cemmap

centre for microdata methods and practice

Intertemporal Collective Household Models: Identification in Short Panels with Unobserved Heterogeneity in Resource Shares

Irene Botosaru
Chris Muris
Krishna Pendakur

The Institute for Fiscal Studies
Department of Economics, UCL

cemmap working paper CWP26/20

Economic and Social Research Council

Intertemporal Collective Household Models: Identification in Short Panels with Unobserved Heterogeneity in Resource Shares

Irene Botosaru, Chris Muris, and Krishna Pendakur*

June 12, 2020

Abstract

We provide a new full-commitment intertemporal collective household model to estimate resource shares, defined as the fraction of household expenditure enjoyed by household members. Our model implies nonlinear time-varying household quantity demand functions that depend on fixed effects.

We provide new econometric results showing identification of a large class of models that includes our household model. We cover fixed-T panel models where the response variable is an unknown monotonic function of a linear latent variable with fixed effects, regressors, and a nonparametric error term. The function may be weakly monotonic and time-varying, and the fixed effects are unrestricted. We identify the structural parameters and features of the distribution of fixed effects. In our household model, these correspond to features of the distribution of resource shares.

Using Bangladeshi data, we show: women's resource shares decline with household budgets; and, half the variation in women's resource shares is due to unobserved heterogeneity.

---

*Email: irene.botosaru@bristol.ac.uk, muerisc@mcmaster.ca, pendakur@sfu.ca. We would like to thank Stéphane Bonhomme, P.A. Chiappori, Ian Crawford, Iván Fernández-Val, Dalia Ghanem, Hiro Kasahara, Shakeeb Khan, Valerie Lechene, Arthur Lewbel, Corrine Low, Maurizio Mazzocco, Jack Porter and Elie Tamer for useful discussions and suggestions; conference participants at the ASSA 2018, Berkeley-Stanford Jamboree 2017, IWEEE 2018, Seattle-Vancouver Econometrics Conference 2016 and 2017, and seminar participants at Cornell, CREST, Harvard, Oxford, Penn State, Queen's University, Rotterdam, Tilburg, University of Bristol, TSE, UBC, UCL, UC Davis, UCLA, UPenn and Vanderbilt for comments and suggestions. We thank Shirleen Manzur for her research assistance on this project, and gratefully acknowledge the financial support of SSHRC under grants IDG 430-2015-00073 and IG 435-2020-0026 and of the Institute for Advanced Study at the University of Bristol.

JEL codes: C14; C23; C41 **keywords**: panel data, fixed effects, incidental parameter, time-varying transformation function, collective household, full commitment, resource shares, gender inequality

1 Introduction

Standard poverty and inequality measures, based on per-capita household income or expenditure, assume that household resources are distributed equally across household members. These may be misleading if some members—such as women—have poor access to household resources. We study *resource shares*, defined as the fraction of the total expenditure of a household consumed by one of its members. Resource shares are not directly observable, but are important because unequal resource shares across household members signal within-household inequality. In this paper, we provide a new intertemporal collective household model that permits the use of short panel data to estimate resource shares within households.

Compared to many cross-sectional approaches, resource shares in our model may be arbitrarily correlated with observed variables, such as the household budget, and may depend on unobserved household-level heterogeneity, such as unobserved bargaining power shifters. Our household model embeds resource shares inside quantity demand functions, where they appear as fixed effects. The nonlinear quantity demand functions in our model are time-varying because quantity demands depend on prices, and prices are not observed but do vary across the waves of our panel.

We show point-identification of a general class of fixed-T time-varying nonlinear panel models, which includes our household model. This class has a response variable equal to a time-varying weakly monotonic transformation function of a linear index of regressors, fixed effects, and error terms. In contrast, almost all existing results for this class of models require time-invariant transformation functions. Our theorems imply novel identification results for time-varying versions of some commonly used models, e.g., the time-varying ordered logit and multiple-spell GAFT models.

We point-identify regression coefficients, transformation functions, and the mean (up to location) and variance of the distribution of fixed effects. The latter two are of specific relevance to our collective household model.

Ours are the first empirical estimates of a full-commitment collective household model in a short panel, and we demonstrate the importance of accounting for both observed and unobserved heterogeneity in women's resource shares. Using a two-period Bangladeshi panel dataset on household expenditures, we show that less than half of the variation in women's resource shares can be explained by observed covariates. This means that there is much more inequality within households than previously thought. We also find that women's resource shares are negatively correlated with household budgets, so that women in poorer households have larger resource shares. This means that women face less economic inequality than would be revealed by looking at the distribution of their household budgets.

Our micro-economic contribution begins with a new model of an efficient full-commitment intertemporal collective (FIC) household that builds on Browning et al. (2013) and Chiappori and Mazzoco (2017). Collective household models posit that household behavior is driven by the preferences of the individuals who comprise the household. In efficient models, the individuals together reach the Pareto frontier. In full-commitment intertemporal models, the individuals in the household face uncertainty, and are able to insure each other against risk by making state-contingent binding commitments over future actions. We show that quantity demand equations in efficient FIC household models must in general depend on this initial commitment—a time-invariant feature of the household.

We then give a parametric model of individual utilities in our FIC household model that delivers quantity demand equations. We model the demand for food by adult women in the household. These demand equations are time-varying monotonic nonlinear functions of a linear index of fixed effects, logged household budgets, and an error term. Here, the fixed effects have an interpretation: they equal the log of the resource share of the woman in the household. They may be correlated with household budgets, and depend on household-level unobserved heterogeneity.

Many strategies to identify resource shares with cross-sectional data require that they are conditionally independent of the household budget. We relax that restriction, and find evidence that resource shares exhibit a slight negative conditional dependence on household budgets.

Our econometric contribution is to provide sufficient conditions for point-identification in a large class of models, with the outcome Y_{it} given by:

$$Y_{it} = h_t(\alpha_i + X_{it}\beta - U_{it}) \tag{1.1}$$

for $i = 1, \ldots, n$ and time periods $t = 1, \ldots, T$ (with $T \geq 2$), where h_t is an unknown time-varying monotonic function, α_i are fixed effects, X_{it} is a vector of regressors with coefficients β, and U_{it} is an error term drawn from a stationary distribution.

Our setting has the following four features:

1. a fixed-T setting—in fact $T = 2$ is sufficient for our results;
2. the monotonic transformation function h_t can be **time-varying** and **weakly** monotonic;
3. the functions h_t, and the distribution of the error term U_{it}, may be nonparametric;
4. and, the fixed effects α_i are **unrestricted.**

We provide sufficient conditions for the identification of h_t and β for two cases, where the error term U_{it} is distributed as logistic and where it follows an unknown distribution. The panel model literature is very large, with many papers showing identification of β (and sometimes h_t) in models with two or three of these features. However, ours is the first paper to cover all four features. An immediate implication of our work is that extensions to time-varying and/or nonparametric counterparts of well-known models, such as ordered choice, censored regression, and duration modeling, can now be shown to be identified. And, of course, our collective household model is in the class and therefore identified.

For the case where h_t is strictly monotonic we provide additional results, identifying the conditional mean (up to location) and conditional variance of the distribution of fixed effects. This relates directly to our microeconomic model, because in that model fixed effects have a clear economic interpretation: they are logged resource shares. To the best of our knowledge, there are no results in the fixed-T, fixed-effects, nonlinear panel literature that cover this aspect of model identification.

In particular, we show identification of the response of the conditional mean of α_i to observed covariates, and provide additional sufficient conditions for the identification of the conditional variance of α_i. The former corresponds to identification of coefficients in the regression of logged resource shares on covariates; the latter corresponds to identification of inequality in resource shares.

Section 2 provides a review of the related literature. In Sections 3 and 4, we provide our main identification results. Section 5 introduces our collective household model, which uses data described in Section 6. We present estimates of women's resource shares in Section 7. All proofs, descriptive statistics for the data, additional robustness results and estimation details are in the Appendix.

2 Existing Literature

Following the terminology in Abrevaya (1999), we refer to the model in this paper as the *fixed-effects linear transformation* (FELT) model. Note that we can express the outcome equation in (1.1) using the latent variable notation

$$Y_{it} = h_t(Y_{it}^*) = h_t\left(\alpha_i + X_{it}\beta - U_{it}\right), \tag{2.1}$$

where Y_{it}^* is a latent variable, α_i are fixed effects and U_{it} are stationary errors.[1]

Because h_t can be *weakly* monotonic, FELT includes many previously studied discrete choice models, such as binary choice, ordered choice and censored models. When h_t is strictly monotonic, it covers other previously studied models, such as duration and Box-Cox regression models. As described below, our results provide identification of some extensions of these models that were not previously shown to be identified. Since h_t can be time-varying, our framework generalizes typical discrete- and continuous-choice models where the transformation function h_t is fixed over time.

Our approach builds on classic results in binary choice, and extends those results to the entire FELT class. FELT nests binary choice models with time effects as a special case when $h_t(Y_{it}^*) = 1\{Y_{it}^* \geq \lambda_t\}$. In these models, the parameter β is known to be identified for U_{it} parametric or nonparametric, see, e.g., Rasch (1960), Chamberlain (1980), Manski (1987), Magnac (2004), and Chamberlain (2010). We use the insights of Chamberlain (1980) and Manski (1987) about binary choice models where the error U_{it} is logistic or nonparametric, respectively, to show identification of all models nested in FELT.

Our theoretical work connects two sets of classic results with two new contributions. The first established result we invoke comes from the cross-sectional work of Doksum and Gasko (1990) and Chen (2002) that shows that transformation models can in general be *binarized* into a set of related binary choice models. The second established result we invoke comes from Chamberlain (1980) and Manski (1987) who show that fixed effects binary choice models are identified.

We begin by showing that we can binarize in a panel data setting, even if the transformation functions h_t vary with time. Given the results of Chamberlain (1980)

[1] This setting excludes some important types of models: dynamic models (e.g. Honoré and Kyriazidou (2000), Aguirregabiria et al. (ming), Khan et al. (2020)); and models of multinomial choice (e.g., Shi et al. (2018)).

and Manski (1987), each of these binary choice models is identified, yielding the (over)identification of the regression coefficient β in the FELT model. Our first contribution is to show that we can re-assemble the identified binarized models to obtain identification of the transformation functions h_t in the FELT model. Our second contribution is specific to the case where h_t is strictly monotonic. Here, we derive sufficient conditions for the identification of the conditional variance of fixed effects, and for the response of the conditional mean of fixed effects to observed covariates.

2.1 Fixed-T Nonlinear Panel Models with Fixed Effects

The literature on panel data methods is vast. There are excellent reviews of the literature, e.g., Arellano and Honoré (2001), Arellano (2003), and Arellano and Bonhomme (2011). Despite the vastness of this literature, we are not aware of any paper that delivers all four features discussed above, demanded by our empirical application.

We outline connections to the literature in the context of our list of four model features. Below, we highlight the key differences between our approach and approaches in the literature that lack one or more of our key features.

Feature 1: We show identification in fixed-T panel models. The incidental parameter problem occurs in fixed-effect panel models with a finite number of time periods, see Neyman and Scott (1948). Essentially, the problem arises from the fact that the n fixed effects α_i cannot be consistently estimated if T does not tend to infinity. Thus, identification of the parameters common across individuals must be shown in a context where the incidental parameters α_i cannot be identified or consistently estimated.

Ours is a fixed-T approach, with $n \to \infty$ and works even if $T = 2$. A large literature analyzes the behavior of fixed effects procedures under the alternative assumption that the number of time periods goes to infinity, e.g., Hahn and Newey (2004), Arellano and Hahn (2007), Arellano and Bonhomme (2009), Fernández-Val (2009), Fernández-Val and Weidner (2016), and Chernozhukov et al. (2018). In this setting, it is generally possible to identify each fixed effect, and consequently, the distribution of fixed effects. In our model, we show identification of specific moments of this distribution even though the number of time periods is fixed.

Feature 2: We allow weakly monotonic time-varying transformation functions h_t. Abrevaya (1999) provides a consistent estimator of β (the "leapfrog" estimator) in the FELT model under the restriction that the transformation functions are *strictly* monotonic.[2] However, because he differences out the transformation functions h_t, his focus does not extend to their identification. Athey and Imbens (2006) propose a "changes in changes" estimator, which is a generalization of the linear differences-in-differences estimator, in both a cross-sectional and a panel data setting. Their panel data fixed-effects setting is a potential outcomes analog to our model with strictly monotonic transformations. They show identification of the average treatment effect, but not identification of h_t or the distribution of α_i. In comparison, we cover the weakly monotone case, and identify the transformation function h_t. In the strictly monotone case, we additionally identify moments of the distribution of fixed effects α_i.

Abrevaya (2000) considers a model that allows for *weak* monotonicity but restricts the transformation functions to be time-invariant (and allows for nonseparable errors). He provides a consistent estimator for β only.[3] A literature on duration models also considers time-invariant transformation functions that are weakly monotonic due to censoring, e.g., Lee (2008), Khan and Tamer (2007), Chen (2010a,b); Chen and Zhou (2012); Chen (2012), and Chen and Zhou (2012); we review this below.

A more recent literature has focused on identification issues in a class of panel models with potentially *non-monotonic* but time-invariant structural functions (or strong assumptions on how those functions vary over time), e.g., Hoderlein and White (2012), Chernozhukov et al. (2013), Chernozhukov et al. (2015). These papers focus on (partial) identification of partial effects. But, because they don't impose monotonicity, these approaches preclude identification of the structural function(s) or of the distribution of fixed effects.

Feature 3: We allow nonparametric transformation functions and nonparametric errors. Bonhomme (2012) proposes a general-purpose likelihood-based approach to obtain identification for models with parametric h_t and parametric U_{it}, even allowing for dynamics. These results exploit the fact that a likelihood function can be constructed for such models and show identification in the presence of fixed

[2] Chen (2010c) in Remark 6 discusses a version of Abrevaya (1999) that allows for some weak monotonicity due to censoring. He focuses on β and does not discuss identification of h_t, although his Remark 1 sketches an approach for estimation of $h_t = h$ for all t.

[3] Chernozhukov et al. (2018) uses a distribution regression technique that is closely related to our binarization approach, and consequently accommodates weakly monotonic transformation functions. However, theirs is a large-T setting.

effects in a finite-T setting. (Earlier work for the same setting by Lancaster (2002) requires $T \to \infty$, sacrificing feature 1.) Our model requires strictly exogenous regressors, precluding many dynamic structures. But, our Theorems 1 and 2 apply even when h_t is nonparametric, U_{it} is nonparametric, or both are nonparametric.

The setting with parametric transformation functions and parametric errors covers many models previously shown to be identified, including the *time-invariant* fixed-effect panel versions of: binary choice (e.g., Rasch (1960), Chamberlain (1980), Magnac (2004), and Chamberlain (2010)); the linear regression model with normal errors; and the ordered logit model (e.g., Das and van Soest (1999); Baetschmann et al. (2015); Muris (2017)). Application of our results immediately shows identification of the time-varying versions of these models. This result is novel for the ordered logit model, where our results imply identification of time-varying thresholds.

Parametric transformation models with nonparametric errors are widely studied, starting with Manski (1987) for the binary choice fixed effects model. (Aristodemou (ming) provides partial identification results for ordered choice with nonparametric errors.) Parametric panel data censored regression models also fit into our framework, and were studied intensively starting with Honoré (1992) (e.g., Charlier et al. (2000), Honoré and Kyriazidou (2000), Chen (2012)). These papers show identification of the regression coefficient β for the linear model with time-invariant censoring and nonparametric errors. In this context, our results show identification of models that were not previously known to be identified. In particular, the model is identified even if the transformation function is nonparametric (as opposed to linear or Box-Cox) and time-varying and/or where the censoring cutoff is time-varying.[4]

Duration models can be recast as transformation models like ours, with nonparametric transformation functions (see Ridder (1990)). Consequently, the large literature on identification of duration models is related to our work. A very common feature in this literature is the use of error terms following the type 1 extreme value distribution (EV1).

Consider the multiple-spell mixed proportional hazards (MPH) model with spell-

[4] Many papers in the literature on censored regression have focused on endogeneity. For example, Honoré and Hu (2004) allow for endogenous covariates, and Khan et al. (2016) study the case of endogenous censoring cutoffs. Our results do not cover the case of endogenous regressors or cutoffs. Horowitz and Lee (2004) and Lee (2008) consider dependent censoring, where the censoring cutoff depends on observed covariates and the error term follows a parametric distribution. We do not consider dependent censoring.

specific baseline hazard, analyzed in Honoré (1993).[5] This model can be obtained from FELT by letting (i) $h_t^{-1}(v) = \log\left\{\int_0^v \lambda_{0t}(u)\,du\right\}$, where λ_{0t} is the baseline hazard for spell t, (ii) α_i and U_{it} are independent across t, and (iii) U_{it} is independent of X_i and distributed as EV1. Honoré (1993) derives sufficient conditions for the identification of this model (Lee (2008) provides consistent estimators under other parametric error distributions). Our theorems immediately provide the novel result that this model is identified when the error terms are drawn from a nonparametric distribution.

Consider the single-spell generalized accelerated failure time (GAFT) model introduced by Ridder (1990) (see also van den Berg (2001)) that has non EV1 errors, and is consistent with a duration model. Just like the MPH model, it can be extended to a multiple-spell setting (e.g., Evdokimov (2011)). Abrevaya (1999) shows that the common parameter vector β in the multiple-spell GAFT model is consistently estimated. However, he does not show identification of the transformation function h_t, which can be seen as dual to identification of the spell-specific baseline hazard function.[6] Evdokimov (2011) considers identification of a related version of the multiple-spell GAFT with spell-specific baseline hazard, but he requires continuity of α_i and at least 3 spells ($T \geq 3$). Our results show identification of both β and h_t in the multiple-spell GAFT model, imposing no restrictions on α_i and requiring just 2 spells ($T = 2$).

Feature 4: We allow for unrestricted fixed effects. This contrasts with identification strategies based on special regressors and with the literature on the identification of correlated random effects models. Special regressor approaches (see the review in Lewbel (2014)) have identifying power in transformation models with fixed effects. They require the availability of a continuous variable that is independent of the fixed effects. With such a variable, one can show identification of transformation models in the cross-sectional case (Chiappori et al. (2015)) and in the panel data case, e.g., Honoré and Lewbel (2002), Ai and Gan (2010), Lewbel and Yang (2016), Chen et al. (2019). Our results do not invoke a special regressor. Further, we are not aware of any special regressor-based papers that identify time-varying transformation

[5] Horowitz and Lee (2004) show identification of this model under the restriction that the baseline hazard is the same for all spells, analogous to time-invariant h_t. Chen (2010b) considers the same model, but relaxes the restriction that errors are type 1 EV, but shows identification of only the common parameter vector β.

[6] Khan and Tamer (2007) establish consistency of an estimator of the regression coefficient in GAFT under the restriction that the baseline hazard is the same for all spells, analogous to time-invariant h_t.

functions or the distribution of fixed effects.[7]

A related literature considers restrictions on the joint distribution of $(\alpha_i, X_{i1}, ..., X_{iT})$. For example, Altonji and Matzkin (2005) impose exchangeability on this joint distribution, and Bester and Hansen (2009) restricts the dependence of α_i on $(X_{i1}, ..., X_{iT})$ to be finite-dimensional. In our model this joint distribution is unrestricted.

A further group of papers establishes identification of panel models, including the distribution of α_i, by using techniques from the measurement error literature that: (i) impose various assumptions on α_i, such as full support and/or continuous distribution; (ii) assume serial independence of U_{it}; and (iii) restrict the conditional distribution of $(\alpha_i, U_{i1}, \ldots, U_{iT})$ conditional on $(X_{i1}, ..., X_{iT})$, see, Evdokimov (2010), Evdokimov (2011), Wilhelm (2015), and Freyberger (2018). In contrast, our results on the identification of the conditional variance of α_i do not require (i). All our other results, including identification of the dependence of α_i on observed covariates, are free of assumptions like (i), (ii) and (iii).

We also show identification of some aspects of the distribution of fixed effects. As we noted above, correlated random effects models identify the distribution of individual effects, but at the cost of restricting their distribution. To our knowledge, we are the first to show identification of moments of this distribution in a nonlinear panel model, when that distribution is unrestricted.

We show the practical importance of these innovations in our empirical work below. Identification of the conditional mean and variance of the distribution of fixed effects in a context with time-varying transformation functions is essential to our investigation of women's access to household resources in rural Bangladeshi.

2.2 Microeconomic Models of Collective Households

Dating back at least to Becker (1962), collective household models are those in which the household is characterized as a collection of individuals, each of whom has a well-defined objective function, and who interact to generate household level decisions such as consumption expenditures. *Efficient collective household* models are those in which the individuals in the household are assumed to reach the (household) Pareto frontier.

[7] We conjecture that the existence of a special regressor would be sufficient to identify time-varying nonparametric transformation functions, and, with strict monotonicity, the distribution of fixed effects. However, we think that a setting with completely unrestricted fixed effects is useful in a variety of empirical applications, including our own.

Chiappori (1988, 1992) showed that, like in earlier results in general equilibrium theory, the assumption of Pareto efficiency is very strong. Essentially, it implies that the household can be seen as maximizing a weighted sum of individual utilities, where the weights are called *Pareto weights*. This in turn implies that the household-level allocation problem is observationally equivalent to a decentralized, person-level, allocation problem.

In this decentralized allocation, each household member is assigned a *shadow budget*. They then demand a vector of consumption quantities given their preferences and their personal shadow budget, and the household purchases the sum of these demanded quantities (adjusted for shareability/economies of scale and for public goods within the household).

Resource shares, defined as the ratio of each person's shadow budget to the overall household budget, are useful measures of individual consumption expenditures. If there is intra-household inequality, these resource shares would be unequal. Consequently, standard per-capita calculations (assigning equal resource shares to all household members) would yield invalid measures of individual consumption and poverty (see, e.g., Dunbar et al. (2013)). In this paper, we show identification of the conditional mean (up to location) and conditional variance of the distribution of resource shares in a panel data context.

The early literature on these models, including Bourguignon et al. (1993); Browning and Chiappori (1998); Vermeulen (2002); Chiappori and Ekeland (2006), constrains goods to be either purely private or purely public within a household. These papers show that one can generally identify the response of resource shares to changes in observed variables such as distribution factors. Like those papers, we can identify the response of resource shares to observed variables, but we also can account for unobserved household-level variables through the inclusion of fixed effects.

We work with a more general model of sharing and scale economies based on Browning et al. (2013). This model allows some or all goods to be partly or fully shared, and the authors show that there is a one-to-one correspondence between Pareto weights and resource shares. Dunbar et al. (2013) use this model, and show how assignable goods, defined as goods consumed exclusively by a single known household member, may be used to identify resource shares (see also Chiappori and Ekeland (2009)). Like them, we use an assignable good to support identification.

A key identifying assumption in Dunbar et al. (2013) is that resource shares are independent of household budgets in a cross-sectional sense. This identifying restriction has been used to estimate resource shares, within-household inequality and individual-level poverty in many countries (DLP and DLP2 in Malawi; Bargain et al. (2014) in Cote D'Ivoire; Calvi (2019) in India; Vreyer and Lambert (2016) in Senegal; Bargain et al. (2018) in Bangladesh). In our model, we show identification of the response of the conditional mean of resource shares to observed covariates, even if resource shares are correlated with (lifetime) household budgets. Consequently, we can test this identifying restriction.

Using cross-sectional data, Menon et al. (2012) and Cherchye et al. (2015) test the restriction that resource shares are correlated with household budgets, and find no evidence that they are. But, their estimators don't have much power. In our empirical work with panel data, we get a fairly precise estimate of this correlation, allowing us to detect even a small dependence. We find evidence that women's resource shares are slightly negatively correlated with household budgets (conditional on other observed variables). This finding suggests that the cross-sectional identification strategy proposed by Dunbar et al. (2013) may come at a cost that is not faced in a panel data setting.

Dunbar et al. (2013) does not accommodate unobserved heterogeneity in resource shares. Two newer papers, Chiappori and Kim (2017) and Dunbar et al. (2019) consider identification in cross-sectional data with unobserved household-level heterogeneity in resource shares. Like Chiappori and Kim (2017) and Theorem 1 in Dunbar et al. (2019), our work investigates identification of the distribution of resource shares up to an unknown normalization. However, the results in those papers are of the random effects type. That is, the authors impose the restriction that the conditional distribution of resource shares is independent of the household budget. In this paper, we consider a panel data setting with household-level unobserved heterogeneity in resource shares, without any restriction on the distribution of resource shares. Further, we show sufficient conditions for identification of the conditional variance of (logged) resource shares.

Sokullu and Valente (2019) use a one-period micro-economic model similar to Dunbar et al. (2013), and estimate it on three waves of a Mexican panel dataset. In contrast, our micro-economic model considers choice in the presence of unobserved household-level heterogeneity, and over many periods under uncertainty.

The literature cited above considered one-period micro-economic models. But, many interesting questions about households, and the distribution of resources within households, are dynamic in nature. For example: how do household members share risk?; how do household investments relate to individual consumption?; how can we use information from multiple time periods to estimate resource shares when there is unobserved household-level heterogeneity?

Chiappori and Mazzoco (2017) give a lovely review of the literature on collective household models in an intertemporal setting. These models generally come in two flavours--limited commitment or full commitment---depending on whether or not the household can commit to a permanent Pareto weight at the moment of household formation. Full-commitment models answer "yes", and limited-commitment model answer "no". Limited commitment models have commanded the most theoretical attention. Much effort has gone into testing the full-commitment model against a limited-commitment alternative, e.g., Ligon (1998); Mazzocco (2007); Mazzocco et al. (2014); Voena (2015).

Fewer papers study the identification of Pareto weights or resource shares in an intertemporal context. Lise and Yamada (2019) use a long panel of Japanese household consumption data to estimate how Pareto weights (which are dual to resource shares) depend on observed covariates and on unanticipated shocks. They find evidence that the full-commitment model does not hold in Japan. Our model is one of full-commitment and our data are a short (2 period) panel, so we provide a complement to the approach of Lise and Yamada (2019) for cases where the data are not rich enough to estimate a limited-commitment model.

Full commitment models are more restrictive, but may be useful nonetheless. Chiappori and Mazzoco (2017) write "In more traditional environments (such as rural societies in many developing countries), renegotiation may be less frequent since the cost of divorce is relatively high, threats of ending a marriage are therefore less credible, and noncooperation is less appealing since households members are bound to spend a lifetime together." We use a full commitment setting to estimate resource shares for rural Bangladeshi households.

In this paper, we adapt the general full-commitment framework of Chiappori and Mazzoco (2017) to the scale economy and sharing model of Browning et al. (2013). Then, like Dunbar et al. (2013) do in their cross-sectional analysis, we identify resource shares on the basis of household-level demand functions for assignable goods. In our general model, observed household-level quantity demand functions depend on resource shares, and resource shares depend a time-invariant factor (a fixed effect) representing the initial (and permanent) Pareto weights of household members.

We then provide a parametric form for utility functions that results in demand equations for assignable goods that are nonlinear in shadow budgets, and have logged shadow budgets that are linear in logged household budgets and a fixed effect. Further, demand equations are time-varying because prices vary over time. Such demand equations fall into the FELT class, and are therefore identified in our short-panel setting. The parametric model also gives meaning to the fixed effect: it equals a logged resource share, so its distribution is an economically interesting object. So, our microeconomic theory demands an econometric model that allows for time-varying transformation functions and that can identify moments of the conditional distribution of fixed effects.

3 Identification

Dropping the i subscript, let $Y = (Y_1, ..., Y_T)'$ and $X = (X_1', ..., X_T')'$. We write FELT as a latent variable model using the notation in (2.1). For $t = 1, ..., T$ and for all α and X,

$$\begin{aligned} Y_t &= h_t(Y_t^*) = h_t\left(\alpha + X_t\beta - U_t\right), \\ U_t|\alpha, X &\sim F_t(u|\alpha, X). \end{aligned} \tag{3.1}$$

Denote the supports of Y_t, Y_t^*, X_t by $\mathcal{Y} \subseteq \mathbb{R}$, $\mathcal{Y}^* = \mathbb{R}$, and $\mathcal{X} \subseteq \mathbb{R}^K$, respectively.[8]

We provide sufficient conditions for identification of (β, h_t).[9] We consider two non-nested cases. The first case allows for nonparametric $F_t(u|\alpha, X)$, requiring only that it is conditionally stationary. In this case, the idiosyncratic errors may be serially dependent. The second case assumes that U_t, $t = 1, \cdots, T$, are serially independent, standard logistic, and strictly exogenous. Of course, the second case has an error distribution that is nested within that of the first case. But, the second case requires weaker assumptions on the distribution of the regressors (c.f. Assumption 3 below). For both cases, we maintain the assumption below:

[8] The supports may be indexed by t. We omit this index here for the sake of concise notation.

[9] Botosaru and Muris (2017) introduce four estimators, depending on whether the outcome variable is discrete or continuous, and on whether the stationary distribution of the error term is nonparametric or logistic.

Assumption 1. *[Weak monotonicity] For each t, the transformation function $h_t : \mathcal{Y}^* \to \mathcal{Y}$ is unknown, non-decreasing, right continuous, and non-degenerate.*

Define the generalized inverse $h_t^- : \mathcal{Y} \to \mathcal{Y}^*$ as

$$h_t^-(y) \equiv \inf\left\{y^* \in \mathcal{Y}^* : \; y \le h_t(y^*)\right\},$$

with the convention that $\inf(\emptyset) = \inf(\mathcal{Y})$. Additionally, let $\underline{\mathcal{Y}} \equiv \mathcal{Y} \setminus \inf \mathcal{Y}$. For an arbitrary $y \in \underline{\mathcal{Y}}$, define the binary random variable

$$\begin{aligned} D_t(y) &\equiv 1\{Y_t \ge y\} \\ &= 1\left\{U_t \le \alpha + X_t\beta - h_t^-(y)\right\}, \end{aligned} \tag{3.2}$$

where the equality follows from specification ((3.1)) and weak monotonicity. Here, we use $\underline{\mathcal{Y}}$ instead of $\mathcal{Y}$ because $D_t(\inf \mathcal{Y}) = 1$ almost surely for all t.

The key insight of our identification argument is to allow the threshold y in (3.2) to be different across time periods, in addition to allowing it to vary across $\underline{\mathcal{Y}}$, the support of the observed outcome. We thus compare outcomes observed at $t = 1$ with threshold y_1 and outcomes observed at $t = 2$ with threshold y_2, where $y_1 \ne y_2$ and $(y_1, y_2) \in \underline{\mathcal{Y}}^2$. This allows us to group individuals into *switchers* and *non-switchers*, where an individual is a switcher provided that $D_1(y_1) + D_2(y_2) = 1$. It is the existence of switchers that informs our identification of the time-varying transformation function.[10] In other words, in comparison with previous results where the threshold is the same in each time period, using different thresholds in each time period reveals new information about the response functions. It is this new information that enables us to identify the time-dependence of h_t.

3.1 Identification strategy: binarization

Two time periods are sufficient for our identification results, so we let $T = 2$ in what follows.

[10] To the best of our knowledge, previous papers, see, e.g., Chen (2002), Chen (2010a); Chernozhukov et al. (2018), restrict the two thresholds to be equal to each other. This is relevant since this restriction essentially prevents identification of time-effects or time-varying transformation functions h_t.

For any two points $(y_1, y_2) \in \underline{\mathcal{Y}}^2$, define the following vector of binary variables

$$D(y_1, y_2) \equiv (D_1(y_1), D_2(y_2)).$$

Our identification strategy for (β, h_1, h_2), which we call binarization, is based on the observation that the 2-vector $D(y_1, y_2)$ follows a panel data binary choice model for *any* $(y_1, y_2) \in \underline{\mathcal{Y}}^2$. This result is summarized in Lemma 1 below.

The identification proof proceeds in three steps. First, we show identification of β and of $h_2^-(y_2) - h_1^-(y_1)$ for arbitrary $(y_1, y_2) \in \underline{\mathcal{Y}}^2$. In the resulting binary choice model, the difference $h_2^-(y_2) - h_1^-(y_1)$ is the coefficient on the differenced time dummy, and β is the regression coefficient on $X_2 - X_1$. For a given binary choice model, identification of β and of $h_2^-(y_2) - h_1^-(y_1)$ follows Manski (1987) for the nonparametric version of our model, and Chamberlain (2010) for the logistic version. This result is summarized in Theorem 1 below.

Second, we show that varying the pair (y_1, y_2) over $\underline{\mathcal{Y}}^2$ obtains identification of

$$\left\{h_2^-(y_2) - h_1^-(y_1),\, (y_1, y_2) \in \underline{\mathcal{Y}}^2\right\}.$$

Third, we show that identification of this set of differences obtains identification of the functions h_1 and h_2 under a normalization assumption on h_1^-. That is, for an arbitrary $y_0 \in \underline{\mathcal{Y}}$, $h_1^-(y_0) = 0$. This type of assumption is customarily made in the literature on transformation models. This result is presented in Theorem 2.

In summary, we show that FELT can be converted into a collection of binary choice models, which allows us to identify the transformation functions h_t. Omitting the fact that FELT can be transformed into *many* binary choice models obtains identification of β only.

Figures 3.1 and 3.2 illustrate the intuition behind our identification strategy for two arbitrary functions, h_1 and h_2, both accommodated by FELT. The line with kinks and a flat part represents an arbitrary function h_1, while the solid curve represents an arbitrary function h_2. Consider Figure 3.1. Pick a $y_1 \in \underline{\mathcal{Y}}$ on the vertical axis. For all $y \leq y_1$, $h_1(y)$ gets mapped to zero, while for all $y > y_1$, it gets mapped to one. Now pick a $y_2 \in \underline{\mathcal{Y}}$. For all $y \leq y_2$, $h_2(y)$ gets mapped to zero, while for all $y > y_2$, it gets mapped to one. This gives rise to a fixed effects binary model for $(D_1(y_1), D_2(y_2))$, also plotted in the figure as the grey solid lines. Our first result

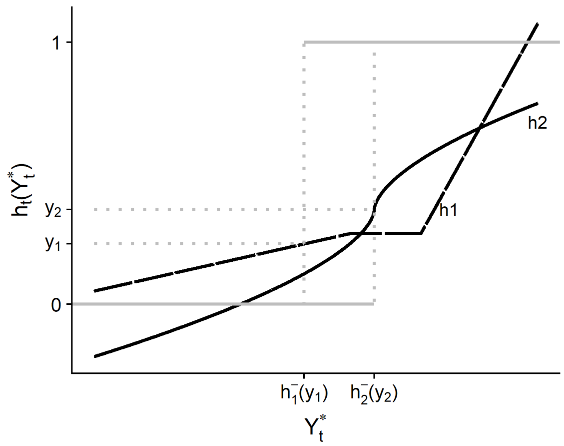

Figure 3.1: FELT functions h_1 and h_2.

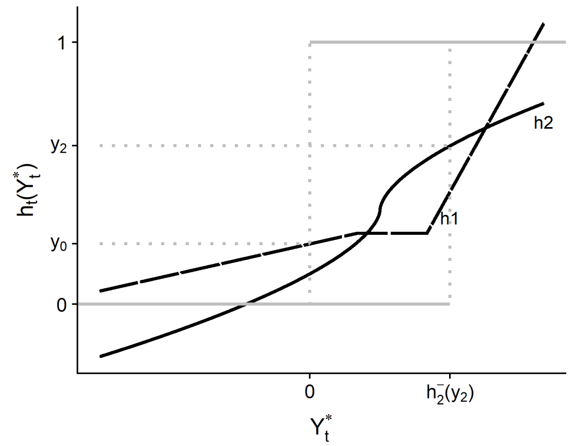

Figure 3.2: Normalization and tracing.

in Theorem 1 identifies the difference $h_1^-(y_1) - h_2^-(y_2)$ at arbitrary points (y_1, y_2), as well as the coefficient β. It is clear that normalizing $h_1^-(.)$ at an arbitrary point identifies the function $h_2^-(y_2)$ at an arbitrary point y_2. This is captured in Figure 3.2. There, for an arbitrary y_0, $h_1^-(y_0) = 0$. Then, as y_2 is arbitrary, Figure 3.2 shows that moving y_2 it on its support traces out the generalized inverse h_2^- on its domain. Theorem 2 wraps up this argument by showing that h_1 and h_2 are identified from their generalized inverses.

3.2 Nonparametric errors

In this section, we provide nonparametric identification results for (β, h_1, h_2). Parts of our identification proof build on Manski (1987), who in turn builds on Manski (1975, 1985).

Assumption 2. *[Error terms]*

(i) $F_1(u|\alpha, X) = F_2(u|\alpha, X) \equiv F(u|\alpha, X)$ *for all* (α, X)*;*

(ii) The support of $F(u|\,\alpha, X)$ *is* $\mathbb{R}$ *for all* (α, X)*.*

Assumption 2 places no parametric distributional restrictions on the distribution of U_{it} and allows the stochastic errors U_{it} to be correlated across time. The first part of the assumption, 2(i), is a stationarity assumption, requiring time-invariance of the distribution of the error terms conditional on the trajectory of the observed regressors and on the unobserved heterogeneity. This assumption excludes lagged dependent variables as covariates. Additionally, as noted by, e.g., Chamberlain (2010), although it allows for heteroskedasticity, it restricts the relationship between the observed regressors and U_{it} by requiring that even when $x_1 \neq x_2$, U_1 and U_2 have equal skedasticities. This type of stationarity assumption is common in linear and nonlinear panel models, e.g., Chernozhukov et al. (2013) and references therein.

Assumption 2(ii) requires full support of the error terms. It guarantees that, for any pair $(y_1, y_2) \in \underline{\mathcal{Y}}^2$, the probability of being a *switcher* is positive. In our context, being a switcher refers to the event $D_1(y_1) + D_2(y_2) = 1$, so that Assumption 2 guarantees that $P(D_1(y_1) + D_2(y_2) = 1) > 0$. This assumption is similar to Assumption 1 in Manski (1987).

Let $\Delta X \equiv X_2 - X_1$ and for an arbitrary pair $(y_1, y_2) \in \underline{\mathcal{Y}}^2$, define

$$\gamma(y_1, y_2) \equiv h_2^-(y_2) - h_1^-(y_1). \tag{3.3}$$

Lemma 1. *Suppose that (Y, X) follows the model in (3.1). Let Assumptions 1 and 2 hold. Then for all $(y_1, y_2) \in \underline{\mathcal{Y}}^2$,*

$$\operatorname{med}(D_2(y_2) - D_1(y_1) \,|X,\ D_1(y_1) + D_2(y_2) = 1) = \operatorname{sgn}(\Delta X\beta - \gamma(y_1, y_2)). \tag{3.4}$$

Proof. The proof builds on Manski (1985, 1987), and is presented in Appendix A.1. $\square$

Let $W \equiv (\Delta X, -1)'$ and $\theta(y_1, y_2) \equiv (\beta, \gamma(y_1, y_2))$, so that (3.4) can be written as

$$\operatorname{med}(D_2(y_2) - D_1(y_1) \,|X, D_1(y_1) + D_2(y_2) = 1) = \operatorname{sgn}(W\theta(y_1, y_2)).$$

For identification of $\theta(y_1, y_2)$ we impose the following additional assumptions.

Assumption 3. *[Covariates]*

(i) The distribution of ΔX is such that at least one component of ΔX has positive Lebesgue density on $\mathbb{R}$ conditional on all the other components of ΔX with probability one. The corresponding component of β is non-zero.

(ii) The support of W is not contained in any proper linear subspace of $\mathbb{R}^{K+1}$.

Assumption 3(i) requires that the change in one of the regressors be continuously distributed conditional on the other components. Assumption 3(ii) is a full rank assumption. These assumptions are standard in the binary choice literature concerned with point identification of the parameters.

Assumption 3 resembles Assumption 2 in Manski (1987), the difference being that our assumption concerns W, which includes a constant that captures a time trend. The presence of this constant requires sufficient variation in X_t over time. No linear combination of the components of X_t can equal the time trend.

Assumption 4. *[Normalization-β] For any $(y_1, y_2) \in \underline{\mathcal{Y}}^2$, $\theta(y_1, y_2) \in \Theta = \mathcal{B} \times \mathbb{R}$, where $\mathcal{B} = \left\{\beta : \beta \in \mathbb{R}^K, \|\beta\| = 1\right\}$.*

Assumption 4 imposes a normalization on β, namely that the norm of the regression coefficient equals 1. Scale normalizations are standard in the binary choice literature, and are necessary for point identification when the distribution of the error terms is not parameterized. Normalizing β (instead of θ) avoids a normalization that would otherwise depend on the choice of (y_1, y_2). In this way, the scale of β remains constant across different choices of (y_1, y_2). Alternatively, one can normalize the coefficient on the continuous covariate (cf. Assumption 3(i)) to be equal to one. In our economic model in Section 5 the latter assumption holds automatically.[11]

Theorem 1. *Suppose that (Y, X) follows the model in (3.1), and let the distribution of (Y, X) be observed. Let Assumptions 1, 2, 3, and 4 hold. Then, for an arbitrary pair $(y_1, y_2) \in \underline{\mathcal{Y}}^2$, $\theta(y_1, y_2)$ is identified.*

Proof. The proof proceeds by showing that FELT can be converted into a binary choice model for an arbitrary pair (y_1, y_2), and then builds on Theorem 1 in Manski (1987), which in turn uses results in Manski (1985). See Appendix A.2. $\square$

So far, we have identified the regression coefficient β and the difference in the generalized inverses at arbitrary pairs $(y_1,\, y_2)$. We consider now identification of the functions h_1 and h_2 on $\underline{\mathcal{Y}}$.

Assumption 5. *[Normalization-h_1] For some $y_0 \in \underline{\mathcal{Y}}$, $h_1^-(y_0) = 0$.*

Such a normalization is standard in transformation models, see, e.g., Horowitz (1996). Without this normalization, all identification results hold up to $h_1^-(y_0)$. We normalize the function in the first time period only, imposing no restrictions on the

[11] There are models with sufficient structure on the transformation function h_t where identification is possible without a normalization on the regression coefficient. Examples include the linear regression model, the censored linear regression model in Honoré (1992), and the interval-censored regression model in Abrevaya and Muris (2020).

function in the second period beyond that of weak monotonicity (cf. Assumption 1). In Section 4, we show that this normalization assumption is not necessary for the identification of the conditional mean or of the conditional variance of the fixed effects conditional on the observed covariates.

Theorem 2. *Suppose that (Y, X) follows the model in (3.1), and let the distribution of (Y, X) be observed. Under Assumptions 1, 2, 3, 4, and 5, the transformation functions h_1 and h_2 are identified.*

Proof. The proof proceeds by identifying the generalized inverses of monotone functions, which obtains identification of the pre-images of h_1 and h_2. This obtains identification of the functions themselves. See Appendix A.3. $\square$

3.3 Logit errors

In this section, we show identification of (β, h_1, h_2) when the error terms are assumed to follow the standard logistic distribution. The logistic case is not nested in the nonparametric case. In particular, when the errors are logistic, we do not require a continuous regressor. However, we require conditional serial independence of the error terms.[12]

Assumption 6. *[Logit] (i) $F_1(u|\alpha, X) = F_2(u|\alpha, X) = \Lambda(u) = \frac{\exp(u)}{1+\exp(u)}$, and U_1 and U_2 are independent; (ii) $E(W'W)$ is invertible.*

Assumption 6(i) strengthens Assumption 2 by requiring the errors to follow the standard logistic distribution and to be serially independent. Note that one consequence of this assumption, which specifies the variance of the error terms to be equal to 1, is to eliminate the need to normalize β. On the other hand, Assumption 6(ii) imposes weaker restrictions on the observed covariates relative to Assumption 3, since it does not require the existence of a continuous covariate. Sufficient variation in ΔX is sufficient to obtain identification of the vector β when the error terms follow the standard logistic distribution.

Theorem 3. *Suppose that (Y, X) follow the model in (3.1), and let the distribution of (Y, X) be observed. Let Assumptions 1 and 6 hold. Then, for an arbitrary pair*

[12] See Chamberlain (2010) and Magnac (2004) for more details about identification under nonparametric versus logistic errors in the panel data binary choice context.

$(y_1, y_2) \in \underline{\mathcal{Y}}^2$, $\theta(y_1, y_2)$ *is identified. Additionally, letting Assumption 5 hold, then the transformation functions* $h_1(\cdot)$ *and* $h_2(\cdot)$ *are identified.*

Proof. See Appendix A.4. □

4 Conditional distribution of fixed effects

If (h_1, h_2) are invertible, we can use the previous identification theorem to identify features of the distribution of the fixed effects conditional on observed regressors. These features are the change in the conditional mean function of α and the conditional variance of α conditional on X_1, X_2. These results are relevant since in our collective household model, the fixed effects represent the log of resource shares, and both the standard deviation of these resource shares and the response of their conditional mean to covariates are key parameters of interest in the empirical literature. As this is relevant to our application, we note here that a normalization assumption, such as 4, on the demand function in the first period is not necessary for these results on the resource shares because, e.g., we only need their deviation with respect to the mean of the fixed effects.

In this section, we provide sufficient conditions for the identification of the change in the conditional mean function of the fixed effects, defined as:

$$\mu(x) \equiv E[\alpha | X = x], \text{ for all } x \in \mathcal{X}, \tag{4.1}$$

as well as for the conditional variance of the fixed effects. For these results, the normalization assumption 4 is not necessary. To provide intuition for this, let

$$c_1 \equiv h_1^{-1}(y_0),$$

at an arbitrary $y_0 \in \underline{\mathcal{Y}}$ and $g_t(y) \equiv h_t^{-1}(y) - c_1$ for all $y \in \underline{\mathcal{Y}}$. Note that Theorem 1 recovers

$$\widetilde{U}_t \equiv \alpha - U_t = h_t^{-1}(Y_t) - X_t\beta,$$

up to c_1, so that the joint distribution of $\left(\widetilde{U}_1, \widetilde{U}_2, X_1, X_2\right)$ is identified up to c_1. By placing restrictions on the distribution of $(\alpha, U_1, U_2, X_1, X_2)$, we can then recover our features of interest.

Theorem 4. *(i) Let the assumptions of Theorem 1 hold, and additionally assume that (4a) (h_1, h_2) are strictly increasing, and (4b) let $m \in \mathbb{R}$ be an unknown constant such that $E(U_t| X = x) = m$, for all $x \in \mathcal{X}$. Then, for any $x, x' \in \mathcal{X}$, the change in the conditional mean function $\mu(x) - \mu(x')$ is identified and given by*

$$\mu(x) - \mu\left(x^{'}\right) = E\left[g_t(Y_t) - X_t\beta\right| X = x] - E\left[g_t(Y_t) - X_t\beta\right| X = x^{'}\Big].$$

Proof. See Appendix A.5. □

Remark 1. As opposed to our main identification result in Theorem 2, Theorem 4 does not use a normalization on the functions (h_1, h_2). If we were to impose the normalization in Assumption 5, the conditional mean function $\mu(x)$ would be identified for all $x \in \mathcal{X}$. This result provides justification for nonparametric regression of α on observables (up to location).

Remark 2. Under slightly weaker conditions, we can obtain the projection coefficients of α on X_t. This is of interest for our empirical application. Recall that the joint distribution of $\left(\widetilde{U}_1, \widetilde{U}_2, X_1, X_2\right)$ is identified up to c_1. Then, assuming $\mathrm{Cov}(U_s, X_t) = 0$, we can identify the projection coefficient of α on X_t from

$$[\mathrm{Var}(X_t)]^{-1}\mathrm{Cov}(\alpha, X_t) = [\mathrm{Var}(X_t)]^{-1}\mathrm{Cov}(\alpha - U_s, X_t) = [\mathrm{Var}(X_t)]^{-1}\mathrm{Cov}\left(\widetilde{U}_s, X_t\right).$$

Second, define the conditional variance of the fixed effects as

$$\sigma_\alpha^2(x) \equiv Var[\alpha| X = x], \text{ for all } x \in \mathcal{X}. \tag{4.2}$$

For this second result, we strengthen our assumptions to include, among others, serial independence of the error term. This allows us to pin the persistence in unit i's time series on α_i instead of on serial dependence in the errors.

Theorem 5. *Let the assumptions of Theorem 1 and assume that (5a) (h_1, h_2) are strictly increasing, and (5b) $Cov[\alpha, U_t| X = x] = 0$ for all $x \in \mathcal{X}$ and t, and (5c) $Cov[U_1, U_2| X = x] = 0$ for all $x \in \mathcal{X}$. Then for all $x \in \mathcal{X}$, the conditional variance function $\sigma_\alpha^2(x)$ is identified and given by:*

$$\sigma_\alpha^2(x) = Cov\left(g_2(Y_2) - X_2\beta, g_1(Y_1) - X_1\beta\right| X = x).$$

Proof. See Appendix A.5. □

It may be possible to obtain the entire conditional distribution of the fixed effects under the assumption that (α, U_1, U_2) are mutually independent by using arguments similar to those in Arellano and Bonhomme (2012).

5 Microeconomic model

In this section, we construct a new model of an efficient full-commitment intertemporal collective (FIC) household. Essentially, we combine the models of Browning et al. (2013) and Chiappori and Mazzoco (2017) to generate an empirically practical model that allows identification of resource shares. Chiappori and Mazzoco (2017) write their model in terms of pure public and pure private goods. We instead adapt that model to the more general sharing model given in the collective household model of Browning et al. (2013).

A feature of efficient models like this is that the household-level problem can be decentralized into an observationally equivalent set of individual decision problems. Each individual problem is to choose demands based on an individual-level constraint defined by a shadow price vector and a shadow budget constraint.

We use subscripts i, j, t. Let $i = 1, ..., n$ index households and assume the household has a time-invariant composition, with N_{ij} members of type j. Let $j = m, f, c$ for men, women and children. Let $t = 1, 2$. Let z be a vector of time-varying household-level demographic characteristics, and let the numbers of household members of each type, N_{im}, N_{if} and N_{ic}, be (time-invariant) elements of z_{it}. Like Chiappori and Mazzoco (2017), this is a model with uncertainty, so we use the superscript $s = 1, 2$ to index states in the second period only.

Indirect utility, $V_j(p, x, z)$, is the maximized value of utility given a budget constraint defined by prices p and budget x, given characteristics z. Let V_j be strictly concave in the budget x. Indirect utility depends on time only through its dependence on the budget constraint and time-varying demographics z. Let $v_{ijt} \equiv V_j(p_t, x_{it}, z_{it})$ denote the utility level of a person of type j in household i in period t.

Browning et al. (2013) model sharing and household scale economies via a household consumption function that reflects the fact that shareable goods feel "cheap" within the household. This is embodied in a shadow price vector for consumption within the household that is weakly smaller than the market price vector p_t faced by single individuals, because singles cannot take advantage of scale economies in household consumption. For example, goods that are not shareable at all—for which there are no scale economies in household consumption—have shadow prices equal to the market price. Goods that are fully shareable, so that each person in the household can enjoy an effective consumption equal to the amount purchased by the household, have a shadow price equal to the market price divided by the number of members.

Let $A_{it} \equiv A(z_{it})$ be a diagonal matrix that gives the shareability of each good, and let it depend on demographics z_{it} (including the numbers of household members). For nonshareable goods, the corresponding element of A_{it} equals 1; for shareable goods, it is less than 1, possibly as small as $1/N_i$ where N_i is the number of household members. Goods may be partly shareable, with an element of A_{it} between $1/N_i$ and 1. With market prices p_t, within-household shadow prices are given by the linear transformation $A_{it}p_t$. Shadow prices are the same for all household members j.

Browning et al. (2013) also allow for inequality in the distribution of household resources. Let η_{ijt} be the *resource share* of type j in household i in time period t. It gives the fraction of the household budget consumed by that type. Each person of the N_{ij} people of type j consumes η_{ijt}/N_{ij} of the household budget x_{it}, so they each have a budget of $\eta_{ijt}x_{it}/N_{ij}$. As we will see below, the resource share is a choice variable for the household.

The resource shares and shadow price vector together define the decentralized shadow budget constraints faced by each household member. The model has each household member facing a shadow budget of $\eta_{ijt}x_{it}/N_{ij}$ and shadow prices of $A_{it}p_t$, so that, within the household, utility v_{ijt} is given by

$$v_{ijt} = V_j(A_{it}p_t, \eta_{ijt}x_{it}/N_{ij}, z_{it}). \tag{5.1}$$

Let $V_{xj}(p, w, z) \equiv \partial V_j(p, w, z)/\partial w$ be the monotonically decreasing marginal utility of person j with respect to their (shadow) budget. Then,

$$V_{xj}(A_{it}p_t, \eta_{ijt}x_{it}/N_{ij}, z_{it})$$

is the value of their marginal utility evaluated at their shadow budget constraint.

Let $p_2^s, z_{i2}^s, \overline{x}_i^s$ for $s = 1, 2$ be the possible realizations of state-dependent variables that occur with household-specific probabilities π_i^1 and π_i^2 (which sum to 1). Here, $\overline{x}_i^s$ is the state-specific lifetime wealth of household i, revealed in period 2. Information about the joint distribution of these unobserved state-dependent variables is embodied in Φ_i, the information set available in period 1 to household i. The information set can also include unobserved time-invariant features of the household. Pareto weights $\phi_{ij} = \phi_j(\Phi_i)$ depend on the information set Φ_i, which varies arbitrarily across households. This household-level time-invariant variable will form the basis of our fixed-effects variation.

Chiappori and Mazzoco (2017) let individuals have expected lifetime utilities given by the sum of period 1 utility and the discounted probability-weighted sum of state-specific period 2 utility.[13] Using the Pareto weights, they then write the Bergson-Samuelson Welfare Function, W_i, for the household as

$$W_i \equiv \sum_{j=1}^{J} N_{ij}\phi_{ij}\left[v_{ij1} + \sum_{s=1}^{2}\rho_i\pi_i^s v_{ij2}^s\right]. \tag{5.2}$$

The term in square brackets is the expected lifetime utility of each member of type j in household i. Each member of type j gets the Pareto weight $\phi_{ij} = \phi_j(\Phi_i)$.

Next, substitute indirect utility (5.1) for utility v_{ij1} and v_{ij2}^s into (5.2),[14] and form the Lagrangian using the intertemporal budget constraint with interest rate τ, $x_{i1} + x_{i2}^s/(1+\tau) = \overline{x}_i^s$, and the adding-up constraints on resource shares, $\sum_j \eta_{ij1} =$

[13]If the membership of the household changed over time, or if the household was choosing its membership, we would need a household welfare function that used some kind of population ethics principle (see Blackorby et al. (2005)). It is for this reason that we focus on households with fixed membership.

Here, we only consider egotistic preferences. However, this is without loss of generality: "It is important to point out, however, that the model with egotistical preferences ... plays a special role. The reason for this is that the solution to the collective model with caring preferences must also be a solution of the collective model ... with egotistical preferences." (Chiappori and Mazzoco (2017), page 21).

The restriction there there are only two periods and only two states is for convenience. None of our conclusions about resource shares depend on it.

[14]In contrast, Chiappori and Mazzoco (2017) substitute *direct* utility for utility v_{ij1} and v_{ij2}^s using a model of pure private and pure public goods. In that model, each individual's utility is given by their direct utility function, which is a function of their (unobserved) consumption of a vector of private goods and their (observed) consumption of a vector of public goods.

$\sum_j \eta_{ij2}^s = 1$. Each household i chooses x_{i1}, η_{ij1} and η_{ij2}^s to maximize

$$W_i = \sum_{j=1}^{J} N_{ij}\phi_{ij} \left[V_j(A_{i1}p_1, \eta_{ij1}x_{i1}/N_{ij}, z_{i1}) + \sum_{s=1}^{2} \rho_i \pi_i^s V_j(A_{i2}p_2^s, \eta_{ij2}^s x_{i2}^s/N_{ij}, z_{i2}^s) \right]$$

$$- \sum_{s=1}^{2} \kappa^s \left(x_{i1} + x_{i2}^s/(1+\tau) - \overline{x}_i^s\right) - \lambda_1 \left(\sum_{j=1}^{J} \eta_{ij1} - 1 \right) - \sum_{s=1}^{2} \lambda_2^s \left(\sum_{j=1}^{J} \eta_{ij2}^s - 1 \right).$$

Because the optimand is additively separable across periods, it can be thought of as a two-stage budgeting problem, where the household first chooses the period budgets, x_{i1} and x_{i2}^s, and then chooses resource shares conditional on this allocation of budget. First-order conditions for η_{ijt} in each period and each state are given by:

$$\phi_{ij} V_{xj}(A_{i1}p_1, \eta_{ij1}x_{i1}/N_{ij}, z_{i1}) x_{i1} - \lambda_1 = 0,$$
$$\rho_i \phi_{ij} V_{xj}(A_{i2}p_2^s, \eta_{ij2}^s x_{i2}^s/N_{ij}, z_{i2}^s) x_{i2}^s - \lambda_2^s = 0,$$

for $s = 1, 2$. Thus for any two types, j and k, we have the following equality:

$$\frac{\phi_{ij}}{\phi_{ik}} = \frac{V_{xk}(A_{i1}p_1, \eta_{ik1}x_{i1}/N_{ik}, z_{i1})}{V_{xj}(A_{i1}p_1, \eta_{ij1}x_{i1}/N_{ij}, z_{i1})} = \frac{V_{xk}(A_{i2}p_2^s, \eta_{ik2}^s x_{i2}^s/N_{ik}, z_{i2}^s)}{V_{xj}(A_{i2}p_2^s, \eta_{ij2}^s x_{i2}^s/N_{ij}, z_{i2}^s)}, \tag{5.3}$$

for $s = 1, 2$ for all i. That is, the household chooses resource shares so as to equate ratios of marginal utilities with ratios of Pareto weights. There is a unique solution to this problem because each person j has a utility function strictly concave in the shadow budget.

Resource shares are implicitly determined by (5.3) and depend on the Pareto-weights $\phi_{i1}, ..., \phi_{iJ}$. Because we are in a full-commitment world, these Pareto-weights are time-invariant. And, because there are both observed and unobserved household-level shifters to Pareto-weights, the Pareto-weights are heterogeneous across observably identical households. Consequently, the Pareto-weights are fixed effects hiding inside the resource share functions.

Household quantity demands given the sharing model of Browning et al. (2013) are very simple: the household purchases the sum of what all the individuals would demand if they faced the within-household shadow price vector $A_{it}p_t$ and had their shadow budget $\eta_{ijt}x/N_{ij}$, adjusted for sharing as defined by A_{it}. A key feature here is that the household demand for a non-shareable good does not have to be adjusted for shareability: it is just the sum of what each individual would demand.

An *assignable good* is one where we observe the consumption of that good by a specific person (or type of person). Assuming the existence of a scalar-value demand function $q_j(p, x, z)$ for an assignable and non-shareable good (e.g., food or clothing) for a person of type j, the household's quantity demand, Q_{ijt}, for the assignable good for each of the N_{ij} people of type j is given by

$$Q_{ijt} = q_j(A_{it}p_t, \eta_{ijt}x_{it}/N_{ij}, z_{it}).$$

Because only people of type j purchase this good, the household does not sum over the demand of other household members. Assuming that the assignable good is a normal good implies that q_j is strictly increasing in its second argument, and is therefore strictly monotonic.

Suppose p_t is unobserved, but varies over time. Then, we may express the household demand for the assignable good of a member of type j as a time-varying function of observed data. Defining $\widetilde{q}_{jt}(A_t p_t, x, z) = q_j(p_t, x, z)$, we have

$$Q_{ijt} = \widetilde{q}_{jt}(\eta_{ijt}x_{it}/N_{ij}, z_{it}). \tag{5.4}$$

This is the structural demand equation that we ultimately bring to the data.

This model is very general. It assumes only that: the household satisfies the intertemporal budget constraint under uncertainty, can fully commit to future actions, reaches the Pareto frontier, and has scale economies embodied in the shareability matrix $A(z)$. It places no additional restrictions on utility functions or the bargaining model. It implies that quantity demands for assignable goods are time-varying functions of resource shares, and that resource shares depend on Pareto weights that are fixed over time (aka: fixed effects).

5.1 PIGL Resource Shares

The model above has resource shares depend on a fixed effect, but expresses those resource shares as a vector of implicit functions, which may be hard to work with. To make the model tractable, we impose sufficient structure on utility functions to find closed forms for resource shares.

In our empirical example below, we work with data that have time-invariant demographic characteristics, so let $z_{it} = z_i$ be fixed over time. This implies that the shareability of goods embodied in A is time-invariant: $A_{it} = A_i = A(z_i)$.

Let indirect utilities be in the price-independent generalized logarithmic (PIGL) class (Muellbauer (1975, 1976)) given by

$$V_j(p, x, z) = C_j(p, z) + (B(p, z)x)^{r(z)} / r(z). \tag{5.5}$$

Here, V_j is homogeneous of degree 1 in p, x if C_j is homogeneous of degree 0 in p and B is homogeneous of degree -1 in p. V is increasing in x if $B(p, z)$ is positive and V is concave in x if $r(z) < 1$. In terms of preferences, this class is reasonably wide. It gives quasihomothetic preferences if $r(z) = 1$, and PIGLOG preferences as $r(z) \to 0$ (this includes the Almost Ideal Demand System of Deaton and Muellbauer (1980)).

The functions C_j vary across types j, and so the model allows for preference heterogeneity between types, e.g., between men and women. The restrictions that $B(p, z)$ and $r(z)$ don't vary across j and that $r(z)$ does not depend on prices p are important: as we see below, they imply that resource shares are constant over time.

Substituting the BCL model, observed demographics and period t budgets, we have the utility of person j in household i in period t as

$$v_{ijt} = V_j(A_i p_t, \eta_{ijt} x_{it}/N_{ij}, z_i) = C_j(A_i p_t, z_i) + B(A_i p_t, z_i)^{r(z_i)} (\eta_{ijt} x_{it}/N_{ij})^{r(z_i)} / r(z_i),$$

and thus marginal utilities are given by

$$V_{xj}(A_{it} p_t, \eta_{ijt} x_{it}/N_{ij}, z_{it}) = B(A_i p_t, z_i)^{r(z_i)} (\eta_{ijt} x_{it}/N_{ij})^{r(z_i)-1}.$$

For $r(z_i) \neq 1$, and for any pair of types j, k, we substitute into (5.3) and cancel terms,

$$\frac{\phi_{ij}}{\phi_{ik}} = \frac{V_{xk}(A_{it} p_t, \eta_{ijt} x_{it}/N_{ij}, z_{it})}{V_{xj}(A_{it} p_t, \eta_{ijt} x_{it}/N_{ij}, z_{it})} = \left(\frac{\eta_{ikt}/N_{ik}}{\eta_{ijt}/N_{ij}}\right)^{r(z_i)-1},$$

Rearranging, we get

$$\left(\frac{\eta_{ikt}}{\eta_{ijt}}\right) = \frac{N_{ik}}{N_{ij}} \left(\frac{\phi_{ij}}{\phi_{ik}}\right)^{1/(r(z_i)-1)}. \tag{5.6}$$

The household chooses resource shares in each period and each state to satisfy (5.6). Since the right-hand side has no variation over time or state, this implies that, given PIGL utilities (5.5), the resource shares in a given household i are independent of period t and state s. However, resource shares do vary with both observed and unobserved variables across households i. Let the fixed resource shares that solve the first-order conditions with PIGL demands be denoted $\overline{\eta}_{ij}$.

5.2 The Demand for Women's Food

We will estimate the demand equation for women's food in nuclear households comprised of 1 man, 1 woman and $1-4$ children, so $N_{if} = N_{im} = 1$. Let the resource share for adult women be $\overline{\eta}_{if}$ and define

$$\alpha_i \equiv \ln\left(\overline{\eta}_{if}/N_{if}\right) = \ln \overline{\eta}_{if}, \tag{5.7}$$

equal to the logged resource share of the woman in the household.

Let there be a multiplicative Berkson (1950) measurement error denoted $\exp(-U)$ which multiplies the budget, so that if we observe x, the actual budget is $x/\exp(U)$. The measurement error is i.i.d. across time and households. Here, the measurement error does not affect resource shares, but does affect the distribution of observed quantity demands. Plugging this measurement error and the PIGL form for resource shares given by (5.7) into the assignable goods demand equation (5.4) yields a household demand for women's food, Q_{ift}, given by

$$Q_{ift} = \widetilde{q}_{ft}(\exp(\alpha_i)\, x_{it}/\exp(U_{it}), z_i).$$

This is a FELT model, conditional on covariates:

$$Y_{it} = h_t(Y_{it}^*, z_i), \tag{5.8}$$

where $h_t(Y_{it}^*, z_i) = \widetilde{q}_{ft}\left(\exp\left(Y_{it}^*\right), z_i\right)$ and

$$Y_{it}^* = \alpha_i + X_{it} - U_{it} \tag{5.9}$$

and $X_{it} = \ln x_{it}$ is the logged household budget.

The assumption that the assignable good is normal means that the time-varying functions h_t are strictly monotonic in Y_{it}^*. One could additionally impose that the demand functions $\widetilde{q}_{ft}$ come from the application of Roy's Identity to the indirect utility function (5.5). These demand functions equal a coefficient times the shadow budget plus a coefficient times the shadow budget raised to a power, where the coefficients are time-varying and depend on z_i.[15] We do not impose that additional structure here; instead, we show in Section 7 that the estimated demand curves given by FELT are close to the PIGL shape restrictions.

Here, the time-dependence of h_t is economically important; it is driven by the price-dependence of preferences and by the fact that prices are common to all households i but vary over time t. Further, the fixed effects α_i are economically meaningful parameters: they are equal to the logged women's resource shares in each household. The standard deviation of the logs is a common inequality measure, and the standard deviation of α_i is identified by FELT given strict monotonicity, as we show in Section 4. Further, the covariation of α_i with observed regressors is identified.

This model is useful to answer two important questions. First, are fixed effects (resource shares) fully explained by observed demographics and budgets, or do we need to appeal to unobserved heterogeneity? Second, are fixed effects correlated with log-budgets X_{it}? Some results concerning identification of collective household models in cross-sectional data rely on the assumption of independence.

6 Data

We use data from the 2012 and 2015 Bangladesh Integrated Household Surveys. This data set is a household survey panel conducted jointly by the International Food Policy Research Institute and the World Bank. In this survey, a detailed questionnaire was administered to a sample of rural Bangladeshi households. This data set has two useful features for our purposes: 1) it includes person-level data on food intakes and household-level data on total household expenditures; and 2) it is a panel, following roughly 6000 households over two (nonconsecutive) years. The former allows us to use food as the assignable good to identify our collective household model parameters.

[15]Individual demands are derived by the application of Roy's Identity to (5.5), and are:

$$q_{jt}(x,z) = c_{jt}(z)\, x^{1-r(z)} + b_t(z)\, x$$

where $c_{jt}(z) = -\frac{\nabla_p C_j(p_t,z)}{B(p_t,z)^r}$, $b_t(z) = -\nabla_p \ln B(p_t,z)$. This notation makes clear that we have time-varying demand functions, due to the fact that prices vary over time. In our application, prices in each period are not observed, so we allow the ($z-$dependent) functions c_{jt} and b_t to vary over time. We require that the assignable good be normal, meaning that its demand function is globally increasing in x. This form for demand functions is globally increasing if $c_{jt}(z)$, $b_t(z)$ and $1 - r(z)$ are all positive.

The latter allows us to model household-level unobserved heterogeneity in women's resource shares.

The questionnaire was initially administered to 6503 households in 2012, drawn from a representative sample frame of all rural Bangladeshi households. Of these, 6436 households remained in the sample in 2015. In these data, expenditures on food include imputed expenditure from home production. We drop households with a discrepancy between people reported present in the household and the personal food consumption record, and households with no daily food diary data. Of the remaining data, 6205 households have total expenditures reported for both 2012 and 2015.

In this paper, we focus on households that do not change members between periods.[16] There are 1920 households whose composition is unchanged between 2012 and 2015. Roughly half of these households have more than one adult man or more than one adult women. To simplify the interpretation of estimated resource shares we focus on nuclear households. This leaves 871 nuclear households comprised of one man, one woman and 1 to 4 children, where children are defined to be 14 years old or younger.

The assignable good, Y_{it}, is annual consumption of food by the woman. The surveys contain 7-day recall data on household-level quantities (measured in kilograms) of food consumption in 7 categories: Cereals, Pulses, Oils; Vegetables; Fruits; Proteins; Drinks and Others. These consumption quantities include home-produced food and purchased food and gifts. They include both food consumed in the home (both cooked at home and prepared ready-to-eat food), as well as food consumed outside the home (at food carts or restaurants). These weekly quantities are grossed up to annual consumption expenditure by multiplying by 52 and multiplying by estimated village-level unit-values (following Deaton (1997)).

Our household-level annual consumption, x_{it}, is the sum of total expenditure on, and imputed home-produced consumption of, the following categories of consumption: rent, food, clothing, footwear, bedding, non-rent housing expense, medical expenses, education, remittances, religious food and other offerings (jakat/ fitra/ daan/ sodka/ kurbani/ milad/ other), entertainment, fines and legal expenses, utensils, furniture, personal items, lights, fuel and lighting energy, personal care, cleaning, transport

[16] That is, we exclude households with births, deaths, new members by marriage or adoption, etc. Although a full-commitment model can accommodate such changes in household composition, it is easier to think through the meaning of a person's resource share if the composition is held constant.

and telecommunication, use-value from assets, and other miscellaneous items. These spending levels derive from one-month and three-month duration recall data in the questionnaire, and are grossed up to the annual level. Estimation uses $X_{it} = \ln x_{it}$, the natural logarithm of annual consumption.

Our model is also conditioned on a set of time-invariant demographic variables z_i. We include several types of observed covariates in z_i that may affect both preferences and resource shares: 1) the age in 2012 of the adult male; 2) the age in 2012 of the adult female; 3) the average age in 2012 of the children; 4) the average education in years of the adult male; 5) the average education in years of the adult female; 6) an indicator that the household has 2 children; 7) an indicator that the household has 3 or 4 children; and 8) the fraction of children that are girls.[17]

For the first five of these demographic variables, in order to reduce the support of the regressors, we top- and bottom-code each variable so that values above (below) the 95^{th} (5^{th}) percentiles equal the 95^{th} (5^{th}) percentile values. For all seven of these variables, we standardize the location and scale so that their support is $[0, 1]$. This support restriction simplifies our monotonicity restrictions when it comes to estimation, as explained in Section 7.

We do not trim the data for outliers in the budget or food quantity demands. Instead, we trim the support of the estimated nonparametric regression functions to account for fact that these estimators are high-variance near their boundaries.

Table 2 in the Appendix, Section C gives summary statistics on these data.

7 Estimation of Resource Shares

Following our identification results, estimation could be based on composite versions of the maximum score estimator or the conditional logit estimator (see Botosaru and Muris (2017)). Here, we instead follow a sieve GMM approach that facilitates the inclusion of a large vector of demographic conditioning variables z and the imposition of strict monotonicity on the demand functions (aka: normality of the assignable good) .

[17]Since household membership is fixed for all households in our sample, age, number and gender composition are time-invariant by construction. However, education level of men and women are time-varying in roughly 20% of households. For our time-invariant education variables, we use the average education across the two observed years.

The women's food demand equation (5.8) is a FELT model, conditional on observed covariates z. Denote the inverse demand functions $g_t(Y_{it}, z_{it}) = h_t^{-1}(\cdot, z_{it})$. Given (5.8) a two-period setting with $t = 1, 2$, and time-invariant demographics $z_{it} = z_i$, we have

$$\alpha_i + X_{it} - U_{it} = g_t(Y_{it}, z_i), \tag{7.1}$$

implying the conditional moment condition

$$E\left[g_2(Y_{i2}, z_i) - g_1(Y_{i1}, z_i) - \triangle X_{it} | \, X_{i1}, X_{i2}, z_i\right] = 0. \tag{7.2}$$

We provide a detailed description of our GMM estimator in the Appendix. Briefly, we approximate the inverse demand functions, g_t, $t = 1, 2$, using Bernstein polynomials. In the main text, we use 8th order Bernstein polynomials restricted so that estimated demand curves are strictly monotonically increasing. In the Appendix, we provide estimates for other orders.

We characterize several interesting features of the distribution of resource shares. Recall from Theorems 4 and 5 that identification of features of this distribution does not impose a normalization on assignable good demand functions, and only identifies the distribution of logged resource shares (fixed effects) up to location. Consequently, we only identify features of the resource share distribution up to a scale normalization.

Let $\widehat{g}_{it} = \widehat{g}_t(Y_{it}, z_i)$ equal the predicted values of the inverse demand functions at the observed data. Recall that $g_t\left(Y_{it}, z_i\right) = \alpha_i + X_{it} - U_{it}$, so we can think of $\widehat{g}_{it} - X_{it}$ as a prediction of $\alpha_i - U_{it}$. We then compute the following summary statistics of interest, leaving the dependence of $\hat{g}_{it}$, $t = 1, 2$, on z_i implicit:

1. an estimate of the standard deviation of α_i given by

   $$\hat{std}\left(\alpha\right) = \sqrt{\hat{cov}\left(\left(\widehat{g}_{i2} - X_{i2}\right), \left(\widehat{g}_{i1} - X_{i1}\right)\right)},$$

   where $\hat{cov}$ denotes the sample covariance. The standard deviation of logs is a standard (scale-free) inequality measure. So this gives a direct measure of inter-household variation in women's resource shares.

2. an estimate of the standard deviation of the projection error, e_i, of α_i on $\bar{X}_i =$

$\frac{1}{2}(X_{i1}+X_{i2})$ and Z_i. Consider the projection

$$\alpha_i = \gamma_1 \bar{X}_i + \gamma_2 Z_i + e_i,$$

where Z_i contains a constant. We are interested in the standard deviation of e_i. To obtain this parameter, we compute estimators for γ_1, γ_2 from the pooled linear regression of $\hat{g}_{it} - X_{it}$ on $\bar{X}_i$ and Z_i. Call these estimators $\hat{\gamma}_1, \hat{\gamma}_2$. Then, as in $\hat{std}(\alpha)$ in (1), an estimate of the standard deviation of e_i is given by:

$$\hat{std}(e_i) = \sqrt{\hat{cov}\left(\left(\widehat{g}_{i2} - X_{i2} - \hat{\gamma}_1 \bar{X}_i - \hat{\gamma}_2 Z_i\right), \left(\widehat{g}_{i1} - X_{i1} - \hat{\gamma}_1 \bar{X}_i - \hat{\gamma}_2 Z_i\right)\right)}.$$

This object measures the amount of variation in α_i that cannot be explained with observed regressors. If it is zero, then we don't really need to account for household-level unobserved heterogeneity in resource shares. It is much larger than zero, then accounting for household-level unobserved heterogeneity is important.

3. an estimate of the standard deviation of $\alpha_i + X_{it}$ for $t = 1, 2$, computed as

$$\hat{std}(\alpha_i + X_{it}) = \sqrt{\hat{var}(\alpha_i) + \hat{var}(X_{it}) + 2\hat{cov}(\alpha_i, X_{it})},$$

where

$$\hat{cov}(\alpha_i, X_{i1}) = cov(\hat{g}_{i1} - X_{i1}, X_{i1}),$$
$$\hat{cov}(\alpha_i, X_{i2}) = cov(\hat{g}_{i2} - X_{i2}, X_{i2}),$$

and $\hat{var}(\alpha_i) = \left(\hat{std}(\alpha)\right)^2$ and $\hat{var}(X_{it})$, $t = 1, 2$, is observed in the data. Since $\alpha_i + X_{it}$ is a measure of the woman's shadow budget, $\hat{std}(\alpha_i + X_{it})$ is a measure of inter-household inequality in women's shadow budgets. This inequality measure is directly comparable to the standard deviation of X_i (shown in Table 1), which measures inequality in household budgets.

4. an estimate of the covariance of α_i, X_{it} for $t = 1, 2$, denoted $\hat{cov}(\alpha_i, X_{it})$. This object is of direct interest to applied researchers using cross-sectional data to identify resource shares. If this covariance is non-zero, then the independence of resource shares and household budgets is cast into doubt, and identification

strategies based on this restriction are threatened.

Of these, the first 2 summary statistics are about the variance of fixed effects, and are computed using data from both years. Their validity requires serial independence of the measurement errors U_{it}. In contrast, the second 2 summary statistics are about the correlation of fixed effects with the household budget, and are computed at the year level. They are valid with stationary U_{it}, even in the presence of serial correlation.

We also consider the multivariate relationship between resource shares, household budgets and demographics. Recall that the fixed effect α_i subject to a location normalization; this means that resource shares are subject to a scale normalization. So, we construct an estimate of the woman's resource share in each household as $\widehat{\eta}_i = \exp\left(\frac{1}{2}\left(\hat{g}_{i1} - X_{i1}\right) + \left(\hat{g}_{i1} - X_{i1}\right)\right)$, normalized to have an average value of 0.33. Then, we regress estimated resource shares $\widehat{\eta}_i$ on $\bar{X}_i$ and Z_i, and present the estimated regression coefficients, which may be directly compared with similar estimates in the cross-sectional literature.

The estimated coefficient on X_i gives the conditional dependence of resource shares on household budgets, and therefore speaks to the reasonableness of the restriction that resource shares are independent of those budgets (an identifying restriction used in the cross-sectional literature). Finally, using the estimate of the variance of fixed effects, we construct an estimate of R^2 in the regression of resource shares on observed covariates. This provides an estimate of how much unobserved heterogeneity matters in the overall variation of resource shares.

8 Results

Figure 8 shows our estimates of h_1 and h_2 (or, equivalently, of g_1 and g_2) for $K = 8$, for a family with two children with mean values, $\overline{z}$, of the other demographics. The figures have food quantities q_t on the vertical axis and $\widehat{g}_t\left(q_t\right)$ on the horizontal axis, so the horizontal axis is like a predicted logged household budget. Solid lines give the nonparametric estimates, and 95% pointwise confidence bands for the nonparametric estimates are denoted by dotted lines. Additionally, to provide reassurance that the PIGL utility model—which implies the FELT demand curves—fits the data adequately, we display the PIGL demand curve closest to the FELT estimates in each

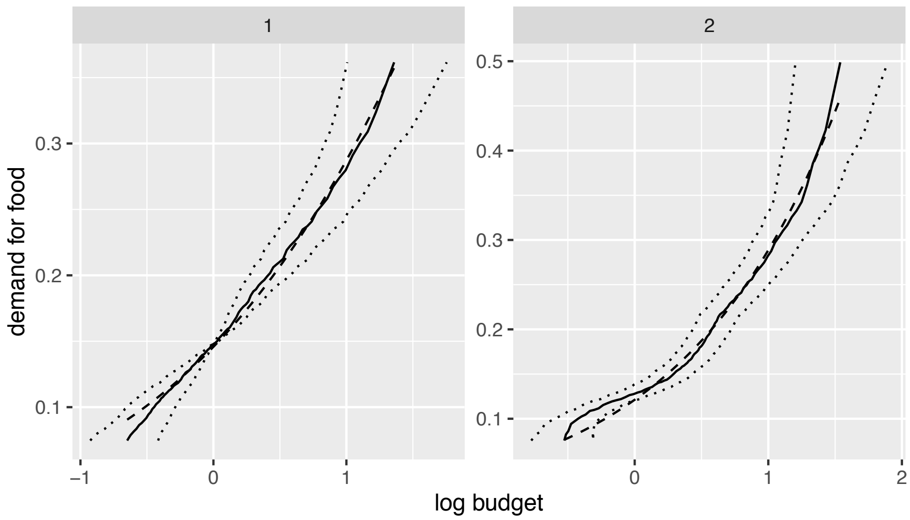

Figure 8.1: Estimated demand functions. Solid line is the nonparametric estimate, evaluated at the mean value of the demographics. The dotted lines indicate the 95% confidence interval. The dashed line is the PIGL closest to the nonparametric estimate. Left panel is for period 1, right panel is for period 2.

time period with dashed lines.[18]

Note that since g_t are identified only up to location (of g_1), we normalize the average of $\widehat{g}_t$ to half the geometric mean of household budgets at t =1, $\overline{x}_1$. Because estimated nonparametric regression functions can be ill-behaved near their boundaries, we truncate the estimated functions at the 5th and 95th percentiles of the distribution of q_t in each t. The key message from 8 is that these estimated demand curves are somewhat nonlinear, estimated reasonably precisely, and not too far from PIGL. The estimated PIGL curvature parameter is $r(\overline{z}) = 0.06$, which means that food demands are close to PIGLOG (as in Banks et al. (1997)). Table 2 gives our summary statistics (items 1-4 above), with bootstrapped 95% confidence intervals, for our estimates with 8 Bernstein polynomials (see the Appendix for other lengths of the Bernstein sieve). In the lower panel, we provide estimated regression coefficients, also with bootstrapped 95% confidence intervals, where we regress estimated resource shares $\widehat{\eta}_i$ on log-budgets $\overline{X}_i$ and demographics z_i. Starting with the top panel of Table

[18] We compute these PIGL demand curves by nonlinear least squares estimation of a pooled q_{it} on $\widehat{g}_{it}$, where the demand curves have the form $q_{it} = c_t x^{1-r} + b_t x$. We estimate the model on a grid of 198 points, one for each interior percentile of q_{it} in each period $t = 1, 2$.

	Estimate	q(2.5)	q(97.5)
Variability of fixed effects α_i			
$\hat{std}\,(\alpha_i)$	0.2647	0.1518	0.3707
$\hat{std}\,(e_i)$	0.1637	0.1262	0.1931
$\hat{cov}\,(\alpha_i, X_{i1})$	-0.0901	-0.1292	-0.0429
$\hat{cov}\,(\alpha_i, X_{i2})$	-0.1034	-0.1418	-0.0561
$\hat{std}\,(\alpha_i + X_{i1})$	0.3537	0.2720	0.4605
$\hat{std}\,(\alpha_i + X_{i2})$	0.3763	0.3010	0.4760
Regression estimates			
R^2: $\overline{X}_i, z_i$ on η_i	0.5205	0.3361	0.6508
$\overline{X}_i$	-0.0452	-0.0700	-0.0196
age–woman	-0.2513	-0.4437	-0.0545
age–man	0.2845	0.1389	0.4191
2 children	-0.0502	-0.1694	0.0413
3 or 4 children	-0.1248	-0.2407	0.0140
avg age of children	-0.0095	-0.1979	0.2151
fraction girl children	0.0136	-0.0881	0.1246
education–woman	-0.0325	-0.1519	0.1198
education–men	-0.1330	-0.2589	-0.0015

Table 1: Estimates.

1, the standard deviation of α_i is a measure of inter-household dispersion in women's resource shares. If this dispersion is very small, then variation in resource shares does not induce much inequality, and we can reasonably use the household-level income distribution as a proxy for person-level inequality. However, if the dispersion is large, then household-level measures of inequality leave out a lot of the action.

The estimated value is roughly 0.26, with a 95% confidence interval covering roughly 0.15 to 0.37. To get a sense of the magnitude for the standard deviation of logged resource shares, suppose that women's resource shares were lognormally distributed. Then our estimated standard deviation of 0.26 is consistent with 95% of the distribution of the resource shares lying in the range $[0.25, 0.75]$, which represents quite a bit of heterogeneity across households.

The next row of Table 1 considers how much of the variation in α_i we can explain with observed covariates. The standard deviation of e_i gives a measure of the unexplained variation, and gives us an idea of whether household-level unobserved heterogeneity is an important feature of the data. If the standard deviation of e_i is very small, then fixed effects are not needed—conditioning on observed covariates would be sufficient. Our estimate of the standard deviation of the unexplained variation in α_i is about 0.16. This is large relative to the overall estimated standard deviation of 0.26, and suggests that accounting for household-level unobserved heterogeneity is quite important.

The next two rows give the covariance of α_i and X_{it}. Here, we see that log resource shares α_i strongly and statistically significantly negatively covary with observed household budgets (the implied correlation coefficients are close to -0.8). This means that women in poor households are somewhat less poor than they appear (on the basis of their household budget), and women in richer households are somewhat more poor than they appear. This is consistent with households that are closer to subsistence having a more equal distribution of resources.

The next two rows give the estimated standard deviation of women's log shadow budgets. This is a scale-free parameter: it does not depend on the location normalization of α_i (which corresponds to a scale normalization of shadow budgets). The estimated standard deviations are 0.35 and 0.38 in the two periods, respectively. We can compare these with the standard deviation of log-budgets, reported in Table 1, of 0.49 and 0.53. The point estimates suggest that there is *less* inequality in women's shadow budgets than in household budgets. Although the confidence intervals are large, the test of the hypothesis that the standard deviation of log-budgets equals the standard deviation of log-shadow budgets rejects in both years.[19]

Thus, if we take these results at face value, there is *less* consumption inequality among women than household-level analysis would suggest. However, another implication of this is that there is *more* gender inequality than household level data would suggest. The reason is that household-level analysis of gender inequality pins gender inequality on over-representation of one gender in poorer households. In our data, all households have 1 man and 1 woman, so household-level analysis of gender inequality would show zero gender inequality. But, because women in richer households have smaller resource shares, this induces gender inequality even in these data.

Finding correlation between α_i and household budgets is not sufficient to invalidate previous identification strategies for cross-sectional settings that rely on independence between resource shares and household budgets. The reason is that the independence required is conditional on other observed covariates. To get a handle on this, the

[19]For $H_0 : Var\left(X_{it}\right) - Var\left(\alpha_i + X_{it}\right)$, we have the following estimated test statistics and (confidence intervals). Period 1: 0.110 (0.0236,0.165); Period 2: 0.137 (0.0527,0.191).

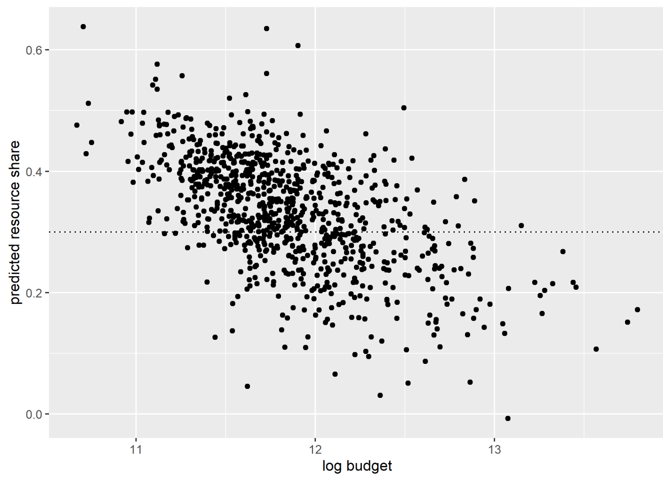

Figure 8.2: Scatterplot of predicted resource shares and log budget.

bottom panel of Table 1 presents estimates of coefficients in a linear regression of normalized (to average 0.33) estimated resource shares $\widehat{\eta}_i$ on log-budgets $\overline{X}_i$ and other covariates z_i.

The figure below shows the scatterplot of predicted resource shares versus the log household budget. Here, we see a lot of variation in resource shares, and it is clearly correlated with household budgets. The overall variation here provides an estimate of the explained sum of squares in an infeasible regression of true resource shares η_i on $\overline{X}_i$ and z_i.[20] We may construct an estimate of the total sum of squares of resource shares from our estimate of the standard deviation of α_i. This yields an estimate of R^2 in the infeasible regression, which we interpret as the fraction of variation in resource shares explained by observables.

In the first row of the bottom panel, we see that observed variables explain roughly half the variation in resource shares (the estimate of R^2 is 0.52). This magnitude of explained variation is very close to that reported in Dunbar et al. (2019) in their random-effects cross-sectional estimate based on Malawian data. This large magnitude of unexplained variation (roughly half) suggests that accounting for unobserved heterogeneity in resource shares is quite important.

[20]This artificial regression is infeasible because we observe (through g_{it}) a prediction of $\alpha_i + U_{it}$, not of α_i itself. However, because we have an estimate of the variance of α_i, we can construct an estimate of the variance of η_i (subject to the scale normalization that it has a mean of 0.33). In our regression, the LHS variable is $\exp \hat{g}_{it}$, which is a prediction of $\eta_i u_{it}$. Since u_{it} are uncorrelated with $\overline{X}_i$ and z by assumption, the explained sum of squares from this regression applies to η_i, and we can use it to form an estimate of R^2, which we report, along with a bootstrapped confidence interval.

Consider first the coefficient on $\overline{X}_i$. The estimated coefficient is -0.045 and is statistically significantly different from 0. This means that, even after conditioning on other covariates (many of which are highly correlated with the budget), we still see a significant relationship between resource shares and household budgets.

However, the magnitude of this effect is small. Conditional on z_i, the standard deviation of X_{it} is 0.43 in year 1 and 0.47 in year 2. Thus, comparing two households with identical z but which are one standard deviation apart in terms the household budget, we would expect the woman in the poorer household to have a resource share 2 percentage points higher than the woman in the richer household. Thus, the bulk of the variation the makes the standard deviation of women's shadow budgets smaller than that of household budgets is not running through the dependence of resource shares on household budgets, but rather through the dependence of resource shares on other covariates that are correlated with household budgets.

We get a very precise estimate of the conditional dependence of resource shares on household budgets. Overall, then, we see that women's resource shares are statistically significantly correlated with household budgets, even conditional on other observed characteristics. But, the estimated difference in resource shares at different household budgets is quite small. So, we take this as evidence that the identifying restrictions used by Dunbar et al. (2013) (and Dunbar et al. (2019)) may be false, though perhaps not *very* false. It does suggest that alternative identifying restrictions—such as those developed here with a panel data model—may be useful.

The rows of Table 1 give several other coefficients that are comparable to other estimates in the literature. Calvi (2019) finds that women's resource shares in India decline with the age of the woman. In these Bangladeshi data, we find evidence that women's resource shares are strongly negatively correlated with the age of women and positively correlated with the age of men.

Dunbar et al. (2013) find that women's resource shares in Malawi decline with the number of children. Here, we also see that pattern: households with 2 children have women's resource shares 5 percentage points less than households with 1 child; households with 3 or 4 children have resource shares 12 percentage points less. Dunbar et al. (2013) also find that Malawian women's resource shares are higher in households with girls than households with boys. We do not see evidence of this in rural Bangladesh: the estimated coefficient on the fraction of children that are girls statistically insignificantly different from 0.

In the Appendix, we also provide estimates analogous to Table 1 using a different assignable good: clothing. Under the model, using different assignable goods should yield the same estimates of resource shares.[21] This is roughly what we find in our estimates using women's clothing.

Our estimates use 8th order Bernstein polynomials to approximate the inverse demand functionsD. In Appendix B, we present estimates using Bernstein polynomials of order $K = 1, 4, 8, 10$ and show that our finite-dimensional parameter estimates have roughly the same value for $K \geq 8$.

In summary, in these rural Bangladeshi households, we find evidence that women's resource shares have substantial dependence on household-level unobserved heterogeneity and are slightly negatively correlated with household budgets. The former suggests that random-effects type approaches to the estimation of resource shares may be inadequate. The latter suggests that consumption inequality faced by women is actually smaller than household-level consumption inequality. It also suggests that cross-sectional identification strategies invoking independence of resource shares from household budgets, such as Dunbar et al. (2013), could be complemented by panel-based identification strategies such as ours.

References

Abrevaya, J. (1999). Leapfrog estimation of a fixed-effects model with unknown transformation of the dependent variable. *Journal of Econometrics*, 93(2):203–228.

Abrevaya, J. (2000). Rank estimation of a generalized fixed-effects regression model. *Journal of Econometrics*, 95(1):1–23.

Abrevaya, J. and Muris, C. (2020). Interval censored regression with fixed effects. *Journal of Applied Econometrics*, 35(2):198–216.

Aguirregabiria, V., Gu, J., and Luo, Y. (forthcoming). Sufficient statistics for unobserved heterogeneity in dynamic structural logit models. *Journal of Econometrics*.

Ai, C. and Gan, L. (2010). An alternative root-n consistent estimator for panel data binary choice models. *Journal of Econometrics*, 157(1):93–100.

Altonji, J. G. and Matzkin, R. L. (2005). Cross section and panel data estimators for nonseparable models with endogenous regressors. *Econometrica*, 73(4):1053–1102.

Andrews, D. W. (1999). Estimation when a parameter is on a boundary. *Econometrica*, 67:1341–1383.

[21]Food is a plausible *assignable* good (because if one person eats it, nobody else can), but it may not be non-shareable (because there may be scale economies in cooking). In contrast, clothing may be plausibly non-shareable, but it may not be assignable (because, e.g., mothers and daughters might wear each others' clothes). See the Appendix for details on clothing estimates.

Andrews, D. W. (2000). Inconsistency of the bootstrap when a parameter is on the boundary of the parameter space. *Econometrica*, 68:399–405.

Arellano, M. (2003). *Panel Data Econometrics*. Oxford University Press.

Arellano, M. and Bonhomme, S. (2009). Robust priors in nonlinear panel data models. *Econometrica*, 77(2):489–536.

Arellano, M. and Bonhomme, S. (2011). Nonlinear panel data analysis. *Annual Review of Economics*, 3(1):395–424.

Arellano, M. and Bonhomme, S. (2012). Identifying distributional characteristics in random coefficients panel data models. *Review of Economic Studies*, 79(3):987–1020.

Arellano, M. and Hahn, J. (2007). Understanding bias in nonlinear panel models: Some recent developments. In Blundell, R., Newey, W., and Persson, T., editors, *Advances in Economics and Econometrics*, pages 381–409.

Arellano, M. and Honoré, B. (2001). Panel data models: Recent developments. In Heckman, J. and Leamer, E., editors, *Handbook of Econometrics*, volume 5, chapter 53, pages 3219–3296. Elsevier.

Aristodemou, E. (forthcoming). Semiparametric identification in panel data discrete response models. *Journal of Econometrics*.

Athey, S. and Imbens, G. (2006). Identification and inference in nonlinear difference-in-differences model. *Econometrica*, 74(2):431–497.

Baetschmann, G., Staub, K. E., and Winkelmann, R. (2015). Consistent estimation of the fixed effects ordered logit model. *Journal of the Royal Statistical Society A*, 178(3):685–703.

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

Bargain, O., Donni, O., and Kwenda, P. (2014). Intrahousehold distribution and poverty: Evidence from cote d'ivoire. *Journal of Development Economics*, 107:262–276.

Bargain, O., Lacroix, G., and Tiberti, L. (2018). Validating the collective model of household consumption using direct evidence on sharing. Partnership for Economic Policy Working Paper 2018-06.

Becker, G. S. (1962). Investment in human capital: A theoretical analysis. *Journal of Political Economy*, 70(5):9–49.

Berkson, J. (1950). Are there two regressions? *Journal of the American Statistical Association*, 45(250):164–180.

Bester, C. A. and Hansen, C. (2009). Identification of marginal effects in a nonparametric correlated random effects model. *Journal of Business and Economic Statistics*, 27(2):235–250.

Blackorby, C., Bossert, W., and Donaldson, D. (2005). *Population Issues in Social Choice Theory, Welfare Economics, and Ethics*. Cambridge University Press.

Bonhomme, S. (2012). Functional differencing. *Econometrica*, 800(4):1337–1385.

Botosaru, I. and Muris, C. (2017). Binarization for panel models with fixed effects.

cemmap working paper CWP31/17.

Bourguignon, F., Browning, M., Chiappori, P. A., and Lechene, V. (1993). Intra household allocation of consumption: A model and some evidence from french data. *Annales d'Économie et de Statistique*, (29):137–156.

Browning, M. and Chiappori, P. A. (1998). Efficient intra-household allocations: a general characterization and empirical tests. *Econometrica*, 66(6):1241–1278.

Browning, M., Chiappori, P. A., and Lewbel, A. (2013). Estimating consumption economies of scale, adult equivalence scales, and household bargaining power. *Review of Economic Studies*, 80(4):1267–1303.

Calvi, R. (2019). Why are older women missing in india? the age profile of bargaining power and poverty. *Journal of Political Economy*. Forthcoming.

Chamberlain, G. (1980). Analysis of covariance with qualitative data. *Review of Economic Studies*, 47(1):225–238.

Chamberlain, G. (2010). Binary response models for panel data: Identification and information. *Econometrica*, 78(1):159–168.

Charlier, E., Melenberg, B., and van Soest, A. (2000). Estimation of a censored regression panel data model using conditional moment restrictions efficiently. *Journal of Econometrics*, 95:25–56.

Chen, S. (2002). Rank estimation of transformation models. *Econometrica*, 70(4):1683–1697.

Chen, S. (2010a). An integrated maximum score estimator for a generalized censored quantile regression model. *Journal of Econometrics*, 155(1):90–98.

Chen, S. (2010b). Non-parametric identification and estimation of truncated regression models. *Review of Economic Studies*, 77(1):127–153.

Chen, S. (2010c). Root-N-consistent estimation of fixed-effect panel data transformation models with censoring. *Journal of Econometrics*, 159(1):222–234.

Chen, S. (2012). Distribution-free estimation of the Box-Cox regression model with censoring. *Econometric Theory*, 28(3):680–695.

Chen, S., Khan, S., and Tang, X. (2019). Exclusion Restrictions in Dynamic Binary Choice Panel Data Models: Comment on âSemiparametric Binary Choice Panel Data Models Without Strictly Exogenous Regressorsâ. *Econometrica*, 87(5):1781–1785.

Chen, S. and Zhou, X. (2012). Semiparametric estimation of a truncated regression model. *Journal of Econometrics*, 167(2):297–304.

Chen, X. (2007). Large sample sieve estimation of semi-nonparametric models. In *Handbook of Econometrics*, volume 6B, chapter 76, pages 5549–5632. Elsevier.

Cherchye, L., De Rock, B., Lewbel, A., and Vermeulen, F. (2015). Sharing rule identification for general collective consumption models. *Econometrica*, 83(5):2001–2041.

Chernozhukov, V., Fernández-Val, I., Hahn, J., and Newey, W. (2013). Average and quantile effects in nonseparable panel models. *Econometrica*, 81(2):535–580.

Chernozhukov, V., Fernández-Val, I., Hoderlein, S., Holzmann, H., and Newey, W.

(2015). Nonparametric identification in panels using quantiles. *Journal of Econometrics*, 188(2):378–392.

Chernozhukov, V., Fernández-Val, I., and Weidner, M. (2018). Network and panel quantile effects via distribution regression. arXiv:1803.08154.

Chiappori, P., Komunjer, I., and Kristensen, D. (2015). Nonparametric identification and estimation of transformation models. *Journal of Econometrics*, 188:22–39.

Chiappori, P. A. (1988). Rational household labor supply. *Econometrica*, 56(1):63–90.

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

Chiappori, P. A. and Ekeland, I. (2006). The micro economics of group behavior: General characterization. *Journal of Economic Theory*, 130(1):1–26.

Chiappori, P. A. and Ekeland, I. (2009). The microeconomics of efficient group behavior: Identification. *Econometrica*, 77(3):763–799.

Chiappori, P. A. and Kim, J. H. (2017). A note on identifying heterogeneous sharing rules. *Quantitative Economics*, 8(1):201–218.

Chiappori, P.-A. and Mazzoco, M. (2017). Static and intertemporal household decisions. *Journal of Economic Literature*, 55(3):985–1045.

Das, M. and van Soest, A. (1999). A panel data model for subjective information on household income growth. *Journal of Economic Behavior & Organization*, 40(4):409–426.

Deaton, A. and Muellbauer, J. (1980). *Economics and consumer behavior*. University Press, Cambridge.

Deaton, A. S. (1997). *The Analysis of Household Surveys : A Microeconometric Approach to Development Policy*. Washington, D.C. : The World Bank.

Doksum, K. A. and Gasko, M. (1990). On a correspondence between models in binary regression analysis and in survival analysis. *International Statistical Review*, 58:243–252.

Dunbar, G. R., Lewbel, A., and Pendakur, K. (2013). Children's resources in collective households: identification, estimation, and an application to child poverty in Malawi. *American Economic Review*, 103(1):438–471.

Dunbar, G. R., Lewbel, A., and Pendakur, K. (2019). Identification of random resource shares in collective households without preference similarity restrictions. *The Journal of Business and Economic Statistics*.

Evdokimov, K. (2010). Identification and estimation of a nonparametric panel data model with unobserved heterogeneity. Working paper.

Evdokimov, K. (2011). Nonparametric identification of a nonlinear panel model with application to duration analysis with multiple spells. Working paper.

Fernández-Val, I. (2009). Fixed effects estimation of structural parameters and marginal effects in panel probit models. *Journal of Econometrics*, 150(1):71–85.

Fernández-Val, I. and Weidner, M. (2016). Individual and time effects in nonlinear panel data models with large N, T. *Journal of Econometrics*, 192(1):291–312.

Freyberger, J. (2018). Nonparametric panel data models with interactive fixed effects.

The Review of Economic Studies, 85(3):1824–1851.

Hahn, J. and Newey, W. (2004). Jackknife and analytical bias reduction for nonlinear panel models. *Econometrica*, 72(4):1295–1319.

Hoderlein, S. and White, H. (2012). Nonparametric identification in nonseparable panel data models with generalized fixed effects. *Journal of Econometrics*, 168(2):300–314.

Honoré, B. (1992). Trimmed lad and least squares estimation of truncated and censored regression models with fixed effects. *Econometrica*, 60(3):533–565.

Honoré, B. (1993). Identification results for duration models with multiple spells. *Review of Economic Studies*, 60(1):241–246.

Honoré, B. and Hu, L. (2004). Estimation of cross sectional and panel data censored regression models with endogeneity. *Journal of Econometrics*, 122(2):293–316.

Honoré, B. and Kyriazidou, E. (2000). Panel discrete choice models with lagged dependent variables. *Econometrica*, 78(4):839–874.

Honoré, B. and Lewbel, A. (2002). Semiparametric binary choice panel data models without strictly exogeneous regressors. *Econometrica*, 70(5):2053 – 2063.

Horowitz, J. L. (1996). Semiparametric estimation of a regression model with an unknown transformation of the dependent variable. *Econometrica*, 64(1):103–137.

Horowitz, J. L. and Lee, S. (2004). Semiparametric estimation of a panel data proportional hazards model with fixed effects. *Journal of Econometrics*, 119(1):155–198.

Khan, S., Ponomareva, M., and Tamer, E. (2016). Identification of panel data models with endogenous censoring. *Journal of Econometrics*, 194(1):57–75.

Khan, S., Ponomareva, M., and Tamer, E. (2020). Identification of dynamic binary response models. Working paper.

Khan, S. and Tamer, E. (2007). Partial rank estimation of duration models with general forms of censoring. *Journal of Econometrics*, 136(1):251–280.

Lancaster, T. (2002). Orthogonal parameters and panel data. *The Review of Economic Studies*, 69(3):647–666.

Lee, S. (2008). Estimating panel data duration models with censored data. *Econometric Theory*, 24(5):1254–1276.

Lewbel, A. (2014). An overview of the special regressor method. In *Oxford Handbook of Applied Nonparametric and Semiparametric Econometrics and Statistics*, pages 38–62. Oxford University Press.

Lewbel, A. and Yang, T. (2016). Identifying the average treatment effect in ordered treatment models without unconfoundedness. *Journal of Econometrics*, 195:1–22.

Ligon, E. (1998). Risk sharing and information in village economies. *The Review of Economic Studies*, 65(4):847–864.

Lise, J. and Yamada, K. (2019). Household sharing and commitment: Evidence from panel data on individual expenditures and time use. *The Review of Economic Studies*, 86(5):2184–2219.

Magnac, T. (2004). Panel binary variables and suficiency: Generaling conditional logit. *Econometrica*, 72(6):1859–1876.

Manski, C. (1975). Maximum score estimation of the stochastic utility model of choice. *Journal of Econometrics*, 3(3):205–228.

Manski, C. (1985). Semiparametric analysis of discrete response: Asymptotic properties of the maximum score estimator. *Journal of Econometrics*, 27(3):313–333.

Manski, C. F. (1987). Semiparametric analysis of random effects linear models from binary panel data. *Econometrica*, 55(2):357–362.

Mazzocco, M. (2007). Household intertemporal behaviour: A collective characterization and a test of commitment. *The Review of Economic Studies*, 74(3):857–895.

Mazzocco, M., Ruiz, C., and Yamaguchi, S. (2014). Labor supply and household dynamics. *American Economic Review*, 104(5):354–59.

Menon, M., Pendakur, K., and Perali, F. (2012). On the expenditure-dependence of children's resource shares. *Economics Letters*, 117(3):739–742.

Muellbauer, J. (1975). Aggregation, income distribution and consumer demand. *The Review of Economic Studies*, 42(4):525–543.

Muellbauer, J. (1976). Community preferences and the representative consumer. *Econometrica*, (5):979–999.

Muris, C. (2017). Estimation in the fixed effects ordered logit model. *The Review of Economics and Statistics*, 99(3):465–477.

Neyman, J. and Scott, E. (1948). Consistent estimates based on partially consistent observations. *Econometrica*, 16(1):1–32.

Rasch, G. (1960). *Probabilistic Models for Some Intelligence and Attainment Tests.* Copenhagen:Denmark Paedagogiske Institut.

Ridder, G. (1990). The non-parametric identification of generalized accelerated failure-time models. *Review of Economic Studies*, 57(2):167–181.

Shi, X., Shum, M., and Song, W. (2018). Estimating semi-parametric panel multinomial choice models using cyclic monotonicity. *Econometrica*, 86(2):737–761.

Sokullu, S. and Valente, C. (2019). Individual consumption in collective households: Identification using repeated observations with an application to progresa.

van den Berg, G. J. (2001). Duration models: specification, identification and multiple durations. In Heckman, J. and Leamer, E., editors, *Handbook of Econometrics*, volume 5, chapter 55, pages 3381–3460. Elsevier.

Vermeulen, F. (2002). Collective household models: principles and main results. *Journal of Economic Surveys*, 16(4):533–564.

Voena, A. (2015). Yours, mine and ours: Do divorce laws affect the intertemporal behavior of married couples? *American Economic Review*, 105(8):2295–2332.

Vreyer, P. D. and Lambert, S. (2016). Intra-household inequalities and poverty in senegal. mimeo, Paris School of Economics.

Wang, J. and Ghosh, S. (2012). Shape restricted nonparametric regression with bernstein polynomials. *Computational Statistics and Data Analysis*, 56:2729–2741.

Wilhelm, D. (2015). Identification and estimation of nonparametric panel data regressions with measurement error. Cemmap working paper CWP34/15.

ONLINE APPENDICES

A Proofs

A.1 Proof of Lemma (1)

Proof. Define $\overline{D} = 1\{D_1(y_1) + D_2(y_2) = 1\}$. The proof consists in showing the following:

$$\begin{aligned}
& \operatorname{med}\left(D_2(y_2) - D_1(y_1) \,|X, \overline{D} = 1\right) && \text{(A.1)} \\
= \;& \operatorname{sgn}\left(P\left(D(y_1, y_2) = (0,1)\,|X, \overline{D} = 1\right) - P\left(D(y_1, y_2) = (1,0)\,|X, \overline{D} = 1\right)\right) && \text{(A.2)} \\
= \;& \operatorname{sgn}\left(\frac{P\left(D(y_1, y_2) = (0,1), \overline{D} = 1|X\right)}{P(\overline{D} = 1|X)} - \frac{P\left(D(y_1, y_2) = (1,0), \overline{D} = 1|X\right)}{P(\overline{D} = 1|X)}\right) && \text{(A.3)} \\
= \;& \operatorname{sgn}\left(P\left(D(y_1, y_2) = (0,1), \overline{D} = 1|X\right) - P\left(D(y_1, y_2) = (1,0), \overline{D} = 1|X\right)\right) && \text{(A.4)} \\
= \;& \operatorname{sgn}\left(P\left(D(y_1, y_2) = (0,1)\,|X\right) - P\left(D(y_1, y_2) = (1,0)\,|X\right)\right) && \text{(A.5)} \\
= \;& \operatorname{sgn}\left(P\left(D_2(y_2) = 1|X\right) - P\left(D_1(y_1) = 1|X\right)\right) && \text{(A.6)} \\
= \;& \operatorname{sgn}\left(\Delta X\beta - \gamma(y_1, y_2)\right) && \text{(A.7)}
\end{aligned}$$

where ($A.2$) follows since the random variable $D_2(y_2) - D_1(y_1) \in \{-1, 1\}$, which implies that

$$\begin{aligned}
& \operatorname{med}\left(D_2(y_2) - D_1(y_1)\,|X, \overline{D} = 1\right) \\
= & \begin{cases} 1 \text{ if } P\left(D(y_1, y_2) = (0,1)\,|X, \overline{D} = 1\right) > P\left(D(y_1, y_2) = (1,0)\,|X, \overline{D} = 1\right) \\ -1 \text{ if } P\left(D(y_1, y_2) = (0,1)\,|X, \overline{D} = 1\right) < P\left(D(y_1, y_2) = (1,0)\,|X, \overline{D} = 1\right) \end{cases},
\end{aligned}$$

($A.3$) follows from the definition of conditional probability, ($A.4$) follows since the sign function is not affected by scaling both quantities by the same positive factor (the denominator), ($A.5$) follows by the definition of $\overline{D}$, and ($A.6$) follows since:

$$\begin{aligned}
P\left(D_2(y_2) = 1|X\right) &= P\left(D(y_1, y_2) = (0,1)\,|X\right) + P\left(D(y_1, y_2) = (1,1)\,|X\right) \\
P\left(D_1(y_1) = 1|X\right) &= P\left(D(y_1, y_2) = (1,0)\,|X\right) + P\left(D(y_1, y_2) = (1,1)\,|X\right)
\end{aligned}$$

Finally, ($A.7$) follows from Assumption 2(ii), which implies that, e.g.,

$$P\left(D_2(y_2) = 1|\alpha, X\right) > P\left(D_1(y_1) = 1|\alpha, X\right) \Leftrightarrow \alpha + X_2\beta - h_2^-(y_2) > \alpha + X_1\beta - h_1^-(y_1).$$

Integrating both sides over the conditional distribution of α given X obtains:

$$\begin{aligned}
P\left(D_2(y_2) = 1|X\right) > P\left(D_1(y_1) = 1|X\right) &\Leftrightarrow X_2\beta - h_2^-(y_2) > X_1\beta - h_1^-(y_1) \\
&\Leftrightarrow \Delta X\beta - \gamma(y_1, y_2) > 0.
\end{aligned}$$

Result (3.4) now follows. $\square$

A.2 Proof of Theorem 1

Proof. Following Manski (1985), it suffices to show that for an arbitrary $\theta \in \Theta$, $\theta \neq \theta_0 \equiv \theta_0(y_1, y_2)$,

$$P(W\theta < 0 \leq W\theta_0) + P(W\theta_0 < 0 \leq W\theta) > 0. \tag{A.8}$$

Our proof follows very closely that in Manski (1985), with $W\theta$ taking the role of xb and $W\theta_0$ taking the role of $x\beta$. However, our scale normalization is different.

Without loss of generality, let X_K be the continuous regressor in Assumption 3(i). Separate $\Delta X = (\Delta X_{-K}, \Delta X_K)$ where the first component ΔX_{-K} represents all covariates except the K-th one. Similarly, for any $\theta = (\beta, \gamma) \in \Theta$, separate $\beta = (\beta_{-K}, \beta_K)$. Furthermore denote $W_{-K} = (\Delta X_{-K}, -1)$ and $\theta_{-K} = (\beta_{-K}, \gamma)$.

Assume that the associated regression coefficient $\beta_{0,K} > 0$. The case $\beta_{0,K} < 0$ follows similarly. Let $\theta = (\beta, \gamma) \in \Theta$, $\theta \neq \theta_0$. As in Manski (1985, p. 318), consider three cases: (i) $\beta_K < 0$; (ii) $\beta_K = 0$; (iii) $\beta_K > 0$.

Cases (i) and (ii). $\beta_K \leq 0$. The proof is identical to that in Manski (1985), with $X\beta$ replaced by $W\theta$. The fact that we use a different scale normalization does not come into play.

Case (iii). $\beta_K > 0$. note that

$$P(W\theta < 0 \leq W\theta_0) = P\left(-\frac{W_{-K}\theta_{0,-K}}{\beta_{0,K}} < \Delta X_K < -\frac{W_{-K}\theta_{-K}}{\beta_K}\right).$$
$$P(W\theta_0 < 0 \leq W\theta) = P\left(-\frac{W_{-K}\theta_{-K}}{\beta_K} < \Delta X_K < -\frac{W_{-K}\theta_{0,-K}}{\beta_{0,K}}\right).$$

By assumption 4, $\frac{\beta_{-K}}{\beta_K} \neq \frac{\beta_{0,-K}}{\beta_{0,K}}$, which shows that the first K components of the vector θ are not a scalar multiple of the first K components of the vector θ_0. Therefore, θ is not a scalar multiple of θ_0. In particular, $\frac{\theta_{0,-K}}{\beta_{0,K}} \neq \frac{\theta_{-K}}{\beta_K}$. Additionally, assumption 3(ii) implies that $P\left(\frac{W_{-K}\theta_{0,-K}}{\beta_{0,K}} \neq \frac{W_{-K}\theta_{-K}}{\beta_K}\right) > 0$. Hence at least one of the two probabilities above is positive so that (A.8) holds. $\square$

A.3 Proof of Theorem 2

Proof. Under Assumption 5, $h_1^-(y_0) = 0$. Using the pair (y_0, y_2) for binarization thus obtains identification of

$$\begin{aligned}\gamma(y_0, y_2) &= h_2^-(y_2) - h_1^-(y_0)\\ &= h_2^-(y_2).\end{aligned}$$

By varying $y_2 \in \underline{\mathcal{Y}}$, we identify the function h_2^- from the binary choice models associated with $\{D(y_0, y_2) = (D_1(y_0), D_2(y_2)), y_2 \in \underline{\mathcal{Y}}\}$.

The pairs (y_0, y_2) and (y_1, y_2) identify the difference

$$\begin{aligned}\gamma(y_0, y_2) - \gamma(y_1, y_2) &= (h_2^-(y_2) - h_1^-(y_0)) - (h_2^-(y_2) - h_1^-(y_1)) \\ &= h_1^-(y_1).\end{aligned}$$

By varying $y_1 \in \underline{\mathcal{Y}}$ we therefore identify h_1^-.

Thus, the functions h_1^- and h_2^- are identified. Because of monotonicity of h_t (Assumption 1), and because $\mathcal{Y}$ is known, h_t^- contains all the information about the pre-image of h_t. Knowledge of the pre-image of a function is equivalent to knowledge of the function itself. Therefore, h_t can be identified from h_t^-. □

A.4 Proof of Theorem (3)

Proof. For the panel data binary choice model with logit errors, we obtain

$$P(D_2(y_2) = 1|\overline{D}(y_1, y_2) = 1, X, \alpha) \tag{A.9}$$

$$= \frac{P(D_2(y_2) = 1, \overline{D}(y_1, y_2) = 1|X, \alpha)}{P\left(\overline{D}(y_1, y_2) = 1|X, \alpha\right)} \tag{A.10}$$

$$= \frac{P(D_1(y_1) = 0, D_2(y_2) = 1|X, \alpha)}{P\left(\overline{D}(y_1, y_2) = 1|X, \alpha\right)} \tag{A.11}$$

$$= \frac{P(D_1(y_1) = 0, D_2(y_2) = 1|X, \alpha)}{P(D_1(y_1) = 0, D_2(y_2) = 1|X, \alpha) + P(D_1(y_1) = 1, D_2(y_2) = 0|X, \alpha)} \tag{A.12}$$

$$= \frac{1}{1 + \frac{P(D_1(y_1)=1, D_2(y_2)=0|X,\alpha)}{P(D_1(y_1)=0, D_2(y_2)=1|X,\alpha)}} \tag{A.13}$$

$$= \Lambda(\Delta X\beta - \gamma(y_1, y_2)) \tag{A.14}$$

where A.10 follows from the definition of a conditional probability; A.11 follows because $D_2 = 1$ and $\bar{D} = 1$ are equivalent to $D_1 = 0$ and $D_2 = 1$; A.12 follows because $D_1 + D_2 = 1$ happens precisely when either $(D_1, D_2) = (1, 0)$ or $(D_1, D_2) = (0, 1)$; A.13 follows by dividing by the numerator; and the final expression follows by the argument below.

Note that $\frac{P(D_1(y_1)=1, D_2(y_2)=0|X,\alpha)}{P(D_1(y_1)=0, D_2(y_2)=1|X,\alpha)}$ equals

$$\frac{P(D_1(y_1)=1|X,\alpha)P(D_2(y_2)=0|X,\alpha)}{P(D_1(y_1)=0|X,\alpha)P(D_2(y_2)=1|X,\alpha)} \tag{A.15}$$

$$= \frac{\Lambda\left(\alpha+X_1\beta-h_1^-(y_1)\right)\left[1-\Lambda\left(\alpha+X_2\beta-h_2^-(y_2)\right)\right]}{\left[1-\Lambda\left(\alpha+X_1\beta-h_1^-(y_1)\right)\right]\Lambda\left(\alpha+X_2\beta-h_2^-(y_2)\right)} \tag{A.16}$$

$$= \frac{\exp\left(\alpha+X_1\beta-h_1^-(y_1)\right)}{\exp\left(\alpha+X_2\beta-h_2^-(y_2)\right)} \tag{A.17}$$

$$= \exp\left((X_1-X_2)\beta-\left(h_1^-(y_1)-h_2^-(y_2)\right)\right), \tag{A.18}$$

where A.15 follows from serial independence of (U_1,U_2) conditional on (X,α); A.16 from the logit model specification; and A.17 follows from

$$\Lambda(u)/(1-\Lambda(u)) = \exp(u).$$

The discussion above implies that A.9 does not depend on α. Hence,

$$\begin{aligned} p(X,y_1,y_2) &\equiv P(D_2(y_2)=1|\overline{D}(y_1,y_2)=1,X) \\ &= \Lambda(\Delta X\beta-\gamma(y_1,y_2)) \\ &= \Lambda(W\theta(y_1,y_2)). \end{aligned}$$

and note that $p(X,y_1,y_2)$ is identified from the distribution of (Y,X), which is assumed to be observed. Then

$$\theta(y_1,y_2) = [E(W'W)]^{-1}E(W'\Lambda^{-1}(p(X,y_1,y_2)))$$

by invertibility of Λ and the full rank assumption on $E[W'W]$. This establishes identification of β and $\gamma(y_1,y_2)$. The proof in Section A.3 applies, which shows the identification of h_1 and h_2. □

A.5 Proof of Theorem 4

Proof. (a) Note that, without Assumption 5, it follows immediately from the proof of Theorem 2 that we can only identify $\left\{h_t^{-1}(y)-c_1, y\in\mathcal{Y}, t=1,2\right\}$, i.e. we identify the functions $g_t(y)$, $t=1,2$.

Because the functions g_1, g_2 are identified, and because the distribution of (Y,X) is observable, we can identify the distribution of the left hand side of the relation below:

$$\begin{aligned} g_t(Y_t)-X_t\beta &= h_t^{-1}(Y_t)-X_t\beta-c_1 \\ &= \alpha-U_t-c_1, \end{aligned}$$

for $t = 1, 2$.

It follows that for all $x \in \mathcal{X}$,

$$\begin{aligned}\mu(x) &= E[\alpha | X = x] \\ &= E[\alpha - U_t | X = x] + E[U_t | X = x] \\ &= E\left[h_t^{-1}(Y_t) - X_t\beta \middle| X = x\right] + m \\ &= E[g_t(Y_t) - X_t\beta | X = x] + c_1 + m.\end{aligned}$$

is identified up to the constants c_1 and m.

The difference in conditional means at any two values $x, x' \in \mathcal{X}$ is therefore identified and given by:

$$\begin{aligned}&\mu(x) - \mu\left(x^{'}\right) = \\ &(E[g_t(Y_t) - X_t\beta | X = x] + c_1 + m) - \left(E\left[g_t(Y_t) - X_t\beta | X = x^{'}\right] + c_1 + m\right) \\ &= E[g_t(Y_t) - X_t\beta | X = x] - E\left[g_t(Y_t) - X_t\beta | X = x^{'}\right].\end{aligned}$$

(b) To see that the conditional variance is identified, note that for all $x \in \mathcal{X}$,

$$\begin{aligned}&Cov(g_2(Y_2) - X_2\beta, g_1(Y_1) - X_1\beta | X = x) = \\ &= Cov\left(h_2^{-1}(Y_2) - X_2\beta - c_1, h_1^{-1}(Y_1) - X_1\beta - c_1 \middle| X = x\right) \\ &= Cov(\alpha - U_1 - c_1, \alpha - U_2 - c_1 | X = x) \\ &= Var(\alpha | X = x) - Cov(\alpha, U_1 | X = x) - Cov(\alpha, U_2 | X = x) + Cov(U_1, U_2 | X = x) \\ &= Var(\alpha | X = x) = \sigma_\alpha^2(x),\end{aligned}$$

where the first equality follows from the definition of g_t; the second from the model; the third equality follows from the linearity of the covariance; and the fourth equality uses assumption (4c) and (4d). □

B GMM Estimator

The women's food demand equation (5.8) is a FELT model, conditional on observed covariates z. Denote the inverse demand functions $g_t(Y_{it}, z_{it}) = h_t^{-1}(\cdot, z_{it})$. Given (5.8) and a two-period setting with $t = 1, 2$, and time-invariant demographics $z_{it} = z_i$, we have

$$\alpha_i + X_{it} - U_{it} = g_t(Y_{it}, z_i), \tag{B.1}$$

implying the conditional moment condition

$$E[g_2(Y_{i2}, z_i) - g_1(Y_{i1}, z_i) - \triangle X_{it} | X_{i1}, X_{i2}, z_i] = 0. \tag{B.2}$$

Sieve estimators. We approximate the inverse demand functions, g_t, $t = 1, 2$, using Bernstein polynomials. This allows us to impose monotonicity in a straightforward way, see, e.g., Wang and Ghosh (2012). Let $k = 0, ..., K$ index univariate Bernstein functions denoted as $\mathcal{B}_k(\cdot, K)$, where K is the degree of the Bernstein polynomial and where the Bernstein functions are given by:

$$\mathcal{B}_k(u, K) = \binom{K}{k} u^k (1-u)^{K-k}, \, u \in [0, 1].$$

Let $l = 0, ..., L$ index the elements of z_i, and let the first (index 0) element of z_i be a constant equal to 1, that is $z_i = [1, z_{i1}, \ldots, z_{iL}]$.

Our approximation to $g_t(Y_{it}, z_i)$ is given by:

$$g_t(Y_{it}, z_i) \approx \sum_{k=0}^{K} \beta_{kt}(z_i) \mathcal{B}_k(Y_{it}, K) \approx \sum_{l=0}^{L} \sum_{k=0}^{K} z_i^{(l)} \beta_{kt}^{(l)} \mathcal{B}_k(Y_{it}, K).$$

For example, when there are no covariates $L = 0$, the expression above reduces to the standard Bernstein polynomial approximation $g_t(Y_{it}) \approx \sum_{k=0}^{K} \beta_{kt}^{(0)} \mathcal{B}_k(Y_{it}, K)$. The Bernstein coefficients are linear functions of the demographics, and the dependence of the Bernstein coefficients on the demographics is allowed to vary with time. In this way, the relationship between the (nonlinear) demand Y_{it} and the latent budget Y_{it}^* depends on demographic characteristics and on prices, through t. Since Bernstein polynomials are defined on the unit interval, we normalize Y_{it} to be uniform on $[0, 1]$ by applying its empirical distribution function.[22]

Unconditional moments. To form GMM estimators, we construct the following unconditional moments:

$$E\left[\left(\sum_{l=0}^{L} \sum_{k=0}^{K} z_i^{(l)} \left(\beta_{k2}^{(l)} \mathcal{B}_k(Y_{i2}, K) - \beta_{k1}^{(l)} \mathcal{B}_k(Y_{i1}, K)\right) - \triangle X_{it}\right) \mathcal{B}_{k'}(x_{it}, K) z_i^{(l')}\right] = 0$$

for $k' = 0, \ldots, K$, $l' = 0, \ldots, L$, and $t = 1, 2$, and

$$E\left[\left(\sum_{l=0}^{L} \sum_{k=0}^{K} z_i^{(l)} \left(\beta_{k2}^{(l)} \mathcal{B}_k(Y_{i2}, K) - \beta_{k1}^{(l)} \mathcal{B}_k(Y_{i1}, K)\right) - \triangle X_{it}\right) X_{it} z_i^{(l')}\right] = 0,$$

for $l' = 0, \ldots, L$ and $t = 1, 2$. We include the second condition (where the logged household budget X_{it} is exogenous) because we ultimately wish to consider the correlation of α_i and X_{it}. For a given order of the sieve approximation, K, the equations

[22] We use this normalization for the estimation of the Bernstein coefficients, but we present our results in terms of untransformed Y_{it}. These results are obtained by applying the inverse transformation to the function estimated with transformed data.

above amount to a parametric linear GMM problem.

We impose increasingness on the estimates of the functions g_1 and g_2 by imposing $\beta_{kt}(z_i) \geq \beta_{k-1,t}(z_i)$, for all z_i, for all t, and for $k \geq 2$. This results in a quadratic programming problem with linear inequality restrictions, which we implement in `R` using the `quadprog` package.

Degree of Bernstein polynomial. Implementing this method requires the selection of the degree of the Bernstein polynomial, K. While developing a formal selection rule for this parameter would be desirable, it is beyond the scope of the present paper. Nonetheless, we adopt an informal selection rule for the number of Bernstein basis functions – the smoothing parameter – based on the following observation. In our semiparametric setting, the estimators are known to have the same asymptotic distribution for a range of smoothing parameters (see, e.g., Chen (2007)). When the number of Bernstein basis functions is small, the bias dominates and the estimates exhibit a decreasing bias as the number of terms increases. On the other hand, when the number of basis functions is large, the statistical noise dominates. We implement our estimation method over a range of smoothing parameter values, that is, $K \in \{1, 2, \ldots, 12\}$, in search of a region where the estimates are not very sensitive to small variations in the smoothing parameter. We select the mid-point of that region, so our main results use $K = 8$. We present results for $K \in \{1, 4, 8, 10\}$ in the Appendix. In the Appendix table, "X" denotes a case where an estimated variance is negative.

Confidence bands. We compute confidence bands via the nonparametric bootstrap, although we do not provide a formal justification for it in this setting.[23] We use 1,000 bootstrap replications for all our results. We report pointwise 95% confidence bands. All reported estimates in the main text are bias-corrected using the bootstrap.

Estimates and parameters of interest. We present estimates for the demand functions, $\widehat{h}_t(x, \overline{z}) = \widehat{g_t}^{-1}(\cdot, \overline{z})$ for $t = 1, 2$, where $\bar{z}$ represents a household with 1 child and has other observed characteristics that are the average of those of 1-child households. The functions h_t (demand functions) are not of direct interest, but identification of the h_t supports identification of moments of the distribution of fixed effects α_i. In our context, α_i equals the log of the resource share of the woman in household i.

We characterize the following interesting features of the distribution of resource shares. Recall from Theorems 4 and 5 that identification of features of this distribution does not impose a normalization on assignable good demand functions, and only identifies the distribution of logged resource shares (fixed effects) up to location. Consequently, we only identify features of the resource share distribution up to a scale normalization.

Let $\widehat{g}_{it} = \widehat{g}_t(Y_{it}, z_i)$ equal the predicted values of the inverse demand functions at the observed data. Recall that $g_t(Y_{it}, z_i) = \alpha_i + X_{it} - U_{it}$, so we can think of $\widehat{g}_{it} - X_{it}$ as a prediction of $\alpha_i - U_{it}$. We then compute the following summary statistics of

[23] When the parameter is on the boundary and the bootstrap is not consistent, it is common in the literature to perform an m-out-of-n bootstrap, see, e.g., Andrews (1999, 2000).

interest, leaving the dependence of $\hat{g}_{it}$, $t = 1, 2$, on z_i implicit:

1. an estimate of the standard deviation of α_i given by

$$\hat{std}(\alpha) = \sqrt{\hat{cov}\left((\widehat{g}_{i2} - X_{i2}), (\widehat{g}_{i1} - X_{i1})\right)},$$

   where $\hat{cov}$ denotes the sample covariance. The standard deviation of logs is a standard (scale-free) inequality measure. So this gives a direct measure of inter-household variation in women's resource shares.

2. an estimate of the standard deviation of the projection error, e_i, of α_i on $\bar{X}_i = \frac{1}{2}(X_{i1} + X_{i2})$ and Z_i. Consider the projection

$$\alpha_i = \gamma_1 \bar{X}_i + \gamma_2 Z_i + e_i,$$

   where Z_i contains a constant. We are interested in the standard deviation of e_i. To obtain this parameter, we compute estimators for γ_1, γ_2 from the pooled linear regression of $\hat{g}_{it} - X_{it}$ on $\bar{X}_i$ and Z_i. Call these estimators $\hat{\gamma}_1, \hat{\gamma}_2$. Then, as in $\hat{std}(\alpha)$ in (1), an estimate of the standard deviation of e_i is given by:

$$\hat{std}(e_i) = \sqrt{\hat{cov}\left(\left(\widehat{g}_{i2} - X_{i2} - \hat{\gamma}_1 \bar{X}_i - \hat{\gamma}_2 Z_i\right), \left(\widehat{g}_{i1} - X_{i1} - \hat{\gamma}_1 \bar{X}_i - \hat{\gamma}_2 Z_i\right)\right)}.$$

   This object measures the amount of variation in α_i that cannot be explained with observed regressors. If it is zero, then we don't really need to account for household-level unobserved heterogeneity in resource shares. It is much larger than zero, then accounting for household-level unobserved heterogeneity is important.

3. an estimate of the standard deviation of $\alpha_i + X_{it}$ for $t = 1, 2$, computed as

$$\hat{std}(\alpha_i + X_{it}) = \sqrt{\hat{var}(\alpha_i) + \hat{var}(X_{it}) + 2\hat{cov}(\alpha_i, X_{it})},$$

   where

$$\hat{cov}(\alpha_i, X_{i1}) = cov(\hat{g}_{i1} - X_{i1}, X_{i1}),$$
$$\hat{cov}(\alpha_i, X_{i2}) = cov(\hat{g}_{i2} - X_{i2}, X_{i2}),$$

   and $\hat{var}(\alpha_i) = \left(\hat{std}(\alpha)\right)^2$ and $\hat{var}(X_{it})$, $t = 1, 2$, is observed in the data. Since $\alpha_i + X_{it}$ is a measure of the woman's shadow budget, $\hat{std}(\alpha_i + X_{it})$ is a measure of inter-household inequality in women's shadow budgets. This inequality measure is directly comparable to the standard deviation of X_i (shown in Table 1), which measures inequality in household budgets.

4. an estimate of the covariance of α_i, X_{it} for $t = 1, 2$, denoted $\hat{cov}(\alpha_i, X_{it})$. This

object is of direct interest to applied researchers using cross-sectional data to identify resource shares. If this covariance is non-zero, then the independence of resource shares and household budgets is cast into doubt, and identification strategies based on this restriction are threatened.

Of these, the first 2 summary statistics are about the variance of fixed effects, and are computed using data from both years. Their validity requires serial independence of the measurement errors U_{it}. In contrast, the second 2 summary statistics are about the correlation of fixed effects with the household budget, and are computed at the year level. They are valid with stationary U_{it}, even in the presence of serial correlation.

We also consider the multivariate relationship between resource shares, household budgets and demographics. Recall that the fixed effect α_i subject to a location normalization; this means that resource shares are subject to a scale normalization. So, we construct an estimate of the woman's resource share in each household as $\widehat{\eta}_i = \exp\left(\frac{1}{2}\left(\hat{g}_{i1} - X_{i1}\right) + \left(\hat{g}_{i1} - X_{i1}\right)\right)$, normalized to have an average value of 0.33. Then, we regress estimated resource shares $\widehat{\eta}_i$ on $\bar{X}_i$ and Z_i, and present the estimated regression coefficients, which may be directly compared with similar estimates in the cross-sectional literature.

The estimated coefficient on X_i gives the conditional dependence of resource shares on household budgets, and therefore speaks to the reasonableness of the restriction that resource shares are independent of those budgets (an identifying restriction used in the cross-sectional literature). Finally, using the estimate of the variance of fixed effects, we construct an estimate of R^2 in the regression of resource shares on observed covariates. This provides an estimate of how much unobserved heterogeneity matters in the overall variation of resource shares.

We also provide estimates that use women's clothing as the assignable non-shareable good. Here, clothing expenditure is equal to four times the reported three-month recall expenditure on the following female-specific clothing items: Saree; Blouse/ petticoat; Salwar kameez; and Orna. We note that although clothing is a semi-durable good, there are 3 years between the waves of the panel. Consequently, we do not think that the demands across periods will be strongly correlated due to the durability of clothing purchased.

Appendix table "Additional Estimates" gives results for $K \in \{1, 4, 8, 10\}$ for food (top panel) and clothing (bottom panel). In the table, "X" denotes a case where an estimated variance is negative.

C Descriptive Statistics

Table 2: Descriptive Statistics

	raw data				top- and bottom-coded and normalized	
Variable	Mean	Std Dev	Min	Max	Mean	Std Dev
X_{i1}	11.68	0.49	10.24	13.68		
X_{i2}	12.02	0.53	10.70	14.08		
$\triangle X_i$	0.34	0.43	-0.94	2.31		
Y_{i1}	1561	1277	0	13500		
Y_{i2}	1912	1727	0	15000		
age of woman	32.76	7.64	19.00	90.00	0.35	0.25
age of man	39.77	10.76	15.00	105.00	0.37	0.25
2 children	0.47	0.50	0.00	1.00	0.47	0.50
3 or 4 children	0.19	0.39	0.00	1.00	0.19	0.39
age of children	8.37	2.75	1.00	14.00	0.50	0.25
fraction girl children	0.46	0.39	0.00	1.00	0.46	0.39
education of woman	4.10	3.43	0.00	10.00	0.44	0.37
education of man	3.41	3.74	0.00	10.00	0.34	0.37

Table 2: Summary statistics.

D Additional estimates

	1			4			8			10		
	Estimate	q(2.5)	q(97.5)	Estimate	q(2.5)	q(97.5)	Estimate	q(2.5)	q(97.5)	Estimate	q(2.5)	q(97.5)
Variability in unobservables: Food												
sd_a	0.3574	0.2258	0.5025	0.3197	0.2210	0.4376	0.2647	0.1518	0.3707	0.2286	0.0462	0.3592
sd_e	0.1947	0.1401	0.2377	0.1845	0.1413	0.2202	0.1637	0.1262	0.1931	0.1631	0.1308	0.1908
sd_ax1	-0.0506	-0.0934	-0.0085	-0.0805	-0.1300	-0.0246	-0.0901	-0.1292	-0.0429	-0.0781	-0.1166	-0.0339
sd_ax2	-0.0549	-0.0979	-0.0129	-0.0832	-0.1313	-0.0277	-0.1034	-0.1418	-0.0561	-0.0930	-0.1310	-0.0497
cov_ax1	0.5159	0.4067	0.6465	0.4272	0.3279	0.5763	0.3537	0.2720	0.4605	0.3635	0.2846	0.4583
cov_ax2	0.5444	0.4489	0.6637	0.4636	0.3757	0.6107	0.3763	0.3010	0.4760	0.3822	0.3104	0.4781
Regression coefficients: Food												
$R^2 : \overline{X}_i, z_i$ on η_i	0.6749	0.5292	0.8088	0.6341	0.4179	0.8026	0.5205	0.3361	0.6508	0.5369	0.3797	0.6447
$\overline{X}_i$	-0.0353	-0.0643	-0.0083	-0.0237	-0.0534	0.0039	-0.0452	-0.0700	-0.0196	-0.0481	-0.0731	-0.0234
age_women	-0.1095	-0.4028	0.2182	-0.2807	-0.5228	-0.0498	-0.2513	-0.4437	-0.0545	-0.3438	-0.5891	-0.1424
age_men	-0.0284	-0.3205	0.2123	0.1146	-0.0846	0.2811	0.2845	0.1389	0.4191	0.2904	0.1306	0.4392
children2	-0.1948	-0.3296	-0.0860	-0.1955	-0.3180	-0.0884	-0.0502	-0.1694	0.0413	-0.0821	-0.2229	0.0167
children3p	-0.2134	-0.3372	-0.1132	-0.2231	-0.3440	-0.0991	-0.1248	-0.2407	0.0140	-0.1159	-0.2466	0.0197
age_children	-0.3004	-0.5448	-0.0620	-0.0819	-0.2875	0.1477	-0.0095	-0.1979	0.2151	0.0949	-0.0910	0.3209
frac_girl	-0.0338	-0.1724	0.1209	0.0674	-0.0487	0.2049	0.0136	-0.0881	0.1246	-0.0322	-0.1466	0.0854
edu_women	0.0250	-0.1319	0.1803	0.0012	-0.1546	0.2036	-0.0325	-0.1519	0.1198	-0.1096	-0.2278	0.0132
edu_men	0.0079	-0.1213	0.1533	-0.1385	-0.2764	0.0091	-0.1330	-0.2589	-0.0015	-0.0202	-0.1501	0.1298
Variability in unobservables: Cloth												
sd_a	0.4729	X	0.7989	0.4224	0.2622	0.6944	0.2690	0.1472	0.4265	0.2443	0.1142	0.3749
sd_e	0.2350	X	0.3936	0.2131	0.1139	0.2875	X	X	0.1210	X	X	X
sd_ax1	-0.1254	-0.2157	-0.0341	-0.0503	-0.1049	0.0304	-0.1394	-0.1745	-0.0994	-0.1659	-0.1964	-0.1340
sd_ax2	-0.1185	-0.2186	-0.0207	-0.0903	-0.1583	-0.0043	-0.1832	-0.2211	-0.1406	-0.1903	-0.2234	-0.1537
cov_ax1	0.4633	X	0.7828	0.5667	0.4144	0.8386	0.1927	X	0.4568	X	X	0.3213
cov_ax2	0.5203	X	0.8248	0.5319	0.3654	0.8218	X	X	0.4190	X	X	0.3217
Regression coefficients: Cloth												
$R^2 : \overline{X}_i, z_i$ on η_i	11.2428	10.9458	14.5224	0.3753	-0.0344	1.0855	0.8791	0.6307	1.2056	1.1260	0.9142	1.3417
$\overline{X}_i$	0.1716	0.0975	0.2610	0.0015	-0.0509	0.0574	-0.0925	-0.1245	-0.0591	-0.1488	-0.1738	-0.1200
age_women	0.2658	-0.4693	1.1146	-0.3608	-0.6580	-0.0759	-0.2666	-0.4136	-0.0894	-0.1123	-0.2224	0.0056
age_men	0.1706	-0.3577	0.8011	0.5223	0.3033	0.8700	0.2860	0.1753	0.4254	0.1840	0.0961	0.2934
children2	-0.5101	-0.8971	-0.1832	0.0950	-0.0863	0.3863	0.0495	-0.0380	0.1572	0.0623	0.0008	0.1369
children3p	-0.3017	-0.7165	0.2456	-0.1110	-0.2746	0.1671	-0.0377	-0.1448	0.2903	-0.0549	-0.1511	0.2868
age_children	-1.1070	-1.7016	-0.5342	-0.0087	-0.2421	0.2433	0.0449	-0.0820	0.1722	-0.0925	-0.2044	-0.0050
frac_girl	-0.0222	-0.3256	0.3410	-0.3711	-0.5782	-0.2411	0.0926	0.0196	0.1633	0.0620	0.0028	0.1103
edu_women	-0.0821	-0.5592	0.3867	-0.0727	-0.2605	0.1185	-0.2170	-0.3154	-0.1348	-0.1029	-0.1815	-0.0414
edu_men	-0.5244	-0.9658	-0.1072	0.0746	-0.1260	0.4180	0.0420	-0.0497	0.1742	0.0167	-0.0537	0.1036

Table 3: Summary statistics for log resource shares and their projection on demographics and budget.