Inefficient Collective Households: Abuse and Consumption

Arthur Lewbel and Krishna Pendakur

Boston College and Simon Fraser University

2019

Introduction - Relevance of Collective Households

Collective household consumption models address questions like these:

What percentage of a married couple's expenditures are controlled by the husband?

What share of household resources go to children?

Are some or all household members living below a povery line?

How much money does a couple save on consumption goods by living together versus living apart?

How much income would a woman living alone require to attain the same standard of living that she'd have if she were married?

Unitary vs Collective Household Consumption Models

Traditional **Unitary model**:

Treat a household like a single utility maximizing (representative) consumer.

What do household's as consumers actually do? **Collective model**:

Joint decision making, bargaining.

Different household members control and consume different fractions (shares) of household resources.

Household members each have their own preferences.

Some goods are shared, partly public, jointly consumed (e.g., gasoline is shared when riding together in the car).

Poverty Measurement

Traditional (Unitary model) household poverty measurement:

Select a poverty line for an individual.

Scale it up by number of children and adults in the household (equivalence scales).

Implication: everyone in the household is assumed to have same utility/welfare level. All are below the scaled up poverty line, or none are.

Collective poverty measurement: Each member's welfare depends:

On each member's own utility function, each member's own **resource share**, and publicness or privateness of goods (the degree to which each good is jointly consumed).

Implication: Different household members have different utility/welfare levels. some could be above the poverty line and others below.

Example Collective Household Earlier Findings 1

Dunbar Lewbel and Pendakur (2013, 2018) **DLP**. Data from Malawi:

Husbands consume around 40% of household resources, regardless of number of children. Wives consume 30% going down to 20% as add more children.

Define a poverty line of $2.00 US per person per day for men and for women, and $1.20 US for a child (typical OECD equivalence scale of .6).

Percents below the poverty line:

World bank (and unitary model): around 89% of households.

DLP collective model: around 58% of men, 85% of women, and 97% of children.

Traditional household level poverty measurement misses the distribution, particularly child poverty.

Example Collective Household Earlier Findings 2

Anderson and Ray (2010) estimate that in India, 1.7 million woman over age 45 "missing:" die at younger than expected ages.
The number missing increases with age from 45 to 70.
A puzzle - poverty kills, but household poverty rates do not correlate with women's age.

Calvi (2019):
Estimates DLP collective model for India couples.
Finds women's resource shares decrease with age starting at 45.
Calculates separate men's and women's poverty rates by age.
Women's poverty rate has .96 correlation with Anderson and Ray's rate of missing women estimates by age.

Measuring poverty at the individual level virtually completely explains the missing women puzzle!

Other Implications of Collective Households

Not just poverty measurement. Other uses of the collective consumption model:

Resource shares measure **bargaining power** (e.g., Sen's empowering women, welfare of children).

Gains from living together - **economies of scale** to consumption.

Indifference scales.

An **indifference scale** (see BCL) is: income a member would need living alone to reach the same utility level over goods that he/she attains as a household member.

Unlike equivalence scales, indifference scales are identifiable from data.

Relevant for required level of life insurance, wrongful death compensation, marriage stability and marriage markets (e.g., compare resource share to outside option)

Introduction to: "Inefficient Collective Households: Abuse and Consumption"

Our Starting Collective Household Model Framework:

Household members each have utility functions.

Household is Pareto efficient - bargaining process is unspecified.

Goods can be partially shared. Extent of sharing/cooperation is not observed.

Data with or without price variation.

Objectives

Definition: Resource shares. The fraction of resources (budget) each household member consumes.

Objectives - what we estimate:

Resource shares allowing for certain inefficiencies (domestic violence).

Some of the costs associated with those inefficiences.

What's new?

Keep the modeling benefits of efficiency (decentralization, don't need an explicit model of the bargaining process, well developed empirical methodologies) while allowing for realistic inefficiency.

Provide an internally consistent explanation for consumption inefficiency.

Why Resource Shares? Why Inefficiency?

Why identify and estimate levels of resource shares?

- Useful as measures of within household bargaining power

- Needed for individual instead of household level poverty measures.

Example uses (from earlier papers that identify resource shares):

- Dunbar Lewbel and Pendakur (2013) find child poverty rates in Malawi much higher than parent's rates.

- Calvi (2016) finds older wives poverty rates in India much higher than household rates (explains "missing older women" mortality rates).

Why allow for inefficiencies?
- evidence such as money holders, domestic abuse and violence (in our Bangladesh data 42% of married women report at least verbal abuse).

Literature - Unitary Models

The Old days: Unitary models:

- One utility (social welfare) function for the whole household.
- Demographics, household size, appear as preference parameters.
- Quantities are scaled/transformed to account for shared consumption.

Engel (1895), Sydenstricker and King (1921), Rothbarth (1943), Prais and Houthakker (1955), Gorman (1956, 1980), Barten (1964), Pollak and Wales (1981).

Earlier Efficient Collective Household Models

Household decisions made by bargaining or social welfare functions.

Becker (1965, 1981), Chiappori (1988, 1992). Browning, Bourguignon, Chiappori, and Lechene (1994), Bourguignon and Chiappori (1994), Browning and Chiappori (1998), Vermeulen (2002), Chiappori, Fortin and Lacroix (2002), Blundell, Chiappori and Meghir (2005), Chiappori and Ekeland (2009).

Limitations of this earlier work:
- Every good is either purely public within the household, or purely private. No partial sharing.
- Children treated as a public good, they don't have utility.

Distribution Factors

Many results in the early efficient collective household model literature depend on distribution factors. **Distribution Factors:**

- are variables that affect household consumption allocations, but do not affect preferences.
- For efficient households, this means they can affect Pareto weights, resource shares, bargaining power, and/or altruism.
- Examples in the literature: home village, area sex ratios, age, divorce and inheritance laws, local supply of education.
- Since related to bargaining power, are relevant for marriage markets, household formation, stability.
- Can be policy variables for affecting intrahousehold bargaining power and resource shares.
- Effects of changes in distribution factors on resource shares are identified from household level demand data.

Earlier Collective models: Main theoretical results

1. If the household is Pareto efficient, don't need to explicitly model household bargaining process.
- That info is summarized in Pareto weights or in resource shares.

2. Even if one knew complete household demand functions, resource share and Pareto weight levels are not identified.

3. What is identified from household demand functions is how resource shares vary in response to changes in some covariates, like distribution factors.
- is sufficient for, "what if we change a policy covariate?"
- NOT sufficient for individual well being and poverty line questions.

Some newer Efficient Collective Household Models

Lewbel (2003), Lewbel and Pendakur (2008), Bargain and Donni (2009, 2012), Couprie, Peluso and Trannoy (2010), Lise and Seitz (2011), Browning, Chiappori and Lewbel (2013 **BCL**) Dunbar Lewbel and Pendakur (2013 **DLP**), Cherchye, De Rock and Vermeulen (2011), Cherchye, De Rock, Lewbel, and Vermeulen (2015), Chiappori (2016):

Advances include:
- Allow partial sharing of goods (not just purely private or purely public).
- Reducing data requirements for identification and estimation.
- Identification (or partial identification) of levels of resource shares.
- Inclusion of children's utility functions - directly address child poverty.
- Indifference scales (replaces or complements equivalence scales - not relevant for the present paper)

Inefficiency in Collective Households

Empirical Evidence mixed. Above papers (and, e.g., Bobonis 2009) support consumption efficiency, others (Udry 1996, and Dercon and Krishnan 2003) reject.

Information asymmetry (e.g., income hiding): Vogley and Pahl (1994), Ashraf (2009), Castilla and Walker (2013).

Static efficiency, intertemporal inefficiency: Mazzocco (2004), Lise and Yamada (2014), and Chiappori and Mazzocco (2017).

Nesting specific cooperative and noncooperative behaviors: Gutierrez (2018).

Ramos (2016): Like us, violence chosen by men which reduces efficiency. A simpler model where violence is exogenous, doesn't directly affect men's utility, only reduces production of a single home produced good.

Basic BCL

2 member BCL. U_j member j utility function, ω_j member j pareto weight, p price vector, g quantity vector, and $p'g = y$ budget.

$$\max_{g_1, g_2} U_1(g_1)\,\omega_1(p,y) + U_2(g_2)\,\omega_2(p,y)$$

$$\text{such that} \quad p'g = y, \quad g = A(g_1 + g_2)$$

A is consumption technology (sharing) matrix. Example: suppose g^k is gallons of gasoline, and couple ride together in car 1/3 of the time. It's as if member j consumes g_j^k gallons, where $g^k = (2/3)(g_1^k + g_2^k)$. Shadow price of gas is $(2/3)p^k$.

Decentralization: Household is equivalent to each member j doing:

$$\max_{g_j} U_j(g_j) \qquad \text{such that} \quad p'Ag_j = \eta_j(p,y)\,y$$

Solution for each is

$$g_j = h_j\Big(\, p'A, \eta_j(p,y)\,y\Big)$$

DLP simplifies BCL

Estimating BCL has demanding data requirements (extensive price variation, uses data on singles to get information about demand functions h_j for all goods, doesn't identify children's utility).

DLP takes the above BCL model and makes it much simpler by:
1. dropping prices (single cross section Engel curve data).
2. focusing just on estimating resource shares η_j, not A.
3. makes assumptions that permit identification without price variation.
4. makes assumptions that permit identification without singles data (so can consider children's utility and estimate their resource shares).

The main identifying assumptions are

a. resource share functions don't depend on y,
b. semiparametric restrictions on utility functional forms,
c. use of private assignable goods
d. preference similarity restrictions on the private assignable goods.

Private Assignable Goods

Key component of DLP: Private assignable goods.

A good is private if it cannot be jointly consumed at all.

Example: Food. Unlike sharing gasoline in the car that transports both of us, if I share an apple, every bite you eat is a bite less for me.

A good is assignable if the researcher knows which household member consumed the good.

Examples: Many data sets collect separate men's, women's, and children's clothes consumption.

Key DLP insight (ignoring covariates other than y):
Let q_j be quantity of a private assignable good consumed by member j.
Household q_j private assignable demand function $q_j = g_j\left(\eta_j y\right)$.
If we can identify the Engel curve demand function g_j, we can invert and solve for resource share η_j.

Consumption Technology and Efficiency

Think about A some more. BCL identifies and estimates A, DLP has A in it but doesn't identify A

The smaller the diagonal elements of A, the more sharing/cooperation. The more sharing/cooperation, the higher each $U_j(g_j)$ can be - efficiency.

Relatively inefficient households share less. But conditional on a given sharing technology A, each household is efficient (satisfies decentralization, previous theory/methods apply).

So the BCL and DLP models can handle households having different efficiency (cooperation/sharing) levels.

Let f be an observed covariate that affects the efficiency level. Binary for now: $f = 0$ for more sharing/cooperation, $f = 1$ for less. Household consumption technology is A_f.

True inefficiency is when a household can *choose* the efficiency level f, and chooses $f = 1$ instead of $f = 0$.

Modeling Inefficiency - A "Cooperation Factor"

Example: suppose $f = 1$ indicates the husband verbally or physically abuses his wife. Effects of f are:

a. $f = 1$ means reduced cooperation, have A_1 instead of better A_0.
b. f could change Pareto weights (and therefore change resource shares).
 - Both a. and b. affect consumption choices of each household member.
c. Each member j may get direct utility or disutility $u_j(f)$ from abuse and/or from cooperation, separate from consumption.
 - All these effects affect each member's attained utility level.

We call f a "cooperation factor." It's a distribution factor, but it also affects sharing/cooperation A. Earlier collective models couldn't have the concept, because they didn't allow variation in cooperation - all goods were purely public or purely private.

Extending BCL to include a cooperation factor

Let the household have J members. Model is now (notice where f is):

$$\max_{g_1,\ldots g_J} \sum_{j=1}^{J} \left[U_j \left(g_1 \right) + u_j \left(f \right) \right] \omega_j \left(p, y, f \right)$$

$$\text{such that} \quad p'g = y, \quad g = A_f \left(g_1 + \ldots + g_J \right)$$

Decentralization still holds, each member demands are

$$g_j = h_j \left(p' A_f , \eta_j \left(p, y, f \right) y \right)$$

Note f only affects $U_j \left(g_j \right)$ by affecting g_j, so f only affects demands by affecting the shadow prices and shadow budget for member j.

Household demand functions are

$$g = A_f \sum_{j=1}^{J} h_j \left(p' A_f , \eta_j \left(p, y, f \right) y \right)$$

Choosing inefficiency - choosing the cooperation factor

$$g_j = h_j\left(p'A_f, \eta_j(p,y,f)\,y\right)$$

$$g = A_f \sum_{j=1}^{J} h_j\left(p'A_f, \eta_j(p,y,f)\,y\right)$$

Assume f is chosen by household member 1. Chooses f by

$$f = \arg\max U_1\left(h_1\left(p'A_f, \eta_1(p,y,f)\,y\right)\right) + u_1(f).$$

Member 1 commits abuse, choosing the inefficient $f = 1$, if his gain in utility $u_1(f)$ and/or his gain in resource share $\eta_1(p,y,f)$, more than offsets his losses from inefficiency (a larger A_f).

Can estimate dollar cost of inefficiency, the cost of the change in A_f, as $p'A_0^{-1}g - p'A_1^{-1}g$.

We do not identify/estimate the "costs" of changing $u_j(f)$ - losses like sadness.

The Full Model

$$\max_{g_1, q_1, \dots g_J, q_J} \sum_{j=1}^{J} \left[U_j \left(q_j, g_j, z, \widetilde{\varepsilon} \right) + u_j \left(f, v, z, p, \pi, y, \widetilde{e} \right) \right] \omega_j \left(f, z, p, \pi, y \right)$$

$$\text{such that} \quad p' g + \sum_{j=1}^{J} \pi_J q_j = y \quad \text{and} \quad g = A_f \sum_{j=1}^{J} g_j$$

q_j private assignable good consumed only by member j.
g_j vector of quantities of other goods consumed by member j.
g vector of purchased quantities of other goods.
p, π, vectors of prices of g and q.
z other covariates, observed taste and power shifters.
$\widetilde{\varepsilon}$, $\widetilde{e}$, unobserved random utility (taste heterogeneity) parameters,
y total budget.
f cooperation factor, e.g., indicator of domestic violence/abuse.
v observed variables that affects direct utility/disutility of f. Example: thickness of the household's walls (get disutility from neighbors more readily hearing abuse).

Simplify the Model

Above model complicated to specify and estimate. Requires extensive relative price variation. Difficult to identify for children.
To simplify, we impose restrictions similar to DLP. Main restrictions are:

1. A private assignable good q_j for each household member j. Will only estimate the demand functions of these goods.
2. A semiparametric utility function specification (includes Deaton and Muellbauer 1980 Almost Ideal Demand System as a special case). Makes budget shares linear in $\ln y$.
3. The SAP (Similarity Across People) assumption used by DLP. Makes budget share Engel curve coefficient of $\ln y$ not vary by j.
4. Resource shares η_j do not depend on y. DLP list supporting evidence.
5. Resource shares η_j do not depend on random utility parameters (is stronger than not having them in the Pareto weights).
6. All households face the same market prices (cross section Engel Curve type data).
7. Like DLP, some joint restrictions on possible A_f matrices and p vectors.

Engel Curves

Define budget shares $w_j = p_j q_j / y$. After imposing above restrictions, we obtain Engel curve budget shares for private assignable goods of the form

$$w_j = \eta_j(f,z)\left[\gamma_j(z) - \beta(z)\left(\ln y + \ln \eta_j(f,z) + \ln \delta(f,z)\right) + \varepsilon_j\right]$$

Where:

z are covariates affecting tastes and power: demographics like age, education, etc; dowry; wealth.

ε_j is a function of some covariates and random utility parameters $\widetilde{\varepsilon}$. Will get $E(\varepsilon_j \mid z, r) = 0$, where r is a vector of instruments discussed below.

$\eta_j(f,z)$ are **resource shares**.

$\ln \delta(f,z)$ is the **percent change in effective budget due to inefficiency** of level f relative to $f = 0$.

Endogeneity of the Cooperation Factor

Issue: Endogeneity of f in the demand functions.

Each member j's utility is $U_j\left(q_j, g_j, z, \widetilde{\varepsilon}\right) + u_j\left(f, v, z, p, \pi, y, \widetilde{e}\right)$

Member $j = 1$ chooses f to maximize $U_1 + u_1$. Both $\widetilde{\varepsilon}$ and $\widetilde{e}$ affect choice of f.

Conditional on f, the household is efficient - equivalent to choosing q and g to maximize weighted sum of above $U_j + u_j$ utilities. This is same as just maximizing weighted sum of U_j utilities

v and $\widetilde{e}$ only appear in the u_j functions, not in the U_j functions that determine q and g. So v and $\widetilde{e}$ only affect q and g by affecting f. This makes v a valid instrument for f in the demand equations for q and g.

We also let y be endogenous or mismeasured. Assuming standard time separability of consumption/savings from within period budget allocation, income $\widetilde{r}$ is a valid instruments for y.

Semiparametric Identification

For $j = 1, ..., J$, have private assignable demands:

$$w_j = \eta_j(f,z)\left[\gamma_j(z) - \beta(z)\left(\ln y + \ln \eta_j(f,z) + \ln \delta(f,z)\right) + \varepsilon_j\right]$$

$E(\varepsilon_j \mid z, r) = 0$, where $r = (v, \widetilde{r})$ are walls and income.

THEOREM: Functions $\eta_j(f,z)$, $\gamma_j(z)$, $\beta(z)$, $\ln \delta(f,z)$ are nonparametrically identified.

Proof sketch: solve for ε_j, start from $E(\varepsilon_j \mid z, v, \widetilde{r}) = 0$.
Variation in $E(\ln y \mid z, v, \widetilde{r})$ due to $\widetilde{r}$ identifies product $\beta(z) E\left[\eta_j(f,z) \mid z, v\right]$.
Resource shares sum to one separately identifies $\beta(z)$.
Dependence of f on v identifies $\eta_j(f,z)$ from $E\left[\eta_j(f,z) \mid z, v\right]$
Above identifies $\gamma_j(z) - \beta(z) E[\ln \delta(f,z) \mid z, v]$
Normalization $\ln \delta(0,z) = 0$ separately identifies $\gamma_j(z)$.
Dependence of f on v identifies $\delta(f,z)$ from $E[\ln \delta(f,z) \mid z, v]$.

Specification

Due to data limitations and model complexity, we estimate parametrically, parameter vector θ.

Let N_{jh} be the number of members in household h of type $j = m$, f, or c: adult males, adult females, and children. One utility function for each j, scaled by N_{jh}. We have private assignables for each type j.

Do GMM with moments $E\left(\varepsilon_{jh}\phi\left(r_h, z_h\right)\right) = 0$ where

$$\varepsilon_{jh} = \frac{w_{jh}}{\eta_j(f_h, z_h, \theta)} - \gamma_j\left(z_h, \theta\right) - \beta\left(z_h, \theta\right)\Big(\ln y_h - \ln N_{jh} + \ln \eta_j\left(f_h, z_h, \theta\right) + \ln \delta\left(f_h, z_h, \theta\right)\Big)$$

$\eta_j\left(f_h, z_h, \theta\right)$ are resource shares.

$\delta\left(f_h, z_h, \theta\right)$ is the percentage of y that is lost due to inefficiency of level f relative to $f = 0$.

Data: Bangladesh Integrated Household Survey 2015

Rural household survey by International Food Policy Research Institute and the World Bank.

Includes:
1) Household level data on consumption y.
2) Person-level data on food consumption (q_j).
3) Self reported data on the exposure of the primary female spouse to either physical and verbal abuse, f.

Restrict sample to commonest household compositions (1 or 2 men, 1 or 2 women, up to four children) and non-zeros in q_j's. Final sample size $H = 2866$ households.

1 day recall diary of individual food consumption. We use this to disaggregate 7 day recall data on household level food consumption in 7 categories (Cereals, Pulses, Oils; Vegetables; Fruits; Proteins; Drinks and Others).

Expenditures Construction

Multiply food consumption by constructed village (i.e. upazila sub-district) level average prices (village level expenditures divided by village level quantities), scale up to annual food expenditures by household member-type: S_{mh}, S_{fh}, and S_{ch}, e.g., S_{ch} is annual expenditures on food consumed by children (under 14) in each household h.

Total expenditures y_h constructed by one month expenditure recall data on purchases and home-produced values of: rent, food, clothing, footwear, bedding, nonrent housing expense, medical expenses, education, remittances, jakat/daan/sodka, kurbani/milad/other (charity and sacrifices required by Shariah law), entertainment, fines and legal expenses, utensils, furniture, personal items, lights, fuel and lighting energy, personal care, cleaning, transport and telecommunication, use-value from assets, and other.

Budget shares $w_{jh} = S_{jh} / y_h$. Then normalize y_h to make $\ln y_h$ be mean zero.

Covariates

Covariates z_h:

1) the average age in decades of adult males;
2) the average age in decades of adult females;
3) the average age in decades of children;
4) the average education in years of adult males;
5) the average education in years of adult females;
6) the fraction of children that are girls (minus .5);
7) the log of marital wealth (dowry)
8) Log of total household wealth
Plus dummies for each (but one) of 10 household composition types (each number of men, women, children)
(all non-dummies are demeaned),

Reported Abuse

Survey asks head female of the household:

Has any of the following happened to you in the past year?
1) Your husband threatened you with divorce?
2) Your husband, another family member, or household resident verbally abused you?
3) Your husband, another family member, or household resident physically abused you?

Let $f = 1$ if any answer is yes, else $f = 0$. **42% of households have** $f = 1$.

Literature (e.g., Rao 1997 or Krishnan et al 2010) finds three primary correlates of domestic violence in nearby India: **Alcohol consumption; insufficient dowry; and local social acceptance of violence**

We do not have alcohol data, we do include dowry in our covariates.

Structural Instruments - walls, village average f, income

Instruments r for endogeneity of f and y:

Dummy for **walls of Mud and/or Bamboo**. (get disutility from neighbors more readily hearing abuse). Most other walls are Concrete, Tin, and/or Wood.

Village level average f (indicator of local social acceptance of violence). Valid? Can show in theory with food demand errors independent across households. Will check sensitivity and overidentification tests.

Log of total household **annual income** (and squared)

Are **not** claiming these are randomly assigned - just independent of structural model errors, after conditioning on other covariates (like wealth).

Data (excluding z)

.

Table 1b: Summary Statistics (shorter)				
	Mean	Std Dev	Min	Max
ln(total consumption), ln y	0.105	0.554	-1.676	2.769
men's food, w_m	0.161	0.070	0.014	0.514
women's food, w_f	0.145	0.065	0.013	0.534
children's food, w_c	0.131	0.080	0.001	0.488
f, abuse	0.420	0.494	0.000	1.000
f, village average	0.420	0.265	0.000	1.000
Building Mat: Mud, Bamboo	0.156	0.363	0.000	1.000
ln(income)	0.083	1.440	-8.378	3.157

Mean w_j: **food shares each over 10%** (unlike assignable clothing).
Mean f: 42% report abuse.

First stage Regressions

Table 2 shows linear regressions of endogenous f and $\ln y$ on instruments (and on z, not shown). This would be the first stage of linear 2SLS if we did that instead of nonlinear GMM.

Table 2: " First Stage", shorter

		f abuse			$\ln y$, ln(budget)	
		Est	*Std Err*	t	Est	*Std Err*
instruments	ln(income)	0.004	*0.011*	0.32	0.132	*0.014*
	ln(income)^2	-0.001	*0.002*	-0.79	0.023	*0.002*
	Building Material	-0.035	*0.019*	-1.86	-0.098	*0.025*
	village-average f	0.666	*0.033*	20.02	-0.135	*0.048*
R-squared		0.160			0.356	
F-stat		105			34	

f is hard to predict ($R^2 = .16$), instrument F-stats significant.
Building material becomes stronger if drop village-level abuse.

Model Specification

$$\varepsilon_{jh} = \frac{w_{jh}}{\eta_j(f_h, z_h, \theta)} - \gamma_j(z_h, \theta) - \beta(z_h, \theta)\left(\ln y_h - \ln N_{jh} + \ln \eta_j(f_h, z_h, \theta) + \ln \delta(f_h, z_h, \theta)\right)$$

$$\eta_j(f_h, z_h, \theta) = k_{j0} + k_j' z_h + c_j f_h,$$

$$\gamma_j(z_h, \theta) = l_0 + l_j' z_h,$$

$$\ln \delta(f_h, z_h, \theta) = \left(a_0 + a_1' z_h\right) f_h,$$

$$\beta(z_h, \theta) = b_0$$

θ is $a_0, a_1, b_0, c_j, k_{j0}, k_j, l_0$, and l_j' for $j \in \{m, f, c\}$.
Due to multicollinearity, baseline takes $a_1 = 0$. β constant similar to DLP.
Free normalization: $f_h = 0$ makes $\ln \delta = 0$.

Focus on resource shares η_j, on efficiency cost of abuse δ, on c_j (response of η_j to abuse f_h), and on a_0 (reponse of δ to abuse f_h).

GMM moments

GMM moments are

$$E\left(\varepsilon_{jh}\phi\left(r_h, z_h\right)\right) = 0$$

with

$$\phi\left(r_h, z_h\right) = \left(1, z_h, r_h\right) \times \left(1, z_h, r_h\right) \times \left(1, r_h\right)$$

where $\times$ indicates element-wise multiplication, deleting redundant elements.

Resulting instrument vector $\phi\left(r_h, z_h\right)$ has 601 elements (without cubic terms, 185 elements).

Using 601 (or 185) moments to estimate 111 parameters (baseline θ dimension).

Baseline Summary estimates

Table 3: GMM Estimates, Selected Coefficients

function	person	variable	Baseline Estimate	*Std Err*
$\ln \delta$	all	constant, a_0	-0.0534	*0.0179*
η	men	constant, k_{m0}	0.3476	*0.0053*
		f , c_m	0.0141	*0.0020*
	women	constant, k_{f0}	0.3047	*0.0048*
		f , c_f	-0.0086	*0.0018*
$\triangle$ lnmoney metric	men	m1_f1_c2	-0.0047	*0.0064*
	women	m1_f1_c2	-0.0240	*0.0056*
	children	m1_f1_c2	-0.0232	*0.0060*
J-statistic: value [df] p-value			1741 [1692]	0.1988

Summary: valid instruments, inefficiency .05, shares 0 to 1.

J-tests fail to reject instrument validity.

$\ln \delta = -.05$. Inefficiency equivalent to being 5% poorer.

Model specification doesn't impose resource shares $\widehat{\eta}$ between zero and one, but the estimates do lie in that range for every observed value of all covariates.

Summary: shares .35, .30, .35, f effect 1.4, -.09, -.06

Let z_0 be reference household (1 man, 1 women, 2 children, other z elements at mean values)

$\widehat{\eta}_j(0, z_0)$ is resource share when efficient.
$c_j = \widehat{\eta}_j(1, z_0) - \widehat{\eta}_j(0, z_0)$ is change to inefficient.

$\widehat{\eta}_m(0, z_0) = 34.8\%$, $\widehat{\eta}_f(0, z_0) = 30.5\%$, $\widehat{\eta}_c(0, z_0) = 34.7\%$

$c_m = 1.4\%$, $c_f = -0.9\%$, $c_c = -0.5\%$

Each $\widehat{\eta}_j$ and c_j is statistically significant.

Effective shadow budget effects: -.005, .024, .023

Meaning of the $\Delta \ln$ money metric:

With $f = 0$, male resource share η_m is 34.8%, male shadow budget is $0.348y$.

With $f = 1$, η_m goes up by 1.4% to 36.2%, and a loss due to inefficiency equivalent to household y decreasing by $\ln \delta = 5.3\%$.

So with $f = 1$ male shadow budget is $0.362\,(1 - 0.053)\,y = 0.343y$.

Violent men get a larger share of a smaller pie.

Overall effect of $f = 1$ on male shadow budget is near zero: $(0.348 - 0.343)\,y$ (statistically insignificant).

$f = 1$ decreases female and children's shadow budgets by 2.4% and 2.3%, respectively (significant).

Why Choose Inefficiency?

Recall effects of violence $f = 1$ vs $f = 0$ in the model:

a. $f = 1$ means reduced cooperation, have A_1 instead of better A_0.
- equivalent to reducing y by 5.3%.

b. $f = 1$ changes Pareto weights (and therefore change resource shares).
- η_m up 1.4%, η_f down 0.9%, and η_c down 0.5% and .

c. Unknown changes $u_j(0)$ to $u_j(1)$. Direct utility or disutility $u_j(f)$ from abuse and/or from cooperation

Net effect of a. and b. to member 1 was essentially zero. So his motive for choosing $f = 1$ must be in u_1.

Either abusers are deriving utility from abuse, or gets utility from not expending the effort needed for efficient cooperation in joint consumption of goods.

Robustness checks:

Dropping the village average abuse instrument
Dropping cubic instruments
Including household size or all of z in $\ln \delta$
Alternative food share constructions
Including zero share observations
Only using nuclear households

All show the same general patterns and magnitudes. Some show higher efficiency cost δ (up to 7% rather than 5%) for the reference household.

Conclusions

We show identification of resource shares and (some) costs of inefficiency

- keeping the modeling benefits of efficient collectives

- allowing for realistic inefficiency

- providing an internally consistent explanation for consumption inefficiency