Unpacking Household Engel Curves

Philippe De Vreyer
University of Paris-Dauphine, PSL Research University, CNRS, IRD, LEDa, DIAL

Sylvie Lambert
Paris School of Economics, INRAE

Martin Ravallion[1]
Georgetown University and NBER

Abstract: We model sub-household budget allocations using an unusual data set for Senegal. Aggregating to the household level reveals confounding factors in standard Engel curves, including intra-household inequality. Except for education spending, our results are consistent with the separable structures found in two-stage bargaining and collective models of the household. However, we find large discrepancies between the standard household Engel-curve estimates and consistently aggregated sub-household estimates, though in differing degrees and directions depending on the type of commodity. The main source of this discrepancy is a household effect on sub-household spending behavior, which is partially offset by differences in intra-household inequality.

Keywords: Engel curves; intrahousehold inequality; heterogeneity; Senegal
JEL: D12, D13, O12

---

[1] For helpful comments and discussions the authors thank Tanguy Bernard, Isabelle Chort, Vassilis Hajivassiliou, Franco Peracchi and participants at the workshop on “Family and Development” at the University of Bordeaux. Sylvie Lambert thanks the support of the EUR grant ANR-17-EURE-0001.

1. Introduction

Ernst Engel (1857) famously studied household budgets for 200 working class Belgian families, and found that the share devoted to food tends to decline with total household spending—a property that came to be known as Engel's Law. Since then, "Engel curves" for budget shares have been widely used and much studied across the world, with near universal confirmation for Engel's Law.[2] Engel curves have found a wide range of applications, including in the assessment of policies related to agriculture, taxation, trade, industrial organization, housing, and in the measurement of poverty and inequality.

While there have been methodological and computational advances in the specification and estimation of Engel curves, a common and persistent feature has been the reliance on household aggregates that Engel pioneered, along with a degree of imposed homogeneity in the Engel curves, allowing only limited variability in the parameters across and within households. Yet we know that many spending decisions are made by individuals rather than collectively.

This paper studies some neglected but potentially confounding sources of heterogeneity in standard household Engel curves. Three sources of heterogeneity are postulated. First, there can be latent household effects on individual demand behavior. Members of a given household are not autonomous individuals who happen to be living together, but rather they come together selectively, and then interact and influence each other's behavior through the process of consuming (and often working) together. While we may reject the unitary model, it can be expected that there are aspects of the household, and shared local environment, that can have a powerful influence on individual choices. This can happen via individual preferences, which are to some extent formed within a household. Or it can stem from household- or location-specific aspects of the constraints on exercising personal preferences. The role of household influences on individual choices is widely recognized in the literature on consumer behavior outside economics.[3] For the present purpose, this can be thought of as a household effect on individual

[2] At the centenary of the publication of Engel's major work on household budgets, Houthakker (1957) provided estimates across 30 countries largely confirming Engel's Law, though noting that the income elasticity of demand for food tended to be higher in poorer countries. Houthakker reported one estimate for India (from the city of Bhopal) of an elasticity of about unity, which was the only exception he found to Engel's Law.

[3] On the influence of the family and other group memberships on individual preferences and consumer behavior see the surveys by Solomon et al. (2006, Chapters 10 and 11) and Arnould et al. (2005).

Engel curves. That effect is not in general identifiable from standard household-level cross-sectional surveys.

Second, there are differences in individual demand parameters within households. This has received some attention. It is known that Engel's Law may cease to hold at the household level when income gains are assigned to people with different consumption patterns and different preferences over how the extra money should be spent. Attanasio and Lechene (2010, 2014) argue that this is a plausible explanation for why the household food share did not fall due to the income gains to targeted households participating in Mexico's famous PROGRESA program, which provided cash transfers, paid to women. This arrangement for payment appears to have shifted the household Engel curve for food among PROGRESA participants.

Third, there is heterogeneity in the extent of inequality within households. The existence of intra-household inequality is known to be a source of bias in the measurement of poverty and inequality.[4] It is less well known that intra-household inequality can also bias estimates of empirical consumer demand functions, as invariably estimated from household aggregate data. Yet for many goods, and (hence) expenditures, there is a typically an unobserved individual assignment within the household, that may be a source of intra-household inequality, reflecting different reservation utilities outside the household. Furthermore, intra-household inequality can interact with individual parameter heterogeneity in influencing household demands, whereby greater intra-household inequality magnifies the effect of differences in preferences.

There is no obvious reason why these sources of heterogeneity are statistically ignorable when estimating household Engel curves. Latent heterogeneity in demand behavior may well be correlated with household total spending or income. For example, there may well be latent differences in human capital that influence demands, such as when mother's education influences the priority given to nutrition. In turn, human capital is likely to be positively correlated with household consumption or income. Another example relates to the household's social status in the local community of residence. Perceptions of the obligations that come with higher social status may well influence spending patterns; for example, one may feel the need to show off with a TV, or have enough food ready in case someone shows up. Such "status-

[4] Contributions include Haddad and Kanbur (1990), Findlay and Wright (1996), Lise and Seitz (2011), Bargain et al. (2018), De Vreyer and Lambert (2020).

seeking" behavior can be expected at a given level of total spending, but also to be correlated with that spending.

A similar point can be made with regard to intra-household inequality. On a priori grounds, it would seem hard to defend the assumption that such latent inequality is statistically ignorable. One might expect that the extent of intra-household inequality varies with mean consumption or income (and possibly other covariates of household demands). One way that such a correlation can emerge is the existence of non-convexities in utility functions (or household production functions) at low levels of consumption. These can readily generate a negative correlation between the extent of inequality within households and household income in that a form of triage emerges in poor households, whereby resources are allocated to assure that at least one member is fed and clothed adequately. Once household income rises enough, more of that extra income can be shared with those members who had previously been neglected. However, one can make theoretical arguments suggesting that the correlation could go either way. Kanbur and Haddad (1994) show that a (positive or negative) correlation between intra-household inequality and household income can emerge from a model of Nash-type intra-household bargaining in the allocation of resources.

If intra-household inequality falls as household income rises, it may even be that Engel's Law is an artifact of intra-household inequality. It is not difficult to imagine the possibility that all individuals have constant budget shares, but that Engel's Law still emerges in the aggregate household data due to the combined effect of preference heterogeneity and intra-household inequality. Appendix A1 provides an example in which individual Cobb-Douglas preferences can still generate Engel's Law at household level given how differing preferences interact with differing threat points in intra-household bargaining.

The upshot of these observations is that household Engel curves may well be biased by these confounders. Testing this conjecture requires that we unpack the household Engel curves to reveal the potential confounders. That is the aim of this paper. However, we also point to sources of bias in individual Engel curves, even when data are available to permit their estimation. Our primary interest here is in seeing whether the two methods of estimating Engel curves—one using standard household aggregates and another using (rather rare) individualized data—are reasonably consistent with each other, and explaining any revealed inconsistencies.

Past research on Engel curves has been constrained by the lack of data on spending within the household. The focus has thus been on how one might infer aspects of distribution within the household by looking solely at household aggregates. For example, one can gain insights into distribution within the household by studying spending on "adult goods" versus "child goods." A strand of the literature using non-unitary (bargaining and collective) models of the household has provided insights into the measurement of intra-household inequality and poverty, exploiting sharing rules that can be identified under certain assumptions, or using data on the allocation of certain assigned goods within households.[5]

The paper proposes an estimable model of individual Engel curves, with both "internal" (own-spending) and "external" (household spending) covariates. We then aggregate the individual Engel curves to derive a more familiar household model. The aggregation process provides new insights into the structure of the error term of the household Engel curve so as to explore the sources of bias described above. For the purpose of this paper we focus mainly on the longstanding Working-Leser (WL) specification for budget shares linear in log total spending, though we test robustness to allowing budget shares to be nonlinear in log total spending per capita.[6] We assume that the WL model holds at the individual level and then aggregate up to the household level. In doing so we derive a WL specification that can be implemented with standard household-level data sets, but we show that the error term of that model contains a term that reflects inequality within the household. In the special case of common Engel parameters within the household, a Theil T index for intra-household inequality is the relevant omitted variable when one implements the individualized WL model with only standard household level data.[7] Predictions of the aggregate demand responses to household income changes are thus conditional on how changes in spending are distributed within households. The possibility that the level of intra-household inequality varies with total spending thus emerges as a confounding factor in estimating household Engel curves.

[5] Contributions include Kanbur and Haddad (1994), Chiappori and Meghir (2015), Brown et al. (2018), Bargain et al. (2018), Dunbar et al. (2013) and Chavas et al. (2018). The formulation of a collective model in Chiappori (1988) has been influential. On the conditions for identification see Bourguignon et al. (2009).

[6] Working (1943) advocated this functional form, which Leser (1963) found to fit better than some options. On the consumer expenditure function (and hence indirect utility function) implied by such Engel curves see Muellbauer (1975). For an overview of this model and alternatives see Phlips (1983) and Lewbel (2008).

[7] Theil (1967) had proposed a class of functional forms for an inequality index.

In introducing the external effect we relax a common (indeed, universal) assumption in the recent literature on intra-household resource allocation in non-unitary models, which has postulated a two-stage structure, whereby the household first chooses how to allocate total spending across its members, leaving them free (in the second stage) to allocate their assigned share of total spending across (private) commodities (as in, for example, Dunbar et al., 2013). In other words, it is commonly assumed that an individual's budget shares are uncorrelated with household aggregate spending at a given level of the individual's personal allocation of total household spending. Is that a valid assumption? While we may choose to reject the (longstanding) unitary model, we can also recognize that the household is a more complex social institution than the separable two-stage model suggests. Living in richer household may well have an external effect on how a person spends a given individual allocation of total spending, such as through peer effects within the household or inter-temporal consumption behavior.

The paper exploits an unusual data set for Senegal to estimate sub-household Engel curves without imposing a separable (two-stage) structure and to explore the implications for standard household demand analysis. The data allow us to assign total spending within the household. We then focus on how different types of people allocate their spending across commodities. The popular assumption of two-stage separability is generally accepted empirically. However, on aggregating our individualized Engel curves up to the household level we find results that differ markedly from those obtained by only using the aggregate data to estimate the household Engel curves. Substantial and partly predictable heterogeneity in Engel curves is revealed. We find that the conditions for consistent estimation of household Engel curves using standard data are convincingly rejected. The components of bias are partly offsetting, with the bias due to correlated differences in the extent of intra-household inequality working in the opposite direction to that due latent household effects on spending behavior.

2. Inside the aggregate household Engel curve

We assume that individuals can be assigned to one (and only one) of a number of possible types (male head, first wife, second wife,…), with fewer types than people. A key identifying assumption is that Engel parameters vary between types but not within. (A homogeneity assumption of this sort is common for identification purposes in the literature on intra-household resource allocation.[8]) The unique assignment of persons to types is given by a pre-determined discrete function $t(i,j)$, which returns the designated type for each person i=1,..,n_j in household j=1,..,m.

The "unpacked" Engel curve for consumption by person i (of type $t(i,j)$) in household j for an assigned private good of type k (=1,..,K) is assumed to take the form:

$$S_{ij}^k = \alpha_t^k + \beta_t^k \ln \bar{Y}_j + \gamma_t^k \ln Y_{ij} + \pi_t^k X_{ij} + \eta_j^k + \varepsilon_{ij}^k \quad (i=1,..,n_j;\ j=1,..,m;\ t = t(i,j)) \quad (1)$$

where S_{ij}^k is the share of person i's total expenditure Y_{ij} devoted to good k and $\bar{Y}_j$ is total household expenditure per capita, X_{ij} is a vector of control variables (including log household size), η_j^k is a household effect. The data are assumed to include multiple observations for a given type, such that (1) is estimable. Notice that (1) allows the Engel parameters $(\alpha_{t(i,j)}^k, \beta_{t(i,j)}^k, \gamma_{t(i,j)}^k, \pi_{t(i,j)}^k)$ to vary within households.[9]

A key feature of (1) is that we allow both an "own-spending" coefficient ($\gamma_{t(i,j)}^k \neq 0$) and an "external effect" of household total expenditure on individual Engel curves conditional on "own-spending", i.e., we allow $\beta_{t(i,j)}^k \neq 0$. The external effect merits explanation. As noted in the introduction, some economic models of the household postulate a separable, two-stage, process. At the first stage, the household decides an allocation of total spending across its members (and how much to allocate to public goods within the household), while at the second stage each individual chooses how her allocation of total spending is to be assigned across private goods and services. The Appendix A1 provides a simple example of such a model. Another example is found in the set of collective models of the household.[10] In such models, household total spending only matters to individual budget allocations via the agreed individual allocation of that spending. However, an external effect can be postulated. This can stem from one or more of a number of factors, including: (i) future consumption-sharing possibilities in richer households that are not already reflected in an individual's current spending; (ii) peer effects within the household, such that (say) a "poor" individual in a "rich" household tends to adjust her budget shares to conform; or (iii) the effects of measurement errors in individual

[8] See, for example, the empirical implementation of the collective model in Dunbar et al. (2013).
[9] If we impose constant parameters (across person types) as well as constant sharing rules (across households) then we can identify the sharing rules (Dunbar et al., 2013).
[10] Building on earlier work of Chiappori (1988) and Apps and Rees (1988).

expenditures that are at least partially corrected by including household total spending. We treat the existence of an external effect as an empirical issue.

We recognize that it may be considered an overly strong assumption that ε_{ij}^{k} in (1) is an innovation error term. The most worrying threat is probably the possibility of a non-zero correlation between ε_{ij}^{k} and lnY_{ij} . For example, the head of the household may try to incentivize individuals to spend in some preferred way using the allocation of the total budget. We can characterize this by postulating that $\varepsilon_{ij}^{k} = \varphi^{k}\ln Y_{ij} + \tilde{\varepsilon}_{ij}^{k}$ with $\varphi^{k} \neq 0$ and $\tilde{\varepsilon}_{ij}^{k}$ as the innovation error term. Then the “Engel coefficient” estimated from the micro data is interpretable as $\gamma_{t}^{k} + \varphi^{k}$. We do not attempt to remove this source of bias in the sub-household Engel curves, as our main interest lies in the implications for bias in the household level Engel curves.[11]

So, we cannot rule out the possibility that OLS estimates of individual Engel curves based on (1) will be biased even if that is not the case with the household level Engel curves. However (as we will see), even if $Cov(\varepsilon_{ij}^{k}, \ln Y_{ij}) = 0$, new sources of bias in the household Engel curves will emerge when we aggregate up.

It saves notation to interpret the $\beta_{t(i,j)}^{k}$ and $\gamma_{t(i,j)}^{k}$ parameters in (1) as incorporating any correlations between the household effects and the two (log) expenditure variables. This is consistent with the way household-level Engel curves are estimated in which the household-level error term is assumed to be orthogonal to household total expenditures (and other covariates). Our estimation method will address the concern that the η_{j}^{k}’s are potentially correlated with the regressors, and will also allow us to estimate the $\beta_{t(i,j)}^{k}$ despite the fact that we also have a household effect. To simplify the notation we will drop the X_{ij}’s from (1) but they will return to be explicit in the empirical analysis.

By implication, our test of the assumed separable structure in empirical collective models is robust to any (additive) latent household heterogeneity in individual Engel curves, as captured by the η_{j}^{k}’s. This is important given that the latter are potentially correlated with (*inter alia*) household total spending, as discussed in the Introduction.

---

[11] The bias can be removed using either a defensible instrumental variable or a randomized controlled trial (as in Almås et al. 2019).

Aggregation to household-level: Neither S_{ij}^{k} nor Y_{ij} is typically observed. To derive the estimable household-level Engel curves from (1) we can proceed by first eliminating the (unobserved) intra-household values of Y_{ij} by re-writing (1) as:

$$S_{ij}^{k} = \alpha_{t(i,j)}^{k} + (\beta_{t(i,j)}^{k} + \gamma_{t(i,j)}^{k}) \ln \bar{Y}_j + \nu_{ij}^{k} \tag{2}$$

Here the new error term is:

$$\nu_{ij}^{k} \equiv \gamma_{t(i,j)}^{k} \ln \delta_{ij} + \eta_j^k + \varepsilon_{ij}^{k} \tag{3}$$

where $\delta_{ij} \equiv Y_{ij}/\bar{Y}_j$ is the relative income for person *i* in household *j*. Next, on aggregating the budget shares at the household level for good *k* we have:

$$S_j^k = \frac{1}{n_j}\sum_{i=1}^{n} \delta_{ij} S_{ij}^{k} = \alpha_j^k + \left(\beta_j^k + \gamma_j^k\right) \ln\bar{Y}_j + \nu_j^k \tag{4}$$

(Clearly, the household-level analysis cannot distinguish β^k from γ^k.) Here the household-level Engel-curve parameters are related to the unpacked parameters as follows:

$$\alpha_j^k \equiv \frac{1}{n_j}\sum_{i=1}^{n} \delta_{ij}\alpha_{t(i,j)}^{k}\, ; \;\; \beta_j^k + \gamma_j^k \equiv \frac{1}{n_j}\sum_{i=1}^{n} \delta_{ij}(\beta_{t(i,j)}^{k} + \gamma_{t(i,j)}^{k}) \tag{5}$$

(Notice that the household parameters are implicitly weighted more heavily on household members with larger income shares.) The error term of the household Engel curve in (4) is:

$$\nu_j^k = \sum_{i=1}^{n} \delta_{ij}\nu_{ij}^{k} = T_j^k + \eta_j^k + \sum_{i=1}^{n} \delta_{ij}\varepsilon_{ij}^{k} \tag{6}$$

where:

$$T_j^k \equiv \frac{1}{n_j}\sum_{i=1}^{n} \gamma_{t(i,j)}^{k}\delta_{ij} \ln \delta_{ij} \tag{7}$$

We shall refer to T_j^k as the "γ-weighted Theil index" for good *k*. (Note that the weighted Theil index is not bounded below by zero, as in the unweighted index, given that γ's can be negative.) The corresponding ordinary Theil index applied to measuring intra-household inequality is:[12]

$$T_j \equiv \frac{1}{n_j}\sum_{i=1}^{n} \delta_{ij} \ln \delta_{ij} \tag{8}$$

We study the household Engel curves in the form of (4). However, these are still not the standard estimable Engel curve, which has constant-parameters as follows:

$$S_j^k = \alpha^k + (\beta^k + \gamma^k) \ln \bar{Y}_j + \mu_j^k \tag{9}$$

Here:

---

[12] See Theil (1967). For further discussion of the properties of this index see Bourguignon (1979).

$$\alpha^k \equiv \frac{1}{m}\sum_{j=1}^{m}\alpha_j^k\ ;\ \ \beta^k+\gamma^k \equiv \frac{1}{m}\sum_{j=1}^{m}\beta_j^k+\gamma_j^k \tag{10}$$

while the error term in (9) is:

$$\mu_j^k \equiv\ T_j^k+\eta_j^k+\alpha_j^k-\alpha^k+[(\beta_j^k+\gamma_j^k)-(\beta^k+\gamma^k)]\ln\bar{Y}_j+\sum_{i=1}^{n}\delta_{ij}\varepsilon_{ij}^k \tag{11}$$

The standard formula for the total spending elasticities of household demand for WL Engel curves is:

$$\frac{\partial lnY_j^k}{\partial ln\bar{Y}_j}=1+\frac{\beta^k+\gamma^k}{S_j^k} \tag{12}$$

(Here Y_j^k is total household spending on good *k*.) Note, however, that (in our model) this standard formula is only valid when all individuals gain in proportion to their current allocation of the household's total consumption, in which case the effect of the income gain on the intra-household inequality term in (7) can be set to zero. We do not make this assumption in our empirical work, but it will help in interpreting our results.

Conditions for consistent estimation using only household-level data: As usual, the key assumption needed for consistent estimation of the household Engel curve in (10) is that $Cov(\mu_j^k, \ln\bar{Y}_j)=0$. (With other regressors we require of course the same orthogonality condition.) From (11), it can be seen that there are three potential sources of bias when estimating the Engel curve using standard household data, corresponding to the sources of heterogeneity described in the introduction: (i) correlated heterogeneity in the γ-weighted Theil indices (T_j^k's); (ii) correlated household effects (the η_j^k's); and (iii) systematic heterogeneity in the individual Engel curves, such that $Cov[((\beta_j^k+\gamma_j^k)-(\beta^k+\gamma^k)]\ln\bar{Y}_j),\ln\bar{Y}_j]\neq 0$. Given that we allow for a household effect on the sub-household Engel curves, we can treat $\beta_j^k+\gamma_j^k$ as orthogonal to $\ln\bar{Y}_j$ in (11), leaving the η_j^k's to pick up this source of bias in the Engel curve when estimated on household data only.

Thus, two specific orthogonality conditions stand out as problematic. The first is a familiar exogeneity condition, $E(\eta_j^k|\ln\bar{Y}_j)=0$. The second is that $E(T_j^k|\ln\bar{Y}_j)=0$. If both hold then OLS is valid when applied to (9). In the special case of constant Engel parameters across all household members ($\gamma_{t(i,j)}^k=\gamma^k$ for all *i,j*) this requires that the intra-household Theil index is orthogonal to household total spending.

Recall that there is also potential endogeneity of the individual total expenditures in (1), i.e., $Cov(\tilde{\varepsilon}_{ij}^{k}, lnY_{ij}) \neq 0$. So, there is no guarantee that an estimate of the household Engel obtained by consistent aggregation from individual Engel curves is unbiased.

A similar unpacking of the household Engel curves can be conducted under other specification choices. In particular, the case of the quadratic Engel curves is developed in Appendix A2.

3. Data and descriptive statistics

We use the Pauvreté et Structure Familiale (Poverty and Family Structure, henceforth PSF) survey conducted in Senegal in 2006/2007.[13] The survey is described in detail in De Vreyer et al. (2008). The PSF is a nationally representative survey covering about 1800 households spread over 150 clusters drawn randomly from the census districts so as to insure a geographically representative sample. The survey describes a population of which the majority (52%) live in rural areas.

Senegalese households are large, with eight members on average in the PSF. The families are typically multigenerational. Polygamous unions are common, with 25% of married men and 39% of married women engaged in such unions, which mostly comprise a husband and two wives (only 20% of polygamous unions have more than two wives).

The survey collected details on each household's structure and budgetary arrangements. Each household was divided into "cells" according to the following rule: the head of household and unaccompanied dependent members, such as his widowed parent or his children whose mothers do not live in the same household, are grouped together. Then, each wife of the head and her children make up a separate group. Finally, any other family nucleus such as a married child of the household head with his/her spouse and children, forms a separate cell. This disaggregation emerged from field interviews as being the relevant way to split the household into its component groups. Enumerators saw this as a natural way to divide households and had

[13] The PSF survey stems from the cooperation between a team of French researchers and the National Statistical Agency of Senegal. Momar Sylla and Matar Gueye of the Agence Nationale de la Statistique et de la Démographie of Senegal (ANSD) on the one hand and Philippe De Vreyer, Sylvie Lambert and Abla Safir (World Bank) designed the survey. The data collection was conducted by the ANSD thanks to the funding of the IDRC (International Development Research Center), INRA Paris and CEPREMAP.

no difficulty organizing the household in this way and collecting the data accordingly. For the purpose of this study we confine our analysis to households comprising at least two cells, which reduces the sample size to 1430 households and 3979 cells.

Consumption expenditures are recorded in several parts: first, all public expenditures within the household are collected (housing, electricity bills etc). Regarding food expenditures, a detailed account is made of who shares which meal and how much money is specifically used to prepare this meal (the “DQ”, i.e. “dépenses quotidiennes,” which is the name the Senegalese give to the amount of money a woman has at her disposal to buy fresh ingredients for the meals of the day). Then individual consumption is collected at the cell level (such as clothing, mobile phone, transportation, food outside the home), by interviewing directly the cell’s head. Finally, expenditures that are shared between several cells but not the whole household are collected. Hence, a measure of per capita consumption can be constructed at the cell level

Thanks to these features of our data we can construct a relatively individualized measure of consumption, which is rarely available in household surveys, thus allowing us to un-pack the household Engel curve. The measure of cell-specific total spending we use is the amount of expenditures specific to the cell (eventually shared with some other cell of the household) plus the cell’s imputed share of the household’s joint expenditures. For consumption purposes, cell heads are assumed to be the decision makers at the cell level.

When looking at total expenditures, inequalities within the household are evident.[14] The ratio between the expenditures of the richest and the poorest group within a household can be as high as 5.3 even after trimming off the 5% most unequal households. Computing an inequality index for the distribution of cash expenditures in the population, we find a Gini index of 0.471 if we attribute to each person the average per capita consumption level in his or her household. By contrast, the index is 0.497 if instead each individual is attributed the per capita consumption in his cell (i.e. the sum of the per capita expenditures specific to the cell and of the cell’s share of public household expenditures, distributed on a per capita basis within the cell). The Gini index of inequality in the distribution of the cell-specific component of cash expenditures (ignoring the joint consumption within the household) is 0.779.

[14] For further discussion of intra-household inequality in Senegal and its implications for the measurement of poverty see De Vreyer and Lambert (2020).

The cell-specific consumption data reveal a sizeable gender gap. Male headed cells have a per capita consumption 33% higher than that of female headed ones, and this difference is statistically significant at the 1% level (t= 13.12).

Table 1 provides summary statistics for our sample on the main variables. We divide this into household (Table 1a) and cell level by type (Table 1b). Average household food share is 62.5%, though it ranges widely, from 6.9% to 99.8%. The average size is 9 persons (larger than for the full sample given that we exclude single-cell households), ranging from 2 to 44. The average number of cells is 2.8 (ranging from 2 to 12). 36% of cells are those of the head, 37% of the spouse. 86.5% of household heads are male. Notice that spouses have a higher food share than heads (Table 1b).

4. Results comparing household and cell-specific Engel curves

We use these data to estimate household Engel curves in two ways. The first entails estimating the cell-specific Engel curves in equation (1) and then aggregating up to the household level. The second is the standard method using only the household aggregate data, ignoring the fact that we have sub-household data, as in equation (9). We then compare the two. Recall that there are sources of bias in both cases; the individual Engel curves could well be biased by the possibility that the intra-household assignment mechanism is used to incentivize preferred budget allocations, while the standard household-level approach is biased by the sources of heterogeneity discussed above, including intra-household inequality.

Estimates with household fixed effects: Mundlak's (1978) method for estimating panel data models is well suited to our task, whereby the household effect is replaced by a linear-in-parameters function of the mean household characteristics.[15] As shown by Mundlak (1978), this gives the same estimates for the $\gamma_{t(i)}^{k}$'s as one would obtain with the usual within-estimator (allowing for a household fixed effect, correlated with the regressors) but also gives estimates of the β_t^k's. This property can be derived by applying a well-known theorem due to Frisch and Waugh (1933), as demonstrated in Appendix A3, which also makes it clear that the required property of the Mundlak estimator also holds for unbalanced panels.

[15] Also see the discussion of this method and extensions in Wooldridge (2019).

Table 2 summarizes the results. The regressions included urban and regional dummy variables as well as (log) household size and (log) cell size. The complete results for all Engel curves are found in the Appendix B1. Appendix A4 gives results for an augmented specification where the supplementary control variables are age, sex and education of household head; age and education of cell head; number of children less than 10 in household and household structure. This augmented specification gave similar results to Table 2.

The top two panels of Table 2 give the β_t^k and γ_t^k separately for three types of people: the household head and other members in his/her cell (type 1), head's spouses and their children (type 2) and other household members (type 3). The next panel down gives β_t^k+γ_t^k for the same three types, followed by test statistics. The row "β+γ from HH level regression" gives the coefficient obtained by the standard household level data analysis. In other words, this is an estimate of equation (9) treating ν_j^k as a zero-mean error term orthogonal to the regressors. The row "β+γ from cell regressions" is obtained by consistent aggregation from the cell-level Engel curves in equation (1). Here the means are computed at two levels: first, for each household, we compute the weighted mean of the β_t^k+γ_t^k coefficients, with weights equal to the budget share of each cell. Second, we compute the average over the sample of households. Table 2 also gives the Hausman test, indicating that the fixed effects specification is to be preferred for the cell level Engel curves (against the null hypothesis that the random effect model is the correct specification).

Note that Table 2 provides results for public consumption as well as private consumption. While the public consumption Engel curve makes sense at the household level, that is not so at cell level, since public consumption is assigned on a per capita basis. So an individualized Engel curve for such goods makes little sense. However, for completeness we still give the cell Engel curves for this case, while noting that they do not have the same interpretation. Thus, we label the column for public goods with parentheses, as "(public)", to remind readers of this difference with regard to the cell-level Engel curves.

The main results for the quadratic Engel curves are given in Appendix A5. The coefficients on the squared terms are more often significant for the household-level Engel curves than for the cell-level. Indeed, the coefficients on the squared terms are not significantly different from zero in four out of six cases for the cell-level regressions. When estimated using only the household level data, the coefficients on household total spending are reasonably similar between the linear and quadratic specifications, with the latter estimated at mean points (comparing the estimates of Tables 2 and A4). On the other hand, the corresponding aggregated cell-level Engel coefficients increase substantially in absolute value when we switch to the quadratic specification (even though the underlying quadratic terms are generally not significant) and the coefficients seem unreasonably large. Finally, when switching to the quadratic, we are still led to conclude that the standard household-level Engel curves (estimated on household aggregate data) are deceptive about the underlying cell-level Engel curves when aggregated to household level. Overall, these observations lead us to prefer the standard Working-Leser form. We confine attention to that specification below.

We cannot reject the null hypothesis that β_t^k=0 in most cases. So our results are generally consistent with the separable structure of the two-stage models as discussed in Section 2, which is a necessary but not sufficient condition for these models to be valid. The main exception is education spending, for which a strong external effect is evident. This can be interpreted as a social effect of education, such that, at a given level of own-spending, individuals living in richer households feel social pressure to spend more on their children's schooling. Here the social effect could well be the father exercising influence over the spouse(s) to spend more on his children's schooling (including making conditional monetary transfers to), though some role may also be played by competition among the wives.

In the case of food, Engel's Law ($\beta+\gamma < 0$) is coming primarily from the "own-spending" effect (the γ) rather than the external effect (β), though both go in the same direction, consistent with Engel's Law. For a majority of categories, the two effects go in the same direction. The exceptions are transport and clothing; in the case of transport the positive own-spending effect is partly offset by a negative external effect, with the opposite pattern for clothing (except for that of the head). The own expenditure effects (γ) and total effects ($\beta+\gamma$) are positive for all categories of private goods except food, implying that all except food are luxury goods. As expected, food is a necessity.

Table 3 provides the estimates of the Engel curves constrained to have β_t^k=0. This gives more precise estimates.[16] For most categories, the Engel parameters are similar within the household, for both the β_t^k's and γ_t^k's. It is interesting to note that for food (in general consumed jointly by all household members), as well as for the public goods, the cell-level Engel parameters are very similar across individual types (specifically for the household head and his spouse(s)), as shown by the F-stats in Tables 2 and 3. On the other hand, for goods such as transports and education, as well as for clothing when $\beta_t^k$are constrained to be equal to 0, some cell level heterogeneity in Engel parameters emerge.

Comparison of household and cell-based estimates: We find large discrepancies between the the aggregated individual Engel curves and standard estimates from only household aggregate data. For only two categories of spending (transport and clothing) can we not reject the null hypothesis that the household-level estimation of the Engel curves gives the same parameter estimates as one would obtain by consistent aggregation of the individual-level Engel curves for household members. For education, food and other goods and services the differences in the coefficients are sizeable as well as statistically significant (Table 2). When we constrain the estimates for (non-education) spending such that $\beta = 0$ we find that the null that the coefficients are the same can be rejected for every category of spending (Table 3)

A further insight into the extent of the discrepancies is found in Figure 1, which provides the cell-based density functions of the Engel parameters across households for each category of spending, as derived from Table 2.[17] It is notable how tight the distributions are around their means from Table 2. When we compare these densities with the results from the standard household-level specification in Table 2 we see that for every category, the latter are well outside the bulk of the density function of Engel parameters. Clearly, these are not small differences.

For food, Engel's Law holds either way, although the cell analysis indicates a far steeper decline in the food share as total consumption rises; indeed, the estimated parameter is 2.5 times greater when we use the cell data and aggregate up to the household level. At the mean budget share for food of 0.625, this difference would reduce the implied income elasticity of demand for

[16] Tables A5 to A8 in Appendix A5 reproduces the results presented in Tables 3 to 6 for augmented specifications that include a fuller set of controls.

[17] 20 extreme observations were dropped to make the density functions easier to view. This trimming mostly affected the food Engel coefficients; the mean food coefficient for the trimmed distribution in Figure 1 is -0.254.

food (assuming inequality neutrality) by one third—from 0.82 (using household data only) to 0.55 (using cell data).

Given that the issue of who receives a transfer payment has been prominent in the literature, and in light of the findings of Attanasio and Lechene (2010, 2014) (as discussed in the Introduction), it is of interest to consider the effects on the budget share devoted to food by a household comprising two adults in the following stylized cases:

Case 1: Total spending by the head increases in the amount $\Delta = dY_H$ with no external effect and no change in the spending done directly by the spouse. The share of the head's spending going to food (S_H^F) changes by $\gamma_H \Delta / Y_H$ and the share of total household spending devoted to food (S^F) changes by $dS^F = (S_H^F - S^F + \gamma_H)\Delta / Y$.

Case 2: This is the reverse: the spouse's total spending increases in the amount $\Delta = dY_S$ and (similarly to Case 1) there is no external effect and no change in the spending done directly by the head. The share of the spouse's spending going to food (S_S^F) changes by $\gamma_S \Delta / Y_S$ and the share of total household spending devoted to food changes by $dS^F = (S_S^F - S^F + \gamma_S)\Delta / Y$.

We saw in Table 1b that heads have a lower food share than spouses, namely 57.1% and 66.9% respectively, with a household food share of 62.5% (Table 1a). Thus, Case 1 changes the household food share by $-0.26\Delta / Y$ while it changes by $-0.16\Delta / Y$ for Case 2. The food share falls less when the spending gain is for the spouse. Given the similarity in their Engel parameters, the difference in the impact on aggregate food share between the two cases ($0.10\Delta / Y$) is almost entirely due to the difference in individual food shares between the head and the spouse (at mean points, $S_H^F - S_S^F = -0.10$).

5. Sources of bias in standard household Engel curves

Recall that our model points to two main sources of bias in the standard household-level Engel curves, namely the confounding differences across households in either the household effects (the η_j^k's), and the γ-weighted Theil index (T_j^k's). In this section we use our data to see which confounder matters most, so as to better understand the sources of bias in standard household Engel curves.

We begin in Table 4 by adding one or both of the two confounders to a standard household Engel curve. This suggests that the household effects are the main source of the discrepancies we find between the standard household-level estimates of $\beta+\gamma$ and our aggregated cell-specific coefficients; the gaps are much reduced when we control for household fixed effects, while the impact of controlling for the γ-weighted Theil index is much smaller and goes in the other direction.[18]

As a corroboration of that finding, if the household effects are indeed the main source of the discrepancies we find, then we should get similar results to the standard household Engel curve if we ignore the η_j^k's when estimating our cell-based Engel curves. Table 5 gives the analogous results to Table 2 ignoring the η_j^k's in (1), which is then estimated by OLS. This confirms our intuition. Consider for instance the food Engel curve, for which the estimate of $\beta+\gamma$ in Table 5, based on the cell-model, falls to -0.091, which is fairly close to the household-based estimate (-0.112). The gap between Table 2 and Table 5 is mainly due to the difference in the estimated β's. Using OLS one would (incorrectly) reject the two-stage structure as a model for household decision making.

To further understand the sources of the discrepancies we also provide regressions for both of the confounding components of the error term in the standard Engel curve estimated on household data alone. The total bias in estimates of $\beta+\gamma$ can be decomposed into the bias due to the η's and that due to the T's.[19] In order to assess the relative weight of these two sources of bias, we look at how these confounders correlate with log of total expenditures and log of household size. Table 6 shows the results of the regressions, for each category of goods, of each of these confounders on all variables included in the Engel curve specification.[20] (The Appendix B3 gives full results.) The top panel of Table 6 gives results for the household effects while the lower panel gives results for the γ-weighted Theil index. We have greater explanatory power (higher R^2) for the η's. We see that household total spending is a significant predictor of the

[18] Appendix B2 gives results with the full set of controls, though this makes little difference. It also shows that using the intra-household Theil index directly rather than the γ-weighted Theil gives very similar results.
[19] More precisely, the bias is the regression coefficient of ν_j^k on $\ln \bar{Y}_j$ $(Cov(\nu_j^k, \ln \bar{Y}_j)/Var(\ln \bar{Y}_j))$, which is the sum of the regression coefficient of η_j^k on $\ln \bar{Y}_j$ and that of T_j^k on $\ln \bar{Y}_j$.
[20] To calculate the fixed effect we used the STATA predict command (which is akin to averaging the residuals of the Mundlak estimates of the Engel curves over cells). The estimates in Table 6 are based on the cell-level Engel curves with the basic control variables. Appendix A9 provides results using a fuller set of controls.

household effects in all categories.[21] For food and "public" expenditures, the effect is positive, while it is negative for other categories of spending. (Log household size has a negative effect, and the sum of the coefficients on log total spending and log household size is close to zero for most categories, suggesting that the relevant predictor is log per capita spending.)

Turning next to the lower panel of Table 6, we find that the intra-household γ-weighted Theil index of every category of spending is correlated with one or both of log household expenditure and log household size. This holds with or without the extra controls. The γ-weighted inequality index for transport, clothing, education and other spending tends to rise with household total spending, while falling for food and public spending.

Both confounders are significantly correlated with total household spending, but in opposite directions. And for all categories of spending, the effect through the η_j^k's dominates that via the T_j^k's. For the food Engel curve, the η_j^k's are positively correlated with $\ln \bar{Y}_j$ while the T_j^k's are negatively correlated. On balance, the positive effect through the η_j^k's is dominant. All other consumption categories also have offsetting components of the bias term, though here the sign pattern is reversed, with the η_j^k's negatively correlated with $\ln \bar{Y}_j$ while the T_j^k's are correlated positively. Again, the dominant channel of confounding is the latent household effects in standard Engel curves.

This pattern we find in the directions and sizes of the two sources of bias in $\beta+\gamma$ implies that only adding controls for intra-household inequality to a standard household-level Engel curve will increase the bias.

We also augmented the regressions for the confounders with other control variables; full details are found in the Appendix B3. For one or more categories of spending, the household effect was also correlated with the number of children, and the age, gender and education of the head. While significant in many cases, adding the extra covariates did not much reduce the (absolute value) of the coefficient on (log) total spending.

[21] Note that this is consistent with the fact that we find support for the two-stage structure. Differences in household total spending are correlated with the household effect but individuals are largely basing their own budget allocations on their share of household total spending.

6. Conclusions

Traditional Engel curves can hide considerable heterogeneity in parameters and resource allocation within households. Furthermore, this heterogeneity need not be ignorable from the point of view of reliably estimating household Engel curves. Indeed, once we depart from the unitary model of the household, the interpretation of standard household-level Engel curves becomes unclear.

We have postulated sub-household Engel curves that aggregate up to the widely-used Working-Leser form of the aggregate (household-level) Engel curve. Our model allows both internal and external income effects at individual level. Unpacking household Engel curves this way reveals the conditions required for the standard household Engel curve to be unbiased in the presence of intra-household inequality and latent behavioral heterogeneity across households.

We have applied these ideas to an unusual data set for Senegal in which the household spending choices are identified at sub-household level by exploiting the "cell" structure that is common in Senegalese households. This provides a rare window for unpacking household Engel curves.

Two key lessons emerge. First, the (often-assumed) two-stage structure in bargaining-collective models of the household carries a testable implication with our data, namely that household spending should only matter to individual choices via the intra-household allocation of total spending. This exclusion restriction is generally consistent with our results. The exception is education spending, for which cell-specific budget shares are independently, and significantly, affected by the household's overall standard of living.

Second, our data reveal large discrepancies between the standard household-level Engel curves and that obtained by aggregating up cell-level ones. For example, for the food share Engel curve, the coefficient on log total household spending is -0.11 using only household data but -0.28 when one estimates the Engel curve from the sub-household data and aggregates up to the household data. This is enough to reduce the income elasticity of demand for food (evaluated at mean food share) by one third, from 0.82 to 0.55.

There are three aspects of these findings that call for a note of caution. The first is that we cannot rule out biases in our individualized (cell-based) Engel coefficients. We do know which of the two methods is closer to the truth; what is clear is that they don't agree. It is also notable that they converge considerably when we allow for the identified sources of heterogeneity that can confound household-level Engel curves. The second caveat is that, while the two-stage structure assumed in implementations of the collective model performs well, this does not validate all aspects of that model, such as the assumption of Pareto efficiency within households. The third caveat is that our test of the two-stage structure allows for latent household heterogeneity, in the form of a household fixed effect in Engel curves. Two-stage budgeting does not do so well when we relax this, to allow the possibility that the household preference parameter is itself influenced by total household spending.

In these data we find that the bulk of the gap between the two methods of estimating household Engel curves is accountable to the influence of household fixed effects on sub-household Engel curves. The channel of bias via intra-household inequality partially offsets that due to the latent household effects in standard Engel curves calibrated to only household-level data. Thus, only adding controls to reflect intra-household inequality will tend to increase the bias in household-level Engel curves.

Given that we find (at least for our data) that the bulk of the discrepancy between standard household Engel curves and consistently aggregated micro curves is due to household effects on sub-household Engel curves, it may be expected that the most promising means of removing (or at least attenuating) the bias is to use longitudinal data, assuming that the confounding household effect in individual consumption behavior is time invariant. That conclusion remains to be investigated further in future work.

Table 1a: Summary statistics at household level

Variable	N	Mean	Std. Dev.	Min	Max
Log(household exp. per cap.)	1,430	12.414	0.835	10.344	15.315
Hhold share food exp.	1,430	0.625	0.187	0.069	0.998
Hhold share public exp.	1,430	0.088	0.082	0.000	0.808
Hhold share transport exp.	1,430	0.060	0.096	0.000	0.835
Hhold share clothing exp.	1,430	0.065	0.070	0.000	0.888
Hhold share education exp.	1,430	0.024	0.056	0.000	0.818
Hhold share other exp.	1,430	0.139	0.117	0.000	0.892
HHead is male	1,430	0.857	0.351	0.000	1.000
Head's age	1,429	51.667	13.852	20.000	93.000
Head's educ. 1 to 3 years	1,430	0.034	0.182	0.000	1.000
Head's educ. 4 to 5 years	1,430	0.101	0.302	0.000	1.000
Head's educ. junior school	1,430	0.075	0.263	0.000	1.000
Head's educ. High school or more	1,430	0.090	0.287	0.000	1.000
Head's educ. Koranic	1,430	0.301	0.459	0.000	1.000
Household size	1,430	9.055	5.165	2.000	44.000
Number of cells	1,430	2.790	1.193	2.000	12.000
Two cells: Head + SP	1,430	0.446	0.497	0.000	1.000
Two cells: Head + Oth	1,430	0.110	0.313	0.000	1.000
At least 3 cells: Head + n SP	1,430	0.097	0.295	0.000	1.000
At least 3 cells: Head + n SP + Oth	1,430	0.287	0.453	0.000	1.000
Other kinds	1,430	0.060	0.238	0.000	1.000
Number of kids less than 10	1,430	2.720	2.328	0.000	16.000
Dakar	1,430	0.327	0.469	0.000	1.000
Other urban	1,430	0.201	0.401	0.000	1.000
Rural	1,430	0.473	0.499	0.000	1.000
Diourbel	1,430	0.040	0.196	0.000	1.000
Fatick	1,430	0.082	0.274	0.000	1.000
Kaolack	1,430	0.068	0.252	0.000	1.000
Kolda	1,430	0.052	0.222	0.000	1.000
Louga	1,430	0.087	0.283	0.000	1.000
Matam	1,430	0.125	0.331	0.000	1.000
Saint-Louis	1,430	0.058	0.234	0.000	1.000
Tambacounda	1,430	0.048	0.214	0.000	1.000
Thies	1,430	0.068	0.252	0.000	1.000
Ziguinchor	1,430	0.038	0.192	0.000	1.000

Source: PSF survey, authors’ calculations.

Table 1b: Summary statistics at cell level

	Head					Spouses					Other				
Variable	N	Mean	Std. Dev.	Min	Max	N	Mean	Std. Dev.	Min	Max	N	Mean	Std. Dev.	Min	Max
Log(Cell exp. per cap.)	1,430	12.634	0.935	10.263	16.460	1,485	12.214	0.830	9.932	15.479	1,064	12.265	0.809	10.068	16.049
Cell share food exp.	1,430	0.571	0.222	0.020	1.000	1,485	0.669	0.192	0.082	1.000	1,064	0.611	0.197	0.018	1.000
Cell share public exp.	1,430	0.081	0.089	0.000	0.924	1,485	0.090	0.092	0.000	0.834	1,064	0.098	0.094	0.000	0.740
Cell share transport exp.	1,430	0.089	0.137	0.000	0.939	1,485	0.037	0.079	0.000	0.861	1,064	0.051	0.097	0.000	0.842
Cell share clothing exp.	1,430	0.067	0.084	0.000	0.973	1,485	0.067	0.079	0.000	0.833	1,064	0.079	0.091	0.000	0.978
Cell share education exp.	1,430	0.013	0.056	0.000	0.827	1,485	0.022	0.058	0.000	0.885	1,064	0.017	0.051	0.000	0.706
Cell share other exp.	1,430	0.179	0.165	0.000	0.917	1,485	0.114	0.117	0.000	0.855	1,064	0.144	0.124	0.000	0.944
Cell's head is male	1,430	0.865	0.377	0.000	4.000	1,485	0.018	0.241	0.000	4.000	1,064	0.474	0.602	0.000	4.000
Cell head's age	1,429	51.750	13.946	20.000	95.000	1,485	39.485	13.011	15.00 0	96.000	1,063	35.352	14.192	13.000	98.000
Cell head's educ.: 1 to 3 years	1,430	0.034	0.182	0.000	1.000	1,485	0.043	0.203	0.000	1.000	1,064	0.055	0.229	0.000	1.000
Cell head's educ.: 4 to 5 years	1,430	0.101	0.302	0.000	1.000	1,485	0.091	0.288	0.000	1.000	1,064	0.147	0.354	0.000	1.000
Cell head's educ.: junior school	1,430	0.075	0.263	0.000	1.000	1,485	0.044	0.205	0.000	1.000	1,064	0.068	0.251	0.000	1.000
Cell head's educ.: High school +	1,430	0.091	0.290	0.000	2.000	1,485	0.038	0.194	0.000	2.000	1,064	0.055	0.229	0.000	1.000
Cell head's educ.: Koranic	1,430	0.302	0.461	0.000	2.000	1,485	0.178	0.382	0.000	1.000	1,064	0.184	0.390	0.000	2.000
Cell size	1,430	2.236	1.894	1.000	15.000	1,485	4.166	2.109	1.000	12.000	1,064	3.400	1.798	1.000	14.000

Source: PSF survey, authors' calculations.

Table 2: Mundlak estimates of Engel curves

	Food	(Public)	Transport	Clothing	Education	Other
β for head's cell	-0.072	-0.027	-0.028	0.008	0.085	0.035
(s.e.)	(0.038)	(0.020)	(0.024)	(0.020)	(0.013)	(0.030)
β for spouses' cell type	-0.079	-0.028	-0.024	-0.024	0.087	0.068
(s.e.)	(0.048)	(0.026)	(0.030)	(0.025)	(0.016)	(0.038)
β for o. members' cells	-0.025	-0.025	-0.017	-0.041	0.083	0.024
(s.e.)	(0.046)	(0.025)	(0.029)	(0.024)	(0.015)	(0.037)
γ for head's cell	-0.212	-0.036	0.103	0.018	0.017	0.109
(s.e)	(0.015)	(0.008)	(0.010)	(0.008)	(0.005)	(0.012)
γ for spouses' cells	-0.204	-0.037	0.090	0.055	0.025	0.071
(s.e)	(0.021)	(0.011)	(0.013)	(0.011)	(0.007)	(0.016)
γ for o. members's cells	-0.260	-0.035	0.084	0.071	0.025	0.115
(s.e.)	(0.017)	(0.009)	(0.011)	(0.009)	(0.006)	(0.014)
β+γ for head's cell	-0.284	-0.063	0.076	0.026	0.102	0.144
(s.e)	(0.041)	(0.022)	(0.025)	(0.021)	(0.014)	(0.032)
β+γ for spouses' cells	-0.283	-0.064	0.066	0.031	0.112	0.138
(s.e)	(0.042)	(0.023)	(0.026)	(0.022)	(0.014)	(0.033)
β+γ for o. members' cells	-0.284	-0.060	0.067	0.030	0.108	0.140
(s.e)	(0.043)	(0.023)	(0.027)	(0.022)	(0.014)	(0.034)
F stat. spouses vs head diff.	0.010	0.034	2.672	1.146	10.202	0.582
(p-value)	(0.920)	(0.853)	(0.102)	(0.284)	(0.001)	(0.445)
F stat. other vs head diff.	0.003	0.282	1.356	0.419	2.478	0.214
(p-value)	(0.955)	(0.595)	(0.244)	(0.518)	(0.115)	(0.644)
F stat. spouses vs other diff.	0.016	0.402	0.018	0.039	0.843	0.019
(p-value)	(0.899)	(0.526)	(0.893)	(0.844)	(0.359)	(0.890)
β+γ from hhld regression	-0.112	0.002	0.051	0.002	0.008	0.048
(s.e.)	(0.006)	(0.003)	(0.004)	(0.003)	(0.002)	(0.004)
β+γ from cell regressions	-0.284	-0.063	0.070	0.029	0.108	0.141
(s.e.)	(0.041)	(0.022)	(0.026)	(0.021)	(0.014)	(0.033)
z stat. hhold vs cell reg. diff.	4.050	1.850	-0.500	-0.690	-3.710	-2.450
(p-value)	(0.000)	(0.064)	(0.616)	(0.493)	(0.000)	(0.014)
Chi2 stat. Hausman test f.e.	79.024	49.856	65.660	57.201	121.447	55.813
(p-value)	(0.000)	(0.049)	(0.001)	(0.010)	(0.000)	(0.014)
Number of households	1,430	1,430	1,430	1,430	1,430	1,430
Number of cells	3,979	3,979	3,979	3,979	3,979	3,979

Source: PSF survey, authors' calculations.

Note: Other covariates in regressions include urban and regional dummies, together with (log) household size and (log) cell size. The β are the coefficients of total household expenditures in cell level Engel curves, while the γ are the coefficients of own (cell) expenditures. *β+γ from h'hold regression* is the standard Engel coefficient obtained from regression at the household level, while *β+γ from cell regressions* is the comparable Engel coefficient obtained by consistently aggregating cell level estimates. Standard errors in parentheses. Bootstrapped p-values for the z-stat of the difference between the household and the cell level estimates (50 replications).

Table 3: Mundlak estimates for linear Engel curves, constrained to have *β*=0, save education

	Food	(Public)	Transport	Clothing	Education	Other
β for head's cell	0.000	0.000	0.000	0.000	0.085	0.000
(s.e.)	(0.000)	(0.000)	(0.000)	(0.000)	(0.013)	(0.000)
β for spouses' cell type	0.000	0.000	0.000	0.000	0.087	0.000
(s.e.)	(0.000)	(0.000)	(0.000)	(0.000)	(0.016)	(0.000)
β for o. members' cells	0.000	0.000	0.000	0.000	0.083	0.000
(s.e.)	(0.000)	(0.000)	(0.000)	(0.000)	(0.015)	(0.000)
γ for head's cell	-0.221	-0.037	0.100	0.037	0.017	0.103
(s.e)	(0.009)	(0.005)	(0.006)	(0.005)	(0.005)	(0.007)
γ for spouses' cells	-0.220	-0.038	0.090	0.046	0.025	0.095
(s.e)	(0.011)	(0.006)	(0.007)	(0.006)	(0.007)	(0.009)
γ for o. members' cells	-0.231	-0.034	0.089	0.050	0.025	0.102
(s.e.)	(0.012)	(0.006)	(0.007)	(0.006)	(0.006)	(0.009)
β+γ for head's cell	-0.221	-0.037	0.100	0.037	0.102	0.103
(s.e)	(0.009)	(0.005)	(0.006)	(0.005)	(0.014)	(0.007)
β+γ for spouses' cells	-0.220	-0.038	0.090	0.046	0.112	0.095
(s.e)	(0.011)	(0.006)	(0.007)	(0.006)	(0.014)	(0.009)
β+γ for o. members' cells	-0.231	-0.034	0.089	0.050	0.108	0.102
(s.e)	(0.012)	(0.006)	(0.007)	(0.006)	(0.014)	(0.009)
F stat. spouses vs head diff.	0.004	0.038	3.026	2.728	10.202	1.051
(p-value)	(0.948)	(0.846)	(0.082)	(0.099)	(0.001)	(0.305)
F stat. other vs head diff.	0.875	0.232	2.433	4.482	2.478	0.006
(p-value)	(0.350)	(0.630)	(0.119)	(0.034)	(0.115)	(0.939)
F stat. spouses vs other diff.	0.811	0.346	0.006	0.423	0.843	0.520
(p-value)	(0.368)	(0.557)	(0.940)	(0.516)	(0.359)	(0.471)
β+γ from h'hold regression	-0.112	0.002	0.051	0.002	0.008	0.048
(s.e.)	(0.006)	(0.003)	(0.004)	(0.003)	(0.002)	(0.004)
β+γ from cell regressions	-0.223	-0.037	0.093	0.044	0.108	0.100
(s.e.)	(0.009)	(0.005)	(0.005)	(0.005)	(0.014)	(0.007)
z stat. h'hold vs cell reg. diff.	6.740	4.640	-3.450	-3.390	-3.710	-3.870
(p-value)	(0.000)	(0.000)	(0.001)	(0.001)	(0.000)	(0.000)
Number of households	1,430	1,430	1,430	1,430	1,430	1,430
Number of cells	3,979	3,979	3,979	3,979	3,979	3,979

Source: PSF survey, authors' calculations.

Note: Other covariates in regressions include urban and regional dummies, together with (log) household size and (log) cell size. The β are the coefficients of total household expenditures in cell level Engel curves. They are here constrained to be equal to 0, except for education. The γ are the coefficients of own (cell) expenditures. *β+γ from hhld regression* is the standard Engel coefficient obtained from regression at the household level, while *β+γ from cell regressions* is the comparable Engel coefficient obtained by consistently aggregating cell level estimates. Standard errors in parentheses. Bootstrapped p-values for the z-stat of the difference between the household and the cell level estimates (50 replications).

Table 4: Impact of confounders on log household expenditure per capita coefficient

	Food	(Public)	Transport	Clothing	Education	Other
β+γ from h'hold regression	-0.112	0.002	0.051	0.002	0.008	0.048
No extra covariate	(0.006)	(0.003)	(0.004)	(0.003)	(0.002)	(0.004)
(1) γ-weighted Theil only	-0.097	0.003	0.043	-0.001	0.008	0.046
	(0.006)	(0.003)	(0.004)	(0.003)	(0.002)	(0.005)
(2) Fixed effect only	-0.226	-0.035	0.110	0.026	0.027	0.109
	(0.002)	(0.001)	(0.001)	(0.001)	(0.001)	(0.002)
(1) + (2)	-0.217	-0.034	0.104	0.024	0.028	0.106
	(0.002)	(0.001)	(0.001)	(0.001)	(0.001)	(0.002)
β+γ from cell regression	-0.284	-0.063	0.070	0.029	0.108	0.141
No extra covariate	(0.041)	(0.022)	(0.026)	(0.021)	(0.014)	(0.033)

Source: PSF survey, authors' calculations.

Note: standard errors are between parentheses. Other covariates include (log) household size, and regional and rural/urban dummies. The table shows the values of the log household per capita expenditure coefficient in a household level Engel curve estimation (equation (9)). The first line shows the results obtained when the expenditure for a given class of items is regressed on log household expenditures per capita with controls for log household size and location dummies (as in table 2, line *β+γ from hhld regression*). Lines 2 to 4 show the values of the coefficient when the γ-weighted Theil is added to the regressors (line 2), or the fixed effect (line 3) or both (line 4). The last line displays the values obtained when the coefficient is built by consistently aggregating estimates of the cell level Engel curves following equation (10) (same as table 2, line *β+γ from cell regression)*.

Table 5: OLS estimates

	Food	(Public)	Transport	Clothing	Education	Other
β for head's cell	0.141	0.030	-0.074	-0.032	-0.000	-0.065
(s.e.)	(0.014)	(0.007)	(0.008)	(0.007)	(0.005)	(0.011)
β for spouses' cell type	0.044	0.038	-0.016	-0.032	-0.012	-0.022
(s.e.)	(0.016)	(0.009)	(0.010)	(0.008)	(0.005)	(0.013)
β for o. members' cells	0.141	0.042	-0.038	-0.060	-0.015	-0.070
(s.e.)	(0.016)	(0.009)	(0.010)	(0.008)	(0.005)	(0.013)
γ for head's cell	-0.239	-0.028	0.122	0.027	0.004	0.114
(s.e)	(0.012)	(0.006)	(0.007)	(0.006)	(0.004)	(0.009)
γ for spouses' cells	-0.129	-0.037	0.052	0.030	0.025	0.059
(s.e)	(0.016)	(0.009)	(0.010)	(0.008)	(0.005)	(0.012)
γ for o. members's cells	-0.227	-0.029	0.072	0.055	0.022	0.107
(s.e.)	(0.015)	(0.008)	(0.009)	(0.008)	(0.005)	(0.012)
β+γ for head's cell	-0.098	0.002	0.048	-0.005	0.004	0.049
(s.e)	(0.007)	(0.004)	(0.004)	(0.003)	(0.002)	(0.005)
β+γ for spouses' cells	-0.085	0.002	0.036	-0.002	0.012	0.037
(s.e)	(0.007)	(0.004)	(0.004)	(0.003)	(0.002)	(0.005)
β+γ for o. members' cells	-0.086	0.013	0.034	-0.005	0.007	0.037
(s.e)	(0.008)	(0.004)	(0.005)	(0.004)	(0.003)	(0.006)
F stat. spouses vs head diff.	1.996	0.000	4.812	0.293	7.757	2.575
(p-value)	(0.158)	(0.996)	(0.028)	(0.588)	(0.005)	(0.109)
F stat. other vs head diff.	1.333	4.168	5.309	0.001	0.975	2.229
(p-value)	(0.248)	(0.041)	(0.021)	(0.974)	(0.324)	(0.135)
F stat. spouses vs other diff.	0.014	4.125	0.107	0.203	2.277	0.002
(p-value)	(0.907)	(0.042)	(0.744)	(0.652)	(0.131)	(0.962)
β+γ from hhld level regression	-0.112	0.002	0.051	0.002	0.008	0.048
(s.e.)	(0.006)	(0.003)	(0.004)	(0.003)	(0.002)	(0.004)
β+γ from cell level regressions	-0.090	0.004	0.040	-0.004	0.008	0.041
(s.e.)	(0.004)	(0.002)	(0.003)	(0.002)	(0.001)	(0.003)
z stat. hhold vs cell reg. diff.	-2.630	-0.610	2.010	1.080	-0.080	1.330
(p-value)	(0.009)	(0.542)	(0.044)	(0.279)	(0.936)	(0.184)
Number of households	1,430	1,430	1,430	1,430	1,430	1,430
Number of cells	3,979	3,979	3,979	3,979	3,979	3,979

Source: PSF survey, authors' calculations.
Note: OLS estimates of Engel curves at the household and cell levels. Standard errors in parentheses. Other covariates in regressions include urbanization and regional dummies, together with (log) household size and (log) cell size. Bootstrapped p-values for the z-stat of the difference between the household and the cell level estimates (50 replications).

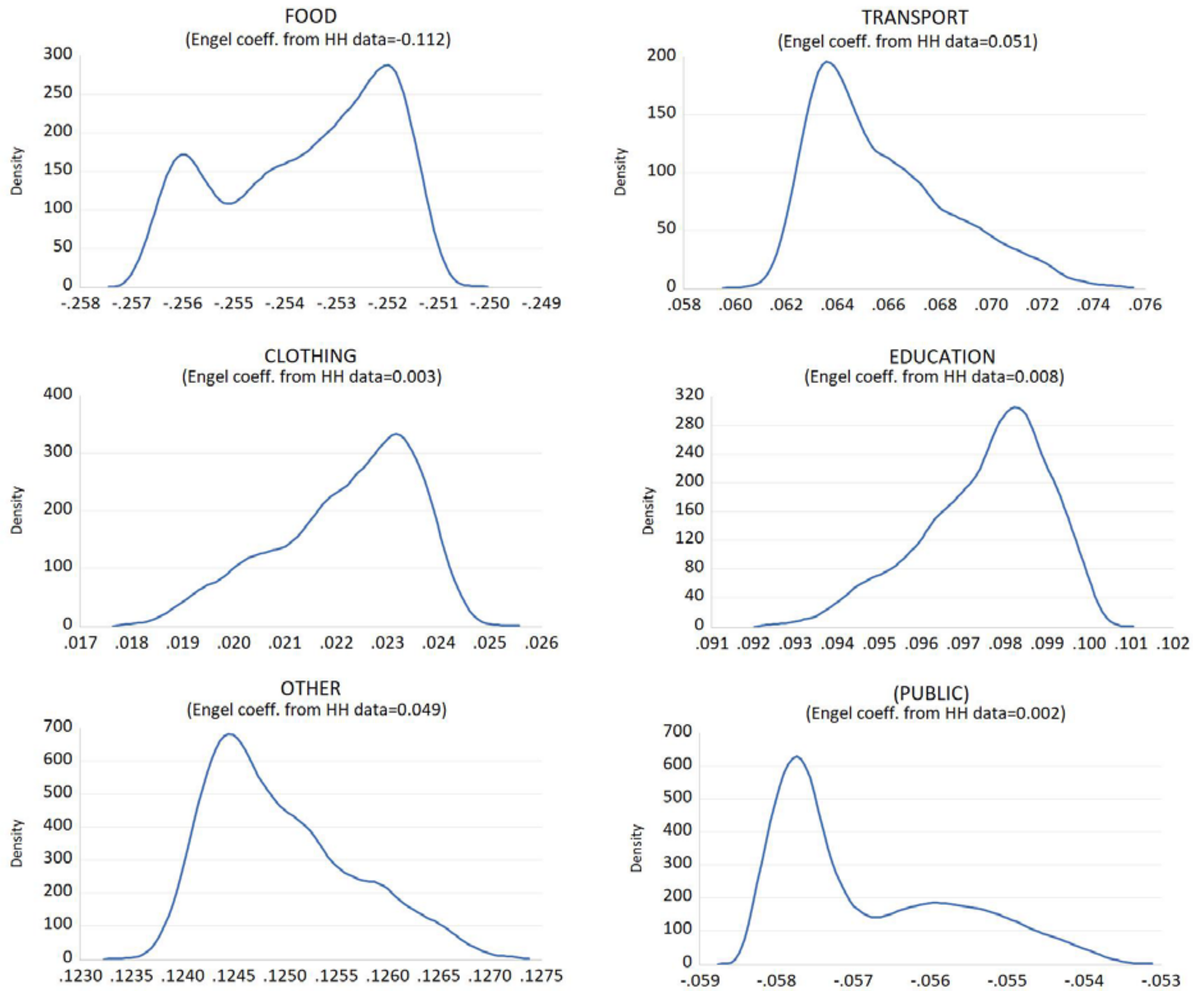
.

Table 6: Regressions for the two confounders

	(1) Food	(2) (Public)	(3) Transport	(4) Clothing	(5) Education	(6) Other
Household effects in cell Engel curves						
Log (total hhold	0.113***	0.0382***	-0.0546***	-0.0224***	-0.0140***	-0.0599***
exp.)	(0.00564)	(0.00313)	(0.00316)	(0.00244)	(0.00152)	(0.00412)
Log (hhold size)	-0.214***	-0.0349***	0.0740***	0.0356***	0.0151***	0.124***
	(0.00755)	(0.00419)	(0.00423)	(0.00326)	(0.00204)	(0.00552)
R^2	0.441	0.289	0.391	0.149	0.105	0.312
Intra-household Theil indices (γ-weighted)						
Log (total hhold	-0.0101***	-0.00104***	0.00493***	0.00127***	0.000158**	0.00482***
exp.)	(0.00131)	(0.000121)	(0.000695)	(0.000161)	(7.88e-05)	(0.000629)
Log (hhold size)	0.00305*	2.76e-05	-0.00219**	0.000167	0.000448***	-0.00149*
	(0.00176)	(0.000162)	(0.000930)	(0.000216)	(0.000105)	(0.000842)
R^2	0.055	0.070	0.050	0.062	0.045	0.055

Source: PSF survey, authors' calculations. Note: N=1,430. OLS estimates. Standard errors in parentheses *** p<0.01, ** p<0.05, * p<0.1. Other covariates in regressions include urban and regional dummies. Household fixed effects were computed from cell level regressions using STATA predict command.

Figure 1: Kernel density functions for estimated Engel parameters for log total spending



References

Almås, Ingvild, Johannes Haushofer and Jeremy P. Shapiro, 2019. "The Income Elasticity for Nutrition: Evidence from Unconditional Cash Transfers in Kenya," NBER WP 25711.

Apps, Patricia F., and Ray Rees. 1988. "Taxation and the Household," *Journal of Public Economics* 35(3): 355–69.

Arnould, Eric, Craig Thompson, 2005, "Consumer Culture Theory (CCT): Twenty Years of Research," *Journal of Consumer Research* 31(4): 868–882.

Attanasio, Orazio P., and Valérie Lechene, 2010. "Conditional Cash Transfers, Women and the Demand for Food." Working Paper no. 10/17, Inst. Fiscal Studies, London.

________________, and _____________, 2014, "Efficient Responses to Targeted Cash Transfers," *Journal of Political Economy* 122(1): 178-222.

Bargain, Olivier, Guy Lacroix, and Luca Tiberti, 2018, "Validating the Collective Model of the Household using Direct Evidence on Sharing," Partnership for Economic Policy Working Paper No. 2018-06.

Bourguignon, François, 1979, "Decomposable Income Inequality Measures," *Econometrica* 47: 901-920.

Bourguignon, François, Martin Browning, and Pierre-André Chiappori, 2009, "Efficient Intra-Household Allocations and Distribution Factors: Implication and Identification," *Review of Economic Studies* 75: 503-528.

Brown, Caitlin and Rossella Calvi and Jacob Penglase, 2018, "Sharing the Pie: Undernutrition, Intra-Household Allocation, and Poverty."

Chavas, Jean-Paul, Martina Menon, Elisa Pagani, and Federico Perali, 2018, "Collective household welfare and intra-household inequality," *Theoretical Economics* 13: 667–696.

Chiappori, Pierre-André, 1988. "Rational Household Labor Supply." *Econometrica* 56:63–89.

Chiappori, Pierre-André, and Costas Meghir, 2015, "Intrahousehold Inequality." In *Handbook of Income Distribution* (Anthony B. Atkinson and François Bourguignon, eds.), 1369–1418, North Holland, Amsterdam, Netherlands.

De Vreyer, Philippe, and Sylvie Lambert, 2020, "Inequality, poverty and the intra-household allocation of consumption in Senegal," *World Bank Economic Review*, in press.

De Vreyer, Philippe, Sylvie Lambert, Abla Safir and Momar B. Sylla, 2008, “Pauvreté et Structure Familiale, Pourquoi une Nouvelle Enquête ?”, *Stateco* 102: 261-275.

Dunbar, Geoffrey R., Arthur Lewbel and Krishna Pendakur, 2013, “Children's Resources in Collective Households: Identification, Estimation, and an Application to Child Poverty in Malawi,” *American Economic Review* 103(1): 438-71.

Engel, Ernst, 1857, “Die Productions- und Consumtionsverhältnisse des Königreichs Sachsen,” *Zeitschrift des statistischen Bureaus des Königlich Sächsischen Ministerium des Inneren* 8–9: 28–29.

Findlay, Jeanette, and Robert E. Wright, 1996, “Gender, Poverty and the Intra -Household Distribution of Resources,” *Review of Income and Wealth* 42(3): 335-351.

Frisch, Ragnar and Frederick Waugh, 1933, “Partial Time Regressions as Compared with Individual Trends,” *Econometrica* 1: 387-401.

Haddad, Lawrence, and Ravi Kanbur, 1990, “How Serious is the Neglect of Intra-Household Inequality?,” *Economic Journal* 100: 866-881.

Houthakker, H. S., 1957, “An International Comparison of Household Expenditure Patterns, Commemorating the Centenary of Engel’s Law,” *Econometrica* 25:532-551.

Kanbur, Ravi, and Lawrence Haddad, 1994, “Are Better off Households More Unequal or Less Unequal?” *Oxford Economic Papers* 46(3): 445–458.

Leser, C. E. V., 1963, “Forms of Engel Functions,” *Econometrica* 31: 694- 703.

Lise, Jeremy, and Shannon Seitz, 2011, “Consumption Inequality and Intra-household Allocations,” *Review of Economic Studies* 78(1): 328-355.

Muellbauer, John, 1975, “Aggregation, Income Distribution and Consumer Demand,” *Review of Economic Studies* 42(4): 525-543.

Mundlak, Yair, 1978, “On the Pooling of Time Series and Cross Section Data,” *Econometrica* 46(1): 69-85.

Phlips, Louis, 1983, *Applied Consumption Analysis.* Amsterdam: North-Holland.

Solomon, Michael, Gary Bamossy, Soren Askegaard, and Margaret Hogg, 2006, *Consumer Behavior: A European Perspective*, Third Edition, Harlow, Essex: Prentice Hall.Theil, Henri, 1967, *Economics and Information Theory*, Amsterdam: North-Holland.

Wooldridge, Jeffrey, 2019, “Correlated Random Effects Models with Unbalanced Panels,” *Journal of Econometrics* 211(1): 137-150.

Working, H., 1943, “Statistical Laws of Family Expenditures,” *Journal of the American Statistical Association* 38: 43-56.