LATE With Mismeasured or Misspecified Treatment: An Application On Woman Empowerment in India

Rossella Calvi*[1], Arthur Lewbel†[2], and Denni Tommasi‡[3]

[1]Department of Economics, Rice University

[2]Department of Economics, Boston College

[3]ECARES, Université Libre de Bruxelles

December 31, 2016

[PRELIMINARY – Do Not Circulate or Cite]

Abstract

We show that a local average treatment effect (LATE) can sometimes be identified and consistently estimated when treatment is mismeasured, or when treatment is estimated using a possibly misspecified structural model. Our associated estimator, which we call Mismeasurement Robust LATE (MR-LATE), is based on differencing two different mismeasures of treatment. In our empirical application, treatment is a measure of empowerment: whether a wife has control of substantial household resources. Due to measurement difficulties and sharing of goods within a household, this treatment cannot be directly observed without error, and so must be estimated. Our outcomes are health indicators of family members. We first estimate a structural model to obtain the otherwise unobserved treatment indicator. Then, using changes in inheritance laws in India as an instrument, we apply our new MR-LATE estimator. We find that women's empowerment substantially decreases their probability of being anemic or underweight, and increases children's likelihood of receiving vaccinations. We find no evidence of negative effects on men's health.

JEL Codes: D13, D11, D12, C31, I32

Keywords: causality, LATE, structural model, collective model, resource shares, bargaining power, health.

*Rice University, P.O. Box 1892, Houston, TX 77251-1892, USA (E-mail: rossella.calvi@rice.edu).

†Boston College, Maloney Hall room 315, 140 Commonwealth Avenue, Chesnut Hill, MA 02467, USA (e-mail: arthur.lewbel@bc.edu).

‡ECARES, Avenue F. D. Roosevelt 50, CP 114, B-1050 Brussels, Belgium (e-mail: dtommasi@ulb.ac.be).

We thank Jonathan De Quidt, Thomas Demuynck, Bram De Rock, Miriam Gensowski, Mette Gørtz, Nikolaj Harmon, Thomas Høgholm Jørgensen, Rasmus Jørgensen, Anna Folke Larsen, Mathieu Parenti, Søren Leth-Petersen, Torben Heien Nielsen, Benjamin Solow, and seminar participants at the University of Copenhagen, ECARES and Stockholm University for their suggestions. We are especially grateful to Yingying Dong for helpful insight. All errors are our own.

1 Introduction

In the estimation of causal effects, treatment is often observed with measurement error, also called misclassification error in the case of discrete treatments. We show that a local average treatment effect (LATE) can sometimes be identified and consistently estimated even though treatment is mismeasured. Our associated estimator, which we call MR-LATE (for Mismeasurement Robust LATE) is based on the difference between two different mismeasures of treatment. Unlike, e.g, Ura (2015), who only obtains bounds, our MR-LATE estimator (which is empirically trivial to implement) can sometimes point identify LATE.

Treatment indicators can be mismeasured for a variety of reasons, such as reporting and recording errors. In addition, if treatment is not directly observed and must itself be estimated based on some model, then both misspecification of the model and estimation error in the estimated treatment can be interpreted as measurement errors in the observed (i.e., estimated) treatment. While our estimator should be useful in many applications where treatment is observed with error, we want to emphasize its potential in applications where treatment is estimated. Due to estimation and potential specification error in treatment estimation, treatment is rarely estimated (particularly structurally estimated) in the causal literature. This is unfortunate, because only considering treatments that are directly observable greatly limits the potential economic significance of estimated causal effects.

For example, in our empirical application we define treatment as a wife having actual (not merely reported) control of substantial household resources. This is difficult to observe, in part due to shared consumption within the household, and so must be estimated. Our outcomes are health measures of household members, and our instrument is a plausibly exogenous change in inheritance laws in India. A typical causal analysis might look directly at the causal impact of the law change on health, but this would tell us nothing about how other changes in women's empowerment might impact health. In contrast, a typically structural analysis of this problem would require not only modeling *resource shares* (i.e., each member's share of total household consumption), but also structurally modeling how inheritance laws affect resource shares, and how resource shares affect health. Many might find these structural models implausible. This summarizes the commonly noted tradeoff that causal reduced form analyses typically answer problems of lesser interest, while structural analyses depend on strong assumptions regarding underlying true behavior.

Our analysis seeks to combine the best of both worlds. We know little about exactly how female power in the form of resource shares may affect health outcomes, and so we address that question using LATE estimation. A great deal more is known about the economics of household consumption, so we employ structural estimation to obtain resource shares, and we employ our new MR-LATE estimator to account for the inevitable estimation and specification errors that arise from employing a structural model of treatment. The end result is that instead of asking what the impact is of a particular randomized intervention (the inheritance law change), we address a more general question: What is the impact of empowering women (by giving them control of household resources)

on family health?

There are many other examples of potential treatments of clear economic significance, that are rarely analyzed causally, because the treatments themselves would need to be estimated. Examples of such treatments are measures of expectations, ability, opportunity, or utility. One might be interested, for example, in quantifying the effect of individuals' future discount rates on their probability of dropping out of school, or the effect of individuals' risk aversion on their investment decisions, or the effect of abilities like non-cognitive skills on future earnings. These are typically estimated using fully structural models. The alternative we propose is using structure to estimate the treatment, and causal methods to estimate the treatment effect, accounting for estimation and specification error in the structural estimation of treatment.

Stemming from the work of Amartya Sen and many others, there is a large literature pointing to the importance of female empowerment in the household. We rely on the collective households framework to structurally recover a measure of bargaining power and, in turn, to estimate our unobserved treatment variable. In this framework, a household is characterized as a collection of individuals, each of whom has a well defined utility function, and who interact to generate Pareto efficient allocations. Recent advances in this literature permit the recovery of resource shares (or *sharing rule*). These measures correspond to intra-household bargaining power or control.

We estimate resource shares for women, men or children in the household using detailed expenditure data from the 2011-2012 National Sample Survey (NSS) of Consumer Expenditure in India. Due to the lack of NSS data on health outcomes, we use these estimates to perform an out-of-sample prediction on the 2005-2006 Indian National Family Health Survey (NFHS-3), which includes the same socio-economic characteristics of individuals and households as the NSS dataset and detailed information about women's (aged 15 to 49), men's (aged 15 to 54) and children's (aged 0 to 5) health indicators. Given this structural estimation of resource shares, we use our new LATE estimator to study the causal effect of living in households with relatively highly empowered women on household members' health status. We define a household to be treated if it falls above the median (or other cutoff) of the distribution of our preferred measure of women's empowerment, i.e., the fraction of parents' resources devoted to the mother. We exploit changes to the Indian inheritance law to construct our instrumental variable. Specifically, we focus on women's exposure to the Hindu Succession Act (HSA) amendments that equalized women's inheritance rights to men's in several Indian states between 1976 and 2005.

We find that accounting for specification, estimation, and/or measurement error in the estimate of treatment is empirically important, with some substantial differences between the standard LATE estimator and our MR-LATE estimator. The results of our empirical analysis indicate that women's empowerment positively affects women's and children's health outcomes, but does not affect men's health outcomes. Our most conservative estimates indicate that the average causal effect on women's body mass index (BMI) is 1.72 and that women in above-median households are 7 percent less likely to be underweight. Moreover, empowered mothers face longer intervals between births, which may improve maternal health and reduce the risk of anemia. Also, mother's

empowerment positively affects children's likelihood to receive vaccinations. The average causal effect on children's probability of receiving a vaccine is an increase of 0.29 in probability, mainly driven by immunization against polio.

The next section is a literature review. This is followed by our identification proof and the derivation of our MR-LATE estimator, including a small Monte-Carlo study of its small sample properties. We then describe our empirical application and results, followed by Conclusions. Proofs and additional material are in an Appendix.

2 Literature Review

Papers empirically documenting substantial measurement (misclassification) errors in observed treatments include Bollinger (1996), Angrist and Krueger (1999), Kane, Rouse, and Staiger (1999), Card, (2001), Black et al. (2003), and Hernandez and Pudney (2007). Homogeneous treatment effects, corresponding to estimation of constant coefficients of a mismeasured binary regressor, have been estimated using instruments by many authors, including Aigner (1973), Kane, Rouse, and Staiger (1999), Black, Berger, and Scott (2000) and Frazis and Loewenstein (2003).

When treatment is mismeasured, point identification (and associated estimators) of average treatment effects (without assuming treatment effects are homogeneous) are provided by Mahajan (2006), Lewbel (2007) and Hu (2008). These papers obtain identification exploiting the assumption that true treatment is exogenous. Under more general conditions bounds on average treatment effects with misclassified treatment are provided by Klepper (1988), Manski (1990), Bollinger (1996), Kreider and Pepper (2007), Molinari (2008), and Imai and Yamamoto (2010), and Kreider, Pepper, Gundersen, and Jollie (2012).

The causal effect we focus on identifying and estimating is the local average treatment effect (LATE) of Imbens and Angrist (1994) and Angrist et al. (1996), which is applicable when the true treatment is endogenous, an exogenous instrument is available, and treatment effects are not homogeneous. Ura (2015) also considers estimation of LATE with mismeasured treatment and standard LATE instrument assumptions[1], but only obtains set identification bounds. We obtain point identification by making use of two rather than a single mismeasured treatment indicator.

The structure of our estimation problem is that a randomized instrument affects mismeasured treatment, and the true treatment then affects an outcome. A similar structure arises in models where outcomes of interest and randomized treatment are not available in the same data set. In these models, a randomized treatment (corresponding to our instrument) affects an intermediate outcome called a *statistical surrogate* (see Prentice (1989)), corresponding to our treatment indicator. The surrogate then effects (or at least strongly correlates with) the outcome of interest. These estimators require that the surrogate satisfy a strong conditional independence assumption.

[1]Specifically, the instrument must satisfy the exclusion restriction and weakly increase the true treatment (Imbens and Angrist, 1994; Angrist et al., 1996). In the case of a continuous misclassified treatment, Lewbel (1998), Song et al. (2015), Hu et al. (2015) and Song (2015) use instruments and further exclusion restrictions to obtain identification and estimation of average marginal effects with classical or nonclassical measurement error in a nonparametric or semiparametric context.

See, e.g., Rosenbaum (1984); Begg and Leung (2000); Frangakis and Rubin (2002); VanderWeele (2015). Athey et al. (2016) overcome this limitation by observing multiple surrogates, that together are used to model the impact on the outcome of an underlying latent surrogate with the desired properties. In a roughly analogous way, we exploit multiple (two) mismeasures of treatment to model the impact on the outcome of an underlying latent (true) treatment that has the desired properties. Beyond this analogy, however, the details of their estimator and underlying assumptions are completely different from ours.

Our structural model of treatment is based on the collective household framework pioneered by Becker (1965, 1981), Apps and Rees (1988) and Chiappori (1988, 1992), in which each household is characterized as a collection of individuals, each of whom has a well defined utility function, and who interact to generate Pareto efficient allocations. Recent advances in this literature permit the recovery of *resource shares* (or *sharing rule*), defined as each member's share of total household consumption (Lewbel and Pendakur (2008), Browning et al. (2013), Dunbar et al. (2013)). Our particular model is based on Dunbar et al. (2013) and Calvi (2016), which obtain resource shares from estimates of Engel curves (demand equations holding prices constant) of clothing items that are consumed exclusively by women, men or children.

From a policy perspective, our results contribute to the general literature estimating the effect of women bargaining power on adults' or children's outcomes. Related to our application, legal reforms aimed at improving women's property rights, inheritance rights in particular, have been used in the literature for this purpose.[2] In the Indian context, Deininger et al. (2013), for example, find evidence of an increase of women's likelihood of inheriting land following the introduction of Hindu Succession Act (HSA) amendments that equalized women's inheritance rights to men's in several Indian states between 1976 and 2005. Moreover, Roy (2008) documents that women's exposure to the HSA reforms improves their bargaining power and autonomy within their marital families, while Roy (2013), Deininger et al. (2013), and Bose and Das (2015) indicate that it increases female education. Jain (2014) shows that HSA reforms mitigate son preference, and might be effective in reducing mortality differences between boys and girls in rural India. Finally, Heath and Tan (2014) argue that the HSA amendments increase women's labor supply, especially into high-paying jobs, while Calvi (2016) shows that they improve adult women's health outcomes and their access to household resources.[3]

Finally, this paper contributes to the long-standing debate on the relative benefits and limitations of causal vs. structural modeling. Proponents of causal methods based on randomization question

[2]Parallel to these studies, an extensive literature studying the effects of Conditional Cash Transfer programs in developing countries has shown that providing women with a large amount of cash in their hands can significantly increase, e.g., the budget shares of expenditures on clothing for children and lower shares of alcohol (Attanasio and Lechene, 2002), increase children's health and education, and livestock (Rubalcava et al., 2009) and improve child development (Macours et al., 2012; Tommasi, 2015). Differently from all this literature, we are able to provide a more accurate assessment of the effect of women's empowerment on family member's outcomes. Indeed, following our approach, we are able to relate a quantitative improvement in women's and other family members' health status, to a quantitative, tangible, improvement in women's bargaining position within the household.

[3]Legal reforms in other countries have been studied as well. La Ferrara and Milazzo (2014), for example, exploit an amendment to Ghana's Intestate Succession Law and compare differential responses of matrilineal and patrilineal ethnic groups, finding that parents substitute land inheritance with children's education. Harari (2014) analyzes a law reform meant to equalize inheritance rights for Kenyan women and shows that women exposed to the reform are more educated, less likely to undergo genital mutilation, and have higher age at marriage and at first child.

the validity of results obtained from complex and questionable structural modeling assumptions (Angrist and Pischke, 2010; Imbens, 2010). Advocates of the structural approach, instead, stress the much richer insights that can be obtained when one allows economic theory to guide the empirical work (Wolpin, 2013). Recent contributions in the econometrics literature have started to formally unify the two camps in order to overcome these divisions (Vytlacil (2002), Heckman et al. (2006); Heckman and Vytlacil (2007); Pearl (2009). Heckman (2010), for example, proposes to combine the best features of both the structural and the causal modeling approaches in what he calls a *third way* of policy analysis, while Lewbel (2016) argues that the best strategy for identification is often to combine the strengths of both approaches. We make a further step in this direction and show that relying on a structure provided by economic theory can significantly expand the set of causal questions that researchers can answer, when combined with statistical techniques that account for the estimation and specification errors of structural models.

3 LATE With Mismeasured or Misspecified Treatment

This section is organized in three parts. First, we describe our theoretical framework and derive our new MR-LATE estimator. Second, we provide a practical example of treatment indicators of the type that allows identification of LATE in this context. Finally, we implement a Monte Carlo experiment to verify that the estimator complies with the predictions of the theory.

3.1 Set Up and Identification

We begin by defining notation. We ignore covariates for now, as everything immediately extends to conditioning on covariates X. Let D be the *unobserved* true binary treatment variable that affects the outcome of interest. Not only is D unobserved, but we also do *not* assume that D can be consistently estimated.

Let Z be a randomized binary instrument that affects the treatment D. Define random binary variables D_0 and D_1 as potential treatments $D_z = D(z)$ for possible realizations z of Z. Then by construction:

$$D = (1-Z)D_0 + ZD_1 = D_0 + (D_1 - D_0)Z \tag{1}$$

Let Y be an observed outcome of interest. Define random variables Y_0 and Y_1 as potential outcomes $Y_d = Y(d)$ for possible realizations d of D. Then:

$$Y = (1-D)Y_0 + DY_1 = Y_0 + (Y_1 - Y_0)D = Y_0 + (Y_1 - Y_0)[(1-Z)D_0 + ZD_1] \tag{2}$$

Assumption 1. *Y and D satisfy the standard Imbens and Angrist (1994) LATE assumptions:*

i. $0 < E(D) < 1$, $0 < E(Z) < 1$ *and* $Z \perp (Y_1, Y_0, D_1, D_0)$.

ii. (Y_1, Y_0, D_1, D_0, Z) *are independent across individuals and have finite means.*

iii. No defiers, that is, $\Pr(D_0 = 1 \text{ and } D_1 = 0) = 0$.

Let C denote a complier, that is, someone who has $D_0 = 0$ and $D_1 = 1$. The Imbens and Angrist (1994) LATE is:

$$\frac{cov(Y,Z)}{cov(D,Z)} = E(Y_1 - Y_0 \mid C) = LATE \tag{3}$$

While we observe a sample of observations of Y and Z, we cannot implement the above LATE estimator because we do not observe D. Assume instead that we observe a binary treatment indicator T (later we will make use of two such indicators, but just consider one for now). T could be a proxy or an estimate of D, or T could be an observation of D that is sometimes mismeasured (i.e., T correctly measures D when $T = D$, and incorrectly measures when $T = 1 - D$). In our empirical application, T will be an estimate of D based on a structural model, so for some individuals T will not equal D either because of estimation error or because the structural model may be misspecified. Our goal is to point identify and estimate LATE, even though D is not directly observed and cannot be consistently estimated.[4]

Define random variables T_0 and T_1 as potential observed treatments, or potential estimated treatments, so $T_d = T(d)$ for possible realizations d of D. Then:

$$T = (1 - D) T_0 + D T_1 \tag{4}$$

The random binary variables T_0 and T_1 can be interpreted as indicators of whether treatment is correctly measured or not. In particular, if $T_0 = 0$ and $T_1 = 1$, then treatment is not mismeasured. This shows the two possible types of measurement or classification error. If $T_0 = 1$, then a true $D = 0$ is misclassified as treated, while if $T_1 = 0$, then a true $D = 1$ is misclassified as untreated.

Assumption 2. *Assumptions involving T:*

i. $Z \perp (Y_1, Y_0, D_1, D_0, T_1, T_0)$

ii. $(T_1, T_0) \perp (Y_1, Y_0) \mid C$

iii. $E(T_1 - T_0 \mid C) \neq 0$

Assumption 2-i just combines the LATE unconfoundedness assumption that $Z \perp (Y_1, Y_0, D_1, D_0)$ with the assumption that the instrument is also independent of the potential measurement errors, and hence of (T_1, T_0). The standard assumption that Z is randomized by experimental or quasi-experimental design is sufficient but stronger than necessary to have 2-i hold. Assumption 2-ii says that, for compliers, the potential mismeasures (T_1, T_0) are independent of the potential outcomes (Y_1, Y_0). Combined with unconfoundedness, this corresponds to the usual classical assumption that measurement errors are unrelated to outcomes.[5] Assumption 2-iii is a minimal relevance condition saying that T provides some information regarding D. This assumption implies that, at least for compliers, the correlation between D and T is nonzero.

[4] Although we observe T and not D, people's behavior is still based on their actual D. So, e.g., introducing measurement error does not change the no defiers assumption. If we incorrectly assumed behavior was based on T, estimating LATE using T in place of D, then what would appear to be defiers could exist. But that would not be the primary source of bias in LATE estimates that ignore the measurement error. Since we never have data from the same person in both a treated and untreated state, we wouldn't see the no defiers assumption being violated in any case.

[5] A sufficient condition for Assumption 2-ii to hold is that $(T_1, T_0) \perp (Y_1, Y_0, D_1, D_0)$, meaning that the measurement errors are independent of the potential outcomes and potential treatments. If this stronger condition holds then $q = E(T_0)/E(T_1 - T_0)$, but we do not require this.

Let $p_d = E(T_d \mid C)$ and define q by:

$$q = \frac{p_1}{p_1 - p_0} \tag{5}$$

By definition, p_1 is the probability that a complier would have their treatment *correctly* observed if they were assigned the true treatment $D = 1$ (i.e., the probability that a complier would have $T = 1$ if they were assigned $D = 1$). Similarly, p_0 is the probability that a complier would have their treatment *incorrectly* observed (meaning $T = 1$) if they were assigned the true treatment $D = 0$. Assumption 2-iii ensures that the denominator of q is nonzero, so q is finite. Define λ by:

$$\lambda = \frac{cov(YT, Z)}{cov(T, Z)} = \frac{E(YT \mid Z = 1) - E(YT \mid Z = 0)}{E(T \mid Z = 1) - E(T \mid Z = 0)} \tag{6}$$

Theorem 1. *Let Assumptions 1 and 2 hold. Then:*

$$\lambda = E[qY_1 + (1-q)Y_0 \mid C] \tag{7}$$

Proof. See Appendix A.1. □

Elements of theorem 1, and results related to theorem 1, appear in some earlier work, including Ura (2015) and references therein. Our primary novelty will be in how we make use of these relationships. There are two ways in which a mismeasured T could differ from the true D: either T might be zero when D is one (i.e., a treated person can be mismeasured as untreated) or T might be one when D is zero (i.e., an untreated person can be mismeasured as treated). The third and fourth possibilities, corresponding to cases that are not mismeasured, are when D and T both equal zero or both equal one. To make use of theorem 1, assume we observe two different mismeasures of treatment, called T^a and T^b. These could be, for instance, two proxies or two different estimates of D.

Recalling that $p_d = E(T_d \mid C)$, let $p_d^a = E\left(T_d^a \mid C\right)$ and $p_d^b = E\left(T_d^b \mid C\right)$, where T_d^a and T_d^b are the potential outcomes associated with treatments T^a and T^b. Similarly, define q^a, q^b, λ^a, λ^b, and our MR-LATE estimator as follows:

$$q^a = \frac{p_1^a}{p_1^a - p_0^a}, \quad q^b = \frac{p_1^b}{p_1^b - p_0^b},$$
$$\lambda^a = \frac{cov(T^a Y, Z)}{cov(T^a, Z)}, \quad \lambda^b = \frac{cov\left(T^b Y, Z\right)}{cov(T^b, Z)}, \text{ and MR-LATE} = \lambda^a - \lambda^b$$

This corollary follows immediately.

Corollary 1. *Let Assumption 1 hold, and let Assumption 2 hold with $T = T^a$ and with $T = T^b$. Then*

$$\textit{MR-LATE} = \left(q^a - q^b\right) E[Y_1 - Y_0 \mid C] = \left(q^a - q^b\right) LATE \tag{8}$$

Corollary 1 has some immediate implications, which are:

Corollary 2. *Let Assumption 1 hold, and let Assumption 2 hold with* $T = T^a$ *and with* $T = T^b$. *Then MR-LATE equals LATE when* $q^a - q^b = 1$. *A sufficient condition for MR-LATE to equal LATE is* $p_0^a = p_1^b = 0$.

The fact that MR-LATE equals LATE when $p_0^a = p_1^b = 0$ follows from equations (8) and (5). More generally, MR-LATE provides a good approximation to LATE when p_0^a and p_1^b are close to zero. Having $p_0^a = 0$ means that, among compliers, the probability that $T_0^a = 1$ is zero, so all compliers who have $D = 0$ must also have $T^a = 0$. So $p_0^a = 0$ if, among compliers, the treatment measure T^a only has one kind of measurement error, never mistaking the actually treated as untreated, even if it is frequently wrong about mistaking the actually untreated as treated. More generally, p_0^a will be close to zero if T^a rarely mistakes the treated as untreated.

Having $p_1^b = 0$ is a little more peculiar, since it says that all compliers who have $D = 1$ must also have $T^b = 0$. This means that, among compliers, it is not T^b but $1 - T^b$ that only has one kind of measurement error, never mistaking the actually untreated as treated, even if it is frequently wrong about mistaking the actually treated as untreated.

A simple summary is this: having MR-LATE equal LATE requires that $D - T^a$ never equals one and $D - T^b$ never equals zero. And having MR-LATE be a good approximation to LATE requires that $D - T^a$ and $D - T^b$ rarely equal one and zero, respectively. It is also possible that Corollary 1 could be used for set identification, in particular, if follows immediately that MR-LATE signs LATE when $0 < q^a - q^b$, and MR-LATE bounds LATE when $0 \leq q^a - q^b \leq 1$.

The estimation of $\lambda^a - \lambda^b$ and hence of MR-LATE is extremely simple. Assume we have iid observations of the vector $\left(Y_i, Z_i, T_i^a, T_i^b\right)$ for individuals $i = 1, ..., n$. Then consider a linear instrumental variables regression of $Y_i T_i^a$ on a constant and on T_i^a, using as instruments a constant and Z_i. Let the estimated coefficient of T_i^a in this regression be $\widehat{\lambda}^a$. Similarly, let $\widehat{\lambda}^b$ be the estimated of coefficient of T_i^b in a linear instrumental variables regression of $Y_i T_i^b$ on a constant and on T_i^b, again using as instruments a constant and Z_i. MR-LATE is then just the difference between the estimated coefficients of treatment in these two 2SLS regressions.

3.2 Illustrative Example

Define the indicator function $I(\cdot)$ to equal one if its argument is true, and zero otherwise. Suppose that D is determined by a threshold crossing model, so $D = I(R^* \geq e)$, meaning that an individual is treated if and only if some variable R^* exceeds a random threshold e. In our empirical application, R^* will be the wife's share of resources (in percentage terms) spent on adults in a household, and treatment D is defined as the wife having relatively high control of household resources.

Suppose we do not observe R^* correctly (either because of measurement error, or because it's estimated using some model, which might itself be misspecified). Instead, what we observe is a variable R, with $R = R^* + \varepsilon$ where ε is some unknown random measurement error, or estimation or

specification error. Assume e and ε are independent of Z and R^*.[6]

Let κ^a and κ^b be two constants chosen by the researcher, with $\kappa^a > \kappa^b$, and define treatment measures T^a and T^b by

$$T^a = I\left(R \geq \kappa^a\right) \quad \text{and} \quad T^b = I\left(R < \kappa^b\right).$$

Note that $R \geq \kappa^a$ implies $R^* \geq \kappa^a - \varepsilon$ so, by construction,

$$p_0^a = \Pr\left(R^* \geq \kappa^a - \varepsilon \mid R^* < e\right)$$

Therefore, if κ^a is larger than the maximum value that $\varepsilon + e$ can take on, then $p_0^a = 0$. More generally, p_0^a is near zero if the chance that $\varepsilon + e$ is greater than κ^a is small. The intuition here is straightforward: An individual is untreated, having $D = 0$, when the true R^* is small. So if we define T^a to equal one only when the observed or estimated R is very large, then the probability of having $T^a = 1$ when $D = 0$ is very small, or zero, meaning that p_0^a is near or equal to zero, as required.

We could guarantee that p_0^a is zero by taking κ^a to be infinite (or just greater than the largest value that R can take on). But then $p_1^a = \Pr\left(R^* \geq \kappa^a - \varepsilon \mid R^* \geq e\right)$ would also equal zero, and having $p_0^a = p_1^a$ violates Assumption 2-iii. More generally, we have a tradeoff, in that the larger κ^a is, the more likely p_1^a is to be close to or equal to zero, but also, the larger κ^a is, the less informative T^a is as a measure of treatment (e.g., the lower is the correlation between T^a and the true D). Equivalently, we want q^a to be as close to one as possible, which ideally means choosing a moderate value of κ^a, to both make p_0^a small and p_1^a large. This tradeoff will be visible in our Monte Carlo analysis.

A comparable construction applies to T^b, where

$$p_1^b = \Pr\left(R^* < \kappa^b - \varepsilon \mid R^* \geq e\right)$$

so $p_1^b = 0$ if κ^b is less than the minimum value that $\varepsilon + e$ can take on, and p_1^b is near zero if the chance of $\varepsilon + e$ being less than κ^b is small.

Point identification, where MR-LATE equals the true LATE in this example, requires that $\varepsilon + e$ be bounded from both above and below, and that κ^b and κ^a are chosen to lie outside these bounds (but still within the range of R^* and R). In some contexts, we may have sufficient information to know these bounds, e.g., the threshold e might be an observable policy variable, and the measurement error ε might be rounding errors of known maximum possible magnitude.[7] In our empirical application, where R^* is the wife's resource share, we cannot be certain that our choices of κ^b and κ^a will point identify LATE, but we can choose large enough values to be confident that p_0^a and p_1^b are at least close to, if not equal to, zero.

The estimation of MR-LATE in this application is as described at the end of the previous section,

[6] The assumption that e and ε are independent of Z and R^* is not necessary. We only assume this here to simplify the exposition. Specifically, we provide probabilities p_0^a and p_1^b that apply to the whole population, not specifically conditioning on compliers. If e or ε correlate with Z or R^*, then the correct expressions for p_0^a and p_1^b will need to condition on compliers, and will therefore be more complicated, though the corresponding intuition regarding identification would be the same.

[7] For example, if the observed R equals R^* rounded to the nearest integer, then $|\varepsilon| \leq 1/2$. More generally, if R is reported interval data, then $|\varepsilon|$ is bounded by half the length of the largest interval.

taking $T^a = I(R \geq \kappa^a)$ and $T^b = I\left(R < \kappa^b\right)$. It is interesting to contrast this with ordinary LATE estimation. In this context, one would typically construct $T = I(R \geq c)$, where c is one's best guess of the midpoint of $\varepsilon + e$, thereby constructing T to be as close as possible to the true unknown D. But estimation of LATE using this T in place of the true unknown D would be typically be substantially biased. The problem with replacing the unknown D with the known T in an ordinary LATE estimation is that compliers who have R close to c are precisely those who are most likely to be misclassified. An alternative would be to apply Ura (2015) to obtain bounds on LATE using this T. However, our MR-LATE estimator is likely to be more informative than Ura's bounds, because we exploit the additional information contained in having, and making use of, the two mismeasures T^a and T^b, instead of just one measure T.

A final question is, when we don't know bounds for $\varepsilon + e$, how should the practitioner choose κ^a and κ^b? One possibility is to make use of distributional assumptions and other available information. For example, the threshold e could be a policy variable, or lie in some range set by theory. If R is estimated, then one could choose κ^a and κ^b to contain a 95% or 99% confidence region of ε, thereby accounting for estimation error (though not specification error). Alternatively, if the difference between R is R^* is measurement error, then usable distribution data might be available from a validation sample.

3.3 Monte Carlo Simulations

In this section, we provide Monte Carlo experiments to check finite sample properties of our estimator. In our data generating process (DGP), true and measured treatment is determined as in the above illustrative example, so

$$D = I(R^* \geq e) \quad \text{and} \quad R = R^* + \varepsilon$$

where the econometrician observes D and R, but not R^*, ε, or e. We let R^* depend linearly on a covariate X, an instrument Z, an unobserved random component S, and a random error U, so

$$R^* = c + \alpha X + \beta Z + \gamma S + U.$$

We let $e = 0$, so $D = I(R^* \geq 0)$. This is without loss of generality, because random variation in e is observationally equivalent to variation in U and ε.

We construct unobserved potential outcomes Y_0 and Y_1, and the corresponding observed outcome Y, as

$$\begin{aligned} Y_0 &= 0.5 + X + S + V_0 \\ Y_1 &= 1.5 + X + S + V_1 \\ Y &= (1 - D)\, Y_0 + D Y_1 \end{aligned}$$

where V_0 and V_1 are additional random unobserved errors. The presence of the unobserved random component S in determining both treatment D and potential outcomes Y_0 and Y_1 makes treatment endogenous. By construction, the average treatment effect equals one.

The exogenously determined variables, errors, and parameter values in our simulations are set as follows.

$$\begin{cases} X \sim N(0,1), \quad S \sim N(0,0.1), \quad Z = \mathbb{1}(\sim U(0,1) > 0.9) \\ (\varepsilon, V_0, V_1) \sim N(\mathbf{0}, \Omega) \\ \Omega = \begin{pmatrix} \sigma & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \\ \alpha = 0.1, \quad \beta = 0.1, \quad \gamma = 0.1, \quad c = 0.4, \quad \sigma = 0.1 \end{cases}$$

These parameter values are chosen to simulate a situation that resembles our empirical application. We also consider drawing the measurement error ε from bounded (truncated normal) distributions, where the support of ε is constrained to lie in the range from $-k$ to k for varying choices of k. When $k = 0$, then ε is identically zero, meaning that R equals R^*. So in the case when $k = 0$, there is no measurement error, and the observed treatment indicator T in that case equals the true treatment indicator D.

Define $T = I(R \geq 0)$. MR-LATE does not depend on this T, but if one were to ignore the measurement error in R, one would apply the usual LATE estimator using T in place of the unobserved D. To calculate MR-LATE we let $T^a = I(R \geq \kappa^a)$ and $T^b = I\left(R < \kappa^b\right)$ where $\kappa^a = \kappa$ and $\kappa^b = -\kappa$, for varying choices of the constant κ.

We compare three different estimators. One is an ordinary least squares (OLS) regression of Y on a constant, T, and X, with the estimated coefficient of T being the OLS estimate of the treatment effect. Second is this same regression, but estimated using linear two stage least squares (2SLS), where Z is the instrument for T. The estimated 2SLS coefficient of T in this regression corresponds to the usual LATE estimator, which ignores the measurement error in T. Third is our MR-LATE estimator, which we obtain as follows: First, we regress YT^a on a constant, T^a, and X, using 2SLS where Z is the instrument for T^a. Then do the same using T^b in place of T^a. Letting $\widehat{\lambda}^a$ and $\widehat{\lambda}^b$ be the estimated 2SLS coefficients of T^a and T^b, respectively, the MR-LATE estimate is then $\widehat{\lambda}^a - \widehat{\lambda}^b$.

Table 1 sets out our first simulation results. For these simulations, we bound ε, and assuming the bound k is known to the econometrician, we set $\kappa = k$. So in this set of simulations for any k MR-LATE point identifies LATE. The Table reports results for three different values of k: 0, 5, and 10. In the case where $k = 0$, the observed treatment T is not mismeasured, so in that case 2SLS consistently estimates LATE, and in that case MR-LATE is numerically identical to LATE[8]. Having k equal 5 and 10 corresponds roughly to 10% and 20% of observations being misclassified, respectively.

The DGP process is simulated 10,000 times, each using a sample size of 10,000. In Table 1, when the measurement error is zero ($k = 0$), both LATE and MR-LATE are identical and almost

[8]More generally, one can readily check that if $T^a = D$ and $T^b = 1 - D$, then MR-LATE becomes identical to the standard LATE estimator.

Table 1: OLS vs 2SLS vs MR-LATE when ϵ is bounded at κ

	True value: $E[Y_1\|C]-E[Y_0\|C]$ = 1.500 - 0.500 = 1.000								
	k = 0			k = 0.05			k = 0.10		
	OLS	2SLS	MR-LATE	OLS	2SLS	MR-LATE	OLS	2SLS	MR-LATE
$E[Y_1\|C]$	1.507	1.500	1.490	1.420	1.515	1.500	1.342	1.560	1.507
$E[Y_0\|C]$	0.492	0.500	0.491	0.593	0.485	0.483	0.686	0.440	0.480
T	1.017	0.999	0.999	0.826	1.030	1.017	0.656	1.120	1.027
sd	0.084	0.098	0.098	0.018	0.104	0.084	0.020	0.117	0.092
Bias	0.017	-0.001	-0.001	-0.174	0.030	0.017	-0.344	0.120	0.027
MSE	0.007	0.010	0.010	0.030	0.012	0.007	0.118	0.028	0.009

Notes: Results are based on 10,000 simulations of 10,000 observations each. We simulate various measurement errors and the availability of a strong and exogenous instrument. When $k = 0$ means that there is no measurement error. Whereas, measurement error with $k = 0.05$ ($k = 0.10$) means that we estimate R that are on average +/- 5% (+/- 10%) of the true value. This means that roughly 10% (20%) of sample is wrongly observed to either treatment or control.

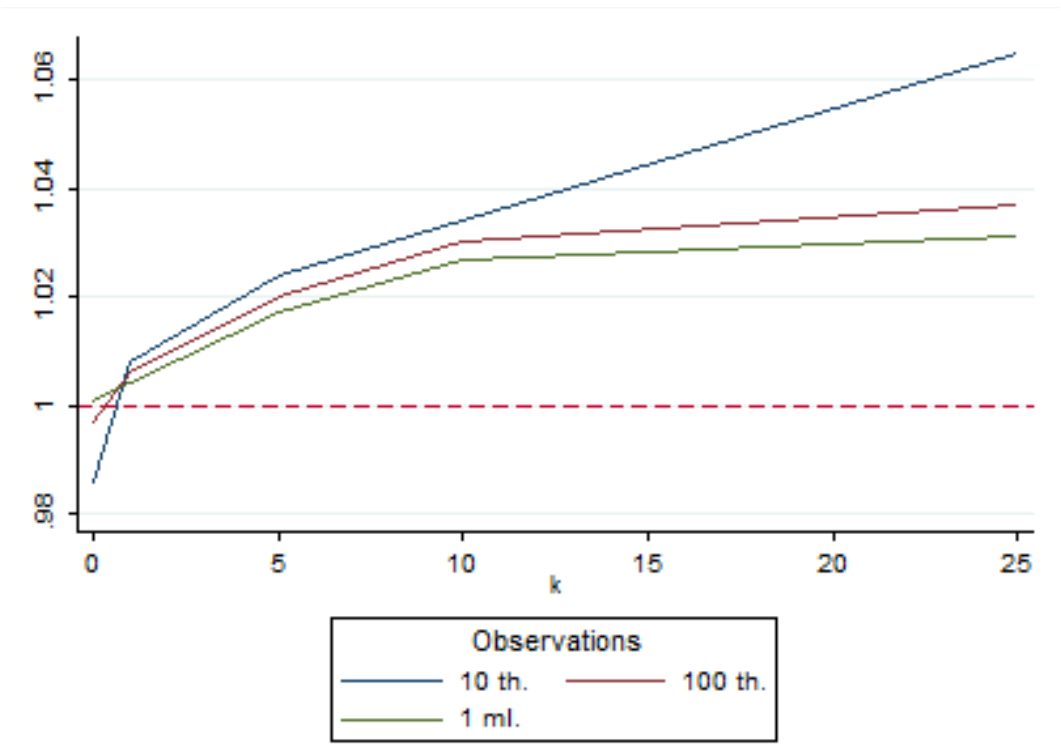


Figure 1: MR-LATE performances as the sample grows

Table 2: Optimal k when ϵ is not bounded

	True value: 1.000							
			MR-LATE					
	OLS	2SLS	k = 0	k = 1	k = 5	k = 10	k = 13	k = 25
T	0.596	1.272	1.272	1.237	1.143	1.095	1.089	1.127
sd	0.021	0.152	0.152	0.146	0.130	0.122	0.121	0.161
Bias	-0.404	0.272	0.272	0.237	0.143	0.090	0.089	0.127
MSE	0.164	0.097	0.097	0.078	0.037	0.024	0.023	0.042

Notes: Results are based on 10,000 simulations of 10,000 observations each.

unbiased, while OLS shows a small upward bias due to the endogeneity of the treatment. With larger k, OLS is substantially downward biased, as the standard attenuation bias from regressor mismeasurement dominates the bias due to endogeneity. In contrast, the table shows that 2SLS is upward biased when $k > 0$. This is also typical in cases of regressor mismeasurement, because the 2SLS denominator term $cov(T, Z)$ is closer to zero than the corresponding denominator in the case without measurement error, $cov(D, Z)$. Our MR-LATE estimator performs much better than the other estimators, with far lower bias and far lower mean squared error.

In Figure 1 we continue to set $\kappa = k$, and examine how the MR-LATE estimator performs as k (and hence also our bounds $\kappa^a = \kappa$ and $\kappa^b = -\kappa$) increases, and as the sample size increases. As discussed in the previous section, the larger these bounds are, the less informative are the treatment measures T^a and T^b. Equivalently, the larger the bounds need to be to guarantee that p_0^a and p_1^b are near zero, the noisier are the estimates of $\widehat{\lambda}^a$ and $\widehat{\lambda}^b$, and so the greater is the amount of data one needs to obtain accurate estimates. This can be seen in Table **??**, which examines k equal to 1, 10, and 25. We find that MR-LATE has an upward bias (far smaller than the bias in OLS and standard 2SLS LATE estimation, as the previous table shows), that increases with k and decreases with sample size.

Finally, Table 2 sets out the simulation results in the case where ε has unbounded support. In this case, MR-LATE cannot consistently estimate the LATE (and the standard LATE estimator is also inconsistent). The variance of ε is also set relatively large here, resulting in T that misclassifies more than 20% of the observations of the true D. The DGP process is again simulated 10,000 times using a sample size of 10,000. As before, we compare our estimator to OLS and 2SLS. For the MR-LATE estimator, we report six different values of the threshold parameter κ: 0, 1, 5, 10, 13, 25. We again have the tradeoff that the larger κ is, the closer p_0^a and p_1^b are to zero and hence the closer the limiting value of MR-LATE is to LATE, but also, the larger κ is, the further $\widehat{\lambda}^a$ and $\widehat{\lambda}^b$ are to their limiting values at any given sample size. In Table 2 we find the best value of κ (in terms of bias and mean square error) is around 13. In real data, one would not be able to search for optimal bounds κ^a and κ^b in this way. However, these numerical results are reassuring in showing that as long as κ is not chosen to be far too small, one obtains a good approximation to the true LATE (much better than the standard 2SLS LATE estimator) over a wide range of possible choices of bounds.

4 Empirical Application: Women's Intra-household Power and Family Health

We apply our estimator to study the impact of women's intra-household empowerment on a variety of measures of family members health status in India. The outcomes Y that we consider are, for each adult in the household, body mass index (BMI), an indicator for being underweight, and an indicator for being anemic. We also consider outcomes for children in the household: height for age, weight for age, and an indicator for having been vaccinated against one or more diseases.

We define a household to be treated, having $D = 1$, if the mother has control of a substantial fraction of the household's resources. Formally, we define $D = I\left(R^* \geq e\right)$ where R^* is the percentage of resources spent on adults in the household that the mother controls, and e is a threshold that could vary across households for unobserved reasons. There exists a number of measures of control over resources, including self reports of power, and observations of income or purchases by women, but these measures are quite crude and imprecise. We will therefore employ a structural model that makes use of both economic theory and detailed household expenditure data to construct R, an estimate R^*. Then, recognizing that any such structural model can suffer from both specification and estimation errors, we apply our MR-LATE estimator.

Unlike the structural model for resource control, which is grounded in economic theory, we do not take a stand on how mother's power D determines health outcomes Y. One could imagine many possible mechanisms linking the two. For example, mothers may have different priorites than fathers regarding expenditures on health related goods, or circumstances that contribute to women having greater power might also affect the health outcomes of family members. Previous research discussed earlier provides indirect evidence of linkages between D and Y, by establishing, e.g., how variables that we believe affect women's power correlate with health outcomes of wives and children. Our goal is therefore to directly quantify treatment effects of D on Y, using a causal rather than structural model, exploiting plausibly exogenous variation in an instrument that correlates with D.

4.1 Estimating the Unobserved Treatment and MR-LATE

We apply the model developed by Dunbar et al. (2013) (hereafter DLP) to obtain and estimate R, the relative resource shares of mothers, which we then use to construct the estimated treatment indicators T^a and T^b required for the MR-LATE estimator.

We consider three types of individuals in households: $t \in \{f, m, c\}$ indicating father, mother, and children. Our data will only include households that consist of one mother, one father, and one to four children aged 0 to 14. The model assumes mothers, fathers, and children each have their own utility function over goods, and the household uses some unmodeled bargaining process or social welfare function to allocate resources to each member in a Pareto efficient way. Let M denote the total expenditures or resources of the household, i.e., the household's total budget.

Let η_t denote the resource share of member t, that is, η_t is the fraction of total household resources M that are spent on goods consumed by member t. Then $R^* = 100 * \eta_f / \left(\eta_f + \eta_m\right)$, which is an objective measure of the mother's power, or control over resources, relative to the father. We will therefore let R be an estimate of $100 * \eta_f / \left(\eta_f + \eta_m\right)$.

A key difficulty in observing or calculating each η_t, and hence R, is that most goods in a household can be shared or consumed jointly to some extent by household members. For example, home heating is almost completely shared, while cooking fuel is jointly consumed just among household members who are eating together. Other goods, like food, are consumed individually, but it is dif-

ficult to track exactly who eats what within the household. To overcome the difficulties raised by these measurement issues, DLP first considers private, assignable goods, one each for the father, mother, and children. What makes a good assignable is that it appears in just one (known) person's utility function, and what makes each a good private is that it is not shared or consumed jointly.

Let scalars w^f, w^m, and w^c denote the household's budget shares (fraction of total expenditures M) of women's clothes, men's clothes, and children's clothes, respectively. These are the private, assignable goods. We can NOT just use w^f and w^m as measures of η_f and η_m, because men and women may have very different tastes for clothes. For example, the wife might consume far fewer household resources than her husband, but still consume more clothes, because she derives far more utility from clothing consumption than her husband does.

What DLP do instead is to identify and estimate the separate clothing Engel curve for each household member. Each Engel curve gives that person's clothing demand w^t as a function of the total resources they control, $\eta_t M$. They then invert each Engel curve to reveal what each person's total resource share η_t must have been.

Let $X = (X_1, ..., X_K)$ denote a vector of observable attributes of households and their members. Household attributes X may affect the preferences of each household member and may also affect the household's bargaining process or social welfare function, and as a result may directly affect resource shares. Let $\widetilde{Z}$ denote a vector of distribution factors. These are defined as variables which may affect resource shares but which do not affect individual preferences. Resource shares are generally a function of both X and $\widetilde{Z}$, so let $\eta_t(X, \widetilde{Z})$ denote the resource share function of member t. Adapting DLP, in an Appendix we derive household demand functions of private assignable goods of the form

$$w^t = \eta_t(X, \widetilde{Z})[\delta^t(X) + (lnM + ln\eta_t(X, \widetilde{Z}))\beta(X)] \quad \text{for } t \in \{f, m, c\} \tag{9}$$

for some functions $\beta(X)$ and $\delta^t(X)$ (note that the demand functions for other goods that aren't private and assignable are more complicated, but are not required to estimate the resource shares).

DLP prove that the functions $\beta(X)$, $\delta^t(X)$, and $\eta_t(X, \widetilde{Z})$ are identified in this model. Some collective household identification results require the presence of distribution factors $\widetilde{Z}$. DLP does not, but the presence of distribution factors can improve the precision of estimates, since they move resource shares separately from other functions.

For our empirical application, we assume the functions $\beta(X)$, $\delta^t(X)$, and $\eta_t(X, \widetilde{Z})$ are all linear in their arguments. In particular, we specify

$$\eta_t(X, Z) = \theta_{t0} + \theta_{t1}X_1 + ... + \theta_{tK}X_K + \theta_{tZ}Z. \tag{10}$$

Equation (10) shows that, in our data we only have one distribution factor $\widetilde{Z}$, which equals our treatment effect instrument Z.[9] For each $t \in \{f, m, c\}$ we append an error term to equation (9),

[9]It should not be surprising that the resource shares η_t contains the instrument Z, since the instrument should correlate with treatment, and treatment is determined by the resource shares. However, one might be concerned that by including Z as a distribution factor in our estimate of each η_t, we might be artificially inducing spurious correlation. We therefore also provide estimates that do not include Z in the model for

yielding a three equation system that we estimate using non-linear Seemingly Unrelated Regressions (SUR).[10] Let $\widehat{\theta}$ denote the estimate of each θ parameter.

We have two different datasets for our empirical analysis. One, the NSS Consumer Expenditure Survey (NSS), contains detailed consumption data for estimating the above model. The other is the third round of the National Family Health Survey (NFHS), which reports the health outcomes we use for our causal treatment effects estimation. Both data sets contain observations of X and Z. For each family member t of the household containing each individual i that we draw from the NFHS data, we use the estimates $\widehat{\theta}$ obtained from the NSS data to construct NFHS estimates of each resource share by

$$\widehat{\eta}_{ti} = \widehat{\theta}_{t0} + \widehat{\theta}_{t1}X_{1i} + ... + \widehat{\theta}_{tK}X_{Ki} + \widehat{\theta}_{tZ}Z_i.$$

Then R for each individual i in the NFHS is given by

$$R_i = 100 * \widehat{\eta}_{fi} / \left(\widehat{\eta}_{fi} + \widehat{\eta}_{mi}\right).$$

Our goal is to estimate a LATE for a range of health outcomes Y. We separately consider health outcomes for mothers, fathers, and children. So, e.g., when i is a child and Y_i is an indicator of whether the child has been vaccinated, the treatment effect we wish to estimate is the increase in probability (among compliers) of the child being vaccinated, where treatment D_i, empowerment of the child's mother, is defined as having $R_i^* \geq e_i$, even though her true relative resource share R_i^* is unobserved.

To apply our MR-LATE estimator to the NFHS data, we let $T_i^a = I\left(R_i \geq \kappa^a\right)$ and $T_i^b = I\left(R_i < \kappa^b\right)$, where (for a few different values of κ) $\kappa^a = c + \kappa$, $\kappa^b = c - \kappa$, and c is the median value of the distribution of R. The MR-LATE estimator then consists of first linearly regressing $Y_i T_i^a$ on a constant, T_i^a, and X_i, using 2SLS where the instruments are a constant, Z_i, and X_i. Then do the same using T_i^b in place of T_i^a. Letting $\widehat{\lambda}^a$ and $\widehat{\lambda}^b$ be the estimated 2SLS coefficients of T^a and T^b, respectively, the MR-LATE estimate is then $\widehat{\lambda}^a - \widehat{\lambda}^b$. We repeat this procedure using a few different health outcome measures Y_i each for men, women, and children, and a few different values of κ.

In the special case of $\kappa = 0$, MR-LATE becomes numerically identical to the standard Imbens and Angrist (1994) 2SLS LATE estimator, using $T_i = I\left(R_i \geq c\right)$ in place of the unobserved $D_i = I\left(R_i^* \geq e_i\right)$. However, this $\kappa = 0$ case will only be consistent if there is no measurement or estimation error in R_i and if e_i exactly equals c. As summarized earlier, for values of $\kappa > 0$, we have generally that larger values of κ are more likely to yield consistent estimates of LATE, but values of κ that are too large will tend to be uninformative, yielding very imprecise estimates.

---

each η_t. DLP show that the resource shares remain identified even without distribution factors.

[10] The non-linear SUR is iterated until the estimated parameters and the covariance matrix settle. The result is asymptotically equivalent to maximum likelihood with multivariate normal errors.

4.2 Data

We implement our empirical analysis using the 2011-2012 India National Sample Survey (NSS) of Consumer Expenditure (68^{th} round) and the 2005-2006 India National Family Health Survey (NFHS). Table A2 in the Appendix presents some descriptive statistics.

NSS data. The 2011-2012 NSS Consumer Expenditure Survey contains detailed data on household expenditure and details about household socio-economic characteristics, and other particulars of household members.[11] We select households consisting of a mother, a father, and one to four children. Among other items, households are asked to report how much they spent on clothing, bedding, and footwear. Given the detailed breakdown of clothing expenditure, it is possible to identify the expenditure on items of clothing that can be assigned to women, men, and children, thereby allowing us to construct expenditures on private, assignable clothing for each.[12] Table A1 in the Appendix contains some descriptive statistics. For clothing items, the survey entry refers to expenditures that occurred in the past 365 days. For simplicity, we convert annual into monthly figures. Budget shares w^f, w^m, and w^c are computed and reported as percentages of total household expenditures M. The vector of characteristics X contains variables that characterize each individual (age and gender), the household (wealth, religion, caste, number of children), and the environment of the household (fixed effects and time trends). Specifically, we include dummies for the number of children, the fraction of female children, women's age and the gender age gap, the average age of children, a measure of household health, regional dummies and indicator variables for being Hindu, Buddhist, Sikh or Jain, for living in rural areas, for female and male higher education, and for being of schedule caste, scheduled tribe, or other backward classes. When focusing on children, we also include the child's age and gender.

NFHS data. The 2005-2006 National Family Health Survey provides a range of health indicators for women (aged 15 to 49), men (aged 15 to 54), and for children born in the 5 years prior to the date of interview. The survey also contains a range of socio-economic characteristics of individuals and households X, comparable to those we observe in the the NSS data. We consider the NHFS women, men and children datasets separately, constructing a few different outcome Y measures for each. The health measures for each include anthropometrics like BMI and measures of anemia.[13] For children, we also observe vaccination records, which we use to construct an ad-

[11] We select a sample of households with one woman and one man above age 15, and up to 4 children under 15. Moreover, we exclude households with no women or no men above 15 years of age, households in the top 1 percent of expenditure, households with a female head and households with head or head of household wife under 15. Finally, we exclude households reporting to have performed any ceremony during the month prior to the survey, as unusual purchases of clothing items and non-standard expenditure patterns may occur for festivities and ceremonies.

[12] As in Calvi (2016), we define expenditure on women assignable clothing as the sum of expenditures on saree, shawls, chaddar, and kurta-pajamas suits for females. For men assignable clothing, we combine expenditure on dhoti, lungi, kurta-pajamas suits for males, pajamas, and salwar. For children, we use expenditure on school uniforms and infant clothing. Tommasi and Wolf (2016) shows that when the data exhibit relatively flat Engel curves in the consumption of the private assignable goods, the DLP model is weakly identified and induces high variability and an implausible pattern in least squares estimates of resource shares. However, the advantage of following Calvi (2016) is that households in her dataset have a large variation in the consumption of private assignable goods, which facilitates identification. Hence the problem of weak identifiability in this context is very mild and negligible.

[13] Body Mass Index (BMI) is defined as weight in kilograms divided by height in meters squared. A cut-off point of 18.5 is used to define thinness or acute undernutrition, and a BMI of 23 or above indicates overweight or obesity for Asian Indians (Shiwaku et al., 2004). Anemia, instead, is a condition in which the number of red blood cells or their oxygen-carrying capacity is insufficient. Although its primary cause is

ditional children's health outcome Y defined as whether a child has ever received any vaccine to prevent diseases.[14]

Hindu Succession Act and its Amendments. We exploit changes in the Indian inheritance law to construct a plausibly unconfounded instrumental variable Z. A woman's right to inherit land and other property is often claimed to play a significant role in determining women's position within the household (World Bank, 2014).

Inheritance rights in India differ by religion and, for most of the population, are governed by the Hindu Succession Act (HSA). The HSA was first introduced in 1956 and applied to all states other than Jammu and Kashmir and only to Hindus, Buddhists, Sikhs, and Jains. Thus, it did not apply to individuals of other religions, such as Muslims, Christians, Parsis, Jews, and other minority communities.[15] It aimed at unifying the traditional Mitakhshara and Dayabhaga systems, which were completely biased in favor of sons (Agarwal, 1995), and established a law of succession whereby sons and daughters would enjoy (almost) equal inheritance rights. Gender inequalities, however, remained even after the introduction of the HSA. On one hand, in case of a Hindu male dying intestate, i.e., without leaving a will, all his *separate* or *self-acquired* property, devolved equally upon sons, daughters, widow, and mother. On the other hand, the deceased's daughters had no direct inheritance rights to *joint family property*, whereas sons were given direct right by birth to belong to the coparcenary.[16] In the decades following the introduction of the HSA, state governments passed amendments that equalized inheritance rights for daughters and sons (Kerala in 1976, Andhra Pradesh in 1986, Tamil Nadu in 1989, and Maharashtra and Karnataka in 1994). A national-level ratification of the amendments occurred in 2005. These amendments only applied to Hindu, Buddhist, Sikh or Jain women, who were not yet married at the time of the amendment.

Both the NSS and the NFHS datasets contain information about women's year of birth, state of residence, and religion. We use this information to construct the variable Z defined as the women's likely eligibility to the HSA amendments. Specifically, Z is defined as the interaction between an indicator variable for being Hindu, Buddhist, Sikh or Jain, and an indicator variable equal to one if a woman was 14 or younger at the time of the amendment in her state and to zero if she was 23 or older. We use 14 and 23 as they are the 10th and 90th percentiles of women's age at marriage in the NFHS sample.[17] As shown in Table A1, in both samples about 80 percent of the sample is

iron deficiency, it coexists frequently with a number of other causes, such as malaria, parasitic infection, and nutritional deficiencies; 90 percent of anaemia sufferers live in developing countries. Prevalence of anaemia in South Asian countries is among the highest in the world (Kaur (2014)).

[14]We specifically observe whether a child received a BCG vaccine (against tuberculosis), one to three DPT vaccines (against diphtheria, pertussis, and tetanus), and one to four polio vaccines (at birth and one to three years after).

[15]While most laws for Christians formally grant equal rights from 1986, gender equality is not the practice, as the Synod of Christian Churches has being arranging legal counsel to help draft wills to disinherit female heirs. The inheritance rights of Muslim women in India are governed by the Muslim Personal Law (Shariat) Application Act of 1937, under which daughters inherit only a portion of what the sons do (Agarwal, 1995).

[16]All persons who acquired interest in the joint family property by birth are said to belong to the *coparcenary*. The Hindu Women's Right to Property Act of 1937 enabled the widow to succeed along with the son and to take a share equal to that of the son. The widow was entitled only to a limited estate in the property of the deceased with a right to claim partition. A daughter, however, had virtually no inheritance rights.

[17]This variable is therefore fully determined by each woman's religion, year of birth and state. Our choice to focus on women's eligibility (rather than their actual exposure) to the inheritance rights amendments has a double motivation. First, while the NFHS data include information about each woman's year at marriage and is therefore suitable for the identification of women who are Hindu, Buddhist, Sikh or Jain women and who were not yet married at the time of the amendment, the NSS does not. Second, as exposure to the HSA amendments is

Hindu, Buddhist, Jain, or Sikh. However, the percentage of households where the woman is eligible to the HSA amendments is much higher in the NSS. This is mainly due to the timing of the surveys, 2011-2012 for the NSS and 2005-2006 for the NFHS, as the former includes a larger number of women who were unmarried at the time of the national amendment in 2005.

5 Empirical Results

In this section, we present summary statistics of the estimated resource shares and treatment, together with our estimates of LATE. A more detailed discussion of the structural estimates can be found in Appendix A.4.

Table 3 contains descriptive statistics for the predicted resource shares obtained in the two samples and for the implied recovered treatment, T. These statistics take into account the empirical distributions of the covariates $(Z, X_1, ..., X_K)$, as they average over all the values of demographic factors observed in the population. That the minima and maxima of estimated resource shares do not fall outside the zero to one range for any person in any household in the two samples is reassuring. We find that the standard deviations of resource shares are larger for men than for women, suggesting that the covariates induce more variation for fathers than for mothers. In both samples, the resource share for mothers is lower than that for fathers. On average, mothers receive about 30 percent of household resources, fathers about 45 percent, and children about 14 percent per child. As the number of children increases, the total allocation to children increases, but the amount allocated per child decreases.[18]

The mean and median of R, i.e. the percentage of parents' resources devoted to the mother, are equal to approximately 40 percent in both data sets (note that we report R as a percentage). We define the threshold c to be the median of R, but our results remain robust to using the mean instead. Mother's are allocated more resources than fathers, meaning $R > 50$, in less than 20 percent of families, almost all of which (95 percent) resides in North-East States.[19] In the NFHS data, women who have $T = 1$ (and hence $R \geq c$) have an average R of 50, while those having $T = 0$ have an average R of 29. So while we can't know these statistics for the truly treated and untreated groups, in our estimated groups the treated women have considerably higher relative resources than the untreated.

Table A3 in the Appendix contains the estimation results of the first stage of the MR-LATE 2SLS estimates. The first stage of the 2SLS is reported separately for women, men, and children. Household level and individual level covariates are included in all specifications. Even conditioning on these several sources of unobserved heterogeneity, the instrument Z is positively and significantly correlated with T^a and T^b (and with T, corresponding to the $\kappa = 0$ case). To check that this is not just due to including Z as a distribution factor in the estimation of R, we re-estimated the NSS data

determined by each woman's year of marriage, we focus on an intent-to-treat effect and thereby mitigate concerns about potential endogeneity of treatment (see for instance Heath and Tan (2014)).

[18] These numbers are in line with the results obtained in Dunbar et al. (2013) using data from the Malawi Integrated Household Survey (IHS2).

[19] This is consistent with the presence of a number of matrilineal societies and cultures in these regions (Khasi and Garo societies, for example).

Table 3: Estimated Resource Shares and Woman's Power

	Obs.	Mean	St. Dev.	Median	Min.	Max.
NSS Sample :						
$\widehat{\eta}_m$	15,166	0.2934	0.0743	0.3088	0.0582	0.5149
$\widehat{\eta}_f$	15,166	0.4425	0.1041	0.4471	0.1107	0.7457
$\widehat{\eta}_m + \widehat{\eta}_f$	15,166	0.7359	0.0795	0.7390	0.4721	1.0000
$\widehat{\eta}_c$	15,166	0.2641	0.0795	0.2610	0.0000	0.5279
$\widehat{\eta}_c/c$	15,166	0.1438	0.0667	0.1285	0.0000	0.4474
$\frac{\widehat{\eta}_m}{\widehat{\eta}_m+\widehat{\eta}_f}$ (%, R)	15,166	40.3358	11.3217	41.4200	8.7477	82.3064
NFHS Sample:						
$\widehat{\eta}_m$	21,111	0.2785	0.0748	0.2897	0.0784	0.4866
$\widehat{\eta}_f$	21,111	0.4439	0.1247	0.4621	0.0975	0.7359
$\widehat{\eta}_m + \widehat{\eta}_f$	21,111	0.7224	0.0878	0.7235	0.4647	0.9450
$\widehat{\eta}_c$	21,111	0.2776	0.0878	0.2765	0.0550	0.5353
$\widehat{\eta}_c/c$	21,111	0.1406	0.0739	0.1182	0.0389	0.4906
$\frac{\widehat{\eta}_m}{\widehat{\eta}_m+\widehat{\eta}_f}$ (%, R)	21,111	39.4617	13.0189	38.9682	10.6291	80.8655

Note: Household level data.

Table 4: Women's Health (MR-LATE Estimates)

	Body Mass Index	Pr(Underweight)	Pr(Anemic)
MR-LATE ($\kappa = 0$)	8.8732	-0.7053	-0.6545
	(3.6625)	(0.3687)	(0.3380)
MR-LATE ($\kappa = 1$)	9.6218	-0.5360	-0.8026
	(13.0836)	(0.4972)	(0.4413)
MR-LATE ($\kappa = 2$)	3.1229	-0.4236	-0.6407
	(12.1546)	(0.3951)	(0.3532)
MR-LATE ($\kappa = 3$)	-1.2389	-0.1798	-0.4294
	(13.6987)	(0.3799)	(0.3238)
MR-LATE ($\kappa = 4$)	-1.2235	-0.2971	-0.7134
	(22.9874)	(0.5245)	(0.4600)
MR-LATE ($\kappa = 5$)	0.4192	-0.3743	-0.6876
	(17.3918)	(0.4342)	(0.3726)
Observations	15,377	15,377	14,286
Mean Dependent Variable	20.96	0.2914	0.1547

Note: NFHS data. Bootstrap standard errors in parentheses. All specifications include individuals and household controls, state, cohort, state-religion fixed effects and state specific linear time trends.

Table 5: Men's Health (MR-LATE Estimates)

	Body Mass Index	Pr(Underweight)	Pr(Anemic)
MR-LATE ($\kappa = 0$)	2.5079	-0.1101	-0.0866
	(3.5123)	(0.4039)	(0.2931)
MR-LATE ($\kappa = 1$)	-1.4072	0.1819	0.0712
	(9.5555)	(0.3804)	(0.3009)
MR-LATE ($\kappa = 2$)	2.0069	0.0299	0.4204
	(19.8799)	(0.5787)	(0.4730)
MR-LATE ($\kappa = 3$)	1.3038	0.0064	0.2789
	(17.2462)	(0.4959)	(0.3957)
MR-LATE ($\kappa = 4$)	-14.9379	0.2367	0.2939
	(24.2349)	(0.6336)	(0.5134)
MR-LATE ($\kappa = 5$)	-3.5089	-0.0216	1.5520
	(86.5589)	(2.0748)	(1.9992)
Observations	7,659	7,659	6,839
Mean Dependent Variable	21.18	0.2341	0.1040

Note: NFHS data. Bootstrap standard errors in parentheses. All specifications include individuals and household controls, state, cohort, state-religion fixed effects and state specific linear time trends.

Table 6: Children's Health (MR-LATE Estimates)

	Height-for-age (z-score)	Weight-for-age (z-score)	Pr(Any Vaccination)
MR-LATE ($\kappa = 0$)	-2.8202	-0.5039	0.7435
	(2.9568)	(2.3122)	(0.3131)
MR-LATE ($\kappa = 1$)	0.2533	2.8110	0.5429
	(4.5152)	(3.6549)	(0.8840)
MR-LATE ($\kappa = 2$)	-5.1043	-1.5836	1.2409
	(7.2577)	(6.3758)	(1.8334)
MR-LATE ($\kappa = 3$)	-28.0820	-31.0721	2.7979
	(14.9145)	(12.2598)	(4.1717)
MR-LATE ($\kappa = 4$)	-49.4842	-63.0806	12.9107
	(41.3689)	(35.8424)	(11.9549)
MR-LATE ($\kappa = 5$)	4.1816	2.8315	-0.5882
	(4.4997)	(4.0924)	(1.5719)
Observations	10,452	10,452	12,109
Mean Dependent Variable	-1.668	-1.782	0.9039

Note: NFHS data. Bootstrap standard errors in parentheses. All specifications include individuals and household controls, state, cohort, state-religion fixed effects and state specific linear time trends.

model omitting Z (that is, setting the coefficient θ_{tZ} to zero) and reran the first stage of the MR-LATE 2SLS estimates, to confirm that Z remains highly correlated with T^a, T^b, and T. We thereby confirm that we do not suffer from a weak instrument problem.[20]

Our MR-LATE estimates are reported in Tables 4, 5 and 6. We provide estimates taking $\kappa = 0$, which corresponds to no measurement error, and $\kappa = 1, 2, 3, 4$, and 5, which can be consistent given a maximum of 2, 4, 6, 8, and 10 percent measurement error, respectively. As expected, increasing κ generally reduces the precision of the estimates. The MR-LATE estimates show positive and significant effects of a women's intra-household power on her own health. Being highly empowered greatly increases her expected body mass index and reduces her likelihood to be underweight or anemic.[21] We do not find a significant of a women's intra-household power on her children's height-for-age or weight-for-height, but we do find that children with highly empowered mothers are more likely to be vaccinated.[22] We also find no significant or systematic effect of a wife's intra-household power on her husband's health outcomes.

These results suggest that policies aimed at empowering women's bargaining position within households leads to improvements in overall family health. In particular, exogenously increasing her control of resources appears to substantially improve her own health, and increases the vaccination rates of her children, without damaging her husband's health.

5.1 Additional Results: Alternative Treatment Proxies

In the Appendix, we describe the model of household behavior that gives rise to R, our estimate of the relative resource share percentage R^*. In this model, R^* is a measure of the wife's bargaining power relative to her husband. Here, we empirically compare our structurally-motivated measure of bargaining power with some more typical proxies of power, namely, women's self-report of control over various household decisions.

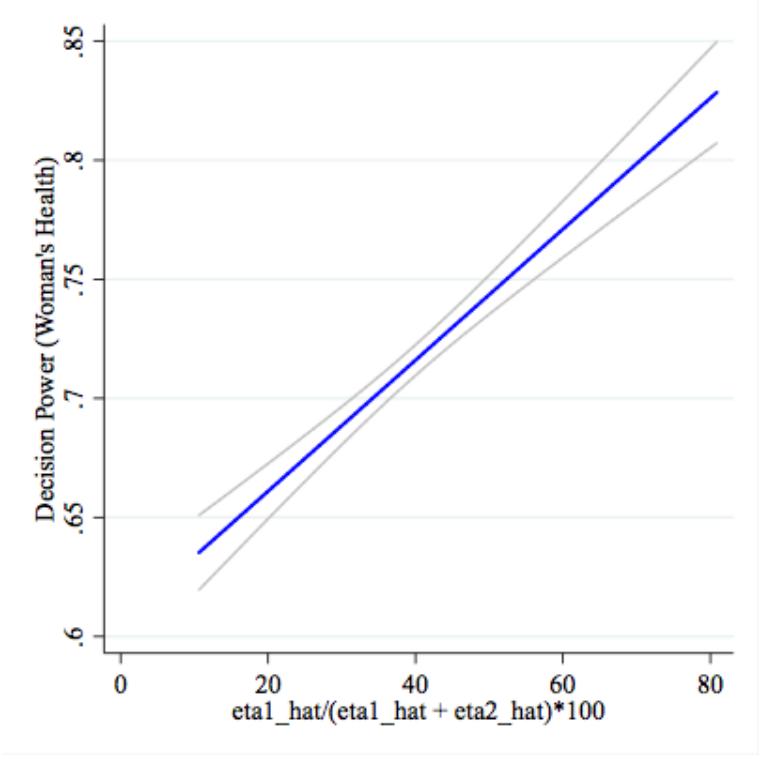
Our NFHS data contains questions of the form, "Who makes decisions about [X] in your household?" Specifically, we focus on decisions about the woman's own health care, large household purchases, visits to family or relatives, and purchases for daily needs. Figure 2 displays the clear positive relationships between women's reported participation in household decisions and our estimates of R. More details are given in Tables A7 and A8 in the Appendix. These findings empirically corroborate the theory underlying our structural model of treatment: the larger is R, the higher is her decision making and bargaining powers within the household.

An alternative to our estimation of LATE would be to use R^* or R itself as a continuous treatment measure. One might then try to estimate the effect of a marginal increase in R^* or R, instead of focusing on the binary treatment effects of D. One advantage of discretizing the way we do is to reduce the impact of measurement, estimation, and specification errors. Even without the use of

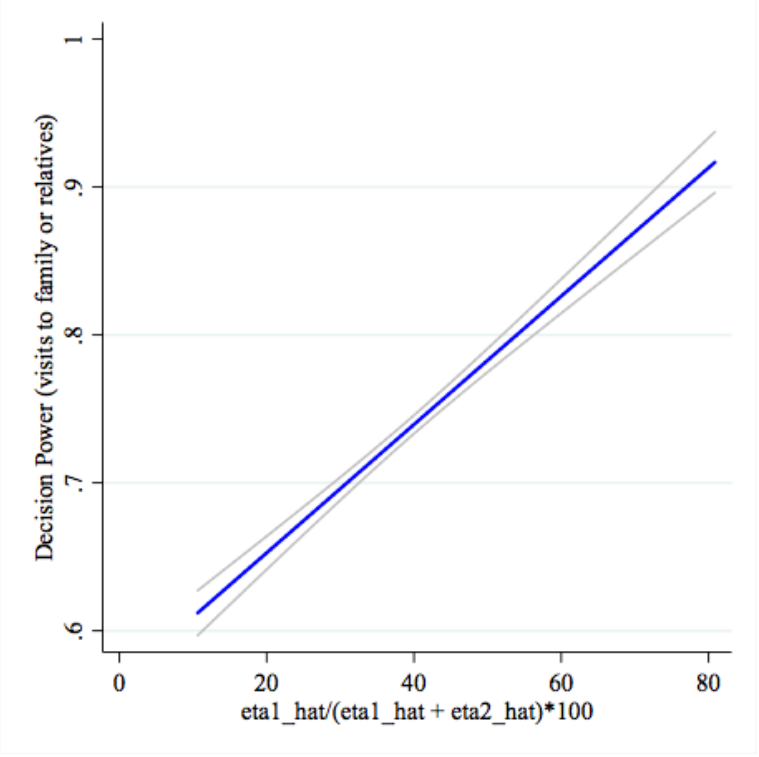
[20] This is also confirmed by the Wald tests for joint significance of the instruments, both included and excluded.

[21] As a rough guide to magnitudes, using the above estimates of the mean of R for the treated and untreated groups, and assuming the effect is linear, these estimates would imply that if a women's relative resource share increased by 1 (out of 100), then her body mass index would increase by an average of 0.44, and her probability of being underweight or anemic would decreases by 3.5 and 3.2 percent, respectively.

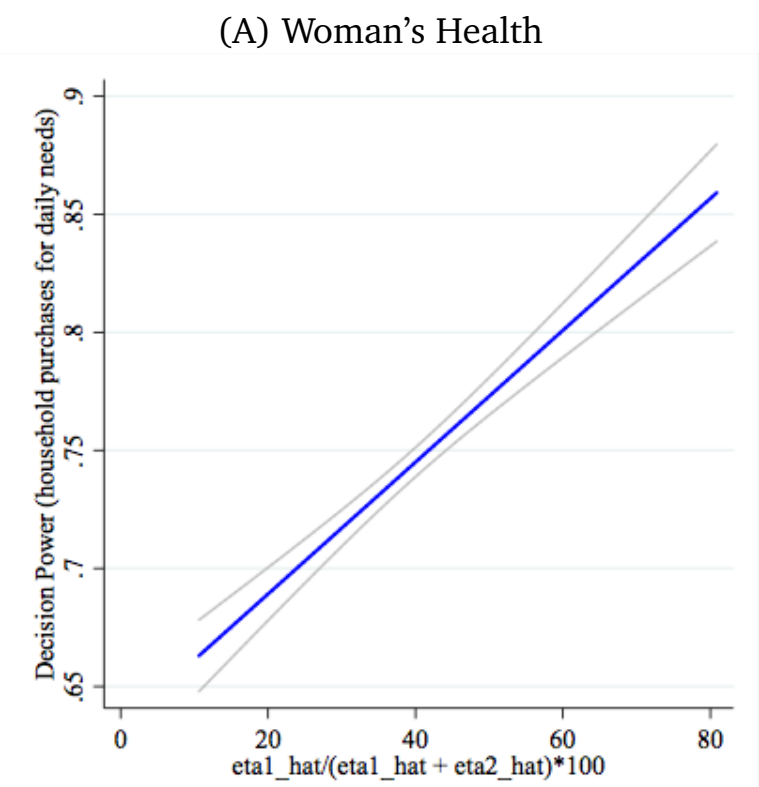
[22] While we do not present the detailed breakdown by vaccinations here, this effect is primarily driven by children's likelihood to be vaccinated against polio, and not by BCG or DPT vaccinations.



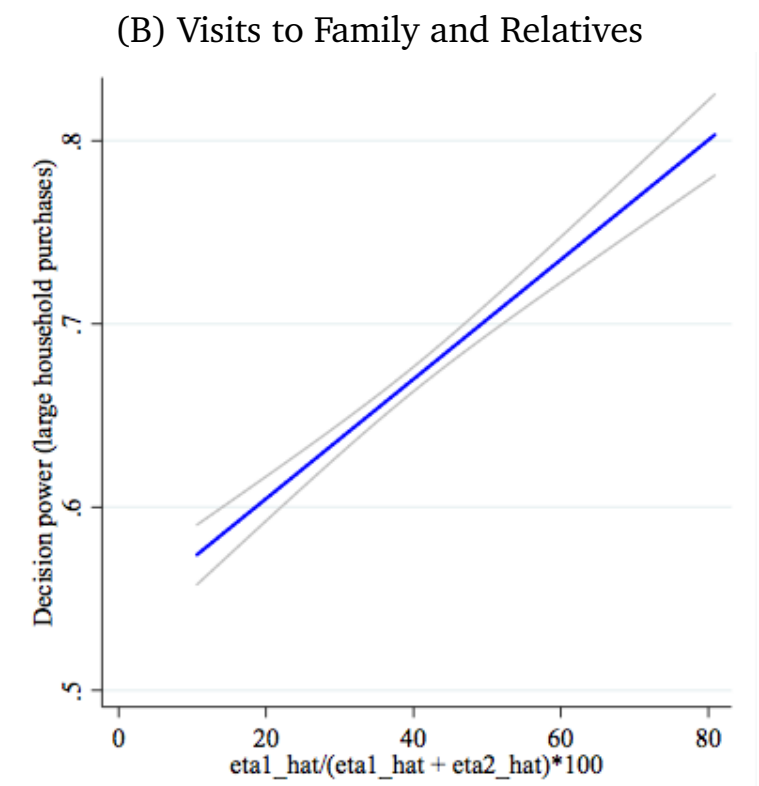
(A) Woman's Health



(B) Visits to Family and Relatives



(C) Small Household Purchases



(D) Large Household Purchases

Figure 2: Structurally Recovered Bargaining Power and Household Decision Making

our MR-LATE estimator, the estimated binary treatment indicator T will frequently equal the true D, even when the underlying continuous variable R is always mismeasured, never equalling the true R^*. Taking advantage of separate T^a and T^b treatment estimates, our MR-LATE can further reduce and sometimes eliminate the impact of mismeasurement.

A second reason for focusing on a binary treatment is that this is sometimes a more appropriate model of behavior. For making discrete decisions (e.g., whether to vaccinate a child or not), what matters might be simply who has the most power, rather than the amount by which one's power exceeds the other. This is analogous to voter models, where what determines the outcome is who has the majority, not how big the majority is or how strongly each voter feels.

We do not know if health decisions are more appropriately modeled as functions of D or as smooth functions of R^*. So for comparison to our MR-LATE estimates, we also consider linearly regressing Y_i on a constant, on R_i, and on X_i, using 2SLS where Z_i is the instrument for R_i. We have no reason to expect that the true relationship of Y to R^* and X is linear, or that estimation errors in R relative to R^* satisfy the assumptions of classical measurement error. But if these assumptions do hold, then the estimated coefficient of R in this regression will consistently estimate the average marginal effect of R^* on Y.

As before, Y_i is a measure of the health status of individual i, X_i is a vector of covariates that include variables that characterize the individual and the household, and the instrument Z_i is again the mother's eligibility for the HSA amendments. Tables A4, A5 and A6 in the Appendix contain the estimation results. These results generally agree with our MR-LATE estimates. We again find positive and significant effects of marginal increases in women's relative power on her body mass index, her likelihood to be underweight or anemic, and on her children's likelihood of being vaccinated, and has no systematic effect on health outcomes for her husband or for other (anthropometric) measures of the health of her children.

6 Conclusion

We apply a novel two-step approach to study the effect of intra-household women's empowerment on the health status of family members in India. Our model looks at the effect on family member health outcomes Y of a treatment D, defined as a wife having relatively high intra-household bargaining power, as measured by an unobservable relative resource share measure R^*. In a first step, we rely on a structural model to recover R, an estimate of R^*. We use this continuous measure R to construct estimated treatment indicators T^a and T^b. Due to measurement, estimation, or specification errors in the structural model for R, neither T^a nor T^b will equal (or consistently estimate) the true treatment indicator D.

We propose a new mismeasurement robust LATE estimator, called MR-LATE, that uses both T^a and T^b to obtain consistent estimates of LATE given an outcome Y and an instrument Z, despite the fact that the true treatment indicator D is unobserved and cannot be consistently estimated. Assuming the instrument Z satisfies the usual properties for LATE estimation, we apply MR-LATE

to study the causal effect of the unobserved D on Y. In our application, we construct our estimated treatment indicators T^a and T^b from an underlying estimated continuous measure R. But this is not necessary to apply our estimator. What we mainly need to have MR-LATE equal the true LATE is just that $D - T^a$ never equals one and $D - T^b$ never equals zero. Having MR-LATE be a good approximation to LATE mainly requires that $D - T^a$ and $D - T^b$ rarely equal one and zero, respectively.

Overall, we find that policies aimed at increasing women's bargaining power within households (such as strengthening their rights to inherit property) should lead to improvements in overall family health. In particular, exogenously increasing a wife's control of resources within a household appears to substantially improve her own health, and increases the vaccination rates of her children, without damaging her husband's health, despite his corresponding descreased share of household resources.

More generally, our analysis highlights the advantages of combining structural and causal features in conducting empirical analysis. Our MR-LATE estimator specifically accounts for the fact that structural estimation generally suffers from multiple errors, including specification errors. An advantage of our procedure is that it allows us to focus on causal effect of direct economic interest and relevance, even if those treatments cannot be directly observed. This may be particularly useful for constructing causal tests and benchmarks of economic theories, since we can directly focus on treatments that are motivated by theory, instead of calculating the treatment effects of less relevant proxies that happen to be directly reported.

References

AGARWAL, B. (1995): "A Field of One's Own," *Cambridge Books*. [19]

ANGRIST, J. D., G. W. IMBENS, AND D. B. RUBIN (1996): "Identification of Causal Effects Using Instrumental Variables," *Journal of the American Statistical Association*, 91, 444–455. [4]

ANGRIST, J. D. AND J.-S. PISCHKE (2010): "The Credibility Revolution in Empirical Economics: How Better Research Design Is Taking the Con out of Econometrics," *Journal of Economic Perspectives*, 24, 3–30. [6]

APPS, P. F. AND R. REES (1988): "Taxation and the Household," *Journal of Public Economics*, 35, 355 – 369. [5]

ATHEY, S., R. CHETTY, G. IMBENS, AND H. KANG (2016): "Estimating Treatment Effects using Multiple Surrogates: The Role of the Surrogate Score and the Surrogate Index," *arXiv preprint arXiv:1603.09326*. [5]

ATTANASIO, O. AND V. LECHENE (2002): "Tests of Income Pooling in Household Decisions," *Review of Economic Dynamics*, 5, 720–748. [5]

BEGG, C. B. AND D. H. LEUNG (2000): "On the use of surrogate end points in randomized trials," *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, 163, 15–28. [5]

BOSE, N. AND S. DAS (2015): "Women's Inheritance Rights, Household Allocation and Gender Bias," *Mimeo*. [5]

BROWNING, M., P.-A. CHIAPPORI, AND A. LEWBEL (2013): "Estimating Consumption Economies of Scale, Adult Equivalence Scales, and Household Bargaining Power," *Review of Economic Studies*, 80, 1267–1303. [5], [32]

CALVI, R. (2016): "Why Are Older Women Missing in India? The Age Profile of Bargaining Power and Poverty," *Mimeo*. [5], [18]

CHIAPPORI, P.-A. (1988): "Rational household labor supply," *Econometrica: Journal of the Econometric Society*, 63–90. [5]

——— (1992): "Collective Labor Supply and Welfare," *Journal of Political Economy*, 100, pp. 437–467. [5]

DEININGER, K., A. GOYAL, AND H. NAGARAJAN (2013): "Women's Inheritance Rights and Intergenerational Transmission of Resources in India," *Journal of Human Resources*, 48, 114–141. [5]

DUNBAR, G. R., A. LEWBEL, AND K. PENDAKUR (2013): "Children's Resources in Collective Households: Identification, Estimation, and an Application to Child Poverty in Malawi," *American Economic Review*, 103, 438–71. [5], [15], [20], [32], [34]

FRANGAKIS, C. E. AND D. B. RUBIN (2002): "Principal stratification in causal inference," *Biometrics*, 58, 21–29. [5]

HARARI, M. (2014): "Women's Inheritance Rights and Bargaining Power: Evidence from Kenya," *Mimeo*. [5]

HEATH, R. AND X. TAN (2014): "Intrahousehold Bargaining, Female Autonomy, and Labor Supply: Theory and Evidence from India," Mimeo, Department of Economics, University of Washington. [5], [20], [34]

HECKMAN, J., S. URZUA, AND E. VYTLACIL (2006): "Understanding Instrumental Variables in Models with Essential Heterogeneity," *Review of Economics and Statistics*, 88, 389–432. [6]

HECKMAN, J. J. (2010): "Building Bridges between Structural and Program Evaluation Approaches to Evaluating Policy," *Journal of Economic Literature*, 48, 356–98. [6]

HECKMAN, J. J. AND E. J. VYTLACIL (2007): "Chapter 70 Econometric Evaluation of Social Programs, Part I: Causal Models, Structural Models and Econometric Policy Evaluation," Elsevier, vol. 6, Part B of *Handbook of Econometrics*, 4779 – 4874. [6]

HU, Y. (2008): "Identification and estimation of nonlinear models with misclassification error using instrumental variables: A general solution," *Journal of Econometrics*, 144, 27 – 61. [4]

HU, Y., J.-L. SHIU, AND T. WOUTERSEN (2015): "Identification and estimation of single-index models with measurement error and endogeneity," *The Econometrics Journal*, 18, 347–362. [4]

IMBENS, G. W. (2010): "Better LATE Than Nothing: Some Comments on Deaton (2009) and Heckman and Urzua (2009)," *Journal of Economic Literature*, 48, 399–423. [6]

IMBENS, G. W. AND J. D. ANGRIST (1994): "Identification and Estimation of Local Average Treatment Effects," *Econometrica*, 62, 467–475. [4], [6], [7], [17]

JAIN, T. (2014): "Where There Is a Will Fertility Behavior and Sex Bias in Large Families," *Journal of Human Resources*, 49, 393–423. [5]

KAUR, K. (2014): "Anaemia 'A Silent Killer' Among Women in India: Present Scenario," *European Journal of Zoological Research*, 3, 32–36. [19]

LA FERRARA, E. AND A. MILAZZO (2014): "Customary Norms, Inheritance, and Human Capital: Evidence from a Reform of the Matrilineal System in Ghana," CEPR Discussion Papers 10159, C.E.P.R. Discussion Papers. [5]

LEWBEL, A. (1998): "Semiparametric Latent Variable Model Estimation with Endogenous or Mismeasured Regressors," *Econometrica*, 66, 105–121. [4]

——— (2007): "Estimation of Average Treatment Effects with Misclassification," *Econometrica*, 75, 537–551. [4]

——— (2016): "The Identification Zoo - Meanings of Identification in Econometrics," *Forthcoming on Journal of Economic Literature*. [6]

LEWBEL, A. AND K. PENDAKUR (2008): "Estimation of Collective Household Models with Engel Curves," *Journal of Econometrics*, 147, 350–358. [5]

MACOURS, K., N. SCHADY, AND R. VAKIS (2012): "Cash Transfers, Behavioral Changes, and Cognitive Development in Early Childhood: Evidence from a Randomized Experiment," *American Economic Journal: Applied Economics*, 4, 247–73. [5]

MAHAJAN, A. (2006): "Identification and Estimation of Regression Models with Misclassification," *Econometrica*, 74, 631–665. [4]

PEARL, J. (2009): "Causal inference in statistics: An Overview," *Statistics Surveys*, 3, 96–146. [6]

PRENTICE, R. L. (1989): "Surrogate endpoints in clinical trials: definition and operational criteria," *Statistics in medicine*, 8, 431–440. [4]

ROSENBAUM, P. R. (1984): "The consquences of adjustment for a concomitant variable that has been affected by the treatment," *Journal of the Royal Statistical Society. Series A (General)*, 656–666. [5]

ROY, S. (2008): "Female Empowerment Through Inheritance Rights: Evidence from India," Mimeo, London School of Economics, London. [5], [34]

——— (2013): "Empowering Women? Inheritance Rights, Female Education and Dowry Payments in India," . [5]

RUBALCAVA, L., G. TERUEL, AND D. THOMAS (2009): "Investments, time preferences and public transfers paid to women," *Economic Development and cultural change*, 57, 507. [5]

SHIWAKU, K., E. ANUURAD, B. ENKHMAA, K. KITAJIMA, AND Y. YAMANE (2004): "Appropriate Body-Mass Index For Asian Populations and Its Implications For Policy and Intervention Strategies," *Lancet*. [18]

SONG, S. (2015): "Semiparametric estimation of models with conditional moment restrictions in the presence of nonclassical measurement errors," *Journal of Econometrics*, 185, 95 – 109. [4]

SONG, S., S. M. SCHENNACH, AND H. WHITE (2015): "Estimating nonseparable models with mismeasured endogenous variables," *Quantitative Economics*, 6, 749–794. [4]

TOMMASI, D. (2015): "How Cash Transfers Improve Child Development," ECARES working paper 2015-19. [5]

TOMMASI, D. AND A. WOLF (2016): "Overcoming Weak Identification in the Estimation of Household Resource Shares," ECARES working paper 2016-12. [18]

URA, T. (2015): "Heterogeneous Treatment Effects with Mismeasured Endogenous Treatment," Mimeo, Duke University. [2], [4], [8], [11]

VANDERWEELE, T. (2015): *Explanation in causal inference: methods for mediation and interaction*, Oxford University Press. [5]

VYTLACIL, E. (2002): "Independence, Monotonicity, and Latent Index Models: An Equivalence Result," *Econometrica*, 70, 331–341. [6]

WOLPIN, K. I. (2013): *The limits of inference without theory*, MIT Press. [6]

WORLD BANK (2014): "Women, Business and the Law 2014. Removing Restrictions to Enhance Gender Equality," *World Bank Report*. [19]

A Appendix

A.1 Proof of Theorem 1

Substituting equation (1) into equation (4) gives

$$T = T_0 + (T_1 - T_0) D = T_0 + (T_1 - T_0)[(1-Z) D_0 + Z D_1] \tag{A1}$$

Multiplying equation (2) by equation (A1) gives

$$YT = [Y_0 + (Y_1 - Y_0)[(1-Z) D_0 + Z D_1]][T_0 + (T_1 - T_0)[(1-Z) D_0 + Z D_1]]$$

Using assumption 2-i, this makes

$$\begin{aligned} E(YT \mid Z=1) &= E[[Y_0 + (Y_1 - Y_0) D_1][T_0 + (T_1 - T_0) D_1]] \\ &= E[T_0 Y_0 + (Y_1 T_1 - Y_0 T_0) D_1] \end{aligned}$$

where the last equality uses $D_1^2 = D_1$. Similarly,

$$E(YT \mid Z=0) = E[T_0 Y_0 + (Y_1 T_1 - Y_0 T_0) D_0]$$

So

$$E(YT \mid Z=1) - E(YT \mid Z=0) = E[(Y_1 T_1 - Y_0 T_0)(D_1 - D_0)]$$

Given the no defiers assumption, either $D_1 - D_0 = 0$ or $D_1 - D_0 = 1$, and someone is a complier if and only if they have $D_1 - D_0 = 1$. The probability of being a complier is $\Pr(D_1 - D_0 = 1) = E[D_1 - D_0]$. We therefore apply the standard LATE logic:

$$\begin{aligned} E(YT \mid Z=1) - E(YT \mid Z=0) &= E[(Y_1 T_1 - Y_0 T_0)(D_1 - D_0)] \\ &= E[Y_1 T_1 - Y_0 T_0 \mid D_1 - D_0 = 1] \Pr(D_1 - D_0 = 1) \\ &= E(Y_1 T_1 - Y_0 T_0 \mid C) E(D_1 - D_0). \end{aligned}$$

Let $t_d = E(T_d \mid C)$. Then, using Assumption 2-ii, the above further simplifies to

$$E(YT \mid Z=1) - E(YT \mid Z=0) = E(t_1 Y_1 - t_0 Y_0 \mid C) E(D_1 - D_0).$$

Replacing Y with one gives

$$\begin{aligned} E(T \mid Z=1) - E(T \mid Z=0) &= E(t_1 - t_0 \mid C) E(D_1 - D_0) \\ &= (t_1 - t_0) E(D_1 - D_0). \end{aligned}$$

And therefore

$$\frac{E(YT \mid Z=1) - E(YT \mid Z=0)}{E(T \mid Z=1) - E(T \mid Z=0)} = \frac{E(Y_1 t_1 - Y_0 t_0 \mid C)}{t_1 - t_0}$$

which equals equation (7), thereby proving the Theorem.

A.2 Treatment indicators when the error term ϵ is bounded: A graphical illustration

Assume that $supp(\varepsilon) \subset (-\kappa, \kappa) \subset supp(R^* - c)$. Then it follows that for $T = T^a$ we have $p_1^a = 1$ with $p_0^a = 0$, and for $T = T^b$ we have $p_1^b = 0$ and $p_0^b = 1$, and so $\lambda^a - \lambda^b = E[Y_1 - Y_0 \mid C]$. Given corollary 2, LATE can be point identified. Figure A1 provides a graphical representation of this. If there was no measurement error, the true treatment and control groups would coincide with the respective observed groups. Hence, all individuals on the green line on the right hand side of c, would have a R larger than the threshold value; otherwise they would be on the green line on the left hand side of c. In case of bounded measurement error, we would need to define two treatment indicators, T^a and T^b, that guarantee the conditions $q^a = 1$ and $q^b = 0$ that we have just pointed out. As we can see, by knowing the bound κ, we are able to define a T^a such that for all individuals on the blue line on the right hand side of $c + \kappa$, $p_1^a = 1$ for sure. That is, with probability 1, these individuals, who are observed in the treatment group, belong to the true treatment group. Following the same reasoning, we are able to define also a T^b such that for all individuals on the red line on the left hand side of $c - \kappa$, $p_0^b = 0$ for sure. That is, with probability 0, these individuals, who are observed in the control group, belong to the true control group.

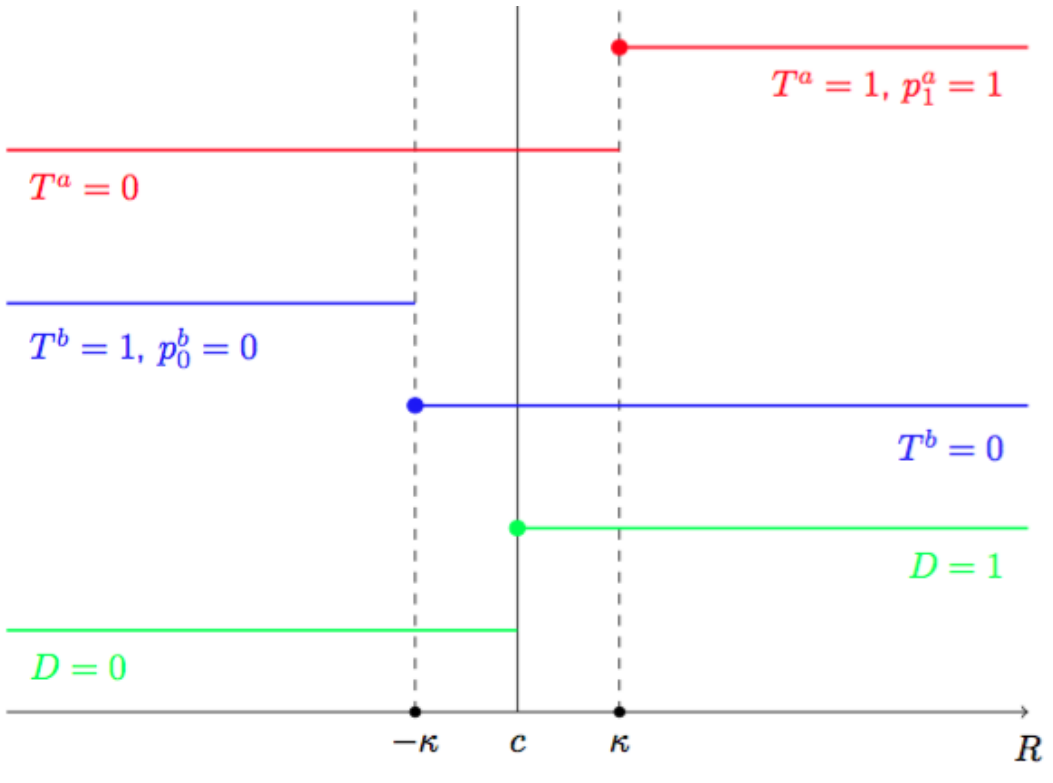


Figure A1: Illustrative Example

A.3 Derivation of household demand equations of private assignable goods

Here we summarize the derivation of our model, based on Browning et al. (2013) and Dunbar et al. (2013), for estimating resource shares from the demand equations for private assignable goods. Let the household be comprised of T types of individuals indexed $t = 1, ..., T$. Recall M is the total expenditures of the household, i.e., the household's total budget, X denotes a vector of observable attributes of households and their members, $\widetilde{Z}$ denotes a vector of distribution factors, and Q_1,...,Q_T are quantities of each assignable good consumed by household member t.

Let P denote the vector of market prices for all of the commodities (goods and services) that the household buys. t There is also a vector S of quantities of other goods the household purchases. Unlike Q_1,...,Q_T, the goods S can be shared and hence jointly consumed to some extent.

Start by assuming the household chooses what to consume using the program

$$\max_{Q_1,...,Q_T,S_1,...,S_T} \widetilde{V}\left[V_1(Q_1,S_1,X),...,V_T(Q_T,S_T,X) \mid \widetilde{Z},X,P/M\right] \tag{A2}$$

$$\text{such that} \quad S = \sum_{t=1}^{T} S_t \quad \text{and} \quad M = P_S' A(X) S + \sum_{t=1}^{T} P_t Q_t$$

where $V_t(Q_t,S_t,X)$ for $t = 1, ..., T$ is the utility function of household member t, and the function $\widetilde{V}$ describes the social welfare function or bargaining process of the household, which exists because the household is pareto efficient.

Household attributes X may affect preferences, and so appear inside the utility functions V_t. These X variables may also affect the bargaining process or social welfare function given by $\widetilde{V}$ (by, e.g., affecting the relative bargaining power of members), and as a result may affect resource shares.

We have scalars Q_1,...,Q_T that are the quantities of private, assignable goods, where member t has quantity Q_t in his or her utility function, and does not have Q_ℓ for all $\ell \neq t$ in his or her utility function. Each member's utility function also depends on a quantity vector of other goods S_t. The market prices of these goods are given by the vector P_S. The square matrix $A(X)$ is what is called by BCL a linear consumption technology function over goods. Having $A(X)$ differ from the identity matrix is what allows goods in S to be partly shared and/or consumed jointly. In particular, $A(X)S$ equals the quantity vector of these goods that the household actually purchases, while $S = \sum_{t=1}^{T} S_t$ is total quantity vector of these goods that the household consumes. These quantities are not the same due to sharing and joint consumption. The smaller an element of $A(X)S$ is relative to the corresponding element of S, the more that good is shared or jointly consumed. See BCL for details. The vector of all prices P includes, P_S, the vector of prices of the elements of S, and P_1,...,P_T, the prices of the private assignable goods Q_1,...,Q_T.

What makes the vector $\widetilde{Z}$ be distribution factors (observed and unobserved, respectively) in the model is that they appear only as arguments of $\widetilde{V}$, and so only affect the allocation of resources within the household, but not the tastes of the individual household members.

Applying duality theory and decentralization welfare theorems, it follows from Browning et al.

(2013) that the household's program above is equivalent to a program where each household member t chooses what to consume using the program

$$\max_{Q_t,S_t} V_t(Q_t,S_t,X) \text{ such that } \eta_t(P,M,X,\widetilde{Z})M = P_S'A(X)S_t + P_tQ_t \tag{A3}$$

where $\eta_t = \eta_t(P,M,X,\widetilde{Z})$ is the resource share of member t, that is, η_t is the fraction of total household resources M that are allocated to member t. This member then chooses quantities Q_t and the vector S_t subject to a linear budget constraint. The vector $P_SA(X)$ equals the vector of shadow prices of goods S. These shadow prices for the household are lower than market prices, due to sharing. Being private and assignable, the shadow price of each Q_t equals its market price P_t. The shadow budget for member t is $\widetilde{M}_t = \eta_t M$. As shown in BCL, the resource share functions $\eta_t(P,M,X,\widetilde{Z})$ for each member t in general depend on the function $\widetilde{V}$ and on the utility functions $V_1...,V_T$.

BCL show that the more bargaining power a household member has (i.e., the greater is the weight of his or her utility function in $\widetilde{V}$), the larger is their resource share η_t. Resource shares η_t all lie between zero and one, and resource shares sum to one, that is, $\sum_{t=1}^{T}\eta_t = 1$.

As in DLP, we will not work with the household demand functions of all goods (which, as shown in BCL, can be rather complicated). Instead, we only make use of the demand functions of the private assignable goods Q_t, which are simpler. Since equation (A3) is an ordinary utility function maximized under a linear budget constraint (linear in shadow prices and a shadow budget), the solution to equation (A3) is a set of Marshallian demand equations for Q_t and S_t.

Let $h_t\left(\widetilde{M}_t,P,X\right)$ be the Marshallian demand function of person t' for their private assignable good, that is, $h_t\left(\widetilde{M}_t,P,X\right)$ is the quantity person t in a household with member attributes X would demand of their assignable good if they had a budget equal to their shadow budget $\widetilde{M}_t$ and faced the within-household shadow price vector that corresponds to the market price vector P. Since each Q_t is private and assignable, the quantity Q_t that member t chooses to consume equals the quantity of this good that the household buys. It therefore follows from the above that the household's quantity demand of each private assignable good Q_t is given by

$$Q_t = h_t\left(\eta_t(P,M,X,\widetilde{Z})M,P,X\right) \quad \text{for } t = 1,...,T. \tag{A4}$$

The interpretation of this equation is that the total resources allocated to member t are $\eta_t M$ (the share η_t of total household budget M) and the function h_t is that member's Marshallian demand function for this good. Since the good is private and assignable, the household's demand for the good just equals that member's own demand for the good. It is important to note that only private assignable goods have the simple form given by equation (A4). The demand functions for other goods are much more complicated, as in BCL.

Let $\widetilde{h}_t\left(\widetilde{M}_t,P,X\right) = P_th_t\left(\widetilde{M}_t,P,X\right)/\widetilde{M}_t$ denote the Marshallian demand function written in budget share form. That is, $\widetilde{h}_t\left(\widetilde{M}_t,P,X\right)$ is the fraction of the total budget $\widetilde{M}_t$ that is spent on the good t.

DLP assume data are single price regime (that is, Engel curve data), so P is a fixed constant that can be dropped from the model. They provide empirical and theoretical evidence that η_t does not depend on M.[23] This allows them to rewrite equation (A4) as $w^t = \eta_t(X,\widetilde{Z})\widetilde{h}_t\left(\eta_t(X,\widetilde{Z})M, X\right)$ for $t = 1,...,T$, where $w^t = P_tQ_t/M$ is the household's budget share of good t, that is, the fraction of the household's total budget M that is spend on buying Q_t. DLP provide a class of functional forms for the utility functions $\widetilde{V}$ that make $\widetilde{h}_t$ linear in the log of its first argument, so $w^t = \eta_t(X,\widetilde{Z})[\delta^t(X) + (lnM + ln\eta_t(X,\widetilde{Z}))\beta(X)]$ for some functions $\delta^t(X)$ and $\beta(X)$. The assumption that $\beta(X)$ does not depend on t is what DLP call the SAP (similar across people) assumption.

A.4 Structural Estimation Details

Table A9 reports the estimated coefficients of the covariates $(X_1,...X_n,Z)$ for mother's and father's resource shares.[24] The model is estimated using the NSS sample. The most important results in this Table are threefold.

First, the higher the number of children, the lower is a mother's resource share. This is holds true for fathers as well, but at a much lesser extent. The fraction of female children is positively related to η_f: if all children are girls, women's resource shares are 1.1 percentage points larger. This result is in line with the findings in Dunbar et al. (2013) and can be attributed to the fact that adult women may be willing (or expected) to forgo a higher fraction of household resources in presence of male children, due to son preference. Although the coefficients are not statistically significant, the higher is women's age the lower is the fraction of household's total expenditure devoted to women.

Second, household socio-economic characteristics play an important role, too. In particular, being part of Scheduled Caste, Scheduled Tribes, and other disadvantaged social classes is associated with higher women's bargaining power. The same holds true for residing in the North-East states, which is consistent with the presence of a number of matrilineal societies and cultures in these regions (Khasi and Garo societies, for example). In contrast, North Indian women and women living in rural areas seem to have a much lower bargaining power. Finally, mothers' resource shares are higher in households belonging to Scheduled Castes, Scheduled Tribes and other backward classes, while fathers' resource shares are larger in families of Hindu, Jain, Sikh and Buddhist religions.

Finally, the estimated model confirms the importance of the HSA amendments (HSAAs) in shaping women's bargaining position within the household. In households where mothers are eligible to these reforms, their resource shares are larger.[25]

[23] Lise and Seitz (2007), Lewbel and Pendakur (2008), Bargain and Donni (2009, 2012) and DLP all use this restriction in their identification results, and supply some theoretical arguments for it. De Rock et al (2013) and Menon, Perali and Pendakur (2013) provide empirical support for this restriction.

[24] The estimated coefficients of the covariates for the preference parameters $\tilde{\alpha}_{food}$, α_t, $t = f,m,c$, $\tilde{\beta}_{food}$, and β are available upon request.

[25] These results align with the findings in Roy (2008) and Heath and Tan (2014) on the effects of HSAA on self-reported measures of women's autonomy and bargaining power. Specifically, Heath and Tan (2014) find that exposure to HSAA decreases the probability that a woman has no say in household decisions (by 6.6 percentage point) and increases the probability that a woman can go alone to the market (by 8.2 p.p.), to a health facility (by 6.9 p.p.) and to places outside of the village (by 8.3 p.p.).

A.5 Additional Tables

Table A1: NSS Consumer Expenditure Data and NFHS Household Data

	NSS Sample				NFHS Sample			
	Obs.	Mean	Median	St. Dev.	Obs.	Mean	Median	St. Dev.
Total Expenditure (Rupees)	15,175	7,332.6730	6,002.0000	5,927.5110				
Woman's Assignable Clothing Budget Share	15,166	1.2225	1.0617	0.9891				
Man's Assignable Clothing Budget Share	15,166	1.4191	1.1826	1.2636				
Children's Assignable Clothing Budget Share	15,166	0.7466	0.5882	0.8257				
$\mathbb{I}$(1 child)	15,175	0.2655	0.0000	0.4416	24,746	0.2485	0.0000	0.4322
$\mathbb{I}$(2 children)	15,175	0.4397	0.0000	0.4964	24,746	0.3761	0.0000	0.4844
$\mathbb{I}$(3 children)	15,175	0.2129	0.0000	0.4094	24,746	0.2265	0.0000	0.4186
$\mathbb{I}$(4 children)	15,175	0.0818	0.0000	0.2741	24,746	0.1489	0.0000	0.3560
Fraction of Female Children	15,175	0.4388	0.5000	0.3660	24,746	0.4725	0.5000	0.3619
$\mathbb{I}$(Hindu, Buddhist, Jain, Sikh)	15,175	0.7860	1.0000	0.4102	24,742	0.7614	1.0000	0.4263
Woman's Age	15,175	32.8019	32.0000	6.5469	24,743	30.2858	30.0000	7.2378
Gender Age Gap	15,175	4.5094	5.0000	5.9424	24,708	5.8047	5.0000	4.9976
Children's Avg. Age	15,175	7.6991	8.0000	3.4585	24,746	6.4912	6.5000	3.5270
Wealth Index	15,175	0.3827	0.3695	0.1729	24,746	0.3808	0.3478	0.2327
$\mathbb{I}$(SC, ST, OBC)	15,175	0.7242	1.0000	0.4470	23,714	0.6834	1.0000	0.4652
$\mathbb{I}$(Woman's Higher Education)	15,175	0.1897	0.0000	0.3921	24,746	0.0820	0.0000	0.2744
$\mathbb{I}$(Man's Higher Education)	15,175	0.2789	0.0000	0.4485	24,746	0.1282	0.0000	0.3343
$\mathbb{I}$(Rural)	15,175	0.5934	1.0000	0.4912	24,746	0.5451	1.0000	0.4980
$\mathbb{I}$(North India)	15,175	0.2659	0.0000	0.4418	24,746	0.3082	0.0000	0.4617
$\mathbb{I}$(East India)	15,175	0.1842	0.0000	0.3876	24,746	0.1616	0.0000	0.3681
$\mathbb{I}$(North-East India)	15,175	0.1593	0.0000	0.3659	24,746	0.2016	0.0000	0.4012
$\mathbb{I}$(South India)	15,175	0.2799	0.0000	0.4489	24,746	0.2017	0.0000	0.4013
$\mathbb{I}$(West India)	15,175	0.1108	0.0000	0.3139	24,746	0.1269	0.0000	0.3329
$\mathbb{I}$(HSA Eligible)	15,175	0.2946	0.0000	0.4559	22,122	0.0566	0.0000	0.2312

Notes: Budget shares are multiplied by 100. Woman's assignable clothing includes expenditures on saree, shawls, chaddar, and kurta-pajamas suits for females; man's assignable clothing includes expenditures on dhoti, lungi, kurta-pajamas suits for males, pajamas, salwar, and cloth for coats, trousers, and suit and for shirt, pajama, kurta, and salwar; children's assignable clothing includes expenditures on expenditure on school uniforms and infant clothing. The household wealth index is obtained using principle component analysis. $\mathbb{I}$(Higher Education Women) and $\mathbb{I}$(Higher Education Men) are indicator variable for higher education (diploma or college) completed by at least one woman or man in the household. North India includes Jammu & Kashmir, Himachal Pradesh, Punjab, Chandigarh, Uttaranchal, Haryana, Delhi, Rajasthan, Uttar Pradesh, and Madhya Pradesh. East India includes West Bengal, Bihar, Jharkhand, Orissa, A & N Islands, and Chattisgarh. North-East India includes Sikkim, Arunachal Pradesh, Assam, Manipur, Meghalaya, Mizoram, Nagaland, and Tripura. South India includes Karnataka, Tamil Nadu, Andhra Pradesh, Kerala, Lakshadweep, and Pondicherry. West India includes Gujarat, Goa, Maharashtra, Daman & Diu, and D & N Haveli.

Table A2: NFHS Individual Data

	Women (age 15 to 49)			Men (age 15 to 54)			Children (age 0 to 5)		
	Obs.	Mean	St. Dev.	Obs.	Mean	St. Dev.	Obs.	Mean	St. Dev.
BMI (kg/m^2)	18,149	20.8923	3.9301	9,366	21.1331	3.4932			
$\mathbb{I}$(BMI $\leq$18.5)	18,149	0.2980	0.4574	9,366	0.2431	0.4290			
$\mathbb{I}$(Anemic)	16,941	0.1599	0.3665	8,607	0.1565	0.3634			
Avg. Birth Gap (months)	15,503	34.6275	17.4954						
Weight (kg)							11,789	10.6885	3.2873
Height (cm)							11,761	84.0554	13.927
$\mathbb{I}$(BCG)							12,814	0.7666	0.4230
$\mathbb{I}$(DPT)							12,749	0.7375	0.4400
$\mathbb{I}$(Polio)							12,860	0.9002	0.2997
$\mathbb{I}$(Any Vaccination)							12,875	0.9095	0.2869
Fraction of Female Children	18,865	0.4711	0.3551	9,895	0.7564	0.4293	48,683	0.4794	0.3240
$\mathbb{I}$(Hindu, Buddhist, Jain, Sikh)	18,865	0.7782	0.4155	9,895	29.8708	5.8405	48,683	0.7615	0.4262
Woman's Age	18,865	29.9274	5.8851	9,895	5.6242	3.7887	48,683	30.5198	5.7048
Gender Age Gap	18,865	5.8685	4.4122	9,895	6.4730	3.5123	48,683	5.8044	4.3917
Children's Avg. Age	18,865	6.4880	3.4559	9,895	0.3792	0.2273	48,683	6.8046	3.1937
Wealth Index	18,865	0.3771	0.2328	9,895	0.7092	0.4541	48,683	0.3440	0.2242
$\mathbb{I}$(SC, ST, OBC)	18,865	0.6774	0.4675	9,895	0.0820	0.2743	48,683	0.7091	0.4542
$\mathbb{I}$(Woman's Higher Education)	18,865	0.0804	0.2719	9,895	0.1307	0.3371	48,683	0.0539	0.2257
$\mathbb{I}$(Man's Higher Education)	18,865	0.1287	0.3348	9,895	0.5325	0.4990	48,683	0.0985	0.2980
$\mathbb{I}$(Rural)	18,865	0.5525	0.4972	9,895	0.3075	0.4615	48,683	0.5960	0.4907
$\mathbb{I}$(North)	18,865	0.3362	0.4724	9,895	0.1187	0.3235	48,683	0.3630	0.4809
$\mathbb{I}$(East)	18,865	0.1807	0.3847	9,895	0.2281	0.4196	48,683	0.1837	0.3873
$\mathbb{I}$(North-East)	18,865	0.2230	0.4162	9,895	0.2226	0.4160	48,683	0.2190	0.4136
$\mathbb{I}$(South)	18,865	0.1446	0.3517	9,895	0.1230	0.3284	48,683	0.1253	0.3310
$\mathbb{I}$(West)	18,865	0.1157	0.3198	9,894	0.4359	0.4959	48,683	0.1089	0.3116
$\mathbb{I}$(HSA Eligible)	15,966	0.0847	0.2785	8,078	0.1378	0.3447	42,495	0.0635	0.2438
Child's Age							43,910	7.5695	4.8679
$\mathbb{I}$(Child is Female)							48,683	0.4845	0.4998

Notes: Age is defined in years for men and women, and in months for children.

Table A3: First Stage Estimates (MR-LATE 2SLS, $\kappa = 0$)

	$\mathbb{1}(\frac{\widehat{\eta}_m}{\widehat{\eta}_f+\widehat{\eta}_m} \geq c)$					
	Women (15-59)		Men (15-54)		Children (0-5)	
	(1) OLS	(2) OLS	(3) OLS	(4) OLS	(5) OLS	(6) OLS
$\mathbb{1}$(HSAA Eligible) (Z)	0.0684	0.142	0.0676	0.141	0.0236	0.0937
	(0.0152)	(0.0213)	(0.0178)	(0.0249)	(0.0169)	(0.0400)
Observations	15,966	15,966	8,078	8,078	12,815	12,815
Mean Dependent Variable	0.4891	0.4891	0.5276	0.5276	0.5341	0.5341

Note: NFHS data. Bootstrap standard errors in parentheses. Median of the distribution of R used as threshold. All specifications include individuals and household controls. Specifications in columns (2), (4) and (6) include state, cohort, state-religion fixed effects and state specific linear time trends.

Table A4: Women's Health (continuous R)

	Body Mass Index	Pr(Underweight)	Pr(Anemic)
	2SLS	2SLS	2SLS
Woman's Relative Power (R)	1.347	-0.107	-0.0993
	(0.552)	(0.0546)	(0.0513)
Observations	15,377	15,377	14,286
Mean Dependent Variable	20.96	0.2914	0.1547

Note: NFHS data. Bootstrap standard errors in parentheses. All specifications include individuals and household controls, state, cohort, state-religion fixed effects and state specific linear time trends.

Table A5: Men's Health (continuous R)

	Body Mass Index	Pr(Underweight)	Pr(Anemic)
	2SLS	2SLS	2SLS
Woman's Relative Power (R)	0.376	-0.0165	-0.0130
	(0.547)	(0.0598)	(0.0443)
Observations	7,659	7,659	6,839
Mean Dependent Variable	21.18	0.2341	0.1040

Note: NFHS data. Bootstrap standard errors in parentheses. All specifications include individuals and household controls, state, cohort, state-religion fixed effects and state specific linear time trends.

Table A6: Children's Health (continuous R)

	Height-for-age (z-score)	Weight-for-age (z-score)	Pr(Any Vaccination)
	2SLS	2SLS	2SLS
Woman's Relative Power (R)	-0.739	-0.132	0.195
	(0.761)	(0.597)	(0.0873)
Observations	10,452	10,452	12,109
Mean Dependent Variable	-1.668	-1.782	0.9039

Note: NFHS data. Bootstrap standard errors in parentheses. All specifications include individuals and household controls, state, cohort, state-religion fixed effects and state specific linear time trends.

Table A7: Structural and Self-reported Bargaining Power

	𝕀(Woman Participate in Final Decisions on)			
	Her Health	Large Household Purchases	Small Household Purchases	Visits to Family and Relatives
Women's Power R	0.00793	0.0152	0.0110	0.0137
	(0.00212)	(0.00220)	(0.00210)	(0.00203)
Observations	18,850	18,848	18,849	18,865

Note: NFHS data. Bootstrap standard errors in parentheses. All specifications include individuals and household controls, state, cohort, state-religion fixed effects and state specific linear time trends.

Table A8: Structural and Self-reported Bargaining Power

	𝕀(Woman Participate in Final Decisions on)			
	Her Health	Large Household Purchases	Small Household Purchases	Visits to Family and Relatives
Highly Empowered Woman (T)	0.00793	0.0152	0.0110	0.0137
	(0.00212)	(0.00220)	(0.00210)	(0.00203)
Observations	18,850	18,848	18,849	18,865

Note: NFHS data. Bootstrap standard errors in parentheses. All specifications include individuals and household controls, state, cohort, state-religion fixed effects and state specific linear time trends.

Table A9: Determinants of Mother's and the Father's Resource Share ($\widehat{\theta}^m$ and $\widehat{\theta}^f$)

	Mother's Resource Share ($\widehat{\eta}_m$)	Father's Resource Share ($\widehat{\eta}_f$)
𝟙(# Kids=1)	0.0444	0.0226
	(0.0250)	(0.0387)
𝟙(# Kids=2)	0.0155	0.00845
	(0.0217)	(0.0359)
𝟙(# Kids=3)	0.0206	0.00618
	(0.0221)	(0.0367)
Fraction of Female Children	0.0353	-0.0155
	(0.0172)	(0.0206)
Age Gap	0.0391	-0.0861
	(0.137)	(0.215)
Woman's Age	-0.490	-0.0896
	(1.944)	(2.276)
Age Gap^2	-0.288	-0.0888
	(0.419)	(0.619)
Woman's Age^2	1.039	-0.758
	(4.788)	(5.557)
Age Gap^3	0.836	-0.285
	(1.506)	(2.344)
Woman's Age^3	-0.565	1.094
	(3.792)	(4.321)
Children's Avg. Age	-0.278	0.429
	(0.210)	(0.274)
𝟙(Hindu, Buddhist, Jain, Sikh)	-0.0226	0.0482
	(0.0177)	(0.0227)
𝟙(Sch, Caste, Sch. Tribe, OBC)	0.0536	-0.0568
	(0.0184)	(0.0206)
Wealth Index	-0.0505	0.154
	(0.0448)	(0.0624)
𝟙(Highly Educated Woman)	0.0210	-0.00688
	(0.0179)	(0.0241)
𝟙(Highly Educated Man)	0.0148	0.00823
	(0.0146)	(0.0222)
𝟙(Rural)	-0.0389	-0.0177
	(0.0160)	(0.0229)
𝟙(North)	-0.0939	0.0671
	(0.0333)	(0.0384)
𝟙(East)	0.00953	-0.0115
	(0.0384)	(0.0420)
𝟙(North-East)	0.0461	-0.194
	(0.0453)	(0.0451)
𝟙(South)	0.00990	-0.0390
	(0.0336)	(0.0367)
𝟙HSAA Eligible	0.0314	-0.0101
	(0.0105)	(0.0133)
Constant	0.360	0.463
	(0.260)	(0.303)
N	15,166	15,166

Note: NSS data. Robust standard errors in parenthesis.