction of these resources. However, I do not model their choices. For ease of exposition, I present the model with only four key household members, and discuss the role played by other members in the estimation section.

Time allocation

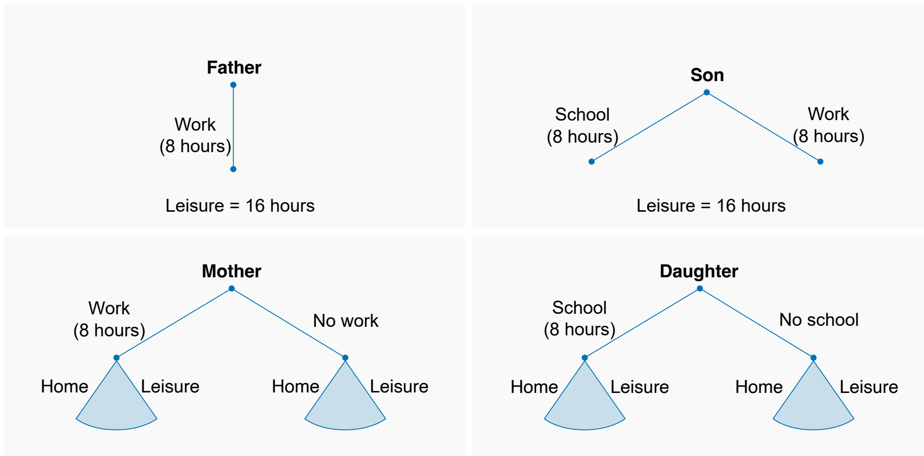
Figure 7 illustrates the time allocation choices available to each household member. The father always works full time and spends the remaining 16 hours of the day on leisure. Thus, there is no time allocation choice for the father. The son may work full time or attend school full time, and he spends the remaining 16 hours of the day on leisure. Motivated by the time-use evidence in Figure 5, neither the father nor the son perform household chores. The mother may work full time and divides the rest of her time between household chores and leisure. Similarly, the daughter may attend school full time and divides the rest of her time between household chores and leisure.

Note that the labor supply and schooling choices are modeled at the extensive margin, whereas the domestic labor supply choice is modeled at the intensive margin. This modeling choice is motivated by the data. Conditional on labor force participation and school enrollment, there is not much variation in time spent on paid work and education. In contrast, there is substantial variation among mothers and daughters in time spent on domestic chores. Moreover, the data suggest that paid work and school are mutually exclusive choices for boys. On the other hand, mothers that engage in paid work also spend time on domestic chores. Similarly, daughters who are enrolled in school often perform household chores.

Home production

Time-use data shows that married women in India spend, on average, six hours a day performing household chores (see Figure 5). This large time investment produces domestic commodities that the household values, such as meals, clean laundry, and a clean home. In the model, these domestic commodities are represented by a composite home-produced good, h. To produce h, the household combines time investments with intermediary goods purchased in the market. Uncooked food

Figure 7: Time use of household members



Note: The figure depicts the possible time allocation choices in the model. The model captures four distinct features of time allocation in Indian households: (i) men do not participate in home production, (ii) daughters never work in the market, (iii) market labor supply and school enrollment are mutually exclusive choices made at the extensive margin, and (iv) domestic labor supply choices are made at the intensive margin.

ingredients are the most important intermediary goods into home production. In fact, household in my sample spend, on average, half of their total monthly budget on uncooked food. Thus, households in the model face a home production function that takes as inputs the time investment of the mother and the daughter (t_m and t_d, respectively) and a vector of uncooked food ingredients ($\mathbf{r}$),

$$h = h\left(t_m, t_d, \mathbf{r}\right). \tag{5}$$

Preferences

The preferences of parent $i \in \{m, f\}$ can be represented by a utility function (u_i) defined over private consumption of a vector of market goods ($\mathbf{c}_i$) and consumption of seven non-market goods: the home-produced good (h), the children's school enrollment status (e_s and e_d), and the leisure of all four members ($\mathbf{z} = (z_m, z_f, z_s, z_d)$). I allow parents to care about their spouse's leisure for two reasons. First, it captures the fact that spouses may enjoy each other's company. Second, it is a way of modeling any social status considerations that may affect the father if his wife works.

Following the collective household literature (Chiappori, 1988, 1992), I assume that the allocation of resources is Pareto efficient. The mother is assigned a Pareto weight (μ) that reflects the relative weight of her utility in the household's social welfare function. Unlike most applications of the collective household model, it is assumed that the Pareto weights do not depend on current conditions such as the offered wage of each spouse. Instead, the Pareto weight is determined by conditions at the time of marriage ($\mathbf{d}$), such as the age and educational attainment of each spouse. I argue that a model with full commitment and no renegotiation is more suitable for a setting like India, where divorce is exceedingly rare and often stigmatized (Jacob and Chattopadhyay, 2016).

Household problem

The household chooses the vectors of private consumption, inputs into home production, and levels of the non-market goods that maximize its social welfare function, subject to the budget constraint, the production technology, and time constraints. That is, the household solves

$$\begin{aligned} \max_{\mathbf{c}_m,\mathbf{c}_f,e_s,e_d,t_d,l_m,t_m,\mathbf{r}} \; & \mu(\mathbf{d})\, u_m(\mathbf{c}_m, h, e_s, e_d, \mathbf{z}) + (1-\mu(\mathbf{d}))\, u_f(\mathbf{c}_f, h, e_s, e_d, \mathbf{z}) \qquad (6) \\ \text{s.t. } \; & \mathbf{p}'\mathbf{A}(\mathbf{c}_m + \mathbf{c}_f) = (1-\phi(N))\left[w_m l_m + w_s(1-e_s) + y - \psi(e_s+e_d) - \mathbf{p}_r'\mathbf{r}\right], \\ & h = h(t_m, t_d, \mathbf{r}), \end{aligned}$$

and subject to each member's time constraint. Total household income comprises the mother's labor income ($w_m l_m$), the son's labor income ($w_s(1-e_s)$), and what I will refer to as "predetermined income" (y), which is the sum of the father's labor income and non-labor income. Household income is spent on schooling ($\psi(e_s+e_d)$), uncooked food ingredients at prices $\mathbf{p}_r$, and final market goods at prices $\mathbf{p}$. The consumption of children is modeled as a tax, specified by the function ϕ. Parents then use the residual income to allocate their private consumption.

The diagonal matrix $\mathbf{A}$ captures economies of scale associated with living as a couple (Barten, 1964). Economies of scale arise because some of the final market goods can be shared or jointly consumed by the two parents. Thus, for shareable goods, the sum of the parents' consumption is greater than the purchased quantity. Specifically, the purchased quantity of good k is equal to the total consumption of good k multiplied by $a_{kk} \in [0.5, 1]$, where $a_{kk} = 0.5$ for fully shareable goods and $a_{kk} = 1$ for nonshareable goods. It is important to model the benefits of scale economies because these benefits are lost upon the death of either parent.

Parental death in the model

Parental death affects the household problem through changes in resources and through shifts in household preferences. Household resources change following the death of a parent for two reasons. First, when a parent dies the household loses that parents' entire time endowment. In the case of paternal death, this implies that the father's labor supply is constrained to zero, which results in a decrease in y. Similarly, maternal death constrains the mother's labor supply to zero, this includes her market labor supply (l_m) as well as her domestic labor supply (t_m). Second, the household loses economies of scale. Specifically, the matrix $\mathbf{A}$ is replaced by the identity matrix.

The death of a parent also affects the household problem through preferences, by changing the objective function. Specifically, the Pareto weight of the deceased parent drops to zero, giving the surviving parent full control over household resources. This assumption is more flexible than it may initially appear. It does not necessarily exclude scenarios where widows or widowers take into account the preferences of their deceased spouse or other family members. These possibilities can be accommodated by reinterpreting the utility functions, u_i. For instance, u_i could incorporate altruistic concerns towards one's spouse or the influence of other family members, as long as u_i

remains unaffected by widowhood. This assumption would be violated if, for example, the extended family were to exert a stronger influence upon the death of one of the spouses.

Setting the Pareto weight of the deceased parent to zero fundamentally changes the household's objective function, especially if mothers and fathers have different preferences. However, even if parents have identical preferences, death will alter the optimal allocation between private and public goods, as discussed in Section 3.3.

4.1 A three-stage representation

The solution to the household problem can be thought of as a three-stage process. This representation decentralizes the social planner problem in (6) into a competitive equilibrium in the household economy (after suitable lump-sum transfers between the parents). The three-stage representation is convenient for the identification argument in the next subsection, and it will guide the estimation strategy in the following section.

The three-stage process is as follows. In stage 1, the household makes time allocation and home production decisions. Specifically, it chooses the school enrollment status of each child, the market labor supply of the mother and the son, the domestic labor supply of the mother and the daughter, and the bundle of uncooked food ingredients. In stage 2, the parents agree upon a division of residual income. This division, which is conditional on the first-stage choices, is known as the "conditional sharing rule" (Blundell et al., 2005). In the final stage, each parent independently chooses a vector of market goods, conditional on the non-market goods chosen in stage 1 and subject to their individual shadow budget allocated in stage 2. The three-stage problem is solved backwards, as described below.

In the third and final stage, parent $i \in \{m, f\}$ chooses a vector of market goods ($\mathbf{c}_i$), subject to market prices ($\mathbf{p}$) and shadow income ($\frac{\eta_i^*}{\Delta_i} X^*$), where X^* is residual household income from the first stage, η_i^* is the sharing rule agreed upon in the second stage, and Δ_i captures the savings due to economies of scale.[12] Each parent takes as given the first-stage choices, $\mathbf{g}^* = (h^*, e_s^*, e_d^*, \mathbf{z}^*)$, and solves the third-stage problem,

$$\max_{\mathbf{c}_i \geq 0} u_i\left(\mathbf{c}_i, \mathbf{g}^*\right) \quad \text{s.t.} \quad \mathbf{p}'\mathbf{c}_i \leq \frac{\eta_i^*}{\Delta_i} X^*. \tag{7}$$

In the second stage, the couple agrees on a division of residual income. Define $V_i\left(X_i, \mathbf{p}; \mathbf{g}^*\right)$ to be the conditional indirect utility function, which is the value of the third-stage problem (7) as a

[12] In the planner's problem (6), the matrix $\mathbf{A}$ captures the economies of scale associated with each market good. For identification purposes, however, it is convenient to summarize the economies of scale by scalars (Δ_m and Δ_f) that inflate the income of each parent in the decentralized problem. This allows identification of the sharing rule without price variation and reduces the computational complexity of the estimation procedure (Lewbel and Pendakur, 2008). Appendix E discusses the assumptions required for the equivalence between these two representations of scale economies.

function of prices, individual income, and non-market goods. The second-stage problem is

$$\begin{aligned} &\max_{\eta_m,\eta_f} \mu(\mathbf{d}) V_m\left(\frac{\eta_m}{\Delta_m}X^*, \mathbf{g}^*\right) + (1-\mu(\mathbf{d})) V_f\left(\frac{\eta_f}{\Delta_f}X^*, \mathbf{g}^*\right) \\ &\qquad\text{s.t. } \eta_m + \eta_f = 1. \end{aligned} \tag{8}$$

In the first stage, the household makes time allocation and home production decisions. When doing so, each parent understands that these choices will have an an impact on the level of residual income and the share of this income that each will receive. The first-stage problem is

$$\begin{aligned} \max_{e_s,e_d,t_d,l_m,t_m,r} \mu(\mathbf{d}) V_m&\left(\frac{\eta_m(h,e_s,e_d,\mathbf{z})}{\Delta_m}X, h, e_s, e_d, \mathbf{z}\right) + (1-\mu(\mathbf{d})) V_f\left(\frac{\eta_f(h,e_s,e_d,\mathbf{z})}{\Delta_f}X, h, e_s, e_d, \mathbf{z}\right) \\ \text{s.t. } X &= (1-\phi(N))\left[y + w_m l_m + w_s(1-e_s) - \psi(e_s+e_d) - \mathbf{p}_r'\mathbf{r}\right], \qquad (9) \\ h &= h(t_m, t_d, \mathbf{r}), \end{aligned}$$

and subject to each member's time constraint.

4.2 Identification

Is it possible to recover the underlying structure of the model—namely, individual preferences, the Pareto weights, economies of scale, and the production function—from observed behavior? Identification of this model is complicated by the fact that the the individual consumption vectors ($\mathbf{c}_i$) and the level of the home-produced good (h) are unobserved. This section shows that under some assumptions, it is possible to recover the underlying structure of the model from observation of itemized household-level expenses and individual time allocations.

Identification of the home production function

The main challenge in identifying the production function is that its output (i.e. the home-produced good) is fundamentally unobserved. Despite this challenge, it is possible to learn about the underlying technology because of the constraints it imposes on the optimizing behavior of households. Specifically, identification relies on examining the choice of inputs as a function of input prices.

The key assumption, implicit in the household's problem (6), is that home production decisions are efficient. Thus, the choice of inputs must be cost-minimizing,

$$(t_m^*, t_d^*, \mathbf{r}^*) \in \arg\min_{t_m,t_d,\mathbf{r}} \{\tilde{w}_m t_m + \tilde{w}_d t_d + \mathbf{p}_r'\mathbf{r} \quad | \quad h(t_m,t_d,\mathbf{r}) \geq \bar{h}\}, \tag{10}$$

where $\tilde{w}_i$ is the price of time of member i. The first-order conditions of (10) require that for any two inputs, the marginal rate of technical substitution must equal the ratio of input prices. These optimality conditions form a system of partial differential equations that can be integrated to recover the production function. Thus, if $(t_m, t_d, \mathbf{r})$ are observed as functions of $(\tilde{w}_m, \tilde{w}_d, \mathbf{p}_r)$, then one can recover h up to a strictly increasing transformation. Note that this strictly increasing transformation yields the same observable behavior and so it only a matter of normalization.

In the empirical application, the price of time is only observed for those individuals working for a wage. Thus, I estimate the home production function using households with married working women and correct for sample selection using the Heckman correction procedure.

Identification of the sharing rule, scale economies, and Pareto weights

The key assumption to achieve identification of the rest of the model is that individual preferences, represented by $u_i(\cdot)$, do not change as a consequence of widowhood. Note that this is an assumption on the stability of the the utility *function*, and it does not imply that an individual's utility *level* is unaffected by the death of his/her spouse. In fact, recall that individuals in the model value their spouse's leisure, and so there is a direct loss in utility associated with widowhood. Moreover, to the extent that preferences for other goods are non-separable from the spouse's leisure, the loss of a spouse may affect the marginal rate of substitution between goods. The assumption does, however, rule out situations such as the following: After losing her husband, a mother cares more about the education of her children because, as a widow, she is more likely to rely on them for old-age support.

The identification result, which follows the methodology developed by Browning et al. (2013) and Lewbel and Pendakur (2008), relies on the decentralized representation of the household's problem. The first step of the argument is that under the assumption that preferences do not change upon widowhood, the preferences for market goods of mothers and fathers can be identified from the demand functions of widows and widowers, respectively. Note that the third-stage problem in (7) gives rise to a system of conditional demand functions for market commodities (Pollak, 1969; Browning and Meghir, 1991). Importantly, total market expenditures and the non-market goods may be endogenous in the budget share equations and so they ought to be instrumented for (Browning and Meghir, 1991). In the empirical application, I assume separability between market and non-market goods (i.e. $u_i(\mathbf{c}_i, \mathbf{g}) = \tilde{u}_i\left(U_i^1(\mathbf{c}_i), U_i^2(\mathbf{g})\right)$), which implies an unconditional demand system.[13]

Having identified the preferences of mothers and fathers for market goods, the second step of the argument is that the sharing rule and scale economies can be identified by comparing the demand functions before and after the death shocks. Specifically, using the decentralized representation of the household's problem, Browning et al. (2013) show that the couple's demand functions for market commodities can be written in terms of the demand functions of individuals, adjusted to reflect the sharing rule and scale economies. By assuming that scale economies can be summarized through a

[13] It is possible to relax the non-separability assumption. For instance, one may allow for non-separability between market commodities and the home-produced good. The conditional system would then be estimated via GMM, using the price of raw food as an instrument for the home-produced good. Although this approach is theoretically appealing, it increases the computational complexity of the sharing rule estimation, leading to nonconvergence issues.

scalar-valued function, Lewbel and Pendakur (2008) recast this framework from a demand system to an Engel curve system that is considerably easier to estimate and does not require price variation. Extending their result to the conditional case, the couple's conditional budget share Engel curve for commodity n can be written as a function of the conditional budget share Engel curve of singles (proof in Appendix E):

$$\omega^{n}\left(\ln X, \mathbf{g}^{*}\right)=\eta_{m}\left(\mathbf{g}^{*}\right)\left[\omega_{m}^{n}\left(\ln \left(\frac{\eta_{m}\left(\mathbf{g}^{*}\right)}{\Delta_{m}\left(\mathbf{g}^{*}\right)} X\right), \mathbf{g}^{*}\right)+\delta_{m}^{n}\left(\mathbf{g}^{*}\right)\right]$$
$$+\eta_{f}\left(\mathbf{g}^{*}\right)\left[\omega_{f}^{n}\left(\ln \left(\frac{\eta_{f}\left(\mathbf{g}^{*}\right)}{\Delta_{f}\left(\mathbf{g}^{*}\right)} X\right), \mathbf{g}^{*}\right)+\delta_{f}^{n}\left(\mathbf{g}^{*}\right)\right]. \tag{11}$$

Lewbel and Pendakur (2008) prove that η, Δ_i, and δ_i are nonparametrically identified from knowledge of $\boldsymbol{\omega}(\cdot)$, $\boldsymbol{\omega}_m(\cdot)$, and $\boldsymbol{\omega}_f(\cdot)$, as long as some of the goods have budget shares that are nonlinear and are sufficiently different across individuals.

Lastly, given a particular cardinalization of the conditional indirect utility function (V_i), the Pareto weights can be recovered from the first-order conditions of problem (8):

$$\mu=\frac{\frac{\partial V_f}{\partial X_f}\frac{1}{\Delta_f}}{\frac{\partial V_m}{\partial X_m}\frac{1}{\Delta_m}+\frac{\partial V_f}{\partial X_f}\frac{1}{\Delta_f}}. \tag{12}$$

Identification of preferences for non-market goods

Having identified the production technology, the bargaining weights, and the economies of scale, the last step is to identify the gender-specific preferences for non-market goods. Conditional on household resources, the effects of paternal death and maternal death on the optimal choice of non-market goods identifies the preferences of mothers and fathers, respectively. In the empirical application, identification relies on the staggered difference-in-differences strategy from Section 2, together with variation in pre-determined income.

A common theme throughout this section is that identification relies on comparing the choices made by widows, widowers, and couples. It is important to note that given the panel structure of the data, these comparisons are not made across different samples of households. In Section 2 I argued that non-orphans are a bad control group to study the causal effect of parental death. Following the same logic, the structural estimation only uses the sample of households in which a child age 12-18 is observed before and after parental death. That is, the households of "couples" are households who will experience a mortality shock at a later wave in the panel.

5 Estimation

Estimation proceeds in three steps. In the first step, I estimate the home production function and predict the offered wages for every mother and son in the sample. The second step uses itemized expenditure data to estimate the sharing rule and scale economies. In the final step, the remaining parameters are estimated via the method of simulated moments (McFadden, 1989; Pakes and Pollard, 1989).

5.1 Estimation "outside the model"

Home production function

Assume the home production function takes a nested CES form,

$$h = A\left(\exp\left(\xi\right)\gamma t^{\rho} + \left(1-\gamma\right)r^{\rho}\right)^{\frac{1}{\rho}}, \tag{13}$$

$$t = \left(t_m^{\tau} + t_d^{\tau} + t_g^{\tau}\right)^{\frac{1}{\tau}}, \tag{14}$$

$$r = A_1\left(\sum_{l=1}^{L}\exp\left(\xi_l\right)\gamma_l r_l^{\rho_1}\right)^{\frac{1}{\rho_1}}, \tag{15}$$

where t and r are the aggregated time and uncooked food inputs into home production, respectively. In the data, households may have a grandmother or a daughter-in-law that contributes to home production. I assume that the household takes these time inputs, denoted by t_g, as given. Lastly, ξ is a household-specific shock to the productivity of time in home production, and the ξ_l account for idiosyncratic preferences for specific food commodities.

As discussed in the previous section, the main challenge in estimating the home production function is that its output is unobserved. Specifically, I do not observe (h, t, r); in fact, the only observables in equations (13)-(15) are $(t_m, t_d, t_g; r_1, ..., r_L)$. Thus, estimation relies on the optimal choice of inputs as a function of input prices.

First, I estimate the parameters of the CES aggregator of uncooked food commodities, γ_l and ρ_1, using itemized food expenditure data from the CPHS and a state-month panel of food commodity price data from the World Food Programme. This exercise delivers a composite uncooked food commodity, r, and a price index for this composite good, p_r. Appendix F.1 describes the estimation procedure and reports the parameter estimates.

Second, I estimate the outer nest of the CES production function. Cost minimization requires that the marginal rate of technical substitution must equal the ratio of input prices. This optimality condition can be rewritten as

$$\ln\left(\frac{t}{r}\right) = \underbrace{\frac{1}{1-\rho}\ln\left(\frac{\gamma}{1-\gamma}\right)}_{b} + \underbrace{\frac{1}{1-\rho}}_{\sigma}\ln\left(\frac{p_r}{\tilde{w}}\right) + \xi, \tag{16}$$

where r and p_r are obtained from the estimation of the inner nest. Ideally, t and $\tilde{w}$ would be obtained in a similar fashion, by estimating the CES aggregator of time inputs. This would require observing the time allocation of multiple women in the household as a function of their wages. However, only a few households in the data have multiple women working for a wage. To get around this issue, I estimate the outer nest of the CES production function using data from households in which the mother is the only woman over the age of 12. In such households, $t = (t_m^\tau + 0^\tau + 0^\tau)^{\frac{1}{\tau}} = t_m$ and $\tilde{w} = \tilde{w}_m$. Thus, the substitution parameter, τ, is unidentified at this point and will be estimated as part of the method of simulated moments.[14]

The main difficulty with estimating (16) is that $\tilde{w}_M$ is only observed for households in which the mother works for a wage (7% of households). This selected sample is more likely to include women with low draws of ξ and exclude women with high draws of ξ. To address this issue I implement Heckman's selection correction for the case of missing data on explanatory variables, where predetermined income (y) and number of children shift the labor force participation choice. As one would expect, the estimates suggest a low elasticity of substitution between time and uncooked food in home production ($\sigma = 0.27$). The details of the estimation procedure and the parameter estimates are presented in Appendix F.1.

Offered wages

When the household chooses the time allocation of each member, it takes the offered wage of the mother and the son as given. In the data, offered wages are only observed for working individuals. Thus, I need to estimate wage equations and predict the offered wage for each mother and son in the sample.

I estimate wage equations using data from wage workers and correct for sample selection using Heckman's correction procedure. Logged offered wages of women are modeled as a linear function of demographic characteristics and the average male wage in the district.[15] Selection into work depends on these variables as well as on predetermined household income and number of children. The model is estimated using adult women from all the CPHS sample (not only those in households with parental death).[16]

Logged offered wages of sons are modeled in a similar fashion, except that educational attainment is excluded from the model because it is an endogenous choice. The model is estimated using data from men between the ages 12 and 18. The selection equation is estimated with data from non-orphans only.

[14] An alternative approach would be to assume perfect substitutability in time inputs (i.e. $\tau = 1$). In that case, the price of time of one woman pins down the price of time of all women in the household who participate in home production (assuming interior solutions). However, allowing for imperfect substitutability improves the model fit in the method of simulated moments.

[15] The demographic characteristics that enter the wage equation are age, age squared, caste, education, rurality, state, and wave.

[16] Since predetermined household income is correlated with widowhood, which is not random, the selection equation is estimated with data from married women only.

5.2 Sharing rule and scale economies

Estimation of the sharing rule and scale economies is achieved by comparing the Engel curves for market goods before and after the parental death shock. Recall that the couple's budget share Engel curve for commodity n can be written as a function of the budget share Engel curve of singles, as previously stated in equation (11). Thus, I first estimate the individual Engel curves, $\omega_i(\cdot)$, using data from widows and widowers, and then I feed those estimates into equation (11) to recover the sharing rule (η_i) and the economies of scale parameters (Δ_i and δ_i^n).

Following Browning et al. (2013) and Lewbel and Pendakur (2008), preferences are parameterized so as to deliver a quadratic almost ideal demand system (Banks et al., 1997). Specifically, the conditional indirect utility function of parent i is assumed to take the form

$$V_i\left(X_i, \mathbf{g}\right) = \left[\left(\frac{\ln X_i - \bar{a}_i}{\bar{b}_i}\right)^{-1} + \bar{\lambda}_i\right]^{-1} + U_i^2\left(\mathbf{g}\right). \tag{17}$$

This parameterization assumes additive separability between market goods and non-market goods in the utility function. That is, it assumes that the marginal rate of substitution between items like clothing and transportation is unaffected by children's school enrollment, levels of leisure, and the quantity of the home-produced good. This assumption simplifies the estimation procedure because it implies that the budget share chosen for each market commodity only depends on total market expenditures, X, and not on the level of the public goods. Furthermore, the separability assumption also implies that the sharing rule is independent of the level of the public goods, since the second stage of the couple's decentralized problem (8) simplifies to

$$\begin{aligned} &\max_{\eta_m, \eta_f} \mu V_m\left(\frac{\eta_m}{\Delta_m} X^*\right) + (1-\mu)\, V_f\left(\frac{\eta_f}{\Delta_f} X^*\right) + \mu U_m^2\left(\mathbf{g}^*\right) + (1-\mu)\, U_f^2\left(\mathbf{g}^*\right) \\ &\quad \text{s.t. } \eta_m + \eta_f = 1. \end{aligned} \tag{18}$$

The specified conditional indirect utility function yields a quadratic almost ideal demand system (QUAIDS), with Engel curves that are quadratic in log expenditure. Adding demographics ($\mathbf{d}_i$) and error terms (ϵ_i^n), the budget share Engel curve for good n is

$$\omega_i^n\left(\ln X_i, \mathbf{d}_i\right) = \beta_{0i}^n + \boldsymbol{\beta}_{0di}^{n\prime}\mathbf{d}_i + \left(\ln X_i - \boldsymbol{\beta}_{3di}'\mathbf{d}_i\right)\beta_{1i}^n + \left(\ln X_i - \boldsymbol{\beta}_{3di}'\mathbf{d}_i\right)^2 \beta_{2i}^n + \epsilon_i^n. \tag{19}$$

I estimate the system (19) using post-death data, separately for widows and widowers. Next, I feed the estimated β's into the budget share Engel curve system of couples, which is given by

$$\begin{aligned}\omega^{n}\left(\ln X, \mathbf{d}_m, \mathbf{d}_f\right) = {} & \eta_m\left(\mathbf{d}_m, \mathbf{d}_f\right)\left[\omega_m^n\left(\ln\left(\frac{\eta_m\left(\mathbf{d}_m, \mathbf{d}_f\right)}{\Delta_m\left(\mathbf{d}_m\right)}X\right), \mathbf{d}_m\right) + \delta_m^n\left(\mathbf{d}_m\right)\right] \\ & + \eta_f\left(\mathbf{d}_m, \mathbf{d}_f\right)\left[\omega_f^n\left(\ln\left(\frac{\eta_f\left(\mathbf{d}_m, \mathbf{d}_f\right)}{\Delta_f\left(\mathbf{d}_f\right)}X\right), \mathbf{d}_f\right) + \delta_f^n\left(\mathbf{d}_f\right)\right] + \epsilon^n, \qquad (20)\end{aligned}$$

$$\eta_m\left(\mathbf{d}_m, \mathbf{d}_f\right) = r_0 + \mathbf{r}_m'\mathbf{d}_m + \mathbf{r}_f'\mathbf{d}_f, \quad \eta_f = 1 - \eta_m, \qquad (21)$$

$$\ln \Delta_i\left(\mathbf{d}_i\right) = q_{0i} + \mathbf{q}_i'\mathbf{d}_i, \qquad (22)$$

$$\delta_i^n\left(\mathbf{d}_i\right) = \delta_i^n, \qquad (23)$$

where the price elasticities of the scale economies are assumed to be constant to ease computation (i.e. $\delta_i^n\left(\mathbf{d}_i\right) = \delta_i^n$). The system (20)-(23) is estimated using pre-death data.

To estimate the model, 11 expenditure categories are considered: ready-to-eat food; intoxicants; restaurants; clothes and toiletries; cosmetics and jewelry; power and fuel; housecare products; bills; transportation; communication and information; and others.[17] The vectors of demographic characteristics each contain the parent's age, the parent's educational attainment, the number of children in the household, and the number of other adults (e.g. grandparents, co-residing adult children) in the household.

I estimate the Engel curve systems by by feasible generalized nonlinear least squares (also known as nonlinear Seemingly Unrelated Regression). The estimator is iterated until the estimated parameters and error covariance matrix settle. This iterative estimator is equivalent to maximum likelihood with multivariate normal errors.

Table 5 shows the estimates of the sharing rule. The sharing rule is estimated to be 0.407, indicating that on average, the mother's resource share is 69% of the father's resource share. The mother's share of resources is increasing in her age and decreasing in her husband's age. The highly unequal intrahousehold distribution of resources implied by these estimates is in line with previous estimates of the sharing rule in India. Calvi (2020), for example, estimates that women' resource shares in Indian households are on average 67% of men's. Appendix F.3 presents the economies of scale estimates and discusses the fit of the model.

[17] Figure F2 in the appendix shows the mean budget shares for widows, widowers, and couples.

Table 5: Estimates of the sharing rule

Parameter estimates	
η_m	0.407***
	(0.069)
$\eta_m \times \text{age}_m$	0.012***
	(0.004)
$\eta_m \times \text{HS}_m$	0.014
	(0.009)
$\eta_m \times \text{age}_f$	-0.011***
	(0.003)
$\eta_m \times \text{HS}_f$	-0.014*
	(0.008)
N (Widows)	19227
N (Widowers)	4816
N (Couples)	5039

Note: η_m is the mother's resource share in households in which both parents are 40 years old and neither have any secondary education. $\eta_m \times \text{age}_i$ shows the effect of parent i's age on the mother's resource share. Similarly, $\eta_m \times \text{HS}_i$ shows the effect of parent i's secondary education on the mother's resource share. Sample consists of households with children that experience parental death by age 18. Standard errors clustered at the household level.

5.3 Method of Simulated Moments

Parameterization

Having estimated the production function, the sharing rule, and scale economies, the final step is to estimate the preference parameters associated with the non-market goods. To do so, I parameterize the conditional indirect utility functions as follows,[18]

$$
\begin{aligned}
V_i\left(X_i^j, h^j, e_s^j, e_d^j, \mathbf{z}^j\right) = & \ln X_i^j + \left(\theta_{hi} + \epsilon_{hi}^j\right)\frac{1}{1-\nu}\left(\frac{h^j}{N^j}\right)^{1-\nu} \\
& + \sum_{n\in\{s,d\}}\left(\theta_{eni} + \delta_0 \text{age}_n + \delta_1 \text{wave} + \epsilon_{eni}^j\right) e_n^j \\
& + \sum_{n\in\{m,f,s,d\}} n^j\left(\theta_{zni} + \epsilon_{zni}^j\right)\ln\left(z_n^j\right),
\end{aligned} \tag{24}
$$

for parent $i \in \{m, f\}$ in household j, where $\boldsymbol{\epsilon_i} \sim \mathcal{N}(\mathbf{0}, \Sigma)$ are independent taste shocks, and $n^j \in \{m, f, s, d\}$ is an indicator for whether there is a mother, a father, a son, or a daughter in household j. Recall that X_i is income for market consumption, $\frac{h}{N}$ is the per-capita home-produced good, e_n is child n's school enrollment status, and z_n is the leisure of member n. Note that preferences for schooling are allowed to vary with the child's age, thereby capturing the fact that the effect of school dropout on child development depends on age. To match the overall increase in school enrollment over the span of the panel, preferences for schooling are also allowed to depend on calendar time.

There are 20 parameters left to estimate: five preference parameters for each parent (θ's),[19] the elasticity coefficient (ν), two parameters that govern how preferences for schooling depend on the child's age and calendar time (δ's), the cost of schooling parameter (ψ), the elasticity of substitution in home production time (τ), and the variance of the five taste shocks (diagonal of Σ).

Estimation sample

The estimation sample includes only households with at least one child that loses a parent between the ages 12 to 18. Most (70%) households in this sample have at most one son and at most one daughter in the relevant age range. In households where there is more than one child of the same gender, the estimator targets the choices of the oldest of these children. Thus, the model explains the choices of the oldest son and oldest daughter, who in most cases are the only son and

[18] For the estimation of the sharing rule and scale economies, the first term of the indirect utility function was assumed to have the QUAIDS form, $\left[\left(\frac{\ln X_i - \bar{a}_i}{\bar{b}_i}\right)^{-1} + \bar{\lambda}_i\right]^{-1}$. The parameters (a, b, λ) are price indices that can be recovered from estimation of the demand system. However, without price variation, these become unidentified constants. If the units of the market commodities are all rescaled so that they have the same price, then $\bar{\lambda}$ is equal to zero. With $\bar{\lambda} = 0$, $\ln X_i$ represents the same preferences as $\left[\left(\frac{\ln X_i - \bar{a}_i}{\bar{b}_i}\right)^{-1} + \bar{\lambda}_i\right]^{-1}$.

[19] Note that θ_f and θ_s are not identified because fathers and sons in the model always spend 16 hours a day on leisure.

only daughter between the ages 12 to 18.

In the estimation sample, 16% households have a grandparent. Importantly, parental death does not affect the probability of living with a grandparent (see Table C3). In the baseline model, grandparents consume household resources and they may contribute to household income and home production, but they are not decisionmakers.[20] However, one may worry that grandparents have distinct preferences that influence household choices. As a robustness check, I estimate the structural model with the subsample of nuclear households. The results presented below are robust to excluding households with grandparents from the sample.

Remarriage of widows is exceedingly rare in India. In my sample, only 5% of paternal orphans live with a step-father at some point in the panel. In contrast, remarriage of widowers is relatively more common. In fact, 25% of maternal orphans in the sample live with a step-mother at some point in the panel. Remarriage, however, does not usually happen immediately after widowhood. Since I do not model the choice of remarriage, the estimation sample excludes observations recorded from more than a year after parental death. To evaluate the implications of this restriction, I compare the short-term and long-term effects of parental death on school enrollment. I find that the effects of parental death within a year of the shock are similar to the long-term effects.

Estimation

Estimation follows the method of simulated moments (McFadden, 1989; Pakes and Pollard, 1989). Specifically, the estimator targets four types of moments. First, it targets the cross-household average of all choices at baseline (i.e. pre-death). Second, I split the sample into six groups according to pre-determined income and parental vital status, and target the average choices in each group. Third, the estimator targets the effects of paternal death and maternal death on all choices, controlling for age, wave, and household random effects. Finally, to identify the variance of the household-specific taste shocks, I target the estimated variance of the household random effects from these regressions. Table F7 lists all targeted moments and Figures F5-F8 depict the link between parameters and targeted moments by plotting the sensitivity matrix proposed by Andrews et al. (2017). The estimator minimizes the distance between simulated moments and data moments, where the weighting matrix is the inverse of the variance-covariance matrix of the data moments.

Table 6 shows the MSM estimates. The estimates indicate that mothers value the schooling of sons and daughters more than fathers (relative to private consumption). Interestingly, both parents have a stronger preference for the schooling of daughters, compared to the schooling of sons. This result may be capturing differential returns to schooling between boys and girls.

Importantly, Table 6 shows that mothers and fathers also differ in how much they value the daughter's leisure, the home-produced good, and the mother's leisure. Thus, to understand how schooling choices depend on parental preferences, it is not enough to examine how much mothers and fathers care about education. Parental preferences for other goods—especially those associated

[20] I assume that the income and home production time provided by grandparents is exogenous.

Table 6: MSM Estimates

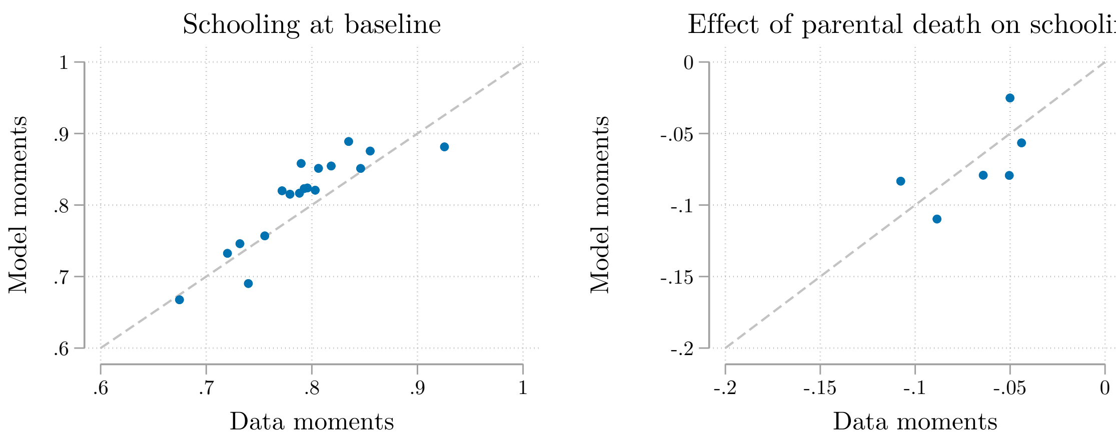
Parameters	Mothers	Fathers	Other
Son's schooling (θ_{es})	1.29	1.08	
	(0.03)	(0.03)	
Daughter's schooling (θ_{ed})	1.92	1.77	
	(0.05)	(0.03)	
Daughter's leisure (θ_{zd})	2.07	2.18	
	(0.04)	(0.04)	
Home-produced good (θ_h)	43.12	36.68	
	(2.88)	(2.33)	
Mother's leisure (θ_{zm})	2.21	1.92	
	(0.03)	(0.02)	
Elasticity coefficient for h (ν)			2.12
			(0.02)
Substitution of home production time (τ)			0.65
			(0.01)
Cost of schooling (ψ)			0.26
			(0.04)
Preference for schooling over time(δ_1)			0.01
			(0.00)
Preference for schooling by age (δ_0)			-0.17
			(0.01)
Variance of taste shock for son's schooling			0.40
			(0.03)
Variance of taste shock for daughter's schooling			0.67
			(0.07)
Variance of taste shock for daughter's leisure			0.49
			(0.04)
Variance of taste shock for home-produced good			4.99
			(4.87)
Variance of taste shock for mother's leisure			0.60
			(0.03)

Note: The estimator minimizes the distance between simulated moments and data moments, where the weighting matrix is the inverse of the variance-covariance matrix of the data moments. Standard errors in parentheses.

with children's opportunity cost of schooling—also shape children's educational investments.

To assess model fit, Figure 8 plots the simulated moments against the data moments for all moments associated with school enrollment. Figure F4 shows the model's performance on non-targeted moments, namely school enrollment as a function of age, of predetermined income, and of offered wages. The model fits targeted and non-targeted moments well, except for the effect of maternal death on daughter's time in home production. The model overpredicts the increase in daughter's time in home production following maternal death (see Table F7). While the death of a mother increases daughters time in home production, the observed effect is moderate compared to what the model predicts given the production function estimates.

Figure 8: Model fit - Targeted moments associated with schooling



Note: The figure plots the simulated moments against the data moments for all moments associated with school enrollment. A model with perfect fit would show all markers along the 45-degree line.

6 Counterfactual Analysis

This section first uses the estimated model to examine the role of resources and preferences in the schooling effects of parental death. The model is then used to evaluate various policies.

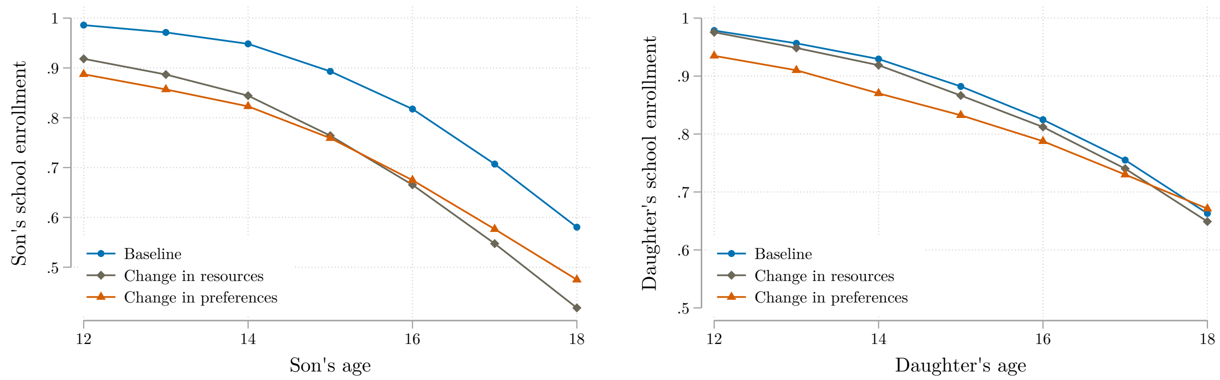
6.1 Decomposing the mechanisms

To assess the role of resources and preferences, I examine the contribution of each mechanism sequentially. The top and bottom panels of Figure 9 decompose the mechanisms associated with paternal death and maternal death, respectively.

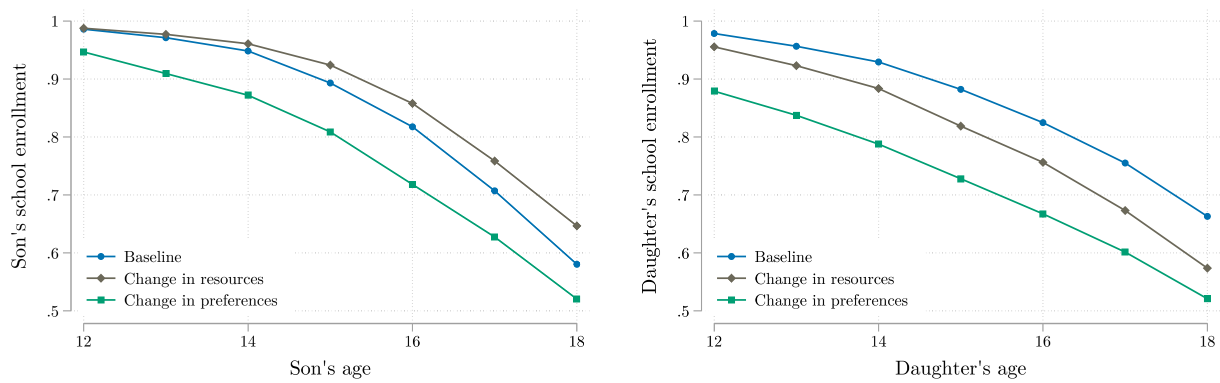
The blue solid dots in Figure 9a plot the school enrollment rates of sons and daughters in the model when both parents are alive. As a first step, I change household resources. Specifically, I subtract the average income loss associated with paternal death and adjust for household size. The resulting schooling choices, illustrated by gray diamonds, capture the role of household resources. The income loss has a large effect on sons, but not on daughters. One might expect that the decline in income would induce mothers to enter the labor force, leading to an increase in daughters' time in home production and a decline in school enrollment. This is not the case because married women rarely participate in the labor market. The model is able to capture this empirical pattern through the husband's preference for his wife's leisure. Thus, the change in resources following paternal death affects daughters' schooling only in combination with the change in household preferences.

Figure 9: Decomposing the mechanisms associated with parental death

(a) Paternal death



(b) Maternal death



Note: The figure sequentially decomposes the effect of parental death on school enrollment into a resource effect and a preference effect. Details of this exercise are provided in the main text.

Next, I change the household's objective function to include only the mother's preferences. The resulting change in enrollment rates reflects both the difference in preferences between mothers and fathers, and the transition from a two-parent household to a single-parent household. The difference in enrollment rates between the gray diamonds and the orange triangles captures the role of preferences. The change in household preferences has a large effect on daughters, but not on sons.

Figure 9b shows the analogous decomposition exercise for the case of maternal death. As before, the blue dots present the pre-shock enrollment choices. Next, the gray diamonds show the change in enrollment due to the loss in household resources. This loss amounts to constraining the mother's domestic and market labor supplies to zero, and adjusting for household size. The change in resources significantly decreases the enrollment rate of daughters, whereas it increases the enrollment rate of sons. The positive effect on sons' schooling results from the reduction in household size, which lowers the marginal utility of income.

Finally, the squares illustrate the effects of modifying the household's objective function to include only the father's preferences. This leads to a large decline in the enrollment rates of both sons and daughters, highlighting the importance of the preference mechanism in the case of maternal death.

6.2 Counterfactual analysis of public policies

This section evaluates the effectiveness of various policies aimed at mitigating the negative impact of parental death on schooling. The first policy under consideration is an unconditional cash transfer. For the case of paternal deaths, such a policy currently exists in India in the form of the Indira Gandhi National Widow Pension Scheme. This is a non-contributory pension for widows in below-poverty-line (BPL) households. It provides a monthly transfer of Rs. 300 to widows aged 40-79 and Rs. 500 monthly pension to older widows. Many state governments supplement the national widow pension by increasing the transfer amount or relaxing the eligibility criteria. In Kerala, for example, widows of all ages with household income below Rs. 100,000 receive a Rs. 1,000 monthly pension. Similarly, eligible widows in Delhi receive Rs. 2,500 a month. In 2016-17, there were 5.7 million beneficiaries of the Indira Gandhi National Widow Pension Scheme (Ministry of Rural Development, 2018).

Figure 10 shows the schooling effects of providing unconditional cash transfers to households of paternal and maternal orphans. The bars depict the regression coefficients of school enrollment on parental death using the econometric specification from Section 2. Unconditional transfers are effective at mitigating the negative effect of parental death on boys' schooling. In contrast, these transfers have no impact on daughters' schooling. This is expected given that for sons, the opportunity cost of schooling is paid work.

The second policy under consideration is the provision of domestic services. Figure 11 shows how the effects of parental death on school enrollment vary with exogenous increases in home production time. Providing 2 hours a day of domestic services to households of widowers reduces the negative effect of maternal death on daughters' schooling by nearly one half. However, fully compensating the household for the loss in home production does not completely eliminate the schooling effects of maternal death. This result is consistent with the importance of the preference mechanism in the case of maternal death.

The previous two policies only address the loss of household resources. One way to counteract the preference mechanism is to change the relative price of schooling. This can be achieved by a cash transfer that is conditional on school enrollment or a schooling subsidy. Figure 12 shows the impact of conditional cash transfers on the enrollment rates of orphans. Compared to unconditional transfers, conditional cash transfers are considerably more effective at mitigating the effects of

Figure 10: The effect of parental death in the model - Unconditional cash transfers

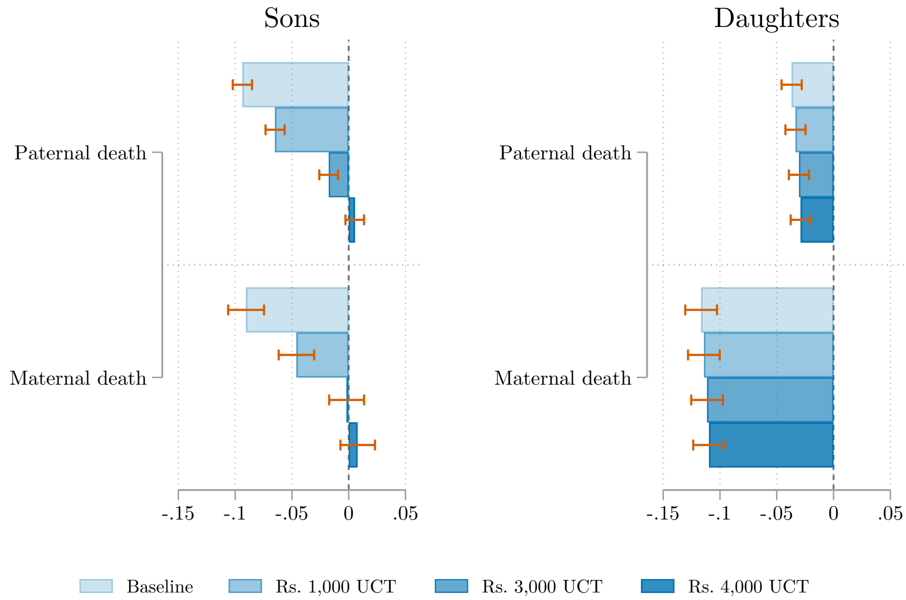
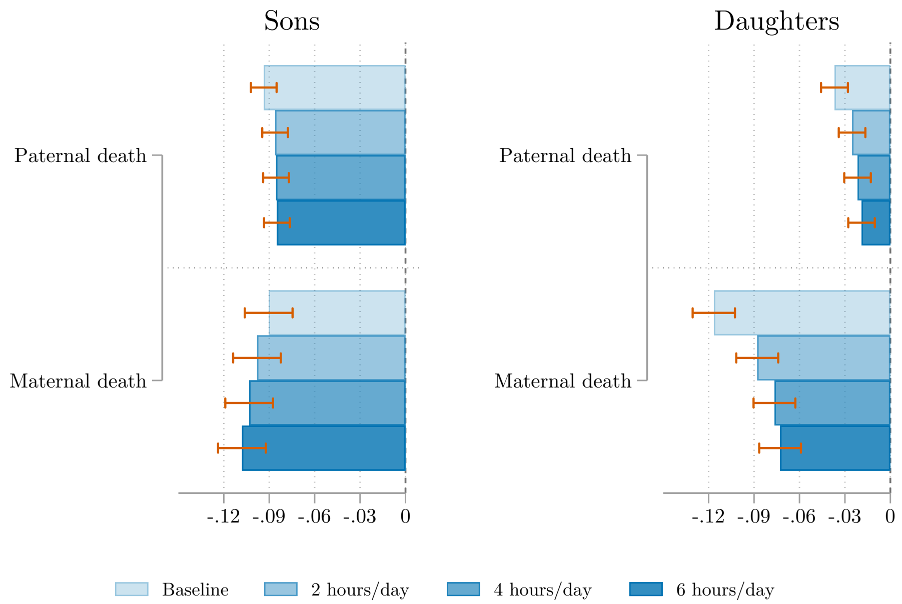
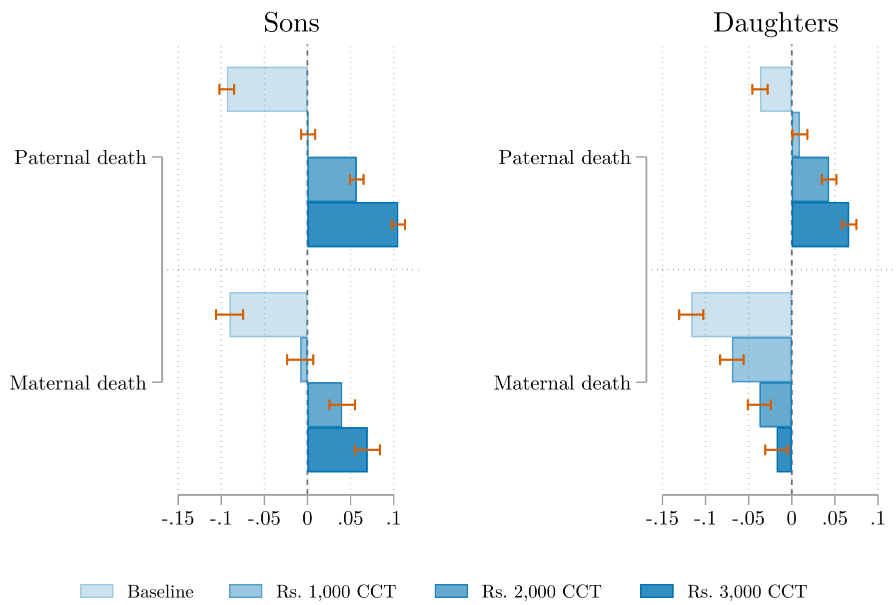


Figure 11: The effect of parental death in the model - Provision of domestic services



parental death. For example, a conditional cash transfer of Rs. 3,000 a month is sufficient to eliminate the schooling effects of maternal death on girls.

Figure 12: The effect of parental death in the model - Conditional cash transfers



Finally, recall that the model estimates indicate that mothers, on average, receive a Pareto weight equal to 0.4 in the household. In the last counterfactual simulation, I examine the effect of equalizing the bargaining weights of fathers and mothers. Figure 13 shows that this shift in intrahousehold bargaining would lead to increases in school enrollment, especially among daughters.

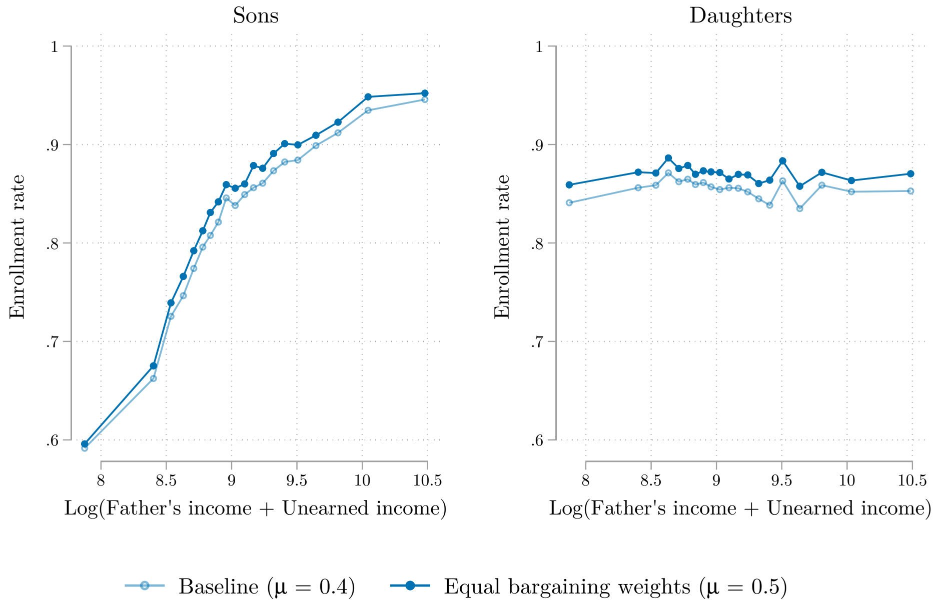
7 Conclusion

Between March 2020 and October 2021, nearly 5 million children around the world lost a father or mother as a result of the COVID-19 pandemic (Unwin et al., 2022). While the pandemic has brought the vulnerable situation of orphans to the forefront of policy discussion, the loss of a parent during childhood has always been a common occurrence in many regions of the world.

This paper contributes to our understanding of how and why the loss of a parent affects educational choices. First, I estimate the causal effects of parental death on schooling in India. Second, the paper disentangles and quantifies the mechanisms underlying these effects. Lastly, I evaluate the effectiveness of various policy interventions aimed at mitigating the negative impact of parental death on schooling.

The paper uses data from a large longitudinal survey of Indian households to answer these questions. This rich survey offers the unique opportunity to observe nearly 10,000 children before and after the death of a parent. Exploiting variation in the timing of parental death, I find marked differences in the effects of parental death, both by the gender of the child and the gender of the

Figure 13: The effect of equalizing bargaining weights



parent. Paternal death mostly affects sons, who are more likely to drop out of school and enter the labor force. Maternal death affects both sons and daughters, increasing the probability that sons enter the labor force and daughters engage in home production.

To disentangle the underlying mechanisms, I first analyze shifts in the time allocation of surviving household members. Then, using these behavioral responses, I estimate a structural model of household consumption and time allocation. The estimated model shows that the schooling effects of paternal death primarily stem from changes in household resources. Conversely, the effects of maternal death are largely explained by changes in household preferences and by the loss of mothers' contribution to home production. Counterfactual simulations suggest that the effectiveness of interventions depend on the gender of the orphan and the deceased parent.

There are several avenues for future research. Future work should further explore the intergenerational consequences of remarriage. Remarriage of widows is still rare in India, but it is common in other settings. Additionally, this paper focuses solely on the cognitive dimension of human capital, yet parental death is likely to affect children's non-cognitive, socio-emotional development. Given suitable data, this model could accommodate the study of this dimension of human capital. Furthermore, while this study primarily relies on parental death for identification, the model can be extended to explore the effects of other forms of parental absence, such as migration, divorce, and incarceration. Though identification might be more challenging in these contexts, this model could provide valuable insights into these scenarios when coupled with the right data.

References

Adda, J., Bjorklund, A., and Holmlund, H. (2011). The role of mothers and fathers in providing skills: Evidence from parental deaths. *IZA discussion paper*.

Ainsworth, M., Beegle, K., and Koda, G. (2005). The impact of adult mortality and parental deaths on primary schooling in North-Western Tanzania. *The Journal of Development Studies*, 41:412–439.

Andrews, I., Gentzkow, M., and Shapiro, J. M. (2017). Measuring the sensitivity of parameter estimates to estimation moments. *The Quarterly Journal of Economics*, 132(4):1553–1592.

Ardington, C. and Leibbrandt, M. (2010). Orphanhood and schooling in South Africa: Trends in the vulnerability of orphans between 1993 and 2005. *Economic Development and Cultural Change*, 58:507–536.

Attanasio, O., Cattan, S., and Meghir, C. (2022). Early childhood development, human capital, and poverty. *Annual Review of Economics*, 14:853–892.

Banks, J., Blundell, R., and Lewbel, A. (1997). Quadratic engel curves and consumer demand. *Review of Economics and statistics*, 79(4):527–539.

Barclay, K. and Hällsten, M. (2022). Does the impact of parental death vary by parental socioeconomic status? A study of children's educational and occupational attainment. *Journal of Marriage and Family*, 84(1):141–164.

Barten, A. (1964). Family composition, prices and expenditure patterns.

Beegle, K., Deweerdt, J., and Dercon, S. (2006). Orphanhood and the long-run impact on children. *American Journal of Agricultural Economics*, 88:1266—-1272.

Beller, A. H. and Chung, S. S. (1992). Family structure and educational attainment of children: Effects of remarriage. *Journal of Population Economics*, 5(1):39–59.

Bicego, G., Rutstein, S., and Johnson, K. (2003). Dimensions of the emerging orphan crisis in sub-Saharan Africa. *Social science & medicine*, 56(6):1235–1247.

Blundell, R., Chiappori, P.-A., and Meghir, C. (2005). Collective labor supply with children. *Journal of Political Economy*, 113:1277–1306.

Browning, M., Chiappori, P.-A., and Lewbel, A. (2013). Estimating consumption economies of scale, adult equivalence scales, and household bargaining power. *The Review of Economic Studies*, 80:1267—1303.

Browning, M. and Meghir, C. (1991). The effects of male and female labor supply on commodity demands. *Econometrica*, pages 925–951.

Callaway, B. and Sant'Anna, P. (2021). Difference-in-differences with multiple time periods. *Journal of Econometrics*, 225:200—230.

Calvi, R. (2020). Why are older women missing in India? The age profile of bargaining power and poverty. *Journal of Political Economy*, 128:2453–2501.

Cas, A., Frankenberg, E., Suriastini, W., and Thomas, D. (2014). The impact of parental death on child well-being: Evidence from the Indian Ocean tsunami. *Demography*, 51:437–457.

Case, A. and Ardington, C. (2006). The impact of parental death on school outcomes: Longitudinal evidence from South Africa. *Demography*, 43:401–420.

Case, A., Paxson, C., and Ableidinger, J. (2004). Orphans in Africa: Parental death, poverty, and school enrollment. *Demography*, 41:483–508.

Chen, J. J. (2012). Dads, disease, and death: Determinants of daughter discrimination. *Journal of Population Economics*, 25:119–149.

Chen, S., Chen, Y., and Liu, J. (2009). The impact of unexpected maternal death on education: First evidence from three national administrative data links. *American Economic Review: Papers & Proceedings*, 99:149–153.

Cherchye, L., Chiappori, P.-A., De Rock, B., Ringdal, C., and Vermeulen, F. (2021). Feed the children. Working paper.

Cherchye, L., Rock, B. D., and Vermeulen, F. (2012). Married with children: A collective labor supply model with detailed time use and intrahousehold expenditure information. *American Economic Review*, 102:3377–3405.

Chiappori, P.-A. (1988). Rational household labor supply. *Econometrica*, pages 63–90.

Chiappori, P.-A. (1992). Collective labor supply and welfare. *Journal of Political Economy*, 100(3):437–467.

Chuong, C. and Operario, D. (2012). Challenging household dynamics: Impact of orphanhood, parental absence, and children's living arrangements on education in South Africa. *Global Public Health*, 7(1):42–57.

Del Boca, D., Flinn, C., and Wiswall, M. (2014). Household choices and child development. *Review of Economic Studies*, 81(1):137–185.

Del Boca, D., Monfardini, C., and Nicoletti, C. (2017). Parental and child time investments and the cognitive development of adolescents. *Journal of Labor Economics*, 35(2):565–608.

Dribe, M., Debiasi, E., and Eriksson, B. (2022). The effect of parental loss on social mobility in early twentieth-century Sweden. *Demography*, 59(3):1093–1115.

Dupraz, Y. and Ferrara, A. (2023). Fatherless: The long-term effects of losing a father in the U.S. Civil War. *Journal of Human Resources*.

Emerson, P. M. and Souza, A. P. (2007). Child labor, school attendance, and intrahousehold gender bias in Brazil. *The World Bank Economic Review*, 21(2):301–316.

Ermisch, J. F. and Francesconi, M. (2001). Family structure and children's achievements. *Journal of population economics*, 14:249–270.

Evans, D. and Miguel, E. (2007). Orphans and schooling in Africa: A longitudinal analysis. *Demography*, 44:35–57.

Gennetian, L. A. (2005). One or two parents? half or step siblings? The effect of family structure on young children's achievement. *Journal of Population Economics*, 18:415–436.

Gertler, P., Levine, D., and Ames, M. (2004a). Schooling and parental death. *The Review of Economics and Statistics*, 86:211–225.

Gertler, P., Martinez, S., Levine, D., and Bertozzi, S. (2004b). Lost presence and presents: How parental death affects children. Working paper.

Gimenez, L., Chou, S.-Y., Liu, J.-T., and Liu, J.-L. (2013). Parental loss and children's well-being. *Journal of Human Resources*, 48(4):1035–1071.

Glick, P. and Sahn, D. E. (2000). Schooling of girls and boys in a West African country: The effects of parental education, income, and household structure. *Economics of education review*, 19(1):63–87.

Global Burden of Disease Collaborative Network (2019). Global burden of disease study 2019.

Goldstein, E. (2022). A mighty toll: Mine accidents and the long-run effect of parental loss. Working paper.

Greenwood, J., Guner, N., and Vandenbroucke, G. (2017). Family economics writ large. *Journal of Economic Literature*, 55(4):1346–1434.

Guarcello, L., Lyon, S., Rosati, F., and Valdivia, C. (2004). The influence of orphanhood on children's schooling and labour: Evidence from Sub Saharan Africa. Understanding Children Work (UCW Project) Working Paper no. 13.

Haveman, R. and Wolfe, B. (1995). The determinants of children's attainments: A review of methods and findings. *Journal of economic literature*, 33(4):1829–1878.

Himaz, R. (2013). Impact of parental death in middle childhood and adolescence on child outcomes. *Journal of African Economies*, 22(3):463–490.

Jacob, S. and Chattopadhyay, S. (2016). Marriage dissolution in India: Evidence from census 2011. *Economic and Political Weekly*, 51(33):25–27.

Kailaheimo-Lönnqvist, S. and Erola, J. (2020). Child's age at parental death and university education. *European Societies*, 22(4):433–455.

Kailaheimo-Lönnqvist, S. and Kotimäki, S. (2020). Cause of parental death and child's health and education: The role of parental resources. *SSM-Population Health*, 11:100632.

Kalil, A., Mogstad, M., Rege, M., and Votruba, M. E. (2016). Father presence and the intergenerational transmission of educational attainment. *Journal of Human Resources*, 51(4):869–899.

Kobiané, J.-F., Calvès, A.-E., and Marcoux, R. (2005). Parental death and children's schooling in Burkina Faso. *Comparative education review*, 49(4):468–489.

Krein, S. F. and Beller, A. H. (1988). Educational attainment of children from single-parent families: Differences by exposure, gender, and race. *Demography*, 25(2):221–234.

Lewbel, A. and Pendakur, K. (2008). Estimation of collective household models with Engel curves. *Journal of Econometrics*, 147:350–358.

Liu, C., Grotta, A., Hiyoshi, A., Berg, L., and Rostila, M. (2022). School outcomes among children following death of a parent. *JAMA Network Open*, 5(4):e223842–e223842.

Lundberg, S. J., Pollak, R. A., and Wales, T. J. (1997). Do husbands and wives pool their resources? evidence from the United Kingdom child benefit. *Journal of Human resources*, pages 463–480.

McFadden, D. (1989). A method of simulated moments for estimation of discrete response models without numerical integration. *Econometrica: Journal of the Econometric Society*, pages 995–1026.

Mencarini, L., Pasqua, S., and Romiti, A. (2019). Single-mother families and the gender gap in children's time investment and non-cognitive skills. *Review of Economics of the Household*, 17:149–176.

Ministry of Rural Development (2018). Report of the standing committee on rural development (2017-2018).

Pakes, A. and Pollard, D. (1989). Simulation and the asymptotics of optimization estimators. *Econometrica: Journal of the Econometric Society*, pages 1027–1057.

Phipps, S. A. and Burton, P. S. (1998). What's mine is yours? The influence of male and female incomes on patterns of household expenditure. *Economica*, 65(260):599–613.

Pollak, R. A. (1969). Conditional demand functions and consumption theory. *The Quarterly Journal of Economics*, 83:60–78.

Ringdal, C. and Sjursen, I. (2021). Household bargaining and spending on children: Experimental evidence from Tanzania. *Economica*, 88:430–455.

Roth, J., Sant'Anna, P., Bilinskiz, A., and Poe, J. (2022). What's trending in difference-in-differences? A synthesis of the recent econometrics literature.

Senne, A. (2014). Death and schooling decisions over the short and long run in rural Madagascar. *Journal of Population Economics*, 27:497–528.

Somanchi, A. (2021). Missing the poor, big time: A critical assessment of the consumer pyramids household survey. Working Paper.

Thomas, D. (1990). Intra-household resource allocation: An inferential approach. *Journal of human resources*, pages 635–664.

United Nations Population Division (2022). World population prospects 2022.

Unwin, J., Hillis, S., Cluver, L., Flaxman, S., Goldman, P., Butchart, A., Bachman, G., Rawlings, L., Donnelly, C., Ratmann, O., Green, P., Nelson, C., Blenkinsop, A., Bhatt, S., Desmond, C., Villaveces, A., and Sherr, L. (2022). Global, regional, and national minimum estimates of children affected by COVID-19-associated orphanhood and caregiver death, by age and family circumstance up to oct 31, 2021: An updated modelling study. *The Lancet Child Adolescent Health*, In Press:1–11.

Weiss, Y. and Willis, R. J. (1985). Children as collective goods and divorce settlements. *Journal of Labor Economics*, 3(3):268–292.

Wong, S. Y. (2023). When should parents migrate? The relative importance of parental time and money investments from middle childhood to early adolescence. Working paper.

Wooldridge, J. (2010). *Econometric analysis of cross section and panel data.* MIT Press.

A The Causal Effects of Parental Death

Table A1 shows that among orphans, the timing of parental death is uncorrelated with a series of relevant characteristics at baseline. This result suggests that the timing of parental death, among children who experience the shock between ages 12 and 18, is as good as random.

Table A1: Balance in the sample of orphans

D.V.	(1) Urban	(2) Upper Caste	(3) HH size	(4) Dad HS	(5) Mom HS
Child's age when father died	0.00	-0.00	-0.02	-0.00	-0.01***
	(0.00)	(0.00)	(0.02)	(0.00)	(0.00)
D.V. mean	0.26	0.19	5.27	0.14	0.06
N	4612	4548	4612	4612	4610
Child's age when mother died	0.01*	0.00	-0.03	-0.01	-0.00
	(0.00)	(0.01)	(0.03)	(0.01)	(0.00)
D.V. mean	0.20	0.17	5.41	0.14	0.05
N	1707	1697	1707	1707	1707

Note: Sample of children who experienced parental death between the age of 12 and 18. Each point estimate corresponds to a separate regression.

Modified Callaway and Sant'Anna (2021) estimator

Recent studies have shown two-way fixed effect (TWFE) specification yields a sensible estimand only if (i) all units have the same treatment effect, and (ii) the treatment has the same effect regardless of how long it has been since it started. If either of these assumptions fails, the TWFE estimates are difficult to interpret and, in extreme cases, may even have the wrong sign. Intuitively, this happens because the TWFE specification makes "forbidden comparisons" between already-treated units. In my application, children experience parental death at different times and, most importantly, at different ages. One might expect that the treatment effects of parental death vary with the age of the child at the time of shock.

To alleviate this concern, I implement a modified version of the estimator developed by Callaway and Sant'Anna (2021). The modification is that the treatment effect depends on the individual's age at the time of treatment. Formally, let G_i be child i's age at the time of the shock, and let $Y_{ia}(g)$ denote i's potential outcome at age a if he/she experienced parental death at age g. Then, the estimands of interest are $ATT(g,a) = E[Y_{ia}(g) - Y_{ia}(\infty)|G_i = g]$, which gives the average treatment effect at age a for the group who experienced a parental death at age g. These parameters are identified under straightforward generalizations of the parallel-trends assumption and the no anticipation assumption. Identification is achieved by comparing the expected change in schooling for group g, between ages $g-1$ and a to that for a control group of never-treated children:

$$ATT(g,a) = E[Y_{ia} - Y_{i,g-1}|G_i = g] - E[Y_{ia} - Y_{i,g-1}|G_i > 18], \tag{25}$$

for $a \geq g$, where the control group consists of children who will experience parental death after age 18. Following the aggregation methods proposed by Callaway and Sant'Anna (2021), I report the aggregate ATT in column (3) of Table 2. The estimates are qualitatively similar to the TWFE estimates.

B The Mechanisms

Effect on children's time use

Table B1 mirrors the baseline analysis of Table 4 but employing time use data instead. The estimates derived from time-use data are less precise due to a smaller sample size (as time-use data collection in the CPHS commenced only in November 2019) and measurement error associated with the dependent variable. However, the main findings are qualitatively similar. On average, boys increase their labor supply by 37 minutes per day following the death of their father, while girls increase their time spent on home production by 28 minutes per day following the death of their mother.

Table B1: Effect of parental death on children's time use

D.V.: Hours a day spent on...	(1) Education	(2) Paid work	(3) Home production	(4) Leisure
β^{PB}: Paternal death × Boy	-0.36***	0.61***	0.17**	-0.41***
	(0.10)	(0.11)	(0.07)	(0.13)
β^{PG}: Paternal death × Girl	-0.07	-0.16***	0.21**	0.06
	(0.11)	(0.04)	(0.09)	(0.14)
β^{MB}: Maternal death × Boy	-0.05	0.05	0.27*	-0.28
	(0.19)	(0.19)	(0.16)	(0.27)
β^{MG}: Maternal death × Girl	0.04	-0.20***	0.46***	-0.24
	(0.17)	(0.04)	(0.17)	(0.22)
p-val: $\beta^{PB} = \beta^{PG}$	0.02	0.00	0.69	0.00
p-val: $\beta^{MB} = \beta^{MG}$	0.68	0.19	0.35	0.89
p-val: $\beta^{PB} = \beta^{MB}$	0.13	0.01	0.56	0.65
p-val: $\beta^{PG} = \beta^{MG}$	0.61	0.40	0.17	0.23
Control mean: Boys	3.33	0.45	2.07	17.64
Control mean: Girls	3.31	0.03	2.48	17.73
No. of children	3805	3805	3805	3805
N	21860	21860	21860	21860

Note: Sample includes only children that experience parental death by age 18. Standard errors clustered at the household level.

Effect on surviving parent's time use

Table B2 shows the effect of parental death on the occupation of the surviving parent. The death of a father increases the probability that the widowed mother works for pay by 37 percentage points. Thus, the effect of paternal death on daughters' schooling may be attributed to daughters assuming additional domestic responsibilities as they assist their working mothers. Table B3 shows anologous estimates using time-use data.

Table B2: Effect of parental death on surviving parent's occupation

	Mother's occupation		Father's occupation	
D.V.:	Paid worker	Homemaker	Paid worker	Homemaker
Paternal death	0.37***	-0.37***		
	(0.02)	(0.02)		
Maternal death			0.00	0.00
			(0.01)	(0.00)
Control mean	0.13	0.86	0.95	0.01
No. of children	6382	6382	5495	5495
N	62284	62284	43650	43650

Note: Sample includes only children that experience parental death by age 18. Standard errors clustered at the household level.

Table B3: Effect of parental death on parent's time use

	Mother's time use			Father's time use		
D.V.: Hours a day spent on...	Paid work	Home production	Leisure	Paid work	Home production	Leisure
Paternal death	2.23***	-0.90***	-1.47***			
	(0.22)	(0.13)	(0.20)			
Maternal death				-0.13	0.13	-0.04
				(0.19)	(0.14)	(0.25)
Control mean	0.93	5.36	17.37	7.15	2.05	14.04
No. of children	3269	3269	3269	1927	1927	1927
N	19715	19715	19715	10246	10246	10246

Note: Sample includes only children that experience parental death by age 18. Standard errors clustered at the household level.

C The role of siblings and grandparents

Effect on household size and composition

The impact of parental death on household size and composition extends beyond the mere absence of the deceased individual. The death shock may trigger a behavioral response from surviving family members, who may decide to move in, move out, or stay in the household. Table C1 documents the effect of parental death on household size and composition among children who lose a parent by age 18. The two-way fixed effects regressions are at the child level and the standard errors are clustered at the household level.

Table C1: Effect of parental death on household size and compositon

D.V.:	(1) HH size	(2) Step-parents	(3) Grandparents	(4) Other adults (age > 18)	(5) Children (age $\leq$ 18)
β^P: Paternal death	-0.82***	0.01*	0.00	0.09***	0.01
	(0.03)	(0.00)	(0.01)	(0.02)	(0.02)
β^M: Maternal death	-0.65***	0.14***	0.01	0.06*	0.03
	(0.05)	(0.01)	(0.01)	(0.03)	(0.04)
p-val: $\beta^P = \beta^M$	0.00	0.00	0.75	0.45	0.62
Control mean	5.10	0.00	0.18	0.62	2.38
No. of children	7340	7340	7340	7302	7340
N	82860	82860	82860	81651	82860

Note: Sample includes only children that experience parental death by age 18. Standard errors clustered at the household level.

Column (1) of Table C1 shows that parental death decreases household size by less than one. Specifically, paternal death decreases household size by 0.82 and maternal death decreases household size by 0.65. This difference can be explained by the gender gap in remarriage rates documented in column (2). Although it was legalized by the Hindu Widows' Remarriage Act of 1856, remarriage of widows continues to be uncommon in India. In my sample, only 5% of paternal orphans live with a step-father at some point in the panel. In contrast, remarriage of widowers is relatively more common. In my sample, 25% of maternal orphans live with a step-mother at some point in the panel. Remarriage, however, does not usually happen immediately after widowhood. Within one year of the shock, only 2% of widowed mothers and 15% of widowed fathers remarry.

Another aspect to consider when examining shifts in household composition is the issue of cohabitation with grandparents. Cohabitation with grandparents is not uncommon in India, with 18% of orphans in my sample co-residing with at least one grandparent. However, as shown in column (3), cohabitation with grandparents does not exhibit a significant response to parental death.

One concern regarding cohabitation with grandparents is the possibility of widows and widowers returning to live with their parents following the loss of their spouse. Unfortunately, the CPHS does not track households that move. However, it is important to note that this practice is not widespread in India. To assess the frequency of this phenomenon, I utilize data from the India Human Development Survey, which is a nationally representative panel that tracks individuals and households that move. I find no evidence to suggest that widowhood significantly encourages migration. Specifically, between wave 1 (2005) and wave 2 (2012), only 12% of individuals who experienced the loss of a spouse changed their residence, whereas 10% of all households in the sample relocated during the same period.

Finally, as shown in column (4), parental death leads to a modest increase in the number of other adults in the household (excluding parents, step-parents, and grandparents of the child). These individuals mainly comprise adult siblings of the child and their respective spouses. In other words, parental death prompts children to prolong their stay in the parental home. Conversely, the death of a parent has no impact on the number of household members under the age of 18.

Role of siblings

The impact of parental death on a child's school enrollment may depend on sibling composition. For instance, one may expect that the effect of parental death on schooling is strongest for the oldest son and the oldest daughter of the household. This is because the opportunity cost of schooling increases with age and the returns to years of schooling are likely to be concave. Note that birth order may matter even after conditioning for age. That is, a 15-year old boy may be more likely to be pulled out of school after the shock if he is the oldest son, compared to a 15-year old boy who has an older brother.

Table C2 extends the baseline analysis from Table 4 by incorporating interactions with an indicator of whether the child is the oldest son or oldest daughter. The estimates show that negative effects of parental death are larger for the oldest son and the oldest daughter. Paternal death reduces school enrollment among oldest sons by 16 percentage points, compared to 6 percentage points for their younger brothers. Similarly, maternal deaths reduce school enrollment among oldest daughters by 6 percentage points, compared to a null effect for their younger sisters.

Role of grandparents

The analysis presented thus far indicates that in the event of paternal death, sons (especially the oldest son) tend to assume the economic role of their deceased father. Similarly, in the case of maternal death, daughters (particularly the oldest daughter) often step in to fulfill the role of their late mother in home production. With 18% of children in the sample living with a grandparent, the question arises as to whether the presence of grandparents mitigates these effects. To answer this question, I extend the baseline analysis from Table 4 by incorporating interactions with an indicator of whether the child co-resides with a grandfather or grandmother.

One potential concern when examining the role of grandparents is the endogeneity of co-residence with grandparents, which may be a response to parental death. The concern here is that this choice could be correlated with unobserved determinants of schooling. For instance, households that place greater importance on education might be more likely to have a grandparent

Table C2: Effect of parental death on children's occupation, by birth order

D.V.:	(1) Student	(2) Paid worker	(3) Homemaker	(4) Unoccupied
β^{PB}: Paternal death × Boy	-0.06***	0.08***	-0.03***	0.00
	(0.01)	(0.01)	(0.01)	(0.00)
β^{PB}_o: Paternal death × Boy × Oldest son	-0.10***	0.09***	0.01	-0.01
	(0.02)	(0.02)	(0.01)	(0.00)
β^{PG}: Paternal death × Girl	0.01	-0.03***	0.03***	-0.00
	(0.01)	(0.00)	(0.01)	(0.00)
β^{PG}_o: Paternal death × Girl × Oldest daughter	-0.04***	-0.01**	0.05***	0.00
	(0.01)	(0.00)	(0.01)	(0.00)
β^{MB}: Maternal death × Boy	-0.03	0.03*	-0.02	0.02**
	(0.02)	(0.01)	(0.01)	(0.01)
β^{MB}_o: Maternal death × Boy × Oldest son	-0.03	0.02	0.03*	-0.01
	(0.03)	(0.02)	(0.02)	(0.01)
β^{MG}: Maternal death × Girl	0.00	-0.04***	0.05**	-0.00
	(0.02)	(0.01)	(0.02)	(0.00)
β^{MG}_o: Maternal death × Girl × Oldest daughter	-0.06**	-0.01*	0.08***	-0.00
	(0.02)	(0.01)	(0.02)	(0.01)
p-val: $\beta^{PB} = \beta^{PG}$	0.00	0.00	0.00	0.13
p-val: $\beta^{MB} = \beta^{MG}$	0.30	0.00	0.01	0.01
p-val: $\beta^{PB} = \beta^{MB}$	0.25	0.00	0.37	0.12
p-val: $\beta^{PG} = \beta^{MG}$	0.83	0.18	0.53	0.99
Control mean: Boys	0.86	0.06	0.06	0.02
Control mean: Girls	0.87	0.01	0.11	0.01
No. of children	6820	6820	6820	6820
N	70819	70819	70819	70819

Note: All regressions include wave FE, child FE, and child dummies. Sample includes only children that experience parental death by age 18. Standard errors clustered at the household level.

move in after the shock. However, as reported in Table C1, parental death does not affect the probability of co-residing with a grandparent.

Even if the presence of grandparents is not influenced by parental death, it could still be correlated with omitted variables that impact the relationship between parental death and schooling. Suppose, for example, that households that care more about education are more likely to have a grandparent living with them. In such a case, the estimated coefficient on the interaction between parental death and the presence of grandparents would be biased, as it confounds the effect of preferences and grandparents. Ideally, to address these concerns, I would instrument the presence of grandparents using their vital status. However, the CPHS lacks information on family members outside the household.

Proceeding under the assumption that the presence of grandparents is uncorrelated with other determinants of the relationship between parental death and schooling, Table C3 presents the two-way fixed effects estimates. Interestingly, the presence of a grandfather mitigates the effect of paternal death on boys, but it does not alleviate the impact that maternal death has on girls. Conversely, the presence of a grandmother completely eliminates the negative effect of maternal death on girls, but it does not mitigate the the effect of paternal death on boys.

Table C3: Effect of parental death on children's occupation, by grandparent presence

D.V.:	(1) Student	(2) Paid worker	(3) Homemaker	(4) Unoccupied
β^{PB}: Paternal death × Boy	-0.11***	0.13***	-0.03***	0.00
	(0.01)	(0.01)	(0.01)	(0.00)
β^{PB}_f: Paternal death × Boy × Grandfather	0.08**	-0.07**	-0.01	-0.01
	(0.03)	(0.03)	(0.02)	(0.01)
β^{PB}_m: Paternal death × Boy × Grandmother	-0.02	0.01	-0.00	0.01
	(0.03)	(0.03)	(0.01)	(0.01)
β^{PG}: Paternal death × Girl	-0.02**	-0.04***	0.07***	-0.00
	(0.01)	(0.00)	(0.01)	(0.00)
β^{PG}_f: Paternal death × Girl × Grandfather	0.03	-0.02***	-0.01	-0.00
	(0.02)	(0.00)	(0.02)	(0.01)
β^{PG}_m: Paternal death × Girl × Grandmother	0.02	0.01	-0.02	-0.01
	(0.02)	(0.00)	(0.02)	(0.01)
β^{MB}: Maternal death × Boy	-0.05***	0.04***	-0.00	0.01**
	(0.02)	(0.01)	(0.01)	(0.01)
β^{MB}_f: Maternal death × Boy × Grandfather	0.03	-0.03	0.01	-0.01
	(0.03)	(0.03)	(0.02)	(0.00)
β^{MB}_m: Maternal death × Boy × Grandmother	0.03	-0.00	-0.02	-0.01
	(0.03)	(0.03)	(0.02)	(0.01)
β^{MG}: Maternal death × Girl	-0.05***	-0.05***	0.10***	-0.00
	(0.02)	(0.01)	(0.02)	(0.00)
β^{MG}_f: Maternal death × Girl × Grandfather	0.01	0.00	0.01	-0.02
	(0.04)	(0.01)	(0.03)	(0.02)
β^{MG}_m: Maternal death × Girl × Grandmother	0.06*	0.00	-0.06**	-0.00
	(0.03)	(0.01)	(0.03)	(0.01)
p-val: $\beta^{PB} = \beta^{PG}$	0.00	0.00	0.00	0.50
p-val: $\beta^{MB} = \beta^{MG}$	0.92	0.00	0.00	0.01
p-val: $\beta^{PB} = \beta^{MB}$	0.00	0.00	0.03	0.08
p-val: $\beta^{PG} = \beta^{MG}$	0.16	0.21	0.04	0.43
Control mean: Boys	0.86	0.06	0.06	0.02
Control mean: Girls	0.87	0.01	0.11	0.01
No. of children	6825	6825	6825	6825
N	70950	70950	70950	70950

Note: Sample includes only children that experience parental death by age 18. Standard errors clustered at the household level.

D Human capital production function

The goal of this section is to study the direct effect of parental death on children's academic performance, conditional on schooling choices. To do so, I estimate a human capital production function, where current performance is a function of prior performance, school enrollment status, and the parents' vital status. This exercise requires longitudinal data on children's academic performance (i.e. test scores). Since school enrollment is an input in the production function, it is important that the test scores data are available even if the child is not enrolled in school. In practice, this requires that the tests are administered by a survey, and not by schools. For this part of the study I use data from the Young Lives survey in India.

The Young Lives study is a five-wave panel that follows 3,000 children from Andhra Pradesh and Telangana over a period of 14 years. This unique study collects a rich set of measures of children's academic performance. Specifically, the survey administers age-appropriate tests in mathematics, reading comprehension, English, and vocabulary, regardless of the child's school enrollment status. While this survey is the best source of longitudinal data to study child development in India, it is not without its limitations. In particular, the analysis below is limited by the small sample size. In the 14 years covered by the study, 314 children experienced paternal death and 112 children experienced maternal death.

Another limitation of the Young Lives study is that the experience of these 3,000 children from two Indian states may not be generalizable to the broader Indian population. Although the households in the CPHS and the Young Lives survey are sampled from different populations, Table D1 shows that these samples are very similar in terms of the parents' education, the share of households in urban areas, and household size.

Table D1: Comapring the YL and CPHS samples

	Young Lives mean	CPHS mean
Father completed high school	0.13	0.14
Mother completed high school	0.05	0.06
Urban	0.28	0.25
Household size	5.16	5.29
Observations	4737	5456

Next, I take a linear approximation of the human capital production function and estimate the following equation:

$$\underbrace{k_a}_{\text{HC measure}} = \alpha k_{a-1} + \underbrace{\beta_e e_a}_{\text{School enrollment}} + \underbrace{\beta_m m_a + \beta_f f_a}_{\text{Parents' vital status}} + \underbrace{\zeta_a}_{\text{Age dummies}} + u_a, \tag{26}$$

where u_a includes unobserved ability and time-varying shocks. There are three main challenges in estimating (26). First, school enrollment is an endogenous choice; it can be chosen to either reinforce or compensate the effects of unobserved shocks or ability. The second challenge is that test scores are an error-ridden measure of academic performance, and so the lagged dependent variable is measured with error. Finally, if unobserved time-invariant characteristics (for example, genetics) have a direct impact on test scores, then the lagged test score will be positively correlated with the error term.

To address the endogeneity and measurement error issues, I instrument for school enrollment and the lagged test score. I use the opening of a factory in the community as an instrument that moves the contemporaneous decision to enroll the child in school. The idea is that the opening of factories affects the incentives and resources to send a child to school. The exclusion restriction would be violated if the opening of factories is correlated with improvements in school quality or increases in monetary investments in the child's education, both of which are omitted from the production function. As a source of exogenous variation for the lagged test score, I use the cumulative number of agricultural shocks reported by the household from wave 1 up to the previous wave.[21] The idea is that the child's current stock of human capital is the result of the accumulation of all past inputs, and these inputs are affected by economic shocks.

The top panel of Table D2 reports the OLS and IV estimates of the production function, using the standardize math test score as a measure of academic performance.[22] The least-squares estimates of column (1) indicate that being enrolled in school is associated with an increase in .5 standard deviation in the math test score. Moreover, conditional on enrollment, there seems to be a large positive effect of having your parents alive. However, the IV estimates in column (2) show that conditional on school enrollment and the lagged test score, the parents' vital status has no direct impact. In other words, parental death has a large effect on human capital, but this effect operates only through schooling. Consistent with the evidence from the CPHS data, parental death has a strong effect on school enrollment in this sample as well, as seen in column (1) of the the first-stage regression.

[21] These shocks include livestock death, drought, flooding, erosion, frosts, pests affecting crops, crop failures, pests affecting storage, and pests affecting livestock.

[22] I use the standardized math test score because age-appropriate math tests were administered in multiple rounds. Unlike the other tests, math test scores were also administered to the closest-in-age sibling of the Young Lives children. Thus, the use of math test scores results in a larger sample size.

Table D2: Human capital production function estimates

D.V.: Standardized math test	(1) OLS	(2) IV
School enrollment	0.51***	2.18***
	(0.03)	(0.49)
Lagged dependent variable	0.59***	0.72***
	(0.01)	(0.13)
Father alive	0.08**	-0.05
	(0.04)	(0.06)
Mother alive	0.16***	-0.05
	(0.05)	(0.09)
Wave FE	Yes	Yes
Age dummies	Yes	Yes
Age range	[7,20]	[7,20]
Number of waves	3	3
Kleibergen-Paap F statistic		15.65
No. of paternal orphans	314	295
No. of maternal orphans	112	106
No. of children	4013	3863
N	6685	6239

FIRST-STAGE REGRESSIONS Endogenous regressors:	(1) School	(2) LDV
Factory presence	0.05***	-0.02
	(0.01)	(0.03)
Past no. of agricultural shocks	-0.02***	-0.13***
	(0.01)	(0.02)
Father alive	0.07***	0.20***
	(0.02)	(0.06)
Mother alive	0.11***	0.24***
	(0.04)	(0.09)
F test of excluded instruments	23.49	27.67
Sanderson-Windmeijer F test	33.81	37.37

Note: Standard errors clustered at the household level.

E Identification of sharing rule and scale economies

Write the couple's conditional demand as a function of the conditional demand of widows and widowers,

$$\mathbf{x}(\mathbf{p}, X, \mathbf{g}^*) = \mathbf{A}\mathbf{c}_m\left(A'\mathbf{p}, \eta(\mathbf{p}, X, \mathbf{g}^*) X, \mathbf{g}^*\right) + \mathbf{A}\mathbf{c}_f\left(A'\mathbf{p}, (1-\eta(\mathbf{p}, X, \mathbf{g}^*)) X, \mathbf{g}^*\right)$$

where $X \equiv (1-\phi(N-2))(y + w_m l_m^* + w_s(1-e_s^*) - \psi(e_s^* + e_d^*) - \mathbf{p}_{\mathbf{h}}'\mathbf{r}^*)$ is the residual income after the first-stage choices. Hereinafter, it is convenient to rewrite these conditional demands in budget share form and in terms of log prices and log residual income,

$$\begin{aligned}\omega^n(\ln\mathbf{p}, \ln X, \mathbf{g}^*) = {} & \eta(\ln\mathbf{p}, \ln X, \mathbf{g}^*)\,\omega_m^n(\ln\mathbf{a} + \ln\mathbf{p}, \ln(\eta(\ln\mathbf{p}, \ln X, \mathbf{g}^*)) + \ln X) \\ & + (1-\eta(\ln\mathbf{p}, \ln X, \mathbf{g}^*))\,\omega_f^n(\ln\mathbf{a} + \ln\mathbf{p}, \ln(1-\eta(\ln\mathbf{p}, \ln X, \mathbf{g}^*)) + \ln X)\end{aligned} \tag{27}$$

Browning et al. (2013) show that if $N \geq 3$, then $\mathbf{A}$ and $\eta(\mathbf{p}, X, \mathbf{g}^*)$ are generically identified. Lewbel and Pendakur (2008) assume two additional restrictions that allow identification using only Engel curves. First, they assume that there exists a scalar valued function, $\Delta_i(\ln\mathbf{p}, \mathbf{g}^*)$, that measures the cost savings due to scale economies. This assumption says that these scale economies are independent of the base expenditure and utility level at which they are evaluated. Second, they assume that the sharing rule does not depend on residual expenditures, X. With these two additional assumptions, (27) simplifies to

$$\begin{aligned}\omega^n(\ln\mathbf{p}, \ln X, \mathbf{g}^*) = {} & \eta\left[\omega_m^n(\ln\mathbf{p}, \ln X + \ln\eta - \ln\Delta_m, \mathbf{g}^*) + \delta_m^n\right] \\ & + (1-\eta)\left[\omega_f^n(\ln\mathbf{p}, \ln X + \ln(1-\eta) - \ln\Delta_f, \mathbf{g}^*) + \delta_f^n\right]\end{aligned} \tag{28}$$

where η, Δ_i, and $\delta_i \equiv -\frac{\partial\ln\Delta_i}{\partial\ln p^n}$ are functions of $(\ln\mathbf{p}, \mathbf{g}^*)$. The key difference between (27) and (28) is that the former writes the couple's budget shares as a function of the individual budget shares *evaluated at shadow prices*, whereas the latter writes the couple's budget shares as a function of the individual budget shares *evaluated at market prices*. In fact, if prices are constant, (28) simplifies to the following relation of conditional Engel curves:

$$\begin{aligned}\omega^n(\ln X, \mathbf{g}^*) = {} & \eta(\mathbf{g}^*)\left[\omega_m^n(\ln X + \ln\eta(\mathbf{g}^*) - \ln\Delta_m(\mathbf{g}^*), \mathbf{g}^*) + \delta_m^n(\mathbf{g}^*)\right] \\ & + (1-\eta(\mathbf{g}^*))\left[\omega_f^n(\ln X + \ln(1-\eta(\mathbf{g}^*)) - \ln\Delta_f(\mathbf{g}^*), \mathbf{g}^*) + \delta_f^n(\mathbf{g}^*)\right].\end{aligned}$$

Lewbel and Pendakur (2008) prove that η, Δ_i, and δ_i are nonparametrically identified from Engel curves of individuals and couples, as long as some of the goods have budget shares that are nonlinear and are sufficiently different across individuals.

F Structural Estimation

F.1 Home production function

First, I estimate the parameters of the CES aggregator of raw food commodities, γ_l and ρ_1. Cost minimization in home production requires that for any pair of uncooked food commodities (m, l), the marginal rate of technical substitution be equal to the ratio of input prices,

$$\ln\left(\frac{r_m}{r_l}\right) = \underbrace{\frac{1}{1-\rho_1}\ln\left(\frac{\gamma_m}{\gamma_l}\right)}_{b_{ml}} + \underbrace{\frac{1}{1-\rho_1}}_{\sigma_1}\ln\left(\frac{p_l}{p_m}\right). \tag{29}$$

Assuming exogenous prices for raw food commodities, these optimality conditions form a seemingly unrelated regressions (SUR) model,

$$\ln\left(\frac{r_m}{r_l}\right) = b_{ml} + \sigma_1 \ln\left(\frac{p_l}{p_m}\right) + \xi_l, \tag{30}$$

where ξ_l accounts for idiosyncratic preferences for specific food commodities, as well as measurement and optimization errors.

Given estimates for b_{ml} and σ_1, one can compute estimates for the parameters of interest,

$$\hat{\rho}_1 = \frac{\hat{\sigma}_1 - 1}{\hat{\sigma}_1}, \quad \hat{\gamma}_l = \frac{\exp\left(\frac{\hat{b}_{l1}}{\hat{\sigma}_1}\right)}{1 + \sum_{m=2}^{L}\exp\left(\frac{\hat{b}_{m1}}{\hat{\sigma}_1}\right)}. \tag{31}$$

Since the aggregated uncooked food commodity is unobserved, the constant A_1 is unidentified. But this is just a matter of normalization. I set the value of A_1 so that the average vale of r across households is equal to the average amount of all raw food purchased, which is equal to 110 kg per month. Given these parameter estimates, the composite uncooked food commodity and its price index are

$$r = A_1\left(\sum_{l=1}^{L}\hat{\gamma}_l r_l^{\hat{\rho}_1}\right)^{\frac{1}{\hat{\rho}_1}}, \tag{32}$$

$$p_r = \frac{1}{A_1}\left(\sum_{l=1}^{L}\hat{\gamma}_l^{\frac{1}{1-\hat{\rho}_1}} p_l^{-\frac{\hat{\rho}_1}{1-\hat{\rho}_1}}\right)^{-\frac{1-\hat{\rho}_1}{\hat{\rho}_1}}. \tag{33}$$

I estimate (30) using itemized expenditure data from all households in the CPHS sample. Since the CPHS does not collect data on prices, I use a state-month panel of food commodity price data from the World Food Programme. Combining the household expenditure data and the food price data, I calculate the amount of raw food (in kilograms per month) purchased for seven food categories: cereals, pulses, vegetables, dairy, oils, spices, and tea. Table F1 reports the estimates.

Table F1: Estimates of CES aggregator of raw food commodities

	Point estimate (SE)
σ_1	0.77
	(0.018)
ρ_1	-0.29
	(0.031)
γ_{cereals}	0.37
	(0.008)
γ_{pulses}	0.04
	(0.002)
γ_{veg}	0.28
	(0.005)
γ_{dairy}	0.19
	(0.006)
γ_{oils}	0.06
	(0.003)
γ_{spices}	0.03
	(0.001)
γ_{tea}	0.04
	(0.003)
N	2770573

Note: Standard errors computed via delta method, clustered at the state-wave level.

Second, I estimate the outer nest of the home production function. As explained in the main text, this estimation procedure uses data from households in which the mother is the only woman over the age of 12. Importantly, the price of time is only observed for households in which the mother works for a wage (7% of households), and the decision to work may be correlated with her productivity at home. Thus, I correct for sample selection by implementing Heckman's correction procedure for the case of missing data on explanatory variables. The model is

$$\ln\left(\frac{t}{r}\right) = b_0 + \sigma \ln\left(\frac{p_r}{w_M}\right) + \xi, \tag{34}$$

$$\ln\left(\frac{p_r}{w_M}\right) = \mathbf{x}_m\boldsymbol{\delta_1} + u_1, \tag{35}$$

$$l_M = 1\left[\mathbf{x}_m\boldsymbol{\delta_1} + \mathbf{z}\boldsymbol{\delta_2} + u_2 > 0\right], \tag{36}$$

where $\ln\left(\frac{p_h}{w_M}\right)$ is only observed when $l_M = 1$ and $u_2 \sim \text{Normal}(0,1)$.

Even if $\ln\left(\frac{p_h}{w_M}\right)$ is assumed to be exogenous, endogenous sample selection effectively makes $\ln\left(\frac{p_h}{w_M}\right)$ endogenous in the selected sample and so it must be instrumented (Wooldridge, 2010).

Equation (35) is a linear projection for $\ln\left(\frac{p_h}{w_M}\right)$, where $\mathbf{x}_m$ is a vector of determinants of the offered wage: age, age squared, caste, education, rurality, state, wave, and the average male wage in the district. The assumption is that these variables only affect the ratio between time and material inputs into home production by shifting the ratio between the wage and food prices. The vector $\mathbf{z}$ consists of variables that determine selection into work: exogenous household income (y) and number of children.

Table F2 reports the probit estimates of the selection equation. Consistent with the model, exogenous income and number of children are strong predictors of female labor force participation.

Table F3 presents the estimates of the home production function. Column (1) shows the estimates ignoring sample selection and column (2) corrects for selection. As one would expect, the estimates suggest a low elasticity of substitution between time and raw food in home production.

F.2 Offered wages

I estimate wage equations using data from wage workers and correct for sample selection using Heckman's correction procedure. Logged offered wages of women are modeled as a linear function of demographic characteristics and the average male wage in the district.[23] Selection into work depends on these variables as well as on predetermined household income and number of children. The model is estimated using adult women from all the CPHS sample (not only those in households with parental death).[24]

Logged offered wages of sons are modeled in a similar fashion, except that educational attainment is excluded from the model because it is an endogenous choice. The model is estimated using data from men between the ages 12 and 18. The selection equation is estimated with data from non-orphans only.

Tables F4 and F5 report the estimates of the wage equations for mothers and sons, respectively. Figure F1 plots the average observed wages by age, as well as the predicted offered wages for participants and non-participants.

[23] The demographic characteristics that enter the wage equation are age, age squared, caste, education, rurality, state, and wave.

[24] Since predetermined household income is correlated with widowhood, which is not random, the selection equation is estimated with data from married women only.

Table F2: Selection into female wage labor (Probit estimates)

	(1) Probit coef. (SE)
Log(Exogenous HH income)	-0.119***
	(0.005)
Number of children	0.095***
	(0.007)
Avg district male wage	-0.214***
	(0.022)
Age - 30	0.027***
	(0.002)
$(\text{Age} - 30)^2$	-0.001***
	(0.000)
Upper caste	0.000
	(.)
Intermediate caste	-0.109***
	(0.023)
OBC	0.130***
	(0.019)
SC	0.320***
	(0.019)
ST	0.600***
	(0.023)
No primary	0.000
	(.)
Primary	-0.188***
	(0.013)
Jr HS	-0.307***
	(0.019)
Sr HS	-0.207***
	(0.027)
College	0.214***
	(0.022)
Rural	0.000
	(.)
Urban (small)	-0.311***
	(0.015)
Urban (medium)	-0.276***
	(0.015)
Urban (large)	-0.400***
	(0.015)
Urban (very large)	-0.445***
	(0.018)
Constant	1.034***
	(0.158)
N	510647

Table F3: Estimates of equation (16)

	Ignore selection	Correct for selection
	(1)	(2)
	OLS	Heckit-IV
	Point estimate (SE)	Point estimate (SE)
σ_0	0.22	0.27
	(0.01)	(0.02)
ρ_0	-3.52	-2.72
	(0.22)	(0.29)
γ	0.90	0.89
	(0.00)	(0.01)
$1-\gamma$	0.10	0.11
	(0.00)	(0.01)
N	15731	14934

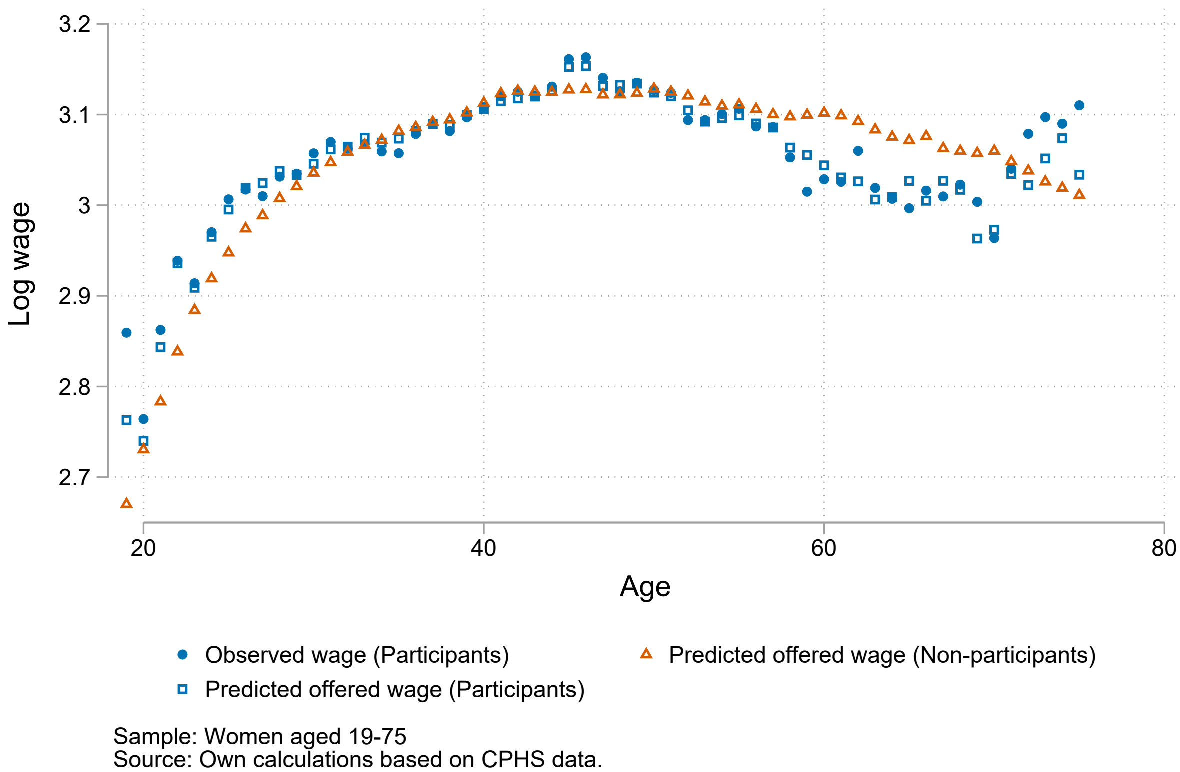
Table F4: Female wage equation

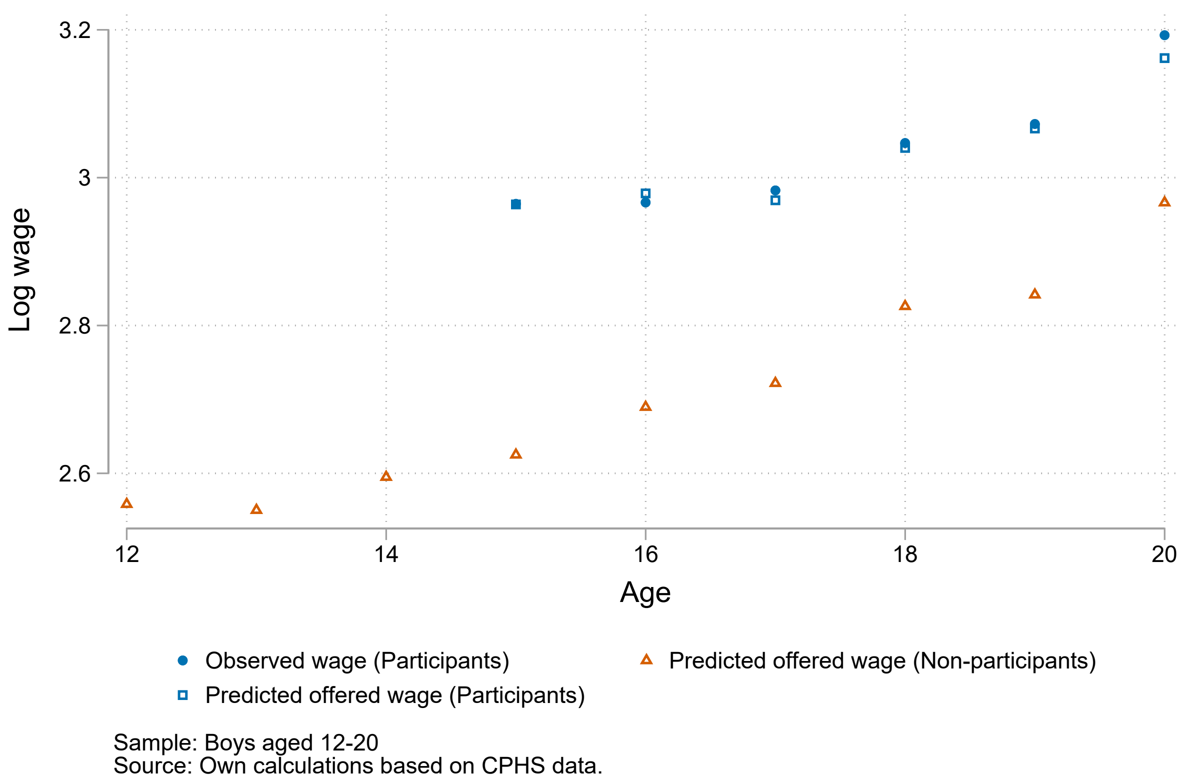
DV: Log(wage)	No correction Coef. (SE)	Selection correction Coef. (SE)
Avg district male wage	0.518***	0.501***
	(0.005)	(0.005)
Age - 30	0.011***	0.013***
	(0.000)	(0.000)
$(\text{Age} - 30)^2$	-0.000***	-0.000***
	(0.000)	(0.000)
Upper caste	0.000	0.000
	(.)	(.)
Intermediate caste	-0.025***	-0.031***
	(0.006)	(0.006)
OBC	-0.090***	-0.078***
	(0.005)	(0.005)
SC	-0.141***	-0.112***
	(0.005)	(0.005)
ST	-0.110***	-0.075***
	(0.005)	(0.006)
No caste	0.054***	0.064***
	(0.013)	(0.013)
No primary	0.000	0.000
	(.)	(.)
Primary	0.075***	0.064***
	(0.003)	(0.003)
Jr HS	0.238***	0.217***
	(0.005)	(0.005)
Sr HS	0.449***	0.424***
	(0.007)	(0.007)
College	0.935***	0.921***
	(0.006)	(0.006)
Rural	0.000	0.000
	(.)	(.)
Urban (small)	0.126***	0.105***
	(0.003)	(0.003)
Urban (medium)	0.192***	0.170***
	(0.003)	(0.004)
Urban (large)	0.162***	0.133***
	(0.003)	(0.004)
Urban (very large)	0.208***	0.175***
	(0.004)	(0.004)
IMR		0.100***
		(0.006)
Constant	1.001***	0.915***
	(0.018)	(0.019)
State FE	Yes	Yes
N	345930	341118

Table F5: Son wage equation

DV: Log(wage)	No correction Coef. (SE)	Selection correction Coef. (SE)
Avg district male wage	0.476***	0.398***
	(0.014)	(0.015)
Age - 15	-0.003	0.042***
	(0.014)	(0.014)
$(\text{Age} - 15)^2$	0.005	0.006*
	(0.004)	(0.004)
Upper caste	0.000	0.000
	(.)	(.)
Intermediate caste	0.048***	0.007
	(0.018)	(0.018)
OBC	-0.018*	0.003
	(0.010)	(0.010)
SC	-0.082***	-0.038***
	(0.010)	(0.010)
ST	-0.134***	-0.081***
	(0.013)	(0.014)
No caste	-0.061*	-0.070**
	(0.035)	(0.035)
Rural	0.000	0.000
	(.)	(.)
Urban (small)	0.012*	0.021***
	(0.007)	(0.007)
Urban (medium)	0.023***	0.029***
	(0.007)	(0.007)
Urban (large)	0.016**	0.019**
	(0.008)	(0.008)
Urban (very large)	0.028***	0.031***
	(0.009)	(0.009)
IMR		0.196***
		(0.018)
Constant	1.276***	1.117***
	(0.052)	(0.056)
State FE	Yes	Yes
N	35012	34979

Figure F1: Predicted offered wages for mothers and sons





F.3 Sharing rule and scale economies

Figure F2 shows the mean budget shares for widows, widowers, and couples in my sample. Note that "couples" are the same households of "widows" and "widowers" observed before the paternal and maternal death shock, respectively.

Figure F2: Budget shares of residual income for widows, widowers, and couples

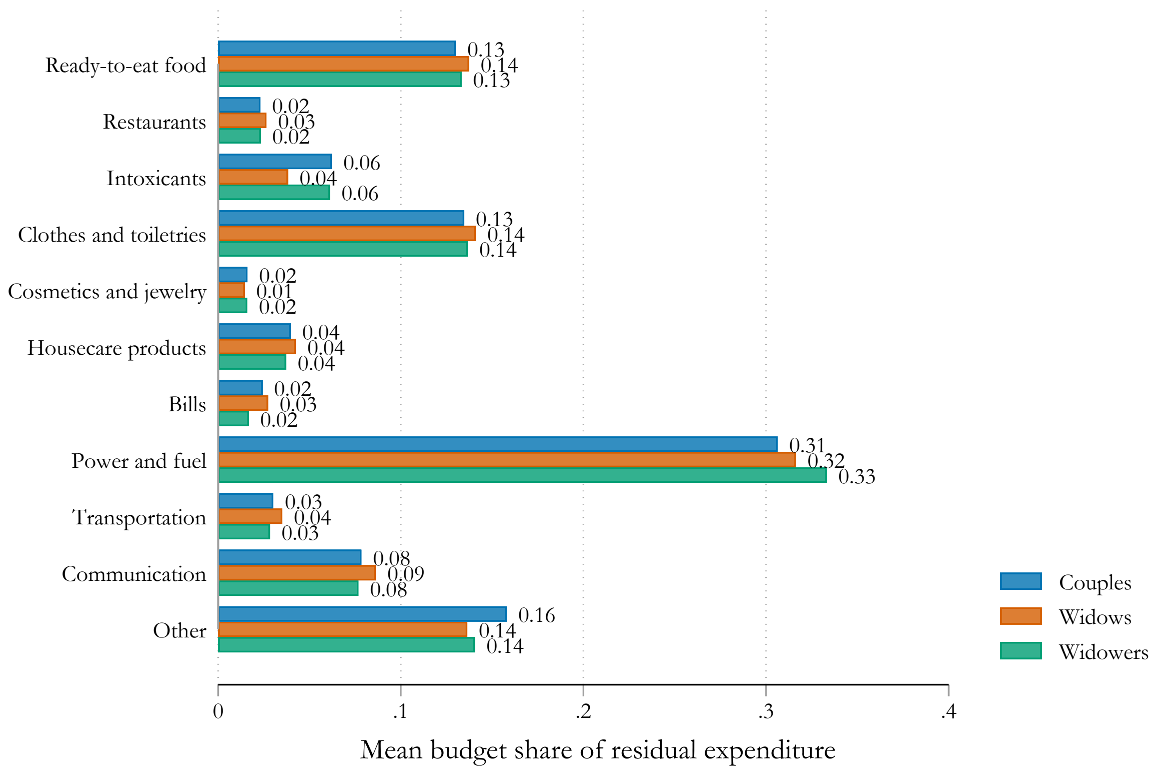


Table F6 shows the estimates of the sharing rule and scale economies. The estimes of the sharing rule are discussed in the main text. To interpret the economies of scale parameters, recall that $\Delta_i(\mathbf{d}_i) = e^{\left(q_{0i}+\mathbf{q}_i'\mathbf{d}_i\right)}$ is a measure of the cost savings due to scale economies experienced by parent i, where Δ is a scalar between 0.5 (all goods are fully shareable) and 1 (all goods are non-shareable). The estimates are largely within the expected range and suggest that mothers enjoy higher economies of scale than fathers ($\Delta_m = e^{-0.780} = 0.46$ and $\Delta_f = e^{-0.285} = 0.75$).

Given the estimates of the sharing rule and scale economies, I can compute indifference scales for each parent, $I_i(\mathbf{d}_i) = \frac{\eta_i(\mathbf{d}_i)}{\Delta_i(\mathbf{d}_i)}$. The indifference scale of parent i is the fraction of the couple's total expenses that i would require in widowhood to reach the same utility from market goods that he/she attained when married. Among couples of age 40 with no post-primary education, the indifference scales are $I_M = \frac{0.407}{0.46} = 0.88$ and $I_F = \frac{1-0.407}{0.75} = 0.79$. That is, a widow, for example,

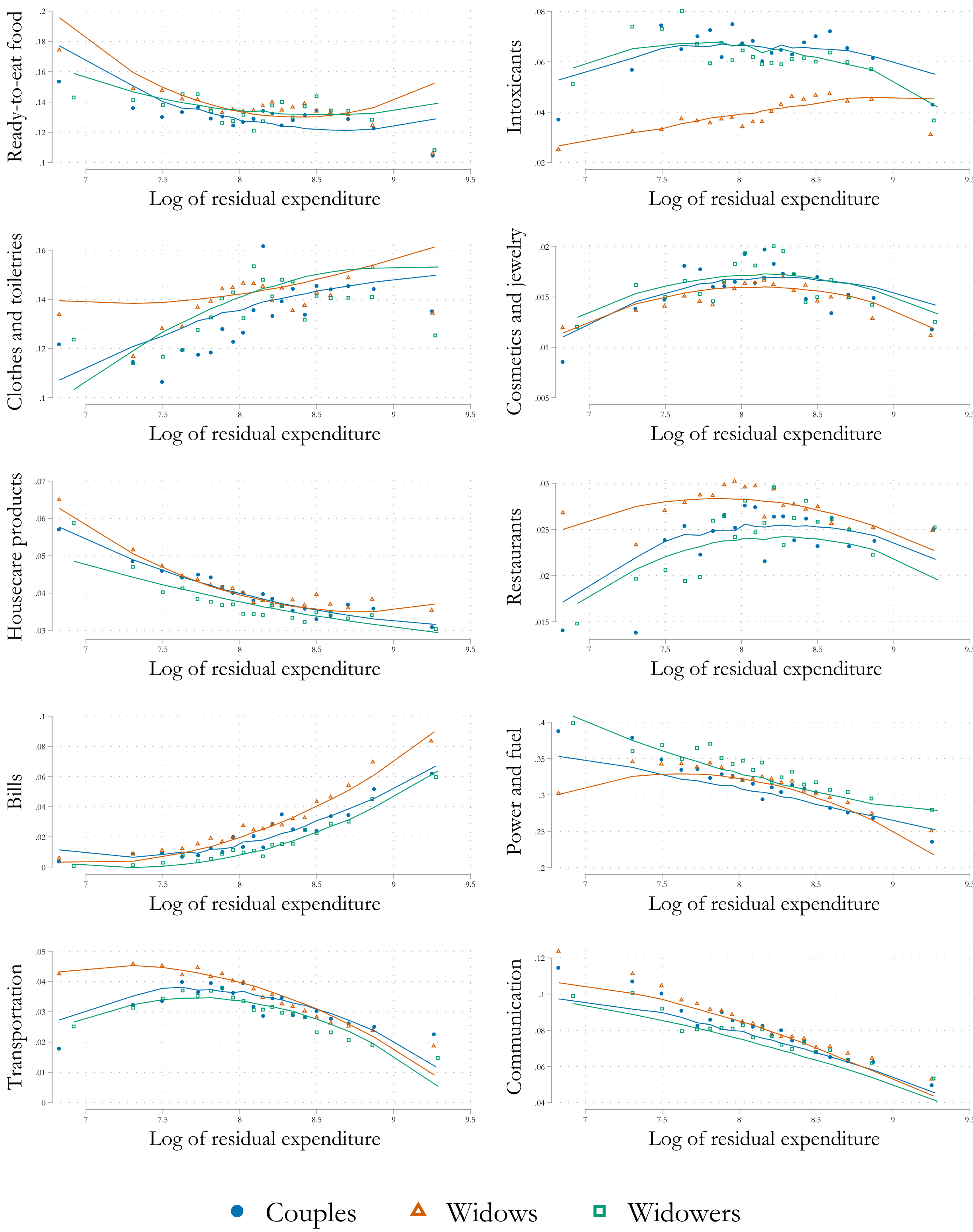
Table F6: Estimates of sharing rule and scale economies

	(1)
r_0	0.407***
	(0.069)
r_{age_m}	0.012***
	(0.004)
r_{hs_m}	0.014
	(0.009)
r_{age_f}	-0.011***
	(0.003)
r_{hs_f}	-0.014*
	(0.008)
q_{0m}	-0.780***
	(0.291)
q_{age_m}	0.030***
	(0.008)
q_{hs_m}	0.493***
	(0.165)
q_{0f}	-0.285*
	(0.154)
q_{age_f}	-0.018***
	(0.005)
q_{hs_f}	-0.063
	(0.063)
N (Widows)	19227
N (Widowers)	4816
N (Couples)	5039

would require about 90% of the couple's total expenses to reach the same utility from market goods she attained while married.

Figure F3 shows the empirical budget share Engel curves (markers) and the model's prediction (solid lines). The model provides a good fit for most commodities.

Figure F3: Budget share Engel curve system - Data and model



F.4 Method of simulated moments

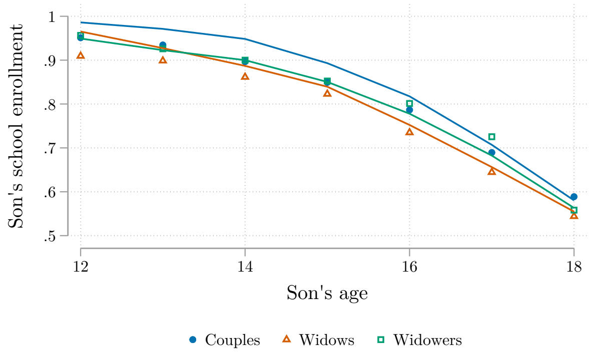
The estimator targets four types of moments. First, it targets the cross-household average of all choices at baseline (i.e. pre-death). Second, I split the sample into six groups according to predetermined income and parental vital status, and target the average choices in each group. Third, the estimator targets the effects of paternal death and maternal death on all choices, controlling for age, wave, and household random effects. Finally, to identify the variance of the household-specific taste shocks, the estimator targets the estimated variance of the household random effects from these regressions. Table F7 lists all targeted moments. The estimator minimizes the distance between simulated moments and data moments, where the weighting matrix is the inverse of the variance-covariance matrix of the data moments.

The model fits targeted well, except for the effect of maternal death on daughter's time in home production. The model overpredicts the increase in daughter's time in home production following maternal death (see Table F7). While the death of a mother increases daughters time in home production, the observed effect is moderate compared to what the model predicts given the production function estimates. Figure F4 shows the model's performance on non-targeted moments, namely school enrollment as a function of age, of predetermined income, and of offered wages.

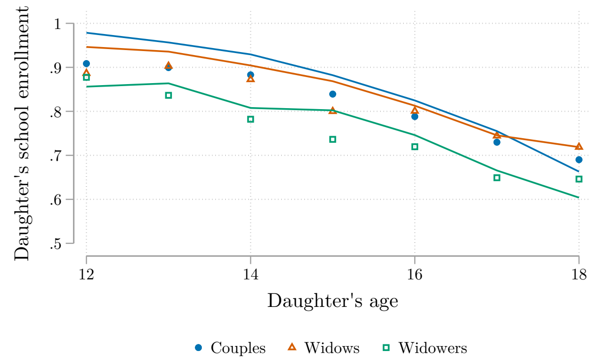
Table F7: Targeted data and simulated moments

	Data	Simulated
Effect of paternal death on son's schooling	-0.11	-0.08
Effect of maternal death on son's schooling	-0.05	-0.08
Effect of age on son's schooling	-0.06	-0.08
Effect of paternal death on daughter's schooling	-0.05	-0.03
Effect of maternal death on daughter's schooling	-0.09	-0.11
Effect of age on daughter's schooling	-0.04	-0.06
Effect of paternal death on uncooked food	-8.62	-14.58
Effect of maternal death on uncooked food	-2.88	-16.96
Effect of paternal death on daughter's home production time	5.74	7.38
Effect of maternal death on daughter's home production time	14.70	38.94
Effect of age on daughter's home production time	3.72	3.45
Effect of paternal death on mother's home production time	-12.36	-19.01
Effect of paternal death on mother's labor participation	0.36	0.34
Schooling of sons at baseline	0.80	0.82
Schooling of daughters at baseline	0.82	0.85
Uncooked food purchased at baseline	94.52	91.52
Daughter's home production time at baseline	29.87	31.82
Mother's home production time at baseline	75.70	87.65
Mother's labor participation at baseline	0.11	0.12
Schooling in waves 1-7 at baseline	0.78	0.82
Schooling in waves 8-14 at baseline	0.81	0.85
Schooling in waves 15-21 at baseline	0.86	0.88
Schooling in waves 22-28 at baseline	0.93	0.88
Baseline - Low y: Schooling of sons	0.76	0.76
Baseline - High y: Schooling of sons	0.83	0.89
Baseline - Low y: Schooling of daughters	0.79	0.86
Baseline - High y: Schooling of daughters	0.85	0.85
Baseline - Low y: Uncooked food purchased	85.14	74.21
Baseline - High y: Uncooked food purchased	104.51	108.83
Baseline - Low y: Daughter's home production time	31.60	33.59
Baseline - High y: Daughter's home production time	28.53	30.46
Baseline - Low y: Mother's home production time	73.22	83.61
Baseline - High y: Mother's home production time	77.55	90.65
Baseline - Low y: Mother's labor participation	0.16	0.20
Baseline - High y: Mother's labor participation	0.06	0.03
Post paternal death - Low y: Schooling of sons	0.67	0.67
Post paternal death - High y: Schooling of sons	0.77	0.82
Post paternal death - Low y: Schooling of daughters	0.79	0.82
Post paternal death - High y: Schooling of daughters	0.80	0.82
Post paternal death - Low y: Uncooked food purchased	64.49	59.49
Post paternal death - High y: Uncooked food purchased	80.47	87.92
Post paternal death - Low y: Daughter's home production time	35.47	44.07
Post paternal death - High y: Daughter's home production time	36.35	34.54
Post paternal death - Low y: Mother's home production time	52.84	61.97
Post paternal death - High y: Mother's home production time	68.91	73.17
Post paternal death - Low y: Mother's labor participation	0.79	0.81
Post paternal death - High y: Mother's labor participation	0.21	0.15
Post maternal death - Low y: Schooling of sons	0.74	0.69
Post maternal death - High y: Schooling of sons	0.79	0.82
Post maternal death - Low y: Schooling of daughters	0.73	0.75
Post maternal death - High y: Schooling of daughters	0.72	0.73
Post maternal death - Low y: Uncooked food purchased	72.78	61.11
Post maternal death - High y: Uncooked food purchased	89.23	89.28
Post maternal death - Low y: Daughter's home production time	47.88	63.99
Post maternal death - High y: Daughter's home production time	39.89	66.81
Variance of household random effect in son's schooling	0.07	0.07
Variance of household random effect in daughter's schooling	0.06	0.08
Variance of household random effect in uncooked food	416.76	504.91
Variance of household random effect in daughter's home production time	267.40	566.97
Variance of household random effect in mother's home production time	380.95	1053.73
Variance of household random effect in mother's labor participation	0.06	0.05
N	49,679	198,716

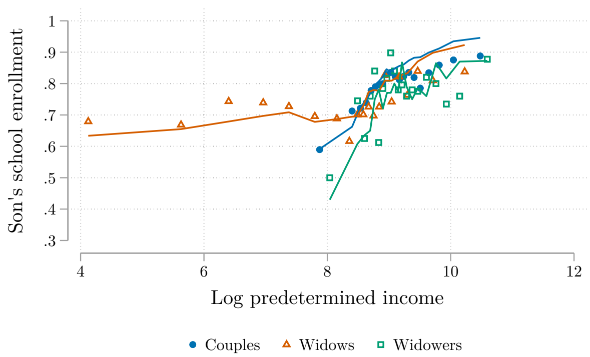
Figure F4: Model fit - Non-targeted moments



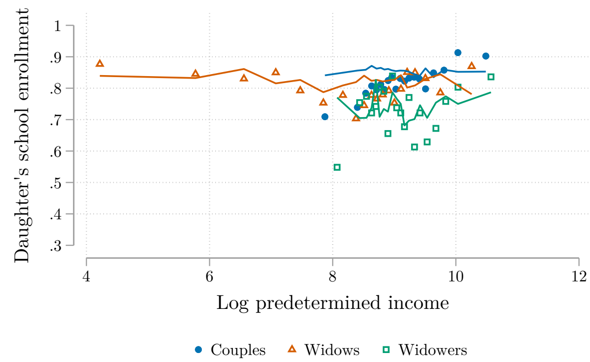
Note: Markers correspond to data and lines correspond to the model.



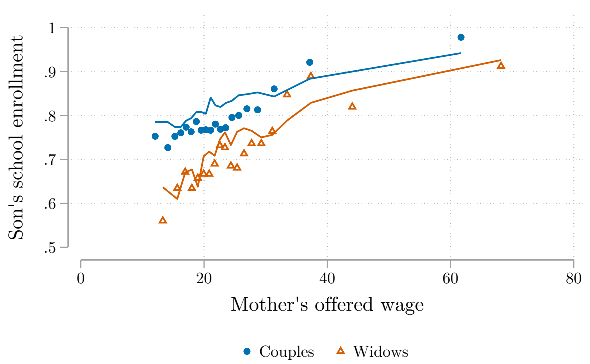
Note: Markers correspond to data and lines correspond to the model.



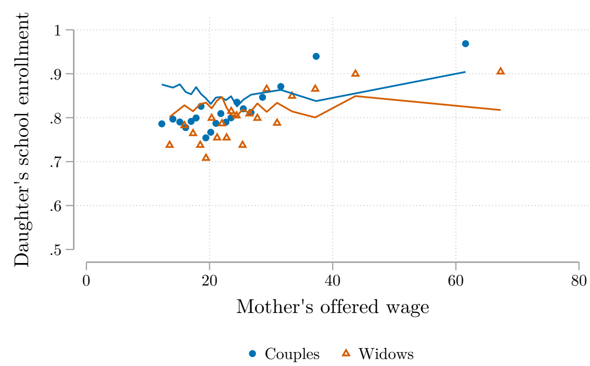
Note: Markers correspond to data and lines correspond to the model.



Note: Markers correspond to data and lines correspond to the model.



Note: Markers correspond to data and lines correspond to the model.



Note: Markers correspond to data and lines correspond to the model.

F.4.1 Sensitivity analysis

Andrews et al. (2017) propose a sensitivity matrix that measures the relationship between parameter estimates and the moments of the data they depend on. I compute this sensitivity matrix and plot the results for each parameter in Figures F5-F8. The 61 targeted moments, organized into 7 groups, are placed in the horizontal axis. Each marker represents the impact of a standard deviation change of the corresponding moment on the estimated parameter. Blue markers indicate a positive effect and red markers indicate a negative effect. For example, the top left graph of Figure F5 shows that the estimate of the mother's preference for son's schooling is mostly driven by the moments associated with son's schooling and mother's labor supply.

Figure F5: Sensitivity (1)

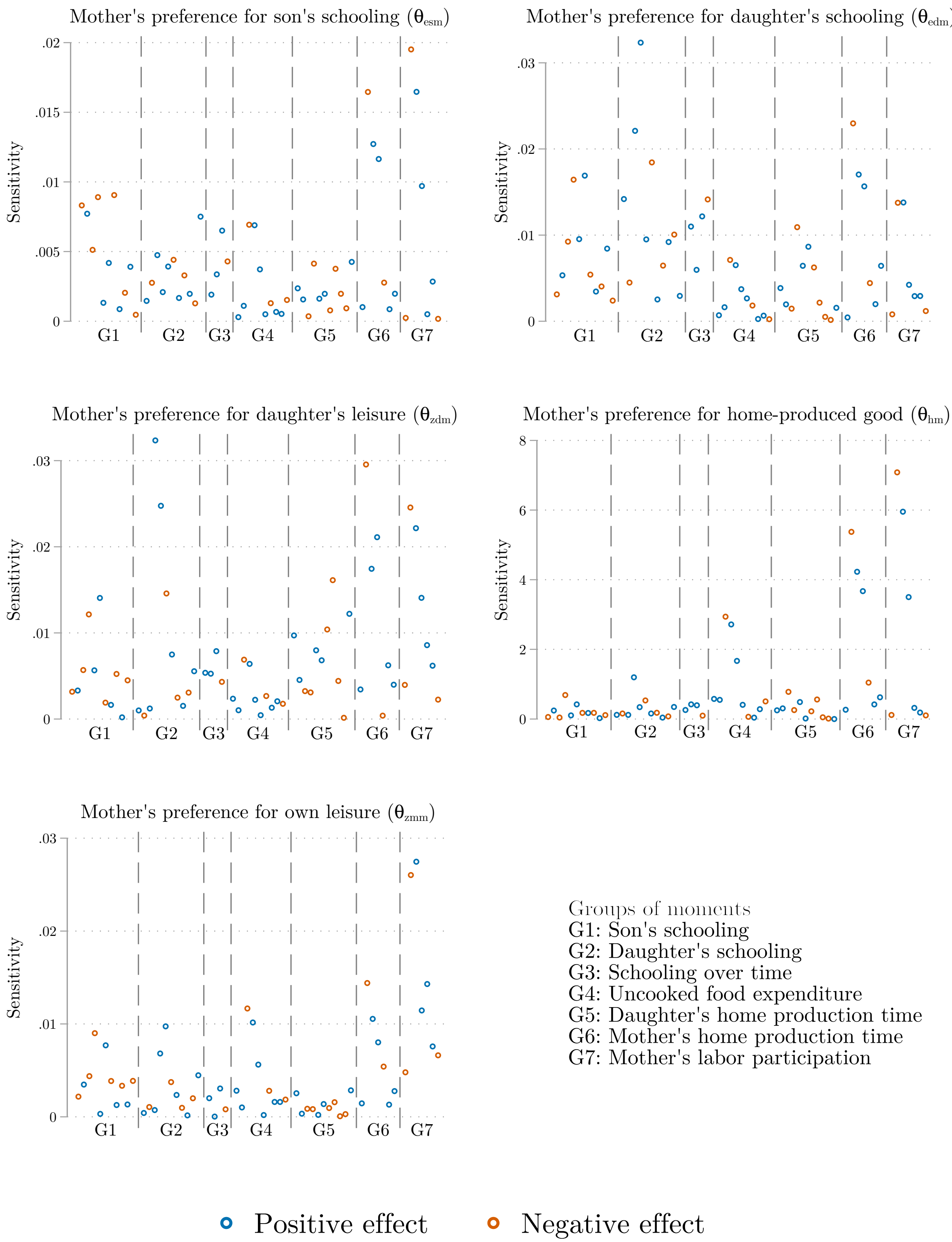


Figure F6: Sensitivity (2)

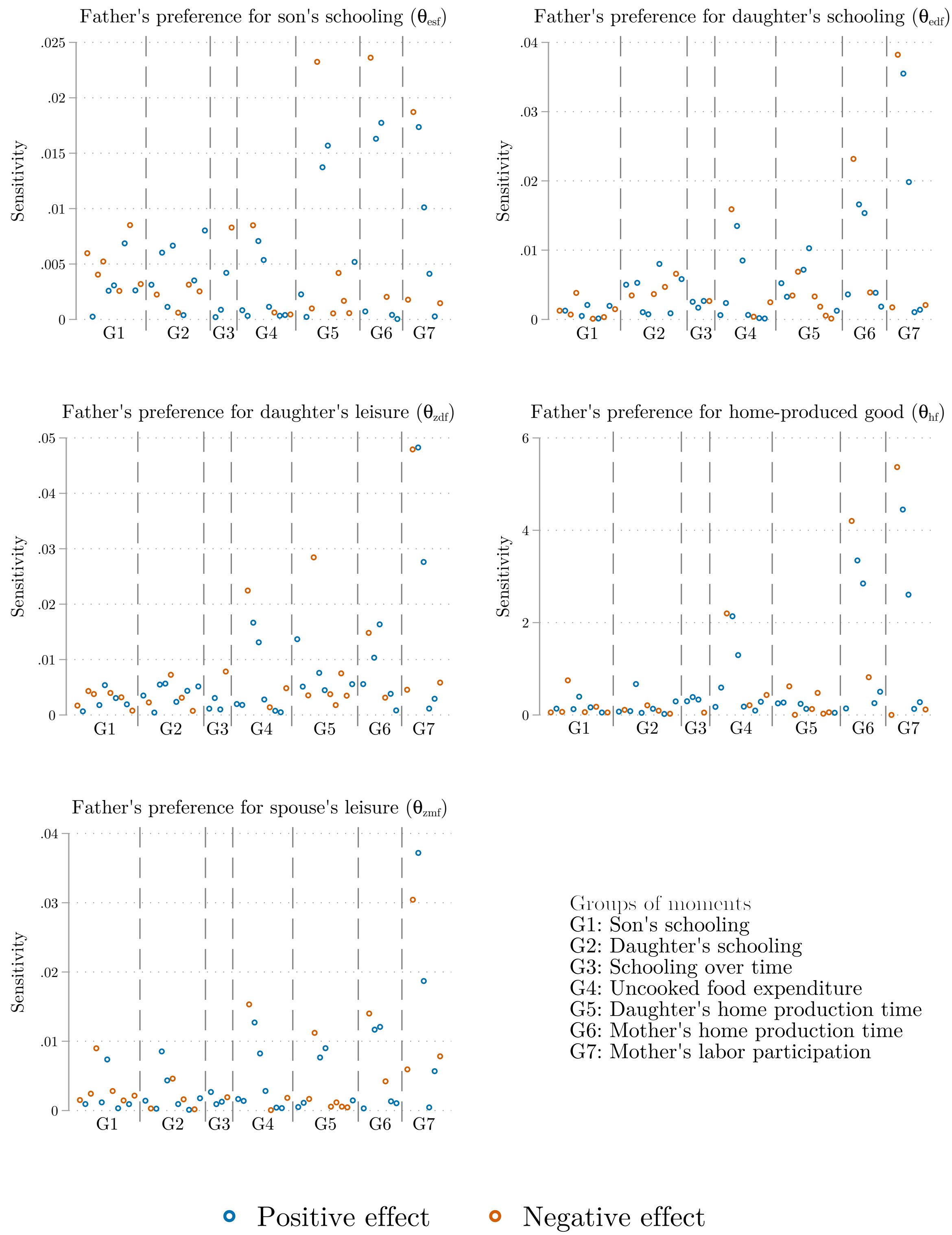


Figure F7: Sensitivity (3)

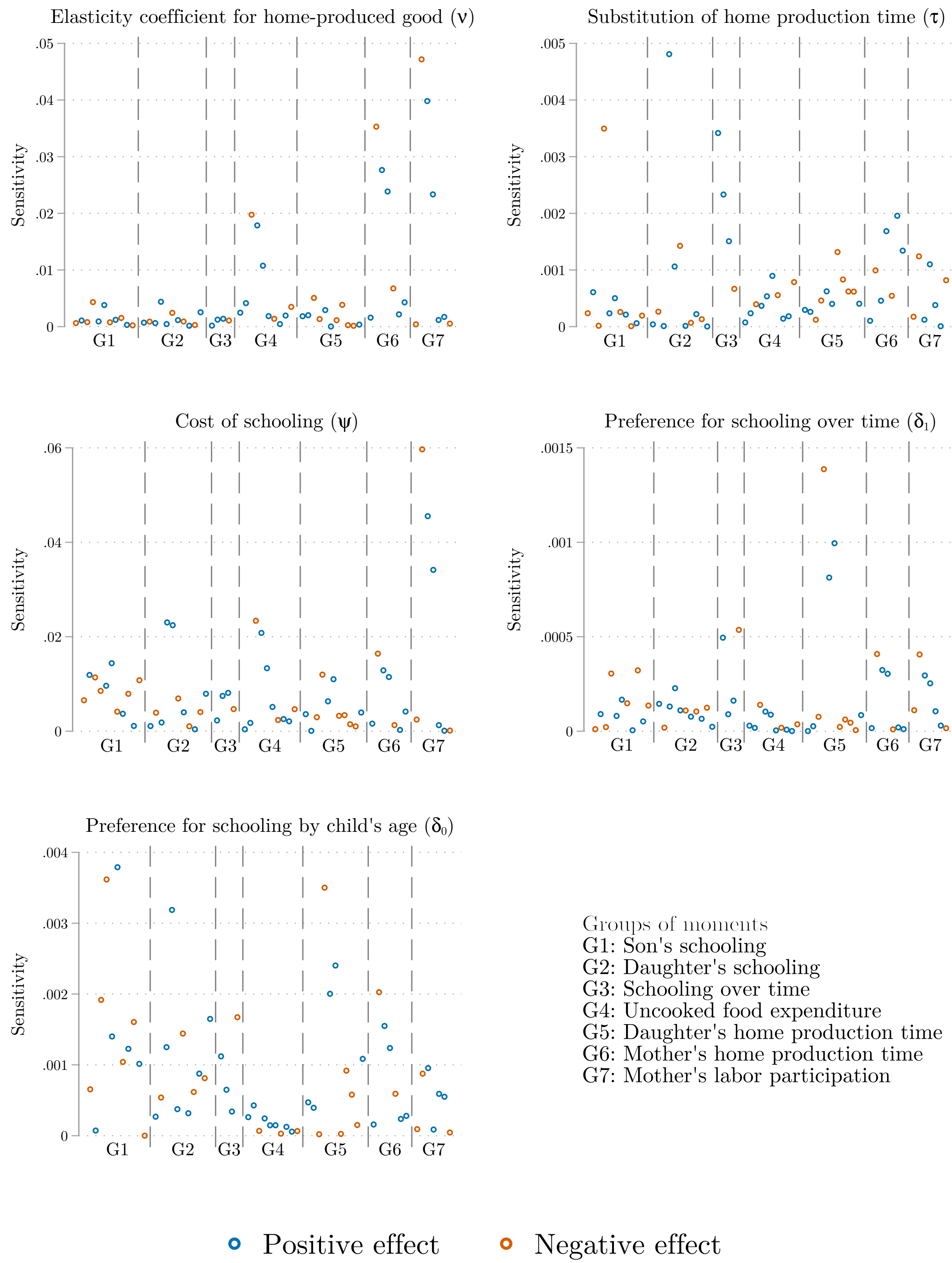


Figure F8: Sensitivity (4)