Taylor & Francis
Taylor & Francis Group

Econometric Reviews

ISSN: 0747-4938 (Print) 1532-4168 (Online) Journal homepage: http://www.tandfonline.com/loi/lecr20

A Composite Generalized Cross-Entropy Formulation in Small Samples Estimation

R. Bernardini Papalia

To cite this article: R. Bernardini Papalia (2008) A Composite Generalized Cross-Entropy Formulation in Small Samples Estimation, Econometric Reviews, 27:4-6, 596-609, DOI: 10.1080/07474930801960469

To link to this article: https://doi.org/10.1080/07474930801960469

Econometric Reviews, 27(4–6):596–609, 2008

ISSN: 0747-4938 print/1532-4168 online
DOI: 10.1080/07474930801960469

A COMPOSITE GENERALIZED CROSS-ENTROPY FORMULATION IN SMALL SAMPLES ESTIMATION

R. Bernardini Papalia

Department of Statistics, University of Bologna, Bologna, Italy

□ *This article introduces a maximum entropy-based estimation methodology that can be used both to represent the uncertainty of a partial-incomplete economic data generation process and to consider the direct influence of learning from repeated samples. Then, a composite cross-entropy estimator, incorporating information from a subpopulation based on a small sample and from a population with a larger sample size, is derived. The proposed estimator is employed to estimate the local area expenditure shares of a sub population of Italian households using a system of censored demand equations.*

Keywords Generalized cross-entropy; Microeconometric models; Repeated samples; Small sample estimation.

JEL Classification C13; C14; C34; D12.

1. INTRODUCTION

This article introduces a maximum entropy-based methodology for modeling incomplete information and learning from repeated small samples, and formulates a Composite Cross-Entropy (CCE) estimator which combines information available at an aggregate population level with information from subpopulations with small sample sizes.

Our first objective is to define a system of restrictions suitable for representing the uncertainty relative to a partial-incomplete economic data generation process and to account for the direct influence of learning from repeated samples. Our second objective is to derive a composite generalized cross-entropy estimator which combines information across (local) subpopulations based on small samples and a (national) major population with a larger sample size. Following Golan (2001), a

Received November 2005; Accepted January 2007
Address correspondence to R. Bernardini Papalia, Department of Statistics, University of Bologna, Bologna, Italy; E-mail: rossella.bernardini@unibo.it

data-weighted prior estimation rule is proposed with the aim of combining two sources of information, one pertaining to the local area subsample (subpopulation) and the other pertaining to the whole national sample major population of interest). The proposed CCE estimator is employed to estimate the local-area expenditure shares of a subgroup of Italian households by means of a nonlinear system of censored demand functions.

The article is organized as follows: Section 2 presents the basic Generalized Maximum Entropy (GME) and Generalized Cross-Entropy (GCE) formulations. A composite estimation procedure based on the maximum entropy principle is developed and discussed in Section 2.2. The results of an application to the estimation of local area expenditure shares, based on the Italian Expenditure Survey, are presented in Section 3. Finally, concluding remarks are provided in Section 4.

2. METHODOLOGY

2.1. The Basic Formulations

With reference to the existing literature (Golan, Judge, and Miller, 1996a, 1998; Golan, Judge, and Perloff, 1997b) the problem we present in this section involves noisy unobserved data pertaining to (i) those units which are not included in the sample, as in the nonrandom samples, or to (ii) the observations from the design matrix which are missing from the data set, as in the censored model. The unobserved variables may be treated as a set of unknown elements in the estimation problem; thus the aim is to estimate the unknown parameters of the statistical model as well as the unknown values of the unobserved variables. We consider the case in which only the dependent variable has missing values, with the missing-data mechanisms either ignorable or specified by latent modeling. A typical example is given by the almost ideal demand system (AIDS) with censored data presented in Golan et al. (2001).

If we let Y indicate the vector of the observable dependent variables $\{y_i\}$ corresponding to the set of selected units in sample C_1 (or to the set of uncensored observations) and Y^* indicate the vector of latent variables $\{y_i^*\}$, where $i \in C_2$ and C_2 represents the set of nonselected units (or the set of censored observations), the vector of response variables $S[(n_1 + n_2) \times 1]$ can be defined as

$$s_i = \begin{cases} y_i & i \in C_1 \\ y_i^* & i \in C_2 \end{cases} \tag{1}$$

with the ith component y_i, if the ith observation corresponds to a selected unit in the sample (or if the ith observation is uncensored), and with the ith component y_i^*, if it corresponds to a nonselected unit (or refers to a censored observation); n_1 and n_2 represent the sample size of the two subsets of the selected (uncensored) and nonselected (censored) sample units, C_1 and C_2, respectively. In this formulation, the variable of interest is not observable for the units subset C_2, which can represent the set of nonselected units (or the set of units with censored values). These models are seemingly unrelated censored models, where negative values of the response variables are always missing (Amemiya, 1984) and where Y^* is the vector consisting of all the "missing/unobserved" components of S.

Within the GME method of Golan, Judge, and Miller (1996a), based on Shannon's entropy measure (Shannon, 1948) and on the maximum entropy principle developed in Jaynes (1957) and Levine (1980), the following model relationship, expressed in matrix terms

$$\begin{bmatrix} Y \\ Y^* \end{bmatrix} = \begin{bmatrix} X^1 & \\ & X^2 \end{bmatrix} \begin{bmatrix} \beta^1 \\ \beta^2 \end{bmatrix} + \begin{bmatrix} \varepsilon^1 \\ \varepsilon^2 \end{bmatrix} \tag{2}$$

may be "reparameterized" as follows:

$$\begin{bmatrix} Y \\ Y^* \end{bmatrix} = \begin{bmatrix} X^1 & \\ & X^2 \end{bmatrix} \begin{bmatrix} Z^1 p^1 \\ Z^2 p^2 \end{bmatrix} + \begin{bmatrix} V^1 w^1 \\ V^2 w^2 \end{bmatrix}, \tag{3}$$

where X^1 and X^2 are the $(n_1 \times K)$ and $(n_2 \times K)$ design matrices of the K explanatory variables for the two subsets of selected (uncensored) and nonselected (censored) units in the sample, respectively, and $\beta^1 = [\beta_1^1, \ldots, \beta_K^1]', \beta^2 = [\beta_1^2, \ldots, \beta_K^2]'$, and $\varepsilon^1, \varepsilon^2$ are the corresponding vectors of the unknown response parameters and errors, respectively. Within this framework, both coefficients and noise components are reparameterized in terms of unknown probabilities, defined over some finite and discrete support spaces, and the estimation procedure is converted into a constrained optimization problem where the objective function consists of the joint entropy $H(\cdot)$ defined below in Eq. (4). Using the M^1-vector Z_k^1 $(Z_k^1 = [Z_{k1}^1, Z_{k2}^1, \ldots, Z_{kM^1}^1]', M^1 \geq 2)$ of supports for the K-coefficients β^1, the M^2-vector Z_k^2 $(Z_k^2 = [Z_{k1}^2, Z_{k2}^2, \ldots, Z_{kM^2}^2]', M^2 \geq 2)$ of supports for the K-coefficients β^2, the J^1-vector V^1 $(V^1 = [V_1^1, V_2^1, \ldots, V_{J^1}^1]', J^1 \geq 2)$ and the J^2-vector $(V^2 = [V_1^2, V_2^2, \ldots, V_{J^2}^2]', J^2 \geq 2)$ of supports for the noise components ε_i^1 and ε_i^2, respectively, along with the associated probability (weights) vectors p_k^1, p_k^2, w_i^1, and w_i^2, the reparameterized coefficients and noise components are defined as $\beta_k^1 = \sum_m Z_{km}^1 p_{km}^1$, $\beta_k^2 = \sum_m Z_{km}^2 p_{km}^2$, $(k = 1, \ldots, K)$, $\varepsilon_i^1 = \sum_j V_j^1 w_{ij}^1$, $(i \in C_1)$, $\varepsilon_i^2 = \sum_j V_j^2 w_{ij}^2$, $(i \in C_2)$. In our GME formulation, we have also specified a finite, discrete support $G = [G_1, G_2, \ldots, G_D]'$ of dimension $D \geq 2$ for each of the values of Y^* and the corresponding probability distribution r so that $y_i^* = \sum_d G_d r_{id}$, $(i \in C_2)$.

Following the existing literature (Golan et al., 1996b, 1997a) and the GME formulation presented in Bernardini and Filippucci (2000), the GME objective function is given by

$$H(p^1, p^2, w^1, w^2, r) = -\sum_k \sum_m p^1_{km} \ln p^1_{km} - \sum_k \sum_m p^2_{km} \ln p^2_{km} - \sum_i \sum_j w^1_{ij} \ln w^1_{ij} - \sum_i \sum_j w^2_{ij} \ln w^2_{ij} - \sum_i \sum_d r_{id} \ln r_{id} \tag{4}$$

subject to:

(i) Data consistency conditions:

$$\begin{aligned} y_i &= \sum_k \sum_m \left(X^1_{ik} p^1_{km} Z^1_{km}\right) + \sum_j w^1_{ij} V^1_j \qquad \text{if } s_i = y_i; \quad i \in C_1 \\ \sum_d G_d r_{id} &\geq \sum_k \sum_m (X^2_{ik} p^2_{km} Z^2_{km}) + \sum_j w^2_{ij} V^2_j \qquad \text{if } s_i = y^*_i; \quad i \in C_2, \end{aligned} \tag{5}$$

(ii) Adding-up constraints:

$$\begin{aligned} &\sum_m p^1_{km} = 1, \quad k = 1, \ldots, K, \qquad \sum_j w^1_{ij} = 1, \quad i \in C_1, \\ &\sum_m p^2_{km} = 1, \quad k = 1, \ldots, K, \qquad \sum_j w^2_{ij} = 1, \qquad \sum_d r_{id} = 1, \quad i \in C_2. \end{aligned} \tag{6}$$

In order to account for the direct influence of learning from repeated small samples, we can transform the GME formulation into a GCE one, which is based on the cross-entropy principle of Kullback (1959). Assuming that (i) both the reference distributions for errors and latent variables are uniform and that (ii) the estimates of parameter probabilities $\hat{p}^1_{km}$ and $\hat{p}^2_{km}$ computed over period $t - 1$, are the priors for p^1_{km} and p^2_{km} over period t, respectively, the resulting GCE formulation is given by

$$\begin{aligned} &I(p^1, p^2, w^1, w^2, r; \hat{p}^1, \hat{p}^2, u^1, u^2, u^r) \\ &\quad = \sum_k \sum_m p^1_{km} \ln(p^1_{km}/\hat{p}^1_{km}) + \sum_k \sum_m p^2_{km} \ln(p^2_{km}/\hat{p}^2_{km}) \\ &\qquad + \sum_i \sum_j w^1_{ij} \ln(w^1_{ij}/u^1_{ij}) + \sum_i \sum_j w^2_{ij} \ln(w^2_{ij}/u^2_{ij}) \\ &\qquad + \sum_i \sum_d r_{id} \ln(r_{id}/u^r_{id}) \end{aligned} \tag{7}$$

subject to constraints (5)–(6).

2.2. The Composite GCE Formulation

Most surveys provide very little information about a specific small area of interest, since surveys are generally designed to produce statistics for larger populations. National surveys yield very precise estimates of a variety of national statistics, but biased estimates of the corresponding local area statistics. A small area (domain) usually refers to a subgroup of a population from which samples are drawn. The subgroup may be a region or a group obtained by cross-classification of demographic factors.

Following Golan (2001), our idea is to define a composite generalized cross-entropy (CCE) estimator combining information across the (local) subpopulation based on a small sample and the (national) major population with a larger sample size. The objective is to introduce both uniform and spike prior alternatives, or some convex combination of the two, and to produce an adaptive statistical method that is data-based and free of subjective choices except for the bounds on the support spaces. The resulting estimation rule is formulated by combining the GCE estimator with uniform priors, and the GCE estimator with spike priors. This estimator simultaneously chooses the mixture of the two alternative priors on a coordinate-by-coordinate basis, and then uses this information, together with the data, to calculate the shrinkage and to provide estimates of the unknown parameters.

Assuming that, with the exception of the point mass prior q_k^1, and q_k^2, all other priors are uniform, the resulting CCE formulation is given by

$$\begin{aligned}
&\min I(p^1, p^2, p^{\gamma^1}, p^{\gamma^2}, w^1, w^2, r; q^1, q^2, q^{\gamma^1}, q^{\gamma^2}, u^1, u^2, u^r) \\
&\quad = \sum_k \left(1-\gamma_k^1\right) \sum_m p_{km}^1 \ln p_{km}^1 / [\left(1-\gamma_k^1\right)\left(-\ln M\right) + \gamma_k^1 \xi^1] \\
&\qquad + \sum_k \gamma_k^1 \sum_m p_{km}^1 \ln\left(p_{km}^1 / q_{km}^1\right) \\
&\qquad + \sum_k \left(1-\gamma_k^2\right) \sum_m p_{km}^2 \ln p_{km}^2 / [\left(1-\gamma_k^2\right)(-\ln M) + \gamma_k^2 \xi^2] \\
&\qquad + \sum_k \gamma_k^2 \sum_m p_{km}^2 \ln\left(p_{km}^2 / q_{km}^2\right) + \sum_k \sum_l p_{kl}^{\gamma^1} \ln\left(p_{kl}^{\gamma^1} / q_{kl}^{\gamma^1}\right) \\
&\qquad + \sum_k \sum_l p_{kl}^{\gamma^2} \ln\left(p_{kl}^{\gamma^2} / q_{kl}^{\gamma^2}\right) + \sum_i \sum_j w_{ij}^1 \ln\left(w_{ij}^1 / u_{ij}^1\right) \\
&\qquad + \sum_i \sum_j w_{ij}^2 \ln\left(w_{ij}^2 / u_{ij}^2\right) + \sum_i \sum_d r_{id} \ln\left(r_{id} / u_{id}^r\right), \qquad (8)
\end{aligned}$$

subject to:

(i) Data consistency conditions:

$$\begin{aligned} y_i &= \sum_k \sum_m \left(X^1_{ik} p^1_{km} Z^1_{km}\right) + \sum_j w^1_{ij} V^1_j \qquad \text{if } s_i = y_i; \quad i \in C_1 \\ \sum_d G_d r_{id} &\geq \sum_k \sum_m \left(X^2_{ik} p^2_{km} Z^2_{km}\right) + \sum_j w^2_{ij} V^2_j \qquad \text{if } s_i = y^*_i; \quad i \in C_2, \end{aligned} \tag{9}$$

(ii) Adding-up constraints:

$$\begin{aligned} &\sum_m p^1_{km} = 1, \quad k = 1, \ldots, K, \qquad \sum_j w^1_{ij} = 1, \qquad \sum_l p^{\gamma 1}_{kl} = 1, \quad i \in C_1, \\ &\sum_m p^2_{km} = 1, \quad k = 1, \ldots, K, \qquad \sum_j w^2_{ij} = 1, \qquad \sum_l p^{\gamma 2}_{kl} = 1, \\ &\sum_d r_{id} = 1, \quad i \in C_2, \end{aligned} \tag{10}$$

where the prior mixtures, γ^1_k and γ^2_k, are $\gamma^1_k = \sum_{l=1}^L Z^{\gamma 1}_l p^{\gamma 1}_{kl}, \gamma^2_k = \sum_{l=1}^L Z^{\gamma 2}_l p^{\gamma 2}_{kl}$ with $Z^{\gamma 1}_0 = 0$, $Z^{\gamma 2}_0 = 0$ and $Z^{\gamma 1}_L = 1$, $Z^{\gamma 2}_L = 1$ always and where $\xi^1 \equiv -\sum_m q^1_{km} \ln q^1_{km}$, $\xi^2 \equiv -\sum_m q^2_{km} \ln q^2_{km}$. The priors u^1, u^2, $q^{\gamma 1}$, $q^{\gamma 2}$, and u^r are specified to be uniform, while the priors q^1 and q^2 for the parameter probabilities, used to produce local area subpopulation estimates, are based on the whole sample (major population group). The support spaces for the coefficients β^1 and β^2 are both of dimension $M \geq 2$.

The solution to the optimization problem (8–10) yields

$$\begin{aligned} \tilde{p}^1_{km} &= \frac{q_{km}^{\tilde{\gamma}^1_k / A^1_k} \exp\left(\left(1/A^1_k\right) \sum_i \tilde{\lambda}_i X^1_{ik} Z^1_{km}\right)}{\sum_m q_{km}^{\tilde{\gamma}^1_k / A^1_k} \exp\left(\left(1/A^1_k\right) \sum_i \tilde{\lambda}_i X^1_{ik} Z^1_{km}\right)} \\ \tilde{p}^2_{km} &= \frac{q_{km}^{\tilde{\gamma}^2_k / A^2_k} \exp\left(\left(1/A^2_k\right) \sum_i \tilde{\lambda}_i X^2_{ik} Z^2_{km}\right)}{\sum_m q_{km}^{\tilde{\gamma}^2_k / A^2_k} \exp\left(\left(1/A^2_k\right) \sum_i \tilde{\lambda}_i X^2_{ik} Z^2_{km}\right)} \\ \tilde{w}^1_{ij} &= \frac{u^1_{ij} \exp \tilde{\lambda}_i V^1_j}{\sum_j u^1_{ij} \exp \tilde{\lambda}_i V^1_j} \\ \tilde{w}^2_{ij} &= \frac{u^2_{ij} \exp \tilde{\lambda}_i V^2_j}{\sum_j u^2_{ij} \exp \tilde{\lambda}_i V^2_j} \end{aligned} \tag{11}$$

$$\tilde{r}_{id} = \frac{u_{id}^r \exp \tilde{\lambda}_i G_d}{\sum_d u_{id}^r \exp \tilde{\lambda}_i G_d}$$

and

$$\tilde{\gamma}_k^1 \equiv \sum_{l=1}^{L} \tilde{p}_{kl}^{\gamma^1} Z_l^{\gamma^1}, \qquad \tilde{\gamma}_k^2 \equiv \sum_{l=1}^{L} \tilde{p}_{kl}^{\gamma^2} Z_l^{\gamma^2}, \tag{12}$$

where λ are the Lagrange multipliers for the data equations (9), and

$$A_k^1 \equiv 1 - \tilde{\gamma}_k^1 / \big((\tilde{\gamma}_k^1 - 1) \ln M + \tilde{\gamma}_k^1 \xi^1 \big) + \tilde{\gamma}_k^1,$$
$$A_k^2 \equiv 1 - \tilde{\gamma}_k^2 / \big((\tilde{\gamma}_k^2 - 1) \ln M + \tilde{\gamma}_k^2 \xi^2 \big) + \tilde{\gamma}_k^2,$$
$$\xi^1 \equiv - \sum_m q_{km}^1 \ln q_{km}^1, \qquad \xi^2 \equiv - \sum_m q_{km}^2 \ln q_{km}^2.$$

The recovered probability estimates are utilized to derive the point estimates:

$$\hat{\beta}_k^1 = \sum_{m=1}^{M^1} Z_{km}^1 \tilde{p}_{km}^1, \qquad \hat{\beta}_k^2 = \sum_{m=1}^{M^2} Z_{km}^2 \tilde{p}_{km}^2,$$

$$\hat{\varepsilon}_i^1 = \sum_{j=1}^{J^1} V_j^1 \tilde{w}_{ij}^1, \qquad \hat{\varepsilon}_i^2 = \sum_{j=1}^{J^2} V_j^2 \tilde{w}_{ij}^2, \qquad \hat{y}_i^* = \sum_{d=1}^{D} G_d \tilde{r}_{id}.$$

The composite GCE estimator is consistent and asymptotically normal; such asymptotic properties are proven on the basis of the results presented by Golan (2001) and Golan, Judge, and Miller (1996a), by following the empirical likelihood literature (Owen, 1991; Qin and Lawless, 1994), or by following Mittelhammer and Cardell (1996).

3. ESTIMATING EXPENDITURE SHARES FOR SMALL DOMAINS

The objective of the application is to produce estimates of expenditure shares for a population subgroup from a sample of households in one particular Italian region (Umbria), by specifying a system of censored expenditure functions. We are only concerned with the subsample of low-income households living in Umbria (households whose total expenditure is below the median of the distribution of all sampled households). We start by considering a complete set of demand equations (Deaton and Muellbauer, 1980), assuming that: (i) demographic household characteristics are used to capture the effect of household preferences on demand; (ii) the household-specific price level is identical for all households. Under these assumptions we may estimate: $s_{ig} = \beta_g^0 + \sum_k \beta_{kg} x_{ik} + \alpha_g \ln e_i + \varepsilon_{ig}$, where s_{ig} indicates the budget share of the class of goods g ($g = 1, \ldots, B$) per household i ($i = 1, \ldots, n$), e_i indicates the logarithm of total expenditure of household i, and x_{ik} ($k = 1, \ldots, K$) represents a set of K household-specific demographic and economic variables (Bernardini Papalia, 2004), β_g^0, β_{kg}, and α_g are unknown parameters, and ε_{ig} are error terms.

Moreover, in order to account for the presence of zero expenditure, according to formulation (1), we then proceed by enlarging the parameter space to consider the latent data corresponding to the censored data (zero expenditure).

The data used for the application have been taken from the Italian Budget Survey for the year 2000. This survey is carried out every year and collects detailed information regarding household expenditure, together with a large number of household characteristics such as size, composition, socioeconomic features, etc. of a sample of 30,000 private households in Italy (collective or institutionalized households are excluded). A stratified random sample is selected in two stages. The first stage refers to Italian municipalities: 150 large municipalities are automatically included in the sample, while 400 smaller municipalities are randomly selected after stratification by region, altitude, and the main economic activity within the area; the annual population registers are used to define the demographic dimension of the municipalities. In the second stage, households are randomly selected from population registers with equal probabilities. The sample is representative at the regional level, and nonresponding households are replaced by households chosen according to dimension and place of residence. The subsample of low-income households within the Umbria region, as previously defined, for the year 2000, consists of a sample of 286 households (observations).

We estimate a system of censored equations for twelve categories of expenditure: Food and Beverages, Alcohol and Tobacco, Clothing and Footwear, Housing, Furnishings, Medical Care, Transportation, Communication, Recreation, Education, Restaurants and Hotels, Other Goods and Services. The explanatory variables have been divided into four groups: (i) demographic variables (number of children, number of family members, households with no children, head of household aged 0–24, head of household aged 45–64, head of household aged >64, households with more than one child, households consisting of a single adult, households with one child, gender of head of household); (ii) the per capita total household expenditure (expressed in logarithmic form); (iii) occupational dummies for two standard classifications (unemployed, retired, and out of the labor force); (iv) indicators of the presence of stocks (home-owning households; households that pay rent; households that make use of a house for which no rent is paid).

TABLE 1 Composite GCE estimates of the system of censored equations (low-income households, Umbria region, 2000)

	Food, beverages	Alcohol, tobacco	Clothing, footwear	Housing	Furnishings	Medical care	Transport	Communic.	Recreation	Education	Restaurants, hotels	Other goods, services
intercept	0.2000*	0.1529*	0.0330	0.2000*	0.0175	0.0315	−0.0733*	0.0883*	0.0364	0.0739*	0.0398	0.2000*
lexp	−0.0024*	−0.0079	0.0030	0.0097*	0.0031	−0.0012*	0.0143*	−0.0041*	0.0000	−0.0049*	0.0012	−0.0112
nkids	−0.0419*	0.0106	0.0048	−0.0204*	0.0191	0.0097	−0.0076	0.0000	0.0000	0.0000	0.0136	0.0063*
ncomp	−0.0034	−0.0008	−0.0040	−0.0066	−0.0048	−0.0051	0.0076*	0.0019	0.0096	0.0007	−0.0002	0.0071*
no-kids	0.0407	0.0000	−0.0008	−0.0170*	−0.0208*	−0.0048	0.0122	−0.0009	0.0043	−0.0009	−0.0016	−0.0070
no-work	0.0071	−0.0019	−0.0031	0.0343	0.0035	0.0014	−0.0265*	0.0032	−0.0140	0.0051	−0.0130	0.0048
age0	−0.0960*	0.0203	0.0769*	−0.0732*	−0.0076	−0.0600*	0.0114	0.0688*	−0.0080	0.0114	0.0055	0.0505*
age2	0.0561*	0.0014	−0.0285*	0.0051*	0.0029	0.0227*	−0.0160	0.0040	−0.0139	−0.0045*	−0.0092*	−0.0201
age3	0.0878*	−0.0088	−0.0346*	0.0285	0.0060	0.0324	−0.0601*	−0.0043	−0.0124	−0.0074*	−0.0130*	−0.0140*
kids1	0.1288*	−0.0065*	−0.0176*	0.0217	−0.0336*	−0.0166*	−0.0125*	−0.0026	0.0033	−0.0037	−0.0343*	−0.0264
kids2	−0.0371*	−0.0108*	−0.0108*	0.0570*	−0.0077	0.0217	−0.0161*	−0.0034	0.0187	−0.0023	0.0030	−0.0121*
kids3	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
sexf	−0.0049	0.0022	0.0255	−0.0396*	0.0103	−0.0016	−0.0249*	0.0120	0.0005	0.0014	0.0000	0.0190*
work1	−0.0219	−0.0027	0.0308	−0.0787	0.0082	0.0126	0.0000	−0.0028	0.0003	0.0072	0.0264	0.0206
work2	−0.0076	0.0000	0.0000	−0.0206	0.0046	0.0061*	0.0279	0.0020	−0.0198	0.0038	0.0037	0.0000
house0	−0.0185	−0.0129	0.0078	0.0621*	−0.0150	0.0091*	0.0036	−0.0055	−0.0137	−0.0019	−0.0286*	0.0136
house2	0.0036	−0.0163	0.0083	0.0655*	−0.0151	−0.0013	0.0202	−0.0061*	−0.0207*	−0.0017	−0.0304*	−0.0061*

*Significantly different from zero at the 5% level.

The local area expenditure shares for the specific small population group in question are estimated by including the adding-up consumer theory condition and the restriction for the shares to add one, with reference to: (i) the GME specification (with uniform priors for parameters and for error probabilities); (ii) the GCE specification with uniform priors for error probabilities and priors based on the whole sample of households for the parameter probabilities, and (iii) the CCE specification with uniform priors for error probabilities and a prior mixture for the parameter probabilities. Five support points that are symmetrically arranged around zero are used for all parameters. For the error supports, $J = 3$ support points for each error are chosen. Since all the dependent variables are shares that lie between 0 and 1, we chose the support vectors for the errors within the interval $(-1, 1)$.

The analysis focuses firstly on a comparison of the parameter estimates and the standard errors across the maximum entropy-based formulations. We then concentrate on the resulting expenditure elasticities produced using these alternative specifications. Table 1 gives: (i) the parameter estimates of the system of censored demand equations; and (ii) the corresponding asymptotic standard errors calculated using the method described in Golan (2001). The results for all the commodity groups seem to be relatively robust in terms of the signs and magnitude of the estimates. The GME parameter estimates do not vary a great deal as the parameter supports is changed, and while they generally have similar coefficients in each equation, the values of their standard errors produce a number of non-significant coefficients which are larger than those for the cross-entropy-based procedures.

There is a greater degree of concordance between the two cross-entropy results and those obtained using GME. Table 2 shows the income elasticity produced by using the three alternative formulations for each category of expenditure. The expenditure elasticity, at the sample mean, $\bar{s}_g$, is obtained as $\eta_g = 1 + \left(\frac{\alpha_g}{\bar{s}_g}\right)$. The choice of support vectors within the intervals $(-100, 100)$ and $(-20, 20)$ for the parameters produces negligible effects on the coefficients and on the estimated elasticities. With regard to estimated expenditure elasticities and total expenditure shares for the categories in question, all the described procedures gave similar results (Table 3). For the subsample of low-income households in Umbria, the largest difference between any of the expenditure elasticities was that for Education. Food, Housing, and Transportation appear to be expenditure elastic showing a consistency with other empirical studies. This is probably due to: (i) the specific subpopulation of households; (ii) the increases in car and/or house ownership together with the propensity of people towards more “luxurious” forms of transport. On the other hand, Medical Care and Communication appear to be expenditure inelastic. It should be noted that these results are evident in all sets of estimates.

TABLE 2 Estimated expenditure elasticities

	Food, beverages	Alcohol, tobacco	Clothing, footwear	Housing	Furnishings	Medical care	Transport	Communic.	Recreation	Education	Restaurants, hotels	Other goods, services
GME	1.038	0.786	0.978	1.053	0.893	0.895	1.022	0.876	0.906	-2.308	0.788	0.943
GCE	1.038	0.786	0.978	1.053	0.893	0.895	1.022	0.876	0.906	-2.306	0.788	0.943
CCE	0.989	0.587	1.055	1.025	1.105	0.967	1.128	0.866	1.001	-1.882	1.074	0.784

TABLE 3 Estimated relative expenditure shares in percentage

	Food, beverages	Alcohol, tobacco	Clothing, footwear	Housing	Furnishings	Medical care	Transport	Communic.	Recreation	Education	Restaurants, hotels	Other goods, services
GME	39.85	2.99	8.78	0.00	4.77	6.47	16.26	5.20	5.20	0.13	2.33	8.01
GCE	39.84	2.99	8.78	0.00	4.77	6.47	16.25	5.20	5.20	0.17	2.33	8.01
CCE	39.81	2.99	8.77	0.00	4.77	6.46	16.24	5.20	5.19	0.24	2.32	8.00

As has already been noted, the GME/GCE approaches produce efficient estimates of a demand system with a large number of censored equations that are robust even if errors are not normal and the exogenous variables are highly correlated. It is also possible to develop demand systems which are consistent with empirical evidence while displaying consistency with consumer-theory assumptions.

4. CONCLUDING REMARKS

The focus of the present article has been the problems associated with modeling incomplete information and learning from repeated small samples, and with estimating small area statistics by combining information across (local) subpopulations, based on small samples, and across a major (national) population with a larger sample size. A CCE model-based estimator for small subgroups of the reference population, which combines information across a small area of interest with additional information available from a larger area, has been proposed. This method combines the GCE with uniform priors (GME) and the GCE with spike priors estimators and represents a new basis to recover the unknown parameters and the unknown variables when: (i) the surveys used to compute the local area estimate are sensitive to sampling and response errors; (ii) the subpopulation sample size is significantly reduced. The working of the proposed estimation procedure has been illustrated by applying it to the computation of local area expenditure shares in a subpopulation of low-income households. Our results also show how prior information, based on the whole sample of households can be included as a part of the CCE problem formulation. The results of the elasticity estimates are interesting, with signs of expenditure elasticity consistent with economic theory and magnitudes within the expected range. The CCE estimation procedure seems to be a promising alternative model-based estimation technique since its implementation involves a minimum outlay on computing, it does not depend on any hypotheses regarding the form of the error distribution in the model, and it produces good results for small samples, especially when local-area level statistics are highly correlated. Finally, estimation for subpopulations based on a small sample can be greatly improved by incorporating additional information from a population with a larger sample size, together with other additional prior information.

REFERENCES

Amemiya, T. (1984). Tobit models: A survey. *Journal of Econometrics* 24:3–61.

Bernardini Papalia, R. (2004). Estimating item weights of consumer price indexes for small reference populations. In: *Proceedings of the European Conference on Quality and Methodology in Official Statistics* (Q2004). Mainz.

Bernardini Papalia, R., Filippucci, C. (2000). Inference from non-random samples: a maximum entropy approach. In: *Proceedings of the Second International Conference on Establishment Surveys.* Buffalo, New York, pp. 1686–1691.

Deaton, A. S., Muellbauer, J. (1980). *Economics and Consumer Behavior.* Cambridge: Cambridge University Press.

Golan, A. (2001). A simultaneous estimation variable selection rule. *Journal of Econometrics* 10: 165–193.

Golan, A., Judge, G., Miller, D. (1996a). *Maximum Entropy Econometrics: Robust Estimation with Limited Data.* New York: Wiley.

Golan, A., Judge, G., Perloff, J. (1996b). A maximum entropy approach to recovering information from multinomial response data. *Journal of the American Statistical Association* 91:841–853.

Golan, A., Judge, G., Miller, D. (1997a). The maximum entropy approach to estimation and inference: An overview. In: Fomby, T. B., Hill, R. C., eds. *Advances in Econometrics. Applying Maximum Entropy to Econometric Problems.* Vol. 12. London, UK: Jai Press Ltd., pp. 3–24.

Golan, A., Judge, G., Perloff, J. (1997b). Estimation and inference with censored and ordered multinomial response data. *Journal of Econometrics* 79:23–51.

Golan, A., Judge, G., Miller, D. (1998). Information recovery in simultaneous-equations statistical models. In: Ullah, A., Giles, D., eds. *Handbook of Applied Economic Statistics.* New York: Marcel Dekker, pp. 365–381.

Golan, A., Perloff, J., Shen, E. (2001). Estimating a demand system with nonnegativity constraints: Mexican meat demand. *The Review of Economics and Statistics* 83(3):541–550.

Jaynes, E. T. (1957). Information theory and statistical mechanics. *Physics Review* 106:620–630.

Kullback, J. (1959). *Information Theory and Statistics.* New York, NY: Wiley.

Levine, R. D. (1980). An information theoretical approach to inversion problems. *Journal of Physics A* 13:91–108.

Mittelhammer, R., Cardell, S. (1996). On the consistency and asymptotic normality of the data constrained GME estimator of the GML. Working paper, Washington State University, Pullman, WA.

Owen, A. B. (1991). Empirical likelihood for linear models. *The Annals of Statistic* 19:1725–1747.

Qin, J., Lawless, J. (1994). Empirical likelihood and general estimating equations. *The Annals of Statistics* 22:300–325.

Shannon, C. E. (1948). A mathematical theory of communication. *Bell System Technical Journal* 27:379–423.