Economics Letters 163 (2018) 75–78

Contents lists available at ScienceDirect

Economics Letters

journal homepage: www.elsevier.com/locate/ecolet

Estimating household resource shares: A shrinkage approach[☆]

Denni Tommasi[a], Alexander Wolf[b,*]

[a] Université libre de Bruxelles, ECARES, Avenue F.D. Roosevelt 50, 1050 Brussels, Belgium
[b] Institute for Fiscal Studies, 7 Ridgmount Street, London, WC1E 7AE, UK

HIGHLIGHTS

- Collective models show great promise in the analysis of intra-household welfare.
- But their empirical application has proven difficult in practice.
- We show how a common feature of these models makes the task so difficult.
- We propose an empirical strategy involving shrinkage to stabilize the estimates.

ARTICLE INFO

Article history:
Received 14 September 2017
Received in revised form 17 November 2017
Accepted 22 November 2017
Available online 6 December 2017

JEL classification:
D13
D11
D12
C31
I32

Keywords:
Collective model
Sharing rule
Resource shares
Demand system
Engel curves
Shrinkage

ABSTRACT

Collective models identifying resource shares are promising tools to analyze intra-household welfare and poverty. However, their empirical application has proven difficult in practice as authors contend with large standard errors and unstable estimates. This paper uses a prominent framework to show how a common feature of the structure of these models makes the task so difficult and proposes an empirical strategy to stabilize the estimates.

1. Introduction

Collective models of the household (Chiappori, 1988, 1992) have become the go-to approach to study intra-household allocations. The ability of models based on Browning et al. (2013, hereafter BCL) to identify *resource shares*, that is, the fraction of household resources devoted to each member, has made them attractive to researchers investigating intra-household welfare and poverty.[1] . However, their estimation has proven difficult in practice. Authors have to contend with large standard errors, unstable estimates and difficult optimization procedures.[2] The source of these difficulties lies in the complexity of the task at hand: to learn about resource allocation among individuals from household-level consumption data. To do this, models have to account for other

☆ This is a substantial revision of an earlier paper that circulated under the title "Overcoming Weak Identification in the Estimation of Household Resource Shares" (ECARES working paper 2016-12). We are grateful to Bram De Rock and Arthur Lewbel for their invaluable comments on this project, as well as participants to several conferences and seminars for useful suggestions. Denni acknowledges financial support from the Fonds National de la Recherche Scientifique (FNRS). All remaining errors are ours.

* Corresponding author.
E-mail addresses: dtommasi@ulb.ac.be (D. Tommasi), alex.wolf@ifs.org.uk (A. Wolf).

1 See Lewbel and Pendakur (2008), Bargain and Donni (2012), Dunbar et al. (2013) and applications by Cherchye et al. (2012), Bargain et al. (2014), Calvi (2016), Calvi et al. (2017), Tommasi (2017).

2 See Wolf (2016) for a more detailed discussion of these problems.

https://doi.org/10.1016/j.econlet.2017.11.030

drivers of patterns in the data, such as preferences and consumption technologies.

Demand systems derived from BCL share the following structure, where demand for a good k is expressed in household budget shares w^k as a function of resource shares η_j for members $j = 1, 2$ of a couple's household and of *desired budget shares* w_j^k, which describe member j's preferences:

$$w^k(p, y) = \sum_{j=1,2} \eta_j(p, y) w_j^k(\pi(p), y\eta_j(p, y)) \quad (1)$$

where y is total household expenditure, p are market prices and $\pi(p)$ are intra-household shadow prices.[3] A key difficulty for estimation can be spotted in Eq. (1): each summand is a product of an individual's resource share and her desired budget share. This *multiplicative feature* induces trade-offs between parameters and makes it hard to pin down the value of the parameter of interest: the resource share η_j.

In this paper, we use the Dunbar et al. (2013, hereafter DLP) model to discuss the consequences for estimation caused by this multiplicative feature and offer a simple solution. The simplified structure of this model has not only made it the most popular approach among practitioners, but also makes it well-suited for our exposition, as the consequences of the above structure emerge clearly.

2. Trade-offs in the model

Starting from (1), DLP make the following identifying assumptions. First, they focus only on household demand for private assignable goods, that is, goods for which we can assume that only one member consumes them and for which there are no economies of scale. Second, they assume that preferences of household members are similar across people (SAP), instead of identical to singles (a common assumption in BCL-type models), and that $\eta_j \perp y$ (Menon et al., 2012). Then, under PIGLOG utility functions, the resulting system maintains the multiplicative feature and takes the following form:

$$\begin{aligned} w^1(y) &= \eta \quad (\delta + \Delta + \beta \ln(\eta y)) \\ w^2(y) &= (1-\eta) \quad (\delta + \beta \ln((1-\eta)y)) \end{aligned} \quad (2)$$

where the desired budget share functions, for each member, are linear Engel curves in log *individual resources* ηy (or $(1-\eta)y$ for member 2), and η, δ, Δ, and β, are parameters to be estimated. In applications, these are typically replaced by linear indexes in characteristics to account for observed heterogeneity. By the SAP assumption, and importantly for us, the constant terms of these curves are allowed to differ by Δ between the two members, whereas the slope β is constrained to be the same.

System (2) allows us to reason fairly straightforwardly about *trade-offs* in the model. Suppose an optimal fit to data has been found. If we now slightly modify Δ, we can obtain a fit that is nearly as good as before by also modifying η in the opposite direction. This is illustrated in Fig. 1, where, for different pairs (η, Δ), we plot minimal values of the root sum of squares (RSS) associated with a toy example of (2). A dashed red line marks the floor of a valley along which the two parameters can be traded off cheaply (in RSS sense), making recovery of either value hard in practice. Though similar trade-offs characterize other BCL-type models, it is especially easy to show in DLP, where it is linear. Put another way, a strong correlation between pseudo-regressors is induced in our system (Greene, 2003, Chapter 17), inducing a corresponding correlation in parameter estimates. This problem can be seen as analogous to multicollinearity, though crucially ours is a feature of the model rather than of the data (See Appendix).

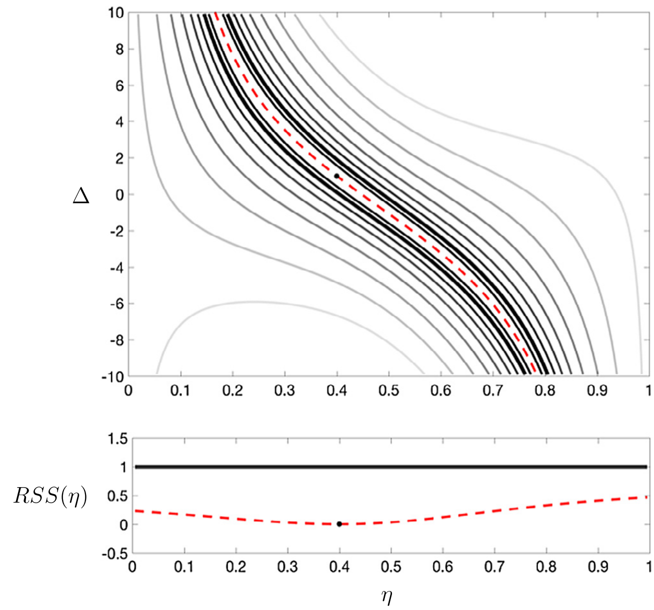


Fig. 1. RSS minima as a function of (fixed) η and Δ.

Parameter estimates are affected in two ways, which are illustrated by the black dots in Fig. 2. First, it makes the location of the sample mean η_0 of the sharing rule more uncertain (Fig. 2(a)).[4] Second, when parameters are replaced by linear indexes in household characteristics to capture observed heterogeneity, the estimated indexes for each household h, $\hat{\eta}_h$ and $\hat{\Delta}_h$, have a strong negative correlation across households (Fig. 2(b)), where $\hat{\eta}_h = \hat{\eta}_0 + \hat{\eta}_1 x_1 + \cdots + \hat{\eta}_k x_k$ and $\hat{\Delta}_h = \hat{\Delta}_0 + \hat{\Delta}_1 x_1 + \cdots + \hat{\Delta}_k x_k$. This occurs reliably even when the true correlation is large and positive. Since the model is identified, estimates by nonlinear least squares are consistent. However, at common sample sizes in household surveys, the issue described here is an important obstacle, yielding large standard errors and unstable estimates.

3. Stabilization

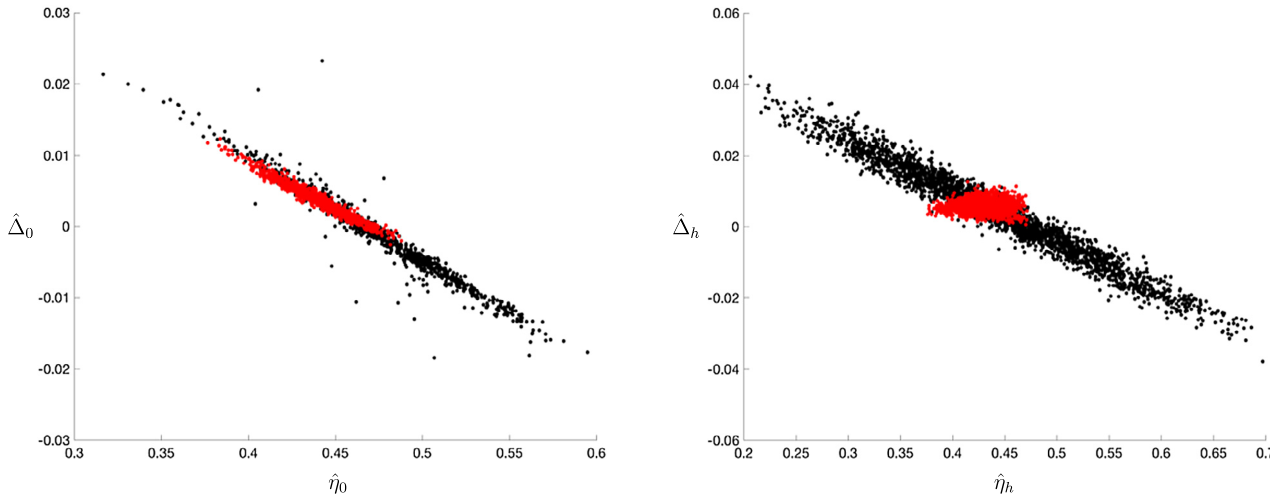
In order to achieve stabilization of the estimates at minimal cost, we proceed in two acts. First, the analogy with multicollinearity suggests that a shrinkage method may be beneficial in reducing the uncertainty around the location of the mean resource share. Second, we restrict the (artificial) correlation between the indexes η_h and Δ_h to address their distortion. These two issues turn out to be independent from one another in the sense that the remedy to one has no effect on the other.

We use data on singles to introduce prior information and design our shrinkage term. We formalize our rationale for shrinkage by assuming similarities between singles' budget shares w_s^j (which are estimated in a first step) and married individuals' desired budget shares. While DLP emphasize that they do not assume such similarities, we will do so in a minimal fashion which bears little resemblance to BCL's assumption of identical preferences. This amounts to using economic theory to motivate and construct

---

[3] Eq. (1) holds only if the shadow consumption z is linear in purchased quantities q, with $z = Aq$, where A is a diagonal matrix describing a linear consumption technology. This technology is common to all BCL-type models and notably disallows complementarities and overheads in scale economies.

[4] This comparison needs a reference point where these effects are attenuated by means of limiting the extent of these trade-offs. Our approach, detailed below, provides this reference point.



(a) Means of $\hat{\eta}_h$ vs $\hat{\Delta}_h$. (b) $[\hat{\eta}_h$ vs $\hat{\Delta}_h]$ in one draw.

Fig. 2. Estimates of resource shares and taste differences.

a shrinkage estimator within our collective framework. This idea provides a middle ground between nested models which is particularly useful in our context (see Fessler and Kasy, 2017 for a related approach).

Assumption. For some 'anchor' value y_a of individual resources and at means of the demographic variables, the ratios

$$R_c = \frac{\delta_h + \Delta_h + \beta_h \ln y_a}{\delta_h + \beta_h \ln y_a} \text{ and } R_s = \frac{w_s^1(\ln y_a)}{w_s^2(\ln y_a)}$$

are such that

$$\tau < \frac{R_c}{R_s} < \frac{1}{\tau}$$

where $\tau \in (0, 1)$.

The researcher's choice of τ is a measure of the strength of this assumption, which may be context or data specific: $\tau = 1$ implies that we assume $R_c = R_s$, whereas $\tau \approx 0$ implies a completely agnostic stance. Note that even the strict case remains considerably weaker than the common assumption that singles and married individuals have the same preferences. Based on this assumption, we modify the objective function by adding a shrinkage term, which consists of a shrinkage parameter λ times a suitable distance between R_c and R_s:

$$\min_{\eta,\delta,\Delta,\beta} n^{-1} RSS(\eta, \delta, \Delta, \beta) + \lambda \left(\exp \left| \ln \left(\frac{\hat{R}_s}{R_c} \right) \right| - 1 \right)^2 \quad (3)$$

where n gives the sample size. The parameter λ needs to be chosen optimally. If it is too large, we bias the estimator too strongly toward R_s, the ratio estimated on singles' data. If it is too small, we do not shrink enough and are faced with the high model-induced variance described above. The optimal value will be a function of the problem at hand as well as of the strength of our assumption, as given by τ.

To choose λ for a given application, we set up a simulation of the model matching moments of the data at hand. That is, we rely on total expenditure and demographic information on couples in the data and generate simulated budget shares using the model, choosing true parameter values and an error term so as to match consumption patterns in the data set. The true parameters are set such that either $\frac{R_c}{R_s} = \tau$ or $\frac{R_c}{R_s} = \frac{1}{\tau}$, where R_s is simply set to one. Then we select the value of λ that minimizes $MSE(\eta_0)$, the

Table 1
Monte Carlo simulation: Estimated root MSE (RMSE) of η_0 for $\tau = 0.7$.

$\log_{10}(\lambda)$	RMSE		Variance share	
	$\frac{R_c}{R_s} = \tau$	$\frac{R_c}{R_s} = \frac{1}{\tau}$	$\frac{R_c}{R_s} = \tau$	$\frac{R_c}{R_s} = \frac{1}{\tau}$
−6.50	0.0415	0.0395	0.8840	0.9520
−6.00	**0.0406**	**0.0376**	**0.5694**	**0.6424**
−5.50	0.0483	0.0458	0.1779	0.1920

Notes: The corresponding root mean squared error (RMSE) for $\lambda = 0$ was 0.0548. Each trial used 500 replications.

loss associated with the estimation of the mean resource share and proceed to estimation on the full data set of singles and couples using the chosen value of the shrinkage parameter.

To address the induced negative correlation shown in Fig. 2(b), we require that the correlation be zero. While our restriction is not implied by theory, values far from zero are highly counterintuitive.[5] The objective function in (3) is thus minimized such that $cov(\eta_h, \Delta_h) = 0$. It is important to note what this restriction *does not* mean: It does not mean that any of the effect sizes must be zero. Rather, the restriction only implies that Δ_h, the difference in *desired budget share functions* between the household members, may not depend on the resource share.

In the earlier version of this paper (Tommasi and Wolf, 2016), we apply the above methodology to a Mexican dataset. We set $\tau = 0.7$ and search a grid of values to find the optimal value of $\lambda = 10^{-6}$ as described in Table 1. Comparing our approach to an unrestricted estimator in a Monte Carlo simulation, we show how substantial reductions in loss are obtained not only for the mean resource share parameter (red dots in Fig. 2(a)), but also for covariate effects (red dots in Fig. 2(b)). Table 2 shows the coefficients associated with each of the twelve covariates used in the trial, all of which see large reductions in loss. These gains in estimator precision imply important gains in the precision with which we can estimate individual poverty rates or individual welfare effects. Consider for instance the effect of a treatment program on resource distribution. To investigate this effect using the DLP model, the researcher would include it as a covariate in the resource share η and be interested in the coefficient $\hat{\eta}_1$, the effect size. The numbers given in 2 show how large the gain in precision on this crucial parameter can quite realistically be, cutting through the model-induced noise, while any bias remains minimal.

5 For instance, a strong negative correlation implies that those men who have very small resource shares also reliably happen to have strong tastes for clothing.

Table 2
True values, MC means, standard deviations and root mean squared errors for the parameter vector η. $\hat{R}_s$ set to R_c/τ.

	True	LS Estimator			Shrinkage Estimator		
		Mean	SD	RMSE	Mean	SD	RMSE
η_0	0.464	0.484	0.039	0.044	0.441	0.014	0.027
η_1	0.010	0.010	0.022	0.022	0.009	0.006	0.006
η_2	0.004	0.004	0.025	0.025	0.003	0.007	0.008
η_3	−0.009	−0.007	0.026	0.026	−0.006	0.007	0.008
η_4	−0.008	−0.004	0.018	0.018	−0.005	0.005	0.005
η_5	−0.002	−0.004	0.024	0.024	−0.002	0.006	0.006
η_6	−0.005	−0.004	0.042	0.042	−0.005	0.015	0.015
η_7	0.011	0.007	0.023	0.023	0.007	0.007	0.008
η_8	−0.010	−0.011	0.024	0.024	−0.005	0.007	0.009
η_9	0.001	−0.002	0.021	0.021	−0.003	0.006	0.007
η_{10}	0.007	0.004	0.029	0.030	0.002	0.011	0.013
η_{11}	0.004	0.004	0.022	0.022	0.002	0.008	0.008
η_{12}	0.001	0.003	0.029	0.029	0.002	0.012	0.012

4. Asymptotic behavior

While no proof of consistency will follow, we do aim to achieve an intuition for what happens in the limit for large n. Under the model assumptions (including the restriction of $cov(\eta_h, \Delta_h)$), the least squares estimator is consistent even for zero-mean heteroscedastic errors (Greene, 2003, Chapter 9 and 10). Consistency of this estimator is equivalent to saying that the MSE of least squares will be zero in the limit for $n \to \infty$. Given our strategy for choosing λ, and since the MSE cannot be negative, our shrinkage parameter must converge to zero as well. In other words, for $n \to \infty$, our strategy becomes irrelevant.

5. Conclusion

We aim to convince the reader of three things. First, while collective models identifying resource shares have great promise as tools to better understand intra-household dynamics, obstacles remain. If they are to succeed empirically, practical issues like the one described here need to be understood and controlled for so as to improve the reliability of the estimates. Second, we show how solid improvements in the estimator's performance can be achieved in the popular DLP model by directly targeting these issues while making minimal additional assumptions. Although we do not explicitly carry over the same analysis to the rest of the BCL-descended models, the same issue can be addressed in similar ways (Wolf, 2016), since they all share the same multiplicative structure. Third, more generally, future research on methods to improve model behavior will surely be necessary to realize the great empirical potential of collective models. Such work rhymes well with recent interest in problems of estimation in structural models with weak identification (Andrews and Mikusheva, 2014) and the issue discussed here is already attracting further theoretical attention (Han and McCloskey, 2017).

Appendix

Correlated pseudo-regressors

Suppose we wish to estimate the simplified model in (A.1) by NLS. We would rely on pseudo-regressors, the derivatives of the model by its parameters.

$$\begin{aligned} w^1(y) &= \eta \quad &(\delta + \Delta \quad &+ \beta \ln(\eta y)) \\ w^2(y) &= (1-\eta) \quad &(\delta \quad &+ \beta \ln((1-\eta) y)). \end{aligned} \tag{A.1}$$

Because this is a system of two equations, the pseudo-regressors for each parameter in the system are vectors of length $2H$ where H is the number of households in the sample. Looking at the pseudo-regressors for η and Δ, each is composed of two sections of equal length, corresponding to the two equations. Rewriting the above system in vector form and as functions of the log expenditure vector $\vec{X}$ and the parameter vector θ:

$$\begin{aligned} W^1(\vec{X}) &= h_1(\theta, \vec{X}) = \eta \quad &(\delta + \Delta \quad &+ \beta(\vec{X} + \ln(\eta))) \\ W^2(\vec{X}) &= h_2(\theta, \vec{X}) = (1-\eta) \quad &(\delta \quad &+ \beta(\vec{X} + \ln(1-\eta))). \end{aligned} \tag{A.2}$$

We can now denote the pseudo-regressors as

$$\frac{\partial}{\partial \Delta} h(\theta, \vec{X}) = \begin{pmatrix} \frac{\partial}{\partial \Delta} h_1(\theta, \vec{X}) \\ \frac{\partial}{\partial \Delta} h_2(\theta, \vec{X}) \end{pmatrix} = \begin{pmatrix} \vec{\eta} \\ \vec{0} \end{pmatrix} \tag{A.3}$$

and

$$\frac{\partial}{\partial \eta} h(\theta, \vec{X}) = \begin{pmatrix} \frac{\partial}{\partial \eta} h_1(\theta, \vec{X}) \\ \frac{\partial}{\partial \eta} h_2(\theta, \vec{X}) \end{pmatrix} = \begin{pmatrix} \delta + \Delta + \beta(\vec{X} + \ln(\eta)) + \beta \\ -\delta - \beta(\vec{X} + \ln(1-\eta)) - \beta \end{pmatrix}. \tag{A.4}$$

It is now easy to see that the two pseudo-regressors will be positively correlated for positive β: Knowing that the desired budget shares (in large parentheses in (A.1)) must be positive for each individual, we can see that $\frac{\partial}{\partial \eta} h_1(\theta, \vec{X}) > 0$ and that $\frac{\partial}{\partial \eta} h_2(\theta, \vec{X}) < 0$. Given that $\eta > 0$ by definition, the covariance between these pseudo-regressors must be positive.

References

Andrews, I., Mikusheva, A., 2014. Weak identification in maximum likelihood: A question of information. Am. Econ. Rev. 104 (5), 195–199.

Bargain, O., Donni, O., 2012. Expenditure on children: A Rothbarth-type method consistent with scale economies and parents' bargaining. Eur. Econ. Rev. 56 (4), 792–813.

Bargain, O., Donni, O., Kwenda, P., 2014. Intrahousehold distribution and poverty: Evidence from côte d'ivoire. J. Dev. Econ. 107, 262–276.

Browning, M., Chiappori, P.-a., Lewbel, A., 2013. Estimating consumption economies of scale, adult equivalence scales, and household bargaining power. Rev. Econom. Stud. 80 (4), 1267–1303.

Calvi, R., 2016, Why are older women missing in india? the age profile of bargaining power and poverty, Working paper, Boston College.

Calvi, R., Lewbel, A., Tommasi, D., 2017. Women's empowerment and family health: Estimating LATE with mismeasured treatment, Working Paper 2017–27, Université libre de Bruxelles, ECARES.

Cherchye, L., De Rock, B., Vermeulen, F., 2012. Economic well-being and poverty among the elderly: An analysis based on a collective consumption model. Eur. Econ. Rev. 56 (6), 985–1000.

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

Chiappori, P., 1992. Collective labor supply and welfare. J. Polit. Econ. 100 (3), 437–467.

Dunbar, G.R., Lewbel, A., Pendakur, K., 2013. Children's resources in collective households: Identification, estimation, and an application to child poverty in Malawi. Am. Econ. Rev. 103 (1), 438–471.

Fessler, P., Kasy, M., 2017. How to use economic theory to improve estimators. Working paper.

Greene, W.H., 2003. Econometric Analysis. Pearson Education.

Han, S., McCloskey, A., 2017. Estimation and inference with a (nearly) singular jacobian, Working paper.

Lewbel, A., Pendakur, K., 2008. Estimation of collective household models with Engel curves. J. Econometrics 147 (2), 350–358.

Menon, M., Pendakur, K., Perali, F., 2012. On the expenditure-dependence of children's resource shares. Econom. Lett. 117 (3), 739–742.

Tommasi, D., 2017. Control of resources and demand for food. Working paper, Université libre de Bruxelles, ECARES.

Tommasi, D., Wolf, A., 2016. Overcoming Weak Identification in the Estimation of Household Resource Shares, Working paper, Université libre de Bruxelles, ECARES.

Wolf, A., 2016. Resource Allocation in Couples A Collective Model with Prior Information, Working paper, Université libre de Bruxelles, ECARES.