Eckernkemper, Tobias; Gribisch, Bastian

Article — Published Version

Classical and Bayesian Inference for Income Distributions using Grouped Data

Oxford Bulletin of Economics and Statistics

Provided in Cooperation with:
John Wiley & Sons

Suggested Citation: Eckernkemper, Tobias; Gribisch, Bastian (2020) : Classical and Bayesian Inference for Income Distributions using Grouped Data, Oxford Bulletin of Economics and Statistics, ISSN 1468-0084, Wiley, Hoboken, NJ, Vol. 83, Iss. 1, pp. 32-65, https://doi.org/10.1111/obes.12396

OXFORD BULLETIN OF ECONOMICS AND STATISTICS, 83, 1 (2021) 0305–9049
doi: 10.1111/obes.12396

Classical and Bayesian Inference for Income Distributions using Grouped Data

Tobias Eckernkemper† and Bastian Gribisch†

†Institute of Econometrics and Statistics,University of Cologne,Universitaetsstr. 22a, D-50937 Cologne, Germany (e-mail: eckernkemper@statistik.uni-koeln.de; bastian.gribisch@statistik.uni-koeln.de)

Abstract

We propose a general framework for Maximum Likelihood (ML) and Bayesian estimation of income distributions based on grouped data information. The asymptotic properties of the ML estimators are derived and Bayesian parameter estimates are obtained by Monte Carlo Markov Chain (MCMC) techniques. A comprehensive simulation experiment shows that obtained estimates of the income distribution are very precise and that the proposed estimation framework improves the statistical precision of parameter estimates relative to the classical multinomial likelihood. The estimation approach is finally applied to a set of countries included in the World Bank database *PovcalNet*.

I. Introduction

The empirical analysis of welfare, income inequality and poverty requires precise estimates of the distribution of income. An overview on the vast and growing literature on statistical inference for income distributions is, for example, provided by Kleiber and Kotz (2003), Chotikapanich (2008) and Bandourian, McDonald and Turley (2003). If the data are fully released, the distribution can be estimated by standard parametric or non-parametric methods like Maximum Likelihood (ML) or kernel density estimation. Especially for developing countries it is, however, common that researchers can only access grouped income data which are, for example, provided by the World Bank and the World Institute for Development Economics Research (WIDER). The data typically consist of population shares and group-specific mean incomes for 10 to 20 income groups, where the group boundaries are not provided. This limited data structure causes problems related to partial identification of unrestricted income distributions and derived inequality measures (see e.g. Cowell, 1991 and Stoye, 2010), and turns the objective to estimating the parameters of prespecified parametric income distributions, which are well known to provide a good fit to observed income data (see e.g. McDonald, 1984, and Hajargasht *et al.*, 2012).

The literature provides a variety of parametric income distributions including, but not limited to Pareto's distribution, the lognormal distribution, Champernowne's

JEL Classification numbers: C21, C51, D31.

distribution, Fisk's distribution, the gamma-, generalized gamma-, Weibull-, Singh–Maddala- and Dagum distribution (see e.g. Kleiber and Kotz, 2003). McDonald (1984) proposed the generalized beta distribution of the second kind (GB2 distribution), which nests the lognormal, generalized gamma, Singh–Maddala, Beta-2 and Dagum distributions. Parker (1999) showed that the GB2 distribution can be derived from microeconomic principles and the distribution has therefore become very popular in applied economic research. An alternative, flexible way of income modelling is based on mixture distributions, which are, for example, analysed by Griffiths and Hajargasht (2012).

Contributions on statistical inference for grouped income data are rare. The traditional and most frequently applied method is ML based on sample proportions using a multinomial likelihood function (see e.g. McDonald, 1984, and Bandourian *et al.*, 2003). This approach is inefficient in the majority of practical applications since it neglects the information content of observed group means and does not account for unknown group boundaries. Subsequent work then focused on nonlinear least squares and GMM estimation, where relative population- and income shares are effectively matched to their theoretical counterparts (see e.g. Wu and Perloff, 2005; Wu, 2006; Chotikapanich, Griffiths, Rao, 2007; Chotikapanich *et al.*, 2012). Hajargasht *et al.* (2012) and Griffiths and Hajargasht (2015) propose GMM frameworks which account for unknown group boundaries and observed group means but lack a solid statistical foundation with respect to the underlying data generating process (DGP). Hajargasht and Griffiths (2020) shift the focus from income distributions to parametric Lorenz curves and provide a GMM framework covering two DGPs of empirical relevance, and Chen (2018) generalizes the GMM framework to incorporate varying data information. Bayesian approaches to the estimation of parametric income distributions are provided by Chotikapanich and Griffiths (2000), Kakamu (2016) and Kakamu and Nishino (2019). All Bayesian methods employ Monte Carlo Markov Chain (MCMC) techniques based on the Metropolis-Hastings (MH) algorithm in order to obtain samples from the parameters' joint posterior distribution. While Chotikapanich and Griffiths (2000) employ the standard multinomial likelihood of McDonald (1984), the recent contributions of Kakamu (2016) and Kakamu and Nishino (2019) employ the joint likelihood of a set of order statistics as proposed by Nishino and Kakamu (2011), which is – however – appropriate for quantile-data only. Moreover, both Bayesian settings do not account for unknown group boundaries and ignore the information of observed group mean incomes.

Interestingly, while those recent contributions which account for the informational content of group mean incomes completely focused on GMM, the early work of Hitomi *et al.* (2008) already developed an asymptotically efficient Quasi-Maximum Likelihood (QML) approach incorporating the information of group means under observed and predetermined group boundaries. Their QML approach is asymptotically equivalent to ML and provides the same asymptotic properties as the subsequent GMM approaches of Hajargasht *et al.* (2012) and Griffiths and Hajargasht (2015). In the present paper we develop a QML estimation scheme which is similar in nature and asymptotically equivalent to the approach of Hitomi *et al.* (2008) and extends the Hitomi framework to unknown group boundaries and two different DGPs of practical relevance, which involve likelihoods containing different data information. Moreover, we find that our QML framework comes particularly close to the true likelihood for reasonable sample sizes, and combining the derived likelihoods

with prior information therefore allows for the implementation of a straight-forward MH sampling scheme for Bayesian inference. Bayesian estimation using MCMC techniques is especially attractive for income distributions, since it directly provides valid inference for nonlinear functions of the distribution parameters, such as the Gini coefficient or the Headcount ratio. Up to our knowledge, the proposed setting is the first to incorporate the information of observed group means into Bayesian estimation of parametric income distributions under grouped data.

We therefore contribute to the literature by offering a comprehensive discussion of classical and Bayesian estimation of parametric income distributions for grouped income data with potentially unknown boundaries while accounting for two methods of grouping observations. The first method (DGP1) builds on proportions of observations in each income group, which have been fixed prior to sampling. As a result the group income means and group boundaries are random. In the second method of grouping (DGP2) the group boundaries are predetermined prior to sampling. Hence both the number of observations and the income means in each group are random. Income data from the World Bank or WIDER typically correspond to DGP1 with unknown group boundaries. Dependent on the type of DGP the likelihood comprises varying data information including group population proportions, group means and group boundaries. The multinomial ML method of McDonald (1984) fits DGP2 with known boundaries and observed population proportions. The informational content of the group means is ignored. The QML approach of Hitomi *et al.* (2008) fits DGP2 with known group boundaries and observed group means and population proportions. Both likelihoods are misspecified in case of DGP1. Finally, the order-statistic based ML approach of Nishino and Kakamu (2011) fits DGP1 with known boundaries but ignores the informative content of observed mean incomes.

Extending the ML approach of McDonald (1984) to incorporate the informational content of the group means requires the derivation of the joint (conditional) density of the mean incomes. This distribution is unknown for all relevant income distributions, but for reasonable sample sizes well approximated by the Gaussian due to standard central limit arguments. We approximate the joint density of the group means by a product of Normals with moments given by their asymptotic counterparts. Under DGP1 the group boundaries constitute random order statistics and can easily be included in the likelihood (known boundaries, comparable to the ML approach of Nishino and Kakamu, 2011). If the boundaries are unknown, we exploit asymptotic results of Beach and Davidson (1983) and maximize the resulting Gaussian likelihood approximation for the group means conditional on the parameters of the income distribution. Under DGP2 both group means and relative population shares are random and the likelihood results from the product of the joint conditional density of group means and the multinomial likelihood. If group boundaries are unknown, we can simply estimate them along with the remaining model parameters. Bayesian estimation is implemented by combining the derived likelihoods with according prior information and sampling the resulting posterior using an independent MH sampler based on a Gaussian approximation to the posterior distribution. Since the proposed likelihood functions are based on Gaussian approximations, they essentially resemble QML functions. However, as our simulation experiments show, the estimation error is of very reduced impact and the QML functions appear close to the true likelihoods.

We proof the consistency of our QML estimation schemes and derive the asymptotic distribution of the QML estimators. By analogy to the results of Hitomi *et al.* (2008) our QML approaches are asymptotically equivalent to ML based on the true (unknown) joint conditional density of the group means and asymptotically efficient under standard regularity conditions. We also find that our QML method under DGP2 has the same asymptotic covariance as the GMM approaches of Hajargasht *et al.* (2012) and Griffiths and Hajargasht (2015).

Taken all together, we provide a comprehensive QML framework for the estimation of income distributions using grouped data. Combining the QML functions with according prior distributions then allows for Bayesian inference using basic MCMC techniques. Our approach is efficient in the sense that all available data information is included in the likelihood, whose characteristics depend on the specific DGP at hand. We further find that the QML estimation is simple and fast to implement with standard asymptotic properties corresponding to asymptotically efficient ML under the usual regularity conditions. With regard to Bayesian inference, the proposed independent MH algorithm for sampling the posterior distribution shows a high degree of efficiency as reflected by high acceptance probabilities and accordingly low numerical standard errors.

We provide an extensive simulation experiment in order to assess the finite sample performance of the proposed estimation schemes. Here we consider the popular GB2 distribution which nests most of the income distributions of practical relevance. The results indicate a sound and stable performance of the QML- and Bayesian estimators under DGP1 and DGP2 and known-/unknown group boundaries. This performance appears to be robust against varying DGPs, parameterizations, sample sizes and numbers of income groups. Our results also indicate significant improvements over the conventional multinomial ML approach and we obtain accurate parameter estimates which come close to those obtained for individual income data. We also find that the precision of the parameter estimates does not suffer significantly if the group boundaries are unknown. This result is of considerable practical relevance, since group boundaries are usually not provided in the World Bank or WIDER data sets.

We finally apply both, the QML- and the Bayesian estimation approach to World Bank data for four countries and find evidence for the GB2 distribution relative to its nested competitors such as the Beta2, Singh–Maddala and the Dagum distribution. The obtained estimates of inequality and poverty measures as well as predictions of income shares show a high degree of accuracy.

The remainder of this paper is organized as follows. Section II gives general definitions and discusses the relevant data generating processes. Section III introduces the QML approach and section IV extends the QML scheme to Bayesian inference using standard MCMC techniques. Section V then provides a simulation experiment in order to assess the finite sample performance of the estimators, and section VI presents the empirical application. Section VII concludes. Proofs are given in the Appendix.

II. Definitions and data generating processes

Let $y_1,\ldots,y_n$ be a random sample from a parametric distribution with density function $f_y(y;\theta)$ $(y>0)$, distribution function $F_y(y;\theta)$ and moment distribution function

$$F_\ell(y;\theta)=\frac{1}{E[y^\ell]}\int_0^y t^\ell f_y(t;\theta)\,dt,\quad \ell=1,2,\ldots \tag{1}$$

where y denotes income and θ comprises the model parameters. In the following we assume that the first and second moments of y exist. For the GB2 distribution we, for example, obtain $\theta=(a,\,b,\,p,\,q)'$ with $a,\,b,\,p,\,q>0$ and

$$\begin{aligned}
f_y(y;\theta)&=\frac{ay^{ap-1}}{b^{ap}B(p,q)(1+(y/b)^a)^{p+q}},\\
F_y(y;\theta)&=B_u(p,q),\\
F_\ell(y;\theta)&=B_u(p+\ell/a,\;q-\ell/a),\\
E[y^\ell]&=b^\ell\frac{B(p+\ell/a,\;q-\ell/a)}{B(p,q)},
\end{aligned}$$

where $u=(y/b)^a/[1+(y/b)^a]$, $B(\cdot)$ denotes the beta function and $B_u(\cdot)$ denotes the beta distribution function evaluated at u. An overview of the GB2 and its nested distributions is provided in Table 1, which has been taken from Hajargasht *et al.* (2012).

The sample is grouped into K income groups where the boundaries are denoted by $\{z_{i-1},z_i\}_{i=1}^K$ with $z_0=0$ and $z_K=\infty$. Let n_i denote the number of observations in income group i such that the sample size obtains as $n=\sum_{i=1}^K n_i$. Typical income data (e.g. World Bank or WIDER) contains information on relative population shares $c_i=n_i/n$ and group-specific mean incomes $\bar{y}_i=(1/n_i)\sum_{j=1}^n y_j g_i(y_j)$, for $i=1,\ldots,K$, where

$$g_i(y)=\begin{cases}1 & \text{if } z_{i-1}<y\leqslant z_i\\ 0 & \text{otherwise.}\end{cases} \tag{2}$$

In some cases we do not have data on mean incomes directly but observe the overall mean income $\bar{y}$ together with income shares $\{s_i\}_{i=1}^K$ instead, where $s_i=(n\bar{y})^{-1}\sum_{j=1}^n y_j g_i(y_j)$. Group-specific mean incomes are then obtained via $\bar{y}_i=s_i\bar{y}/c_i$. Group boundaries $\{z_i\}_{i=1}^{K-1}$ are usually not provided.

The method of grouping individuals into income classes is not unique and likelihood functions for ML or Bayesian estimation of θ must be tailored to the respective DGP in order to enable solid statistical inference. The upcoming subsections therefore define two distinct DGPs which are of particular relevance in practice.

DGP1: Fixed n_i and random z_i, $\bar{y}_i$

Under DGP1 the relative proportions of observations in each income group, $c_i=n_i/n$, are prespecified. This is the case for the majority of the data sets in the World Bank and the WIDER data base. Respective data consist of constant relative population shares corresponding, for example, to deciles or quintiles together with the respective mean incomes.

Denote the cumulative number of group observations by $n_i^c=\sum_{\ell=1}^i n_\ell$. Under DGP1 the group boundary z_i $(i=1,\ldots,K-1)$ corresponds to the n_i^c'th order statistic $y_{[n_i^c]}$ from f_y, which represents a random variable. Note that, strictly speaking, the group boundary can take any value in $[y_{[n_i^c]},y_{[n_i^c+1]})$. Corresponding data generating processes are, however,

TABLE 1

Overview on the GB2 and the nested Beta-2, Singh–Maddala and Dagum distributions

	Density function	Moments	Distribution function	Moment distribution function	Gini coefficient
GB 2	$f_y(y;\theta)=\dfrac{ay^{ap-1}}{b^{ap}B(p,q)(1+(y/b)^a)^{p+q}}$	$E[y^\ell]=\dfrac{b^\ell B(p+\ell/a,\ q-\ell/a)}{B(p,q)}$	$F_y(y;\theta)=B_u(p,q)$, with $u=(y/b)^a/[1+(y/b)^a]$	$F_\ell(y;\theta)=B_u(p+\ell/a,\ q-\ell/a)$, with $u=(y/b)^a/[1+(y/b)^a]$	Integral evaluated numerically
Beta-2 (a = 1)	$f_y(y;\theta)=\dfrac{y^{p-1}}{b^pB(p,q)(1+(y/b))^{p+q}}$	$\mu=bp/(q-1)$ $E[y^2]=bp(p+1)/(q-1)(q-2)$	$F_y(y,\theta)=B_u(p,q)$, with $u=(y/b)/[1+(y/b)]$	$F_\ell(y;\theta)=B_u(p+\ell,q-\ell)$, with $u=(y/b)/[1+(y/b)]$	$G=\dfrac{2B(2p,2q-1)}{pB^2(p,q)}$
Singh-Maddala ($P=1$)	$f_y(y,\theta)=\dfrac{aqy^{a-1}}{b^a(1+(y/b)^a)^{1+q}}$	$E[y^\ell]=\dfrac{b^\ell\Gamma(1+\ell/a,q-\ell/a)}{\Gamma(q)}$	$F_y(y,\theta)=1-\left[1+(\dfrac{y}{b})^a\right]^{-q}$	$F_\ell(y;\theta)=B_u(1+\ell/a,q-\ell/a)$ with $u=(y/b)^a/[1+(y/b)^a]$	$G=1-\dfrac{\Gamma(q)\Gamma(2q-1/a)}{\Gamma(q-1/a)\Gamma(2q)}$
Dagum ($q=1$)	$f_y(y;\theta)=\dfrac{apy^{ap-1}}{b^{ap}(1+(y/b)^a)^{p+1}}$	$E[y^\ell]=\dfrac{b^\ell\Gamma(p+\ell/a,1-\ell/a)}{\Gamma(q)}$	$F_y(y;\theta)=\left[1+(\dfrac{y}{b})^{-a}\right]^{-p}$	$F_\ell(y;\theta)=B_u(p+\ell/a,1-\ell/a)$, with $u=(y/b)^a/[1+(y/b)^a]$	$G=\dfrac{\Gamma(p)\Gamma(2p+1/a)}{\Gamma(p+1/a)\Gamma(2p)}-1$

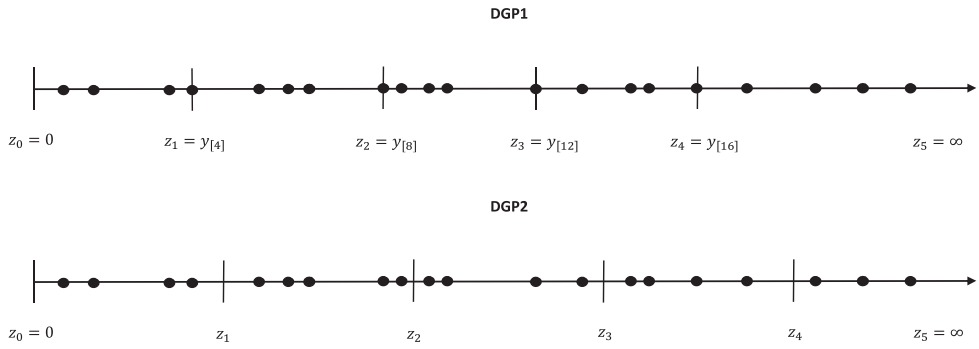

Figure 1. Schematic illustration of the two data generating processes DGP1 and DGP2 for $n = 20$ and $K = 5$ income groups. Black bullets denote individual income y_i on the real line. The example for DGP1 assumes $c_i = 0.2\ \forall\ i$

observationally equivalent and $z_i \mathrel{\widehat{=}} y_{[n_i^c]}$ therefore constitutes an identifying restriction. The upper panel of Figure 1 depicts a schematic illustration of DGP1 for $n = 20$.

We summarize that DGP1 generates random group boundaries and group means, while relative proportions c_i and $n_i = n \cdot c_i$ are preset and therefore deterministic. The non-stochastic nature of the group proportions renders the classical multinomial ML method of McDonald (1984) misspecified and ML estimation for DGP1 can only be based on the information contained in the group boundaries (if available) and the group means.

DGP2: Fixed z_i and random n_i, $\bar{y}_i$

DGP2 assumes prespecified fixed group boundaries resulting in a random number of observations in each income group. Respective data sets contain group means and relative population shares which vary over income groups. Such data are rather infrequently met in practice – a few examples are found in the *PovcalNet* data base of the World Bank for selected countries and years. A schematic illustration of DGP2 is provided in the lower panel of Figure 1. The multinomial ML method of McDonald (1984) and the QML approach of Hitomi *et al.* (2008) are designed under DGP2 with known group boundaries.

We summarize that DGP2 generates random population shares and group means, while group boundaries are preset and therefore deterministic. ML estimation for DGP2 can therefore be based on the information contained in both, the group means and the population shares. Note that the multinomial ML method of McDonald (1984) remains inefficient since the informational content of the group-specific mean incomes is not exploited.

III. Quasi maximum likelihood inference

DGP1: Fixed n_i and random z_i, $\bar{y}_i$

Under DGP1 and known group boundaries (KB) the likelihood for the complete set of observable data obtains as

$$L_{\text{DGP1, KB}}(\theta; \underline{\bar{y}}, \underline{z}) = f(\underline{\bar{y}} \mid \underline{z}; \theta) \cdot f(\underline{z}; \theta), \tag{3}$$

where $\underline{\bar{y}} = \{\bar{y}_i\}_{i=1}^{K}$ and $\underline{z} = \{z_i\}_{i=1}^{K-1}$. Dependence on $\{n_i\}_{i=1}^{K}$ is suppressed for notational convenience.

The i'th group boundary z_i corresponds to the n_i^c'th order statistic of iid random variables from f_y. Exploiting the Markov property of order statistics (see e.g. David and Nagaraja, 2003, Theorem 2.5) we obtain the joint density of group boundaries in (3) as

$$f(\underline{z};\theta) = f(z_1;\theta)\cdot f(z_2 \mid z_1;\theta)\cdot \ldots \cdot f(z_{K-1} \mid z_{K-2};\theta), \tag{4}$$

where standard calculus for order statistics gives

$$f(z_1;\theta) = \frac{n!}{(n_1^c-1)!(n-n_1^c)!}\, F_y(z_1;\theta)^{n_1^c-1}\,[1-F_y(z_1;\theta)]^{n-n_1^c} f_y(z_1;\theta), \tag{5}$$

$$\begin{aligned}(z_i \mid z_{i-1};\theta) = {} & \frac{(n-n_{i-1}^c)!}{(n_i^c-n_{i-1}^c-1)!(n-n_i^c)!} \\ & \cdot \frac{[1-F_y(z_i;\theta)]^{n-n_i^c}}{[1-F_y(z_{i-1};\theta)]^{n-n_{i-1}^c}}\,[F_y(z_i;\theta)-F_y(z_{i-1};\theta)]^{n_i^c-n_{i-1}^c-1} f_y(z_i;\theta).\end{aligned} \tag{6}$$

Note that $f(\underline{z};\theta)$ corresponds to the likelihood analysed by Nishino and Kakamu (2011).

By exploiting conditional independence the joint density of group means in equation (3) can be decomposed into

$$\begin{aligned}f(\underline{\bar{y}} \mid \underline{z};\theta) = {} & f(\bar{y}_1 \mid z_1;\theta)\cdot f(\bar{y}_2 \mid z_1,z_2;\theta) \\ & \cdot \ldots \cdot f(\bar{y}_{K-1} \mid z_{K-2},z_{K-1};\theta)\cdot f(\bar{y}_K \mid z_{K-1};\theta),\end{aligned} \tag{7}$$

where $z_0 = 0$ and $z_K = \infty$. The distribution of the arithmetic mean is unknown for any income distribution of practical relevance (see e.g. Nadarajah, 2005, for the complex derivation of the distribution of the sum of only two GB2 distributed random variables). We therefore replace the individual constituents of $f(\underline{\bar{y}} \mid \underline{z};\theta)$ in equation (7) by approximations, which are consistent in the sense that the resulting approximation error diminishes to zero as $n \to \infty$. Employing the standard Lindeberg Levy Central Limit Theorem (CLT) for iid random variables we obtain

$$\begin{aligned}f(\underline{\bar{y}} \mid \underline{z};\theta) \approx {} & f_N(\bar{y}_1 \mid z_1;\theta)\cdot f_N(\bar{y}_2 \mid z_1,z_2;\theta) \\ & \cdot \ldots \cdot f_N(\bar{y}_{K-1} \mid z_{K-2},z_{K-1};\theta)\cdot f_N(\bar{y}_K \mid z_{K-1};\theta),\end{aligned} \tag{8}$$

where $f_N(\bar{y}_i \mid \cdot)$ denotes the density function of a Gaussian distribution with mean $\mu_i(\theta)$ and variance $\sigma_i^2(\theta)$. Note that this approach is analogous to the QML likelihood approximation of Hitomi *et al.* (2008), who do not refer to CLT arguments but (equivalently) discuss the convergence of the characteristic function of the density of group means. Hitomi *et al.* (2008) then provide a formal proof for the asymptotic equivalence of ML estimates based on $f(\underline{\bar{y}} \mid \underline{z};\theta)$ and QML estimates using the approximation in (8).

Since conditional on the group boundaries (z_{i-1}, z_i), the $n_i - 1$ individual stochastic incomes in group i are independent and identically distributed with density function $f(y \mid z_{i-1}, z_i;\theta) = f(y \mid z_{i-1} < y < z_i;\theta)$ (see David and Nagaraja, 2003, Theorem 2.5) we obtain

$$\mu_i(\theta) = E(\bar{y}_i \mid z_{i-1}, z_i;\theta) = \frac{n_i-1}{n_i}\tilde{\mu}_i(\theta) + \frac{z_i}{n_i} \quad \text{with} \tag{9}$$

$$\begin{aligned}\tilde{\mu}_i(\theta) &= E(y \mid z_{i-1} < y < z_i; \theta) \\ &= \frac{[F_1(z_i;\theta) - F_1(z_{i-1};\theta)] \cdot E(y;\theta)}{F_y(z_i;\theta) - F_y(z_{i-1};\theta)}\end{aligned} \tag{10}$$

and

$$\sigma_i^2(\theta) = Var(\bar{y}_i \mid z_{i-1}, z_i; \theta) = \frac{n_i - 1}{n_i^2}\tilde{\sigma}_i^2(\theta) \quad \text{with} \tag{11}$$

$$\begin{aligned}\tilde{\sigma}_i^2(\theta) &= Var(y \mid z_{i-1} < y < z_i; \theta) \\ &= \left[\frac{(F_2(z_i;\theta) - F_2(z_{i-1};\theta)) \cdot E(y^2;\theta)}{F_y(z_i;\theta) - F_y(z_{i-1};\theta)} - \tilde{\mu}_i(\theta)^2\right],\end{aligned} \tag{12}$$

where $i = 1,\ldots,K-1$. Note that conditional on z_i the last summand in $\bar{y}_i$ is deterministic and given by z_i. For the last income group we obtain $\mu_K = \tilde{\mu}_K$ and $\sigma_K^2 = \tilde{\sigma}_K^2/n_K$.

Inserting the previously derived expressions into equation (3) the resulting approximate log-likelihood under DGP1 and known group boundaries obtains as

$$\begin{aligned}\mathcal{L}_{\text{DGP1, KB}}(\theta; \underline{\bar{y}}, \underline{z}) = \Omega &- \frac{1}{2}\left[\ln(\tilde{\sigma}_K^2(\theta)) - \ln n_K + \tilde{\sigma}_K^{-2}(\theta) n_K (\bar{y}_K - \tilde{\mu}_K(\theta))^2\right] \\ &+ \sum_{i=1}^{K-1}\left\{ -\frac{1}{2}\left[\ln(\sigma_i^2(\theta)) + \frac{(\bar{y}_i - \mu_i(\theta))^2}{\sigma_i^2(\theta)}\right]\right. \\ &\left. + (n_i^c - n_{i-1}^c - 1)\ln[F_y(z_i;\theta) - F_y(z_{i-1};\theta)] + \ln f_y(z_i;\theta)\right\} \\ &+ (n - n_{K-1}^c)\ln\left[1 - F_y(z_{K-1};\theta)\right],\end{aligned} \tag{13}$$

where $n_0^c = F_y(z_0;\theta) = 0$ and

$$\Omega = -\frac{K}{2}\ln(2\pi) + \sum_{i=1}^{K-1}\ln[(n - n_{i-1}^c)!] - \ln[(n - n_i^c)!] - \ln[(n_i^c - n_{i-1}^c - 1)!].$$

Estimation of θ is carried out by maximizing the objective function in equation (13) using numerical techniques routinely available in standard software packages. Note that the unique maximizer $\hat{\theta}$ of (13) can be interpreted as a QML-type estimator where consistency and asymptotic normality follow by the standard regularity conditions provided in Appendix A. We provide results on the consistency and asymptotic normality of the QML estimator for DGP1 and known boundaries in Proposition 1.

Proposition 1. Under the regularity conditions stated in Appendix A, the QML estimator obtained as the unique maximizer of the QML objective in equation (13) is consistent and asymptotically normal with covariance matrix $ACOV_{\text{DGP1, KB}}(\hat{\theta}) = -\frac{1}{n}H(\theta_0)^{-1}$, and

$$H(\theta_0) = \sum_{i=1}^{K} -\left[\frac{\partial\tilde{\mu}_i(\theta_0)}{\partial\theta}\frac{\partial\tilde{\mu}_i(\theta_0)}{\partial\theta'}\right]\frac{c_i}{\tilde{\sigma}_i^2(\theta_0)}$$

$$-\frac{1}{c_i}\left[\frac{\partial F_y(q_y(c_i^c;\theta_0);\theta_0)}{\partial\theta}-\frac{\partial F_y(q_y(c_{i-1}^c;\theta_0);\theta_0)}{\partial\theta}\right]$$
$$\cdot\left[\frac{\partial F_y(q_y(c_i^c;\theta_0);\theta_0)}{\partial\theta}-\frac{\partial F_y(q_y(c_{i-1}^c;\theta_0);\theta_0)}{\partial\theta}\right]',$$

where $\theta_0=\text{plim}(\hat{\theta})$ denotes the true value of θ, $q_y(\cdot;\theta_0)=F_y^{-1}(\cdot;\theta_0)$ denotes the quantile function of y, $c_i^c=\sum_{\ell=1}^{i}c_\ell$ and by definition of the first and the last income group $\partial F_y(q_y(c_0^c,\theta_0);\theta_0)/\partial\theta\,\widehat{=}\,0$ and $\partial F_y(q_y(c_K^c,\theta_0);\theta_0)/\partial\theta\,\widehat{=}\,0$.

Proof: see Appendix B.

The quality of the QML approximation in (8) is further analysed in section V. The results imply overall accurate approximations even for relatively low sample sizes with $n=5,000$ and 10 income groups. In fact the approximation error induced by the Gaussian approximation to the group means appears practically negligible.

The majority of the WIDER and World Bank data sets do not report group boundaries. In order to deal with this situation we have to integrate out the latent group boundaries $\underline{z}$ from the joint (unknown) likelihood in equation (3). It turns out that this problem is solved asymptotically by the results of Beach and Davidson (1983) (see also the related work of Griffiths and Hajargasht, 2015). We obtain the QML objective

$$\begin{aligned}\mathcal{L}_{\text{DGP1, UB}}(\theta;\underline{\bar{y}})=&-0.5[K\ln(2\pi)-K\ln(n)+\ln\mid\Psi(\theta)\mid\\&+n(\underline{\bar{y}}-\mu^*(\theta))'\Psi^{-1}(\theta)(\underline{\bar{y}}-\mu^*(\theta))],\end{aligned}\tag{14}$$

where $\mu^*(\theta)=(1/c_i)\int_{q_y(c_{i-1}^c;\theta)}^{q_y(c_i^c;\theta)}yf_y(y)dy$ and the limiting covariance matrix of $\sqrt{n}\underline{\bar{y}}$, $\Psi(\theta)$, defined in Appendix B.

The results on consistency and asymptotic normality of the QML estimator for DGP1 and unknown group boundaries are provided in Proposition 2.

Proposition 2. Under the regularity conditions stated in Appendix A, the QML estimator obtained as the unique maximizer of the QML objective in equation (14) is consistent and asymptotically normal with covariance matrix $ACOV_{\text{DGP1, UB}}(\hat{\theta})=-\frac{1}{n}H(\theta_0)^{-1}$, where

$$H(\theta_0)=-\frac{\partial\mu^{*'}(\theta_0)}{\partial\theta}\Psi^{-1}(\theta_0)\frac{\partial\mu^{*}(\theta_0)}{\partial\theta'}.$$

Proof: see Appendix B.

DGP2: Fixed z_i and random n_i, $\bar{y}_i$

DGP2 generates random numbers of observations n_i and random mean incomes $\bar{y}_i$ for each group. The likelihood comprising all available data information then obtains as

$$L_{\text{DGP2}}(\theta;\underline{\bar{y}},\underline{n})=f(\underline{\bar{y}}\mid\underline{n};\theta)\cdot f(\underline{n};\theta),\tag{15}$$

where $\underline{n}=\{n_i\}_{i=1}^K$. Dependence on $\underline{z}$ is again suppressed for notational convenience.

The distribution of $\underline{n}$ is multinomial with density function

$$f(\underline{n};\theta)=\frac{n!}{n_1!\cdot\ldots\cdot n_K!}\cdot\pi_1(\theta)^{n_1}\cdot\ldots\cdot\pi_K(\theta)^{n_K}, \tag{16}$$

where

$$\pi_i(\theta)=\Pr(z_{i-1}<y\leqslant z_i;\theta)=F_y(z_i;\theta)-F_y(z_{i-1};\theta),\quad i=1,\ldots,K.$$

We then obtain the QML objective under DGP2 via inserting (16) and the Gaussian approximation (8) in equation (15):

$$\mathcal{L}_{\mathrm{DGP2}}(\theta;\underline{\bar{y}},\underline{n})=\Omega+\sum_{i=1}^{K}\left\{-\frac{1}{2}\left[\ln(\tilde{\sigma}_i(\theta)^2)-\ln n_i+\frac{n_i(\bar{y}_i-\tilde{\mu}_i(\theta))^2}{\tilde{\sigma}_i(\theta)^2}\right]+n_i\ln\pi_i(\theta)\right\}, \tag{17}$$

where

$$\Omega=-\frac{K}{2}\ln(2\pi)+\ln(n!)-\sum_{i=1}^{K}\ln(n_i!). \tag{18}$$

QML estimation of θ is carried out by maximizing the objective in equation (17) over θ (known boundaries) or jointly over $\underline{z}$ and θ (unknown boundaries). Note that the maximization of (16) for known group boundaries corresponds to the multinomial ML method of McDonald (1984).

The results on consistency and asymptotic normality of the QML estimator $\hat{\theta}$ for DGP2 (known and unknown boundaries) are provided in Proposition 3.

Proposition 3. Under the regularity conditions stated in Appendix A, the QML estimator obtained as the unique maximizer of the QML objective in equation (17) is consistent and asymptotically normal with covariance matrix $ACOV_{\mathrm{DGP2}}(\hat{\theta})=-\frac{1}{n}H(\theta_0)^{-1}$, where

$$H(\theta_0)=\sum_{i=1}^{K}\pi_i(\theta_0)\frac{\partial^2\ln\pi_i(\theta_0)}{\partial\theta\partial\theta'}-\left[\frac{\partial\tilde{\mu}_i(\theta_0)}{\partial\theta}\frac{\partial\tilde{\mu}_i(\theta_0)}{\partial\theta'}\right]\frac{\pi_i(\theta_0)}{\tilde{\sigma}_i^2(\theta_0)}.$$

This result holds for both, known and unknown group boundaries (with the parameter vector θ augmented by the set of group boundaries).

Proof: see Appendix C.

Note that the asymptotic covariance matrix of $\hat{\theta}$ under DGP2 corresponds to the one obtained under the QML approach of Hitomi *et al.* (2008) and the GMM estimators of Hajargasht *et al.* (2012) and Griffiths and Hajargasht (2015).

IV. Bayesian inference

Taking the previously derived quasi likelihood functions as close approximations to the true likelihoods (compare our simulation results in section V on the approximation error), and introducing $\Pr(\theta)$ as a joint prior distribution for the parameter vector θ, we obtain the posterior

$$\pi(\theta \mid X) \propto L(\theta; X) \cdot \Pr(\theta), \tag{19}$$

where X denotes the data, that is, the group means $\underline{\bar{y}}$ and/or group boundaries $\underline{z}$ or group counts $\underline{n}$, depending on the DGP. The likelihood $L(\theta; X)$ is chosen from equations (13) and (14 or 17) according to the type of DGP and depending on whether the group boundaries are known or unknown.

We impose independent lognormal priors for the parameters a, b, p and q of the GB2 distribution and sample the posterior by a standard independent MH algorithm similar to the ones applied in Chotikapanich and Griffiths (2000) and Kakamu (2016).[1] We employ a multivariate normal maximum a posteriori proposal for $\ln\theta$, $q(\ln\theta)$, centred at the mode of the posterior density kernel (19) w.r.t. $\ln\theta$. The covariance matrix is obtained by the negative inverse of the log-posterior's hessian at the mode. The mode and the hessian are obtained via numerical maximization of the posterior kernel in (19) over $\ln\theta$ using a standard Quasi-Newton BFGS optimizer.

The MH sampler then proceeds as follows:

(i) Set $\ln\theta^{(1)} = \ln\theta_{\text{mode}}$, where θ_{mode} denotes the posterior mode.
(ii) For $i = 2,\ldots,S$ implement the following MH steps:
  (a) Draw a candidate value $\ln\theta^*$ from the proposal $q(\ln\theta)$.
  (b) Compute the MH acceptance probability

$$\alpha = \min\left\{1,\ \frac{\pi(\theta^* \mid X)\ q(\ln\theta^{(i-1)})}{\pi(\theta^{(i-1)} \mid X)\ q(\ln\theta^*)}\right\}, \tag{20}$$

  where $\theta^* = \exp(\ln\theta^*)$. Note that α can be computed without knowing the integrating constant of the posterior in (19). Also, if θ^* falls outside the feasible parameter region, that is, if $q^* < 2/a^*$ (non-existence of the GB2 variance, see Table 1), set $\alpha = 0$.
  (c) Draw a uniform random variable u from the interval $(0, 1)$.
  (d) If $u \leqslant \alpha$ set $\ln\theta^{(i)} = \ln\theta^*$, else set $\ln\theta^{(i)} = \ln\theta^{(i-1)}$.

The MH algorithm constructs a Markov chain, which converges to the posterior of $\ln\theta$. After a *burnin* of the first M iterations of the sampler, after which convergence is achieved, we obtain with $\{\exp(\ln\theta^{(i)})\}_{i=M+1}^{S}$ a correlated sample from the posterior of θ, from which we can compute an MC approximation to the Bayesian estimate of θ (posterior mean) by the corresponding sample mean. The uncertainty of the estimates is addressed by the posterior standard deviation, approximated by the sample standard deviation of the MC draws. Posterior moments of functions of the model parameters like, for example, the Gini coefficient or the Headcount ratio are easily obtained by evaluating the respective function at each MC draw $\theta^{(i)}$ and then computing the desired sample statistics.

For our simulation experiments in section V and the empirical applications in section VI we choose overall uninformative lognormal priors centred at the posterior mode with

[1] In case of DGP2 and unknown group boundaries we augment θ by the group boundaries and impose additional lognormal prior distributions for the boundaries.

standard deviation equal to 100. For the number of MH simulations and the burnin we choose $S = 120,000$ and $M = 20,000$.

V. Simulation experiment

We now perform a simulation experiment in order to investigate the quality of the likelihood approximation through central limit arguments and the performance of the QML and Bayesian estimation schemes under DGP1 and DGP2 and both known and unknown boundaries in finite samples. We consider a GB2 distribution and four parameter settings of empirical relevance: (i) $a = 1.5$, $b = 106$, $P = 5.1$, $q = 2.9$; (ii) $a = 1.6$, $b = 386$, $P = 1.2$, $q = 1.8$; (iii) $a = 3.1$, $b = 56$, $P = 2.3$, $q = 0.9$; (iv) $a = 4.4$, $b = 69$, $P = 0.7$, $q = 0.6$. The four settings are based on our empirical estimates for income data of India Rural, Peru, Ethiopia and Iraq as discussed in section VI.

We first analyse the quality of the Gaussian approximation to the joint density of the group means. We focus on DGP1 and unknown group boundaries, which is the empirically most realistic scenario, and simulate $N = 100,000$ independent data sets, each of sample size $n = 5,000$. Empirically relevant sample sizes typically amount to $n = 20,000$ or higher (see our empirical application in section VI). We, however, focus on rather low sample sizes in order to tempt the asymptotics of our normality approximation. We then construct $K = 10$ income groups, where the group boundaries are set to the deciles of the simulated data, and compute the K group mean incomes for each of the N data sets. Figure 2 depicts kernel density approximations to the true density of the group means (based on the N simulations) together with the Gaussian approximations with moments given by the corresponding elements of μ^* and Ψ/n in equations (14) and (24). We obtain accurate approximations with some very slight skewness in the last income groups for parameter settings (ii), (iii) and (iv) induced by the strong skewness of the respective income distributions. Note that small approximation errors for the last group cannot be expected to have a significant effect on inference as compared to the hypothetical but unavailable 'true' likelihood, since the likelihood contribution of the mean income in each income group is down weighted by the respective variance. The income variance within the last group, however, is typically exceedingly high if the underlying distribution is heavily skewed to the right. Hence in case of heavily skewed income distributions, the average income in the last income group has practically zero weight.

We now turn to the analysis of the finite sample performance of the QML and Bayesian parameter estimates under both DGP1 and DGP2 and all four parameter settings given above. We consider sample sizes $n \in \{5000, 10000, 20000, 100000\}$ and three different group settings: We construct $K = 10$, $K = 20$, and $K = 50$ income groups, where the group boundaries are set to the respective sample quantiles/theoretical quantiles corresponding to equal group sizes/probabilities under DGP1 and DGP2 respectively. We consider both, known and unknown group boundaries, and compare the performance of the QML- and Bayesian estimators of sections I and IV to the empirically infeasible ML for individual observations (denoted by *ML Raw*), the GMM approach of Hajargasht and Griffiths (2020) labelled *GMM-L*) and the GMM approach of Hajargasht *et al.* for DGP2 (2012, labelled *GMM*). For DGP2 with known group boundaries we also consider the multinomial ML method of McDonald (1984) denoted by *ML Multi*).

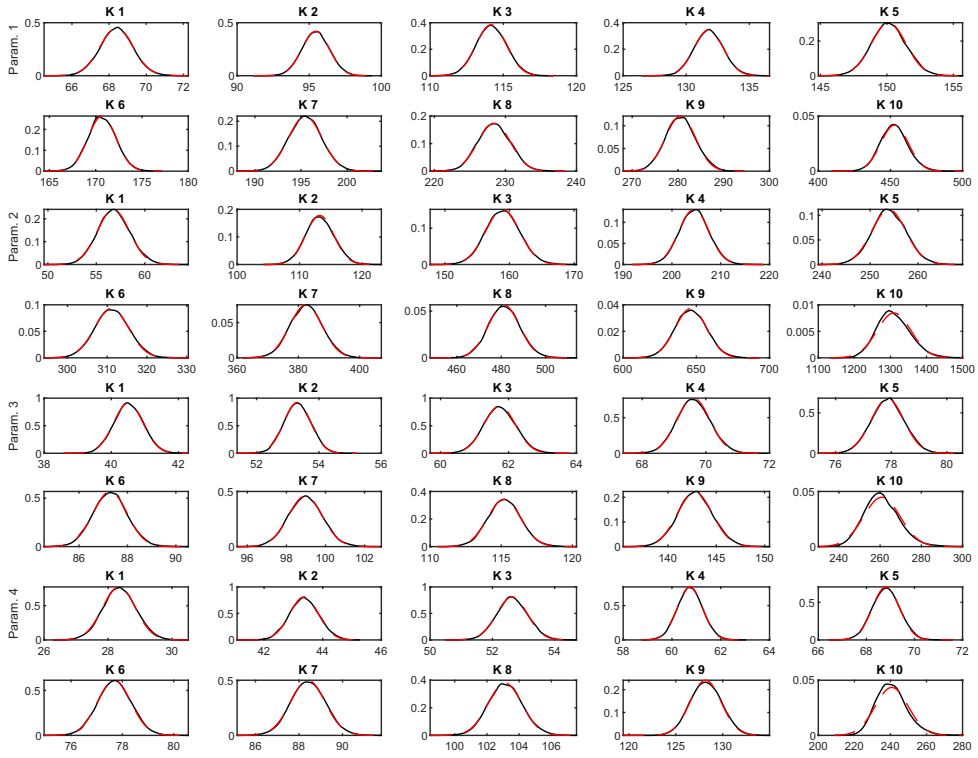

Figure 2. Kernel density estimates of the distribution of $K = 10$ group means together with their Gaussian approximations under DGP1 and unknown group boundaries. Black line: kernel density estimate; Dashed red line: Gaussian approximation with moments given by the corresponding elements of μ^* and Ψ/n in equations (14) and (24). The kernel density estimates are based on 100,000 simulations from a GB2 distribution under DGP1 and unknown boundaries with parameter settings (i), (ii), (iii) and (iv) as illustrated in section V. The sample size is $n = 5,000$

Tables 2 to 5 provide the MSE results for parameter settings (i)–(iv) obtained under 500 independently simulated data sets.[2] The MSEs are decreasing with increasing sample size, reflecting the consistency of the estimates. For parameter settings (ii) to (iv) we observe an overall similar performance of the QML and Bayesian estimators with a tendency to slightly higher MSEs under the Bayesian setting. The performance of QML also appears similar to GMM, which is expected under the asymptotic equivalence of both estimators. The lowest MSEs are typically obtained for ML Raw. Interestingly, under parameter setting (i) and $n \leqslant 10,000$ the best MSE results are obtained under the Bayesian setting, even outperforming ML Raw. For low sample sizes we also find that QML outperforms the GMM approach, in particular for $K = 50$.

The MSEs under known and unknown group boundaries are typically very similar in value. The availability of group boundaries therefore appears to be of limited importance for estimation precision. This is an important finding since the group boundaries are typically

[2]The MH acceptance probabilities α, which determine the efficiency of the MCMC sampling scheme under the Bayesian setting, are ranging between 0.20 and 0.91 with a median value of 0.74. These values indicate a good performance of the MH algorithm and a fast mixing of the generated Markov chains. See also the results on numerical standard errors reported for our empirical estimates in section VI.

TABLE 2

Monte Carlo simulation results (MSEs) for the finite-sample performance of the proposed estimators.

		DGP1							DGP2								
	Parameters	ML Raw	QML KB	GMM-L KB	Bayes KB	QML UB	GMM-L UB	Bayes UB	QML KB	GMM-L KB	GMM KB	ML Multi KB	Bayes KB	QML UB	GMM-L UB	GMM UB	Bayes UB
n = 5,000																	
K = 10	a	0.072	0.113	0.106	0.049	0.122	0.117	0.041	0.115	0.110	0.110	0.514	0.042	0.121	0.117	0.116	0.053
	b	360.850	755.730	631.162	278.064	857.469	763.123	138.312	771.673	675.374	674.520	2846.216	143.058	834.576	758.675	743.490	129.152
	p	23.563	119.537	90.278	55.794	303.943	229.794	2.129	137.254	122.282	116.108	33545.599	2.359	205.380	237.441	232.703	3.485
	q	1.812	3.916	3.134	1.106	4.952	4.045	0.440	3.998	3.272	3.204	79.193	0.495	4.588	4.074	3.949	1.537
K = 20	a	0.062	0.064	0.126	0.040	0.079	0.078	0.036	0.067	0.075	0.063	0.283	0.036	0.079	0.078	0.077	0.050
	b	340.935	354.768	604.974	169.750	518.512	480.859	121.005	393.673	418.755	332.168	2012.441	117.898	493.799	462.944	441.642	184.224
	p	17.951	17.710	72.466	34.925	30.504	26.185	2.137	18.373	20.908	13.463	17135.386	2.124	24.020	24.069	21.327	5.288
	q	1.201	1.319	4.872	0.664	1.977	1.753	0.455	1.489	1.503	1.149	31.235	0.446	1.753	1.619	1.559	1.036
K = 50	a	0.070	0.022	0.071	0.032	0.035	0.055	0.034	0.023	0.056	0.022	0.147	0.032	0.024	0.068	0.041	0.038
	b	366.255	18.899	225.297	134.013	116.845	248.532	131.095	23.039	219.114	37.434	832.430	117.821	29.712	333.565	131.531	133.700
	p	21.605	0.822	10.774	4.000	4.617	8.675	2.494	1.006	7.773	1.221	495.295	2.065	0.901	14.304	4.189	2.939
	q	1.780	0.338	1.435	0.443	0.613	0.937	0.513	0.359	0.897	0.329	7.950	0.450	0.361	1.409	0.577	0.688
n = 10,000																	
K = 10	a	0.037	0.056	0.056	0.022	0.060	0.058	0.021	0.057	0.057	0.055	0.345	0.021	0.059	0.058	0.058	0.018
	b	158.397	309.723	295.625	93.473	346.193	319.824	74.043	318.551	301.128	296.219	2307.602	74.323	339.102	318.430	313.126	56.985
	p	5.086	21.893	23.876	9.852	27.081	18.957	1.167	22.236	20.893	21.036	12579.465	1.247	23.878	21.222	18.830	0.953
	q	0.655	1.382	1.316	0.367	1.529	1.333	0.222	1.393	1.295	1.277	44.738	0.244	1.465	1.341	1.298	0.259
K = 20	a	0.032	0.037	0.069	0.021	0.042	0.041	0.021	0.037	0.040	0.039	0.129	0.021	0.042	0.042	0.041	0.026
	b	123.489	163.099	249.777	68.909	206.185	194.976	72.393	162.786	173.413	162.876	789.042	67.378	195.695	193.164	186.395	95.380
	p	2.621	4.373	17.620	1.307	6.736	5.763	1.331	4.347	4.339	4.194	2831.909	1.312	5.290	5.598	5.277	2.079
	q	0.411	0.550	2.080	0.272	0.713	0.650	0.269	0.570	0.554	0.562	8.554	0.277	0.658	0.644	0.610	0.387
K = 50	a	0.031	0.014	0.033	0.022	0.020	0.030	0.022	0.015	0.030	0.016	0.071	0.022	0.016	0.036	0.026	0.021
	b	122.636	23.771	127.555	71.162	62.344	115.610	78.314	30.665	114.497	33.601	353.434	74.599	33.105	150.508	84.025	78.435
	p	2.525	0.650	4.299	1.282	1.348	2.550	1.381	0.776	2.536	0.762	23.472	1.334	0.757	3.439	1.680	1.350
	q	0.387	0.179	0.585	0.256	0.241	0.346	0.261	0.189	0.353	0.181	1.421	0.259	0.191	0.463	0.282	0.269

continued

TABLE 2

(Continued)

		DGP1							DGP2								
	Parameters	ML Raw	QML KB	GMM-L KB	Bayes KB	QML UB	GMM-L UB	Bayes UB	QML KB	GMM-L KB	GMM KB	ML Multi KB	Bayes KB	QML UB	GMM-L UB	GMM UB	Bayes UB
n = 20,000																	
K = 10	a	0.012	0.022	0.022	0.013	0.024	0.024	0.013	0.022	0.022	0.022	0.198	0.013	0.024	0.024	0.024	0.013
	b	46.026	104.302	100.404	49.556	112.949	109.624	47.068	103.387	99.851	99.568	1406.674	48.433	112.795	109.296	108.813	47.928
	p	0.821	2.016	1.881	0.864	2.159	2.046	0.780	2.001	1.864	1.867	3768.625	0.837	2.141	2.016	2.000	0.814
	q	0.143	0.271	0.255	0.153	0.285	0.273	0.139	0.268	0.252	0.252	16.420	0.149	0.283	0.270	0.268	0.147
K = 20	a	0.013	0.020	0.035	0.015	0.021	0.021	0.015	0.020	0.019	0.019	0.069	0.015	0.022	0.021	0.021	0.014
	b	50.184	81.966	119.626	54.796	89.890	89.135	56.863	80.808	79.090	76.319	373.978	56.206	91.900	89.214	87.682	52.907
	p	1.029	1.707	4.087	1.064	1.834	1.817	1.081	1.705	1.564	1.541	136.964	1.103	1.903	1.803	1.769	0.902
	q	0.173	0.263	0.666	0.193	0.283	0.278	0.193	0.265	0.245	0.243	2.060	0.200	0.291	0.277	0.272	0.160
K = 50	a	0.015	0.010	0.020	0.017	0.012	0.016	0.017	0.009	0.016	0.010	0.036	0.017	0.010	0.019	0.016	0.012
	b	56.412	25.572	63.288	66.339	43.256	64.463	69.444	23.249	61.462	21.626	150.201	66.666	26.205	75.329	55.955	47.762
	p	0.936	0.533	1.277	1.138	0.729	1.046	1.199	0.446	0.968	0.442	3.185	1.142	0.454	1.204	0.927	0.756
	q	0.159	0.115	0.236	0.189	0.131	0.167	0.195	0.106	0.161	0.109	0.420	0.189	0.105	0.188	0.159	0.136
n = 100,000																	
K = 10	a	0.003	0.005	0.005	0.005	0.005	0.005	0.005	0.005	0.005	0.005	0.039	0.005	0.005	0.005	0.005	0.005
	b	9.679	18.892	18.839	20.228	20.274	20.136	21.540	18.917	18.809	18.787	186.515	20.295	20.312	20.195	20.168	18.604
	p	0.151	0.276	0.273	0.304	0.300	0.297	0.330	0.275	0.271	0.272	20.278	0.305	0.301	0.298	0.297	0.275
	q	0.029	0.044	0.044	0.046	0.047	0.047	0.050	0.044	0.043	0.044	0.934	0.046	0.047	0.047	0.047	0.044
K = 20	a	0.003	0.004	0.007	0.004	0.004	0.004	0.004	0.004	0.004	0.004	0.013	0.004	0.004	0.004	0.004	0.004
	b	8.488	13.480	16.444	13.995	14.409	14.211	14.562	13.344	13.586	13.370	45.384	13.903	14.614	14.603	14.544	14.247
	p	0.138	0.206	0.337	0.220	0.222	0.218	0.229	0.202	0.208	0.202	0.834	0.217	0.223	0.223	0.222	0.221
	q	0.028	0.036	0.071	0.038	0.039	0.039	0.039	0.036	0.037	0.036	0.133	0.038	0.039	0.039	0.039	0.038
K = 50	a	0.003	0.002	0.004	0.004	0.003	0.003	0.004	0.002	0.003	0.002	0.007	0.004	0.002	0.004	0.003	0.003
	b	8.645	5.050	9.849	12.791	8.908	10.118	12.852	5.177	9.959	5.459	23.689	12.795	6.831	12.597	10.893	11.969
	p	0.143	0.096	0.193	0.204	0.146	0.170	0.205	0.100	0.166	0.105	0.398	0.204	0.121	0.199	0.177	0.192
	q	0.029	0.023	0.042	0.036	0.029	0.033	0.037	0.024	0.032	0.025	0.069	0.036	0.027	0.036	0.034	0.035

Notes: The results are obtained for 500 Monte Carlo replications and the parameter setting: $a = 1.5$, $b = 106$, $P = 5.1$, $q = 2.9$

TABLE 3

Monte-Carlo simulation results (MSEs) for the finite-sample performance of the proposed estimators.

		DGP1							DGP2								
	Parameters	ML Raw	QML KB	GMM-L KB	Bayes KB	QML UB	GMM-L UB	Bayes UB	QML KB	GMM-L KB	GMM KB	ML Multi KB	Bayes KB	QML UB	GMM-L UB	GMM UB	Bayes UB
n = 5,000																	
K = 10	a	0.030	0.037	0.036	0.039	0.035	0.035	0.039	0.037	0.038	0.042	0.163	0.038	0.035	0.036	0.039	0.056
	b	908.855	1123.503	1002.765	1183.831	1071.157	981.574	1225.352	1117.225	990.557	1037.351	16875320753.750	1197.105	1056.347	937.327	967.892	5103.999
	p	0.044	0.059	0.054	0.071	0.059	0.054	0.074	0.059	0.053	0.055	1.448	0.070	0.058	0.052	0.053	0.132
	q	0.143	0.197	0.174	0.230	0.191	0.170	0.240	0.195	0.170	0.176	264.625	0.231	0.187	0.163	0.167	0.720
K = 20	a	0.026	0.032	0.057	0.032	0.033	0.028	0.031	0.032	0.030	0.037	0.072	0.033	0.033	0.029	0.045	0.031
	b	850.667	1023.509	4052.320	1142.562	1077.169	963.841	1175.545	1040.339	919.945	976.765	6080.959	1152.807	1072.653	922.590	1040.904	1171.449
	p	0.034	0.041	0.083	0.049	0.044	0.040	0.051	0.042	0.037	0.039	0.147	0.049	0.043	0.038	0.041	0.051
	q	0.113	0.142	0.504	0.170	0.153	0.135	0.178	0.145	0.123	0.133	0.843	0.172	0.150	0.126	0.143	0.179
K = 50	a	0.028	0.023	0.034	0.031	0.026	0.029	0.031	0.023	0.029	0.024	0.044	0.031	0.030	0.029	0.030	0.031
	b	897.013	377.505	1201.578	1180.879	795.730	968.697	1208.735	421.993	887.101	436.196	1341.843	1190.560	931.897	913.219	685.259	1139.560
	p	0.037	0.031	0.045	0.046	0.037	0.039	0.049	0.033	0.038	0.031	0.061	0.047	0.039	0.038	0.034	0.045
	q	0.128	0.077	0.175	0.175	0.117	0.137	0.181	0.084	0.130	0.082	0.219	0.177	0.132	0.131	0.102	0.163
n = 10,000																	
K = 10	a	0.015	0.019	0.019	0.020	0.019	0.019	0.020	0.019	0.020	0.019	0.088	0.020	0.019	0.020	0.020	0.021
	b	364.180	479.616	454.253	513.188	492.018	467.375	502.605	478.732	437.189	451.949	400982.373	517.820	488.246	456.642	459.857	569.931
	p	0.021	0.028	0.027	0.031	0.028	0.027	0.031	0.028	0.027	0.027	0.353	0.031	0.028	0.027	0.027	0.034
	q	0.067	0.089	0.084	0.099	0.092	0.087	0.099	0.089	0.083	0.084	6.147	0.101	0.091	0.085	0.085	0.113
K = 20	a	0.012	0.013	0.022	0.014	0.013	0.013	0.014	0.013	0.013	0.013	0.033	0.013	0.013	0.013	0.014	0.014
	b	263.201	363.214	919.640	388.666	370.097	348.508	391.959	361.001	337.577	348.119	1039.653	387.329	365.632	347.019	348.367	399.974
	p	0.016	0.018	0.027	0.020	0.019	0.018	0.021	0.018	0.017	0.017	0.046	0.020	0.018	0.018	0.018	0.021
	q	0.045	0.059	0.123	0.065	0.060	0.056	0.066	0.058	0.053	0.055	0.182	0.065	0.059	0.055	0.055	0.069
K = 50	a	0.012	0.011	0.015	0.014	0.012	0.013	0.014	0.010	0.013	0.012	0.021	0.014	0.014	0.013	0.013	0.014
	b	268.199	188.317	481.684	444.894	305.167	372.576	460.745	218.313	363.252	209.607	586.205	450.021	392.119	374.098	329.919	455.614
	p	0.015	0.014	0.019	0.019	0.016	0.017	0.020	0.014	0.017	0.015	0.028	0.020	0.017	0.017	0.015	0.020
	q	0.045	0.039	0.073	0.069	0.049	0.059	0.074	0.039	0.057	0.042	0.100	0.071	0.059	0.058	0.050	0.072

continued

TABLE 3

(Continued)

		DGP1							DGP2								
	Parameters	ML Raw	QML KB	GMM-L KB	Bayes KB	QML UB	GMM-L UB	Bayes UB	QML KB	GMM-L KB	GMM KB	ML Multi KB	Bayes KB	QML UB	GMM-L UB	GMM UB	Bayes UB
n = 20,000																	
K = 10	a	0.006	0.008	0.008	0.009	0.009	0.009	0.009	0.009	0.008	0.008	0.046	0.009	0.009	0.009	0.009	0.009
	b	127.846	228.994	222.304	237.868	233.490	226.831	243.494	230.133	220.213	222.194	1723.798	238.974	232.730	223.402	224.056	248.885
	p	0.008	0.012	0.012	0.013	0.013	0.012	0.014	0.012	0.011	0.012	0.081	0.013	0.013	0.012	0.012	0.014
	q	0.023	0.037	0.036	0.040	0.040	0.038	0.043	0.038	0.035	0.036	0.354	0.041	0.040	0.038	0.038	0.045
K = 20	a	0.005	0.007	0.012	0.007	0.007	0.007	0.007	0.007	0.007	0.007	0.018	0.007	0.007	0.007	0.007	0.007
	b	100.738	194.207	448.125	200.932	197.042	192.848	203.409	193.221	185.314	189.377	485.841	200.209	196.518	190.123	191.813	203.896
	p	0.007	0.008	0.014	0.009	0.009	0.008	0.009	0.008	0.008	0.008	0.023	0.009	0.009	0.008	0.008	0.009
	q	0.018	0.028	0.062	0.029	0.029	0.028	0.030	0.028	0.026	0.027	0.087	0.029	0.028	0.027	0.028	0.030
K = 50	a	0.006	0.006	0.008	0.007	0.007	0.007	0.007	0.006	0.007	0.006	0.011	0.007	0.007	0.007	0.007	0.007
	b	104.400	113.939	212.969	196.594	156.223	169.969	198.368	116.304	169.876	104.515	268.871	196.700	188.003	176.137	171.081	197.967
	p	0.008	0.008	0.010	0.009	0.009	0.009	0.009	0.008	0.008	0.008	0.013	0.009	0.009	0.008	0.009	0.009
	q	0.021	0.022	0.035	0.031	0.027	0.027	0.031	0.023	0.027	0.021	0.046	0.031	0.029	0.027	0.027	0.031
n = 100,000																	
K = 10	a	0.001	0.002	0.002	0.002	0.002	0.002	0.002	0.002	0.002	0.002	0.008	0.002	0.002	0.002	0.002	0.002
	b	20.327	36.126	35.975	36.418	35.973	35.903	36.269	36.120	35.960	36.021	182.092	36.390	35.953	35.749	35.822	36.315
	p	0.002	0.002	0.002	0.002	0.002	0.002	0.002	0.002	0.002	0.002	0.011	0.002	0.002	0.002	0.002	0.002
	q	0.004	0.006	0.006	0.006	0.006	0.006	0.006	0.006	0.006	0.006	0.040	0.006	0.006	0.006	0.006	0.006
K = 20	a	0.001	0.002	0.002	0.002	0.002	0.002	0.002	0.002	0.001	0.002	0.004	0.002	0.002	0.002	0.002	0.002
	b	21.545	40.339	80.295	41.050	40.795	40.359	41.105	40.729	40.502	40.389	86.571	41.174	40.694	40.340	40.377	41.016
	p	0.001	0.002	0.003	0.002	0.002	0.002	0.002	0.002	0.002	0.002	0.004	0.002	0.002	0.002	0.002	0.002
	q	0.004	0.006	0.011	0.006	0.006	0.006	0.006	0.006	0.006	0.006	0.016	0.006	0.006	0.006	0.006	0.006
K = 50	a	0.001	0.001	0.002	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.002	0.001	0.001	0.001	0.001	0.001
	b	20.134	31.500	45.057	38.980	34.449	34.712	39.066	30.674	36.231	29.872	52.635	39.017	39.799	36.670	35.814	38.997
	p	0.001	0.002	0.002	0.002	0.002	0.002	0.002	0.002	0.002	0.002	0.002	0.002	0.002	0.002	0.002	0.002
	q	0.004	0.005	0.007	0.006	0.005	0.005	0.006	0.005	0.005	0.005	0.008	0.006	0.006	0.005	0.005	0.006

Notes: The results are obtained for 500 Monte Carlo replications and the parameter setting: $a = 1.6$, $b = 386$, $P = 1.2$, $q = 1.8$

TABLE 4

Monte-Carlo simulation results (MSEs) for the finite-sample performance of the proposed estimators.

		DGP1							DGP2								
	Parameters	ML Raw	QML KB	GMM-L KB	Bayes KB	QML UB	GMM-L UB	Bayes UB	QML KB	GMM-L KB	GMM KB	ML Multi KB	Bayes KB	QML UB	GMM-L UB	GMM UB	Bayes UB
n = 5,000																	
K = 10	a	0.106	0.142	0.135	0.148	0.142	0.134	0.146	0.141	0.136	0.142	0.553	0.147	0.142	0.135	0.144	0.130
	b	12.318	17.511	15.769	18.780	18.767	17.454	19.751	17.637	15.973	15.869	158.177	18.765	18.779	17.272	16.750	16.258
	p	0.339	0.511	0.438	0.581	0.563	0.503	0.616	0.524	0.445	0.444	1299.927	0.584	0.565	0.494	0.472	0.483
	q	0.019	0.025	0.023	0.028	0.026	0.024	0.028	0.025	0.022	0.023	0.155	0.028	0.026	0.023	0.023	0.024
K = 20	a	0.099	0.119	0.138	0.123	0.121	0.114	0.126	0.122	0.114	0.121	0.267	0.127	0.120	0.114	0.120	0.104
	b	12.942	16.339	16.260	16.965	16.606	15.463	17.806	16.610	14.826	15.123	47.579	17.566	16.268	15.194	14.628	13.633
	p	0.376	0.496	0.502	0.537	0.502	0.444	0.575	0.500	0.411	0.430	3.247	0.557	0.485	0.432	0.409	0.387
	q	0.019	0.023	0.027	0.025	0.024	0.022	0.027	0.023	0.020	0.021	0.055	0.026	0.023	0.021	0.021	0.018
K = 50	a	0.105	0.063	0.104	0.116	0.100	0.104	0.121	0.071	0.102	0.077	0.153	0.119	0.081	0.106	0.110	0.098
	b	14.626	8.045	13.364	16.940	14.242	14.509	18.461	8.220	13.638	8.847	22.724	17.450	9.856	14.699	13.838	12.956
	p	0.428	0.251	0.379	0.524	0.454	0.465	0.583	0.217	0.391	0.266	0.673	0.542	0.248	0.433	0.359	0.349
	q	0.018	0.013	0.018	0.024	0.019	0.018	0.026	0.013	0.017	0.013	0.025	0.025	0.014	0.018	0.017	0.018
n = 10,000																	
K = 10	a	0.054	0.073	0.072	0.080	0.075	0.074	0.084	0.072	0.071	0.073	0.275	0.080	0.076	0.074	0.076	0.079
	b	5.981	8.813	8.503	10.712	9.582	9.214	11.699	8.915	8.400	8.433	60.243	10.896	9.582	9.171	9.081	10.514
	p	0.137	0.214	0.202	0.306	0.236	0.222	0.336	0.219	0.199	0.201	10.312	0.313	0.236	0.220	0.217	0.299
	q	0.008	0.012	0.011	0.015	0.012	0.012	0.015	0.012	0.011	0.011	0.056	0.015	0.012	0.012	0.012	0.015
K = 20	a	0.047	0.057	0.073	0.058	0.059	0.055	0.061	0.057	0.054	0.057	0.124	0.058	0.059	0.055	0.059	0.056
	b	5.531	6.421	7.330	7.023	6.683	6.542	7.414	6.431	6.183	6.170	18.217	7.027	6.610	6.451	6.333	6.585
	p	0.127	0.156	0.203	0.192	0.164	0.156	0.206	0.156	0.144	0.146	0.565	0.191	0.161	0.153	0.150	0.169
	q	0.008	0.010	0.013	0.011	0.010	0.009	0.011	0.010	0.009	0.009	0.022	0.011	0.010	0.009	0.009	0.010
K = 50	a	0.049	0.041	0.057	0.055	0.053	0.054	0.057	0.041	0.054	0.041	0.080	0.055	0.052	0.055	0.058	0.055
	b	5.639	4.233	6.083	6.494	5.616	6.055	6.721	4.162	5.884	4.067	9.894	6.549	5.562	6.208	6.300	6.258
	p	0.123	0.096	0.137	0.162	0.132	0.136	0.172	0.094	0.129	0.091	0.233	0.165	0.122	0.139	0.138	0.152
	q	0.007	0.006	0.009	0.009	0.008	0.008	0.010	0.006	0.008	0.006	0.012	0.009	0.008	0.008	0.008	0.009

continued

TABLE 4

(Continued)

		DGP1							DGP2								
	Parameters	ML Raw	QML KB	GMM-L KB	Bayes KB	QML UB	GMM-L UB	Bayes UB	QML KB	GMM-L KB	GMM KB	ML Multi KB	Bayes KB	QML UB	GMM-L UB	GMM UB	Bayes UB
n = 20,000																	
K = 10	a	0.029	0.042	0.041	0.042	0.044	0.044	0.044	0.041	0.041	0.041	0.165	0.041	0.044	0.044	0.044	0.042
	b	3.002	4.580	4.486	4.788	4.858	4.766	5.105	4.553	4.428	4.433	30.759	4.766	4.857	4.769	4.727	4.709
	p	0.069	0.111	0.108	0.124	0.119	0.116	0.134	0.111	0.106	0.106	1.805	0.124	0.119	0.116	0.114	0.122
	q	0.004	0.007	0.006	0.007	0.007	0.007	0.007	0.007	0.006	0.006	0.031	0.007	0.007	0.007	0.007	0.007
K = 20	a	0.026	0.031	0.037	0.031	0.032	0.031	0.032	0.031	0.030	0.031	0.068	0.031	0.032	0.031	0.032	0.032
	b	2.781	3.465	3.703	3.564	3.521	3.455	3.634	3.431	3.352	3.363	9.636	3.530	3.502	3.446	3.426	3.614
	p	0.062	0.080	0.092	0.086	0.081	0.079	0.088	0.079	0.076	0.077	0.239	0.085	0.081	0.078	0.078	0.088
	q	0.004	0.005	0.006	0.005	0.005	0.005	0.005	0.005	0.005	0.005	0.011	0.005	0.005	0.005	0.005	0.005
K = 50	a	0.024	0.022	0.028	0.027	0.026	0.026	0.027	0.022	0.026	0.022	0.037	0.026	0.026	0.026	0.027	0.027
	b	2.733	2.315	2.967	3.055	2.903	2.932	3.118	2.235	2.918	2.195	4.792	3.051	2.970	2.969	3.103	3.170
	p	0.058	0.052	0.066	0.069	0.064	0.063	0.071	0.050	0.063	0.050	0.105	0.069	0.064	0.064	0.066	0.073
	q	0.004	0.004	0.004	0.004	0.004	0.004	0.004	0.004	0.004	0.003	0.006	0.004	0.004	0.004	0.004	0.004
n = 100,000																	
K = 10	a	0.005	0.008	0.008	0.008	0.008	0.008	0.008	0.008	0.008	0.008	0.029	0.008	0.008	0.008	0.008	0.008
	b	0.503	0.780	0.777	0.788	0.824	0.822	0.833	0.780	0.776	0.776	4.297	0.788	0.824	0.821	0.819	0.848
	p	0.011	0.018	0.017	0.018	0.018	0.018	0.019	0.018	0.017	0.017	0.097	0.018	0.018	0.018	0.018	0.019
	q	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.004	0.001	0.001	0.001	0.001	0.001
K = 20	a	0.005	0.006	0.007	0.006	0.006	0.006	0.006	0.006	0.006	0.006	0.014	0.006	0.006	0.006	0.006	0.006
	b	0.549	0.644	0.672	0.651	0.684	0.681	0.692	0.646	0.644	0.643	1.865	0.653	0.682	0.681	0.679	0.692
	p	0.011	0.014	0.015	0.014	0.015	0.015	0.015	0.014	0.014	0.014	0.039	0.014	0.015	0.015	0.015	0.015
	q	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.002	0.001	0.001	0.001	0.001	0.001
K = 50	a	0.005	0.005	0.005	0.005	0.005	0.005	0.005	0.005	0.005	0.005	0.008	0.005	0.005	0.005	0.005	0.005
	b	0.550	0.540	0.600	0.588	0.588	0.591	0.601	0.526	0.576	0.534	1.041	0.586	0.649	0.596	0.616	0.601
	p	0.011	0.011	0.013	0.012	0.012	0.012	0.013	0.011	0.012	0.011	0.022	0.012	0.013	0.012	0.013	0.013
	q	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001

Notes: The results are obtained for 500 Monte Carlo replications and the parameter setting: $a = 3.1$, $b = 56$, $P = 2.3$, $q = 0.9$

TABLE 5

Monte-Carlo simulation results (MSEs) for the finite-sample performance of the proposedestimators.

		DGP1							DGP2								
	Parameters	ML Raw	QML KB	GMM-L KB	Bayes KB	QML UB	GMM-L UB	Bayes UB	QML KB	GMM-L KB	GMM KB	ML Multi KB	Bayes KB	QML UB	GMM-L UB	GMM UB	Bayes UB
n = 5,000																	
K = 10	a	0.209	0.249	0.238	0.264	0.255	0.237	0.270	0.258	0.244	0.270	0.644	0.273	0.255	0.238	0.269	0.276
	b	2.150	2.256	2.226	2.267	2.294	2.284	2.310	2.258	2.234	2.240	4.615	2.271	2.287	2.268	2.254	2.309
	p	0.010	0.012	0.011	0.013	0.013	0.012	0.013	0.012	0.011	0.012	0.047	0.013	0.013	0.012	0.012	0.014
	q	0.007	0.009	0.008	0.009	0.009	0.008	0.009	0.009	0.008	0.008	0.028	0.009	0.009	0.008	0.008	0.009
K = 20	a	0.186	0.200	0.231	0.212	0.200	0.194	0.211	0.201	0.199	0.208	0.273	0.213	0.200	0.195	0.210	0.216
	b	1.844	1.893	1.957	1.901	1.910	1.918	1.911	1.891	1.868	1.870	2.403	1.901	1.895	1.889	1.864	1.899
	p	0.008	0.009	0.010	0.009	0.009	0.008	0.009	0.009	0.008	0.008	0.014	0.009	0.009	0.008	0.008	0.009
	q	0.006	0.006	0.008	0.007	0.007	0.006	0.007	0.006	0.006	0.006	0.009	0.007	0.006	0.006	0.006	0.007
K = 50	a	0.203	0.214	0.208	0.220	0.214	0.206	0.223	0.216	0.207	0.218	0.228	0.222	0.211	0.207	0.219	0.226
	b	1.811	1.616	1.832	1.885	1.882	1.837	1.930	1.669	1.801	1.562	1.979	1.880	1.884	1.861	1.825	1.884
	p	0.009	0.009	0.009	0.009	0.009	0.009	0.010	0.009	0.009	0.009	0.010	0.010	0.009	0.009	0.009	0.010
	q	0.006	0.007	0.007	0.007	0.007	0.006	0.007	0.007	0.006	0.006	0.007	0.007	0.007	0.006	0.006	0.007
n = 10,000																	
K = 10	a	0.097	0.111	0.110	0.115	0.111	0.110	0.115	0.111	0.112	0.115	0.279	0.116	0.112	0.111	0.116	0.118
	b	0.912	1.007	0.992	1.009	1.032	1.022	1.035	1.008	0.997	1.003	2.106	1.010	1.032	1.023	1.026	1.037
	p	0.004	0.005	0.005	0.005	0.005	0.005	0.005	0.005	0.005	0.005	0.015	0.005	0.005	0.005	0.005	0.005
	q	0.003	0.003	0.003	0.003	0.003	0.003	0.004	0.003	0.003	0.003	0.009	0.003	0.003	0.003	0.003	0.004
K = 20	a	0.090	0.097	0.110	0.099	0.099	0.095	0.101	0.097	0.095	0.098	0.148	0.099	0.099	0.096	0.100	0.102
	b	0.949	0.993	1.030	0.996	1.002	0.991	1.004	0.996	0.986	0.994	1.269	0.999	1.000	0.998	0.994	1.002
	p	0.004	0.005	0.005	0.005	0.005	0.005	0.005	0.005	0.005	0.005	0.007	0.005	0.005	0.005	0.005	0.005
	q	0.003	0.003	0.004	0.003	0.003	0.003	0.003	0.003	0.003	0.003	0.005	0.003	0.003	0.003	0.003	0.003
K = 50	a	0.093	0.103	0.103	0.103	0.103	0.098	0.105	0.103	0.098	0.103	0.111	0.103	0.102	0.098	0.102	0.104
	b	1.072	1.008	1.092	1.090	1.091	1.079	1.107	1.018	1.065	1.023	1.176	1.085	1.086	1.079	1.074	1.088
	p	0.005	0.005	0.005	0.005	0.005	0.005	0.005	0.005	0.005	0.005	0.006	0.005	0.005	0.005	0.005	0.005
	q	0.003	0.003	0.003	0.003	0.003	0.003	0.003	0.003	0.003	0.003	0.004	0.003	0.003	0.003	0.003	0.003

continued

TABLE 5

(Continued)

		DGP1							DGP2								
	Parameters	ML Raw	QML KB	GMM-L KB	Bayes KB	QML UB	GMM-L UB	Bayes UB	QML KB	GMM-L KB	GMM KB	ML Multi KB	Bayes KB	QML UB	GMM-L UB	GMM UB	Bayes UB
n = 20,000																	
K = 10	a	0.053	0.057	0.057	0.058	0.056	0.055	0.057	0.058	0.057	0.058	0.128	0.059	0.056	0.056	0.057	0.058
	b	0.505	0.553	0.549	0.553	0.563	0.558	0.564	0.555	0.552	0.552	0.819	0.556	0.563	0.559	0.561	0.564
	p	0.002	0.003	0.003	0.003	0.003	0.003	0.003	0.003	0.003	0.003	0.007	0.003	0.003	0.003	0.003	0.003
	q	0.002	0.002	0.002	0.002	0.002	0.002	0.002	0.002	0.002	0.002	0.004	0.002	0.002	0.002	0.002	0.002
K = 20	a	0.048	0.051	0.057	0.052	0.051	0.051	0.052	0.051	0.051	0.051	0.073	0.052	0.051	0.051	0.051	0.052
	b	0.451	0.474	0.514	0.475	0.472	0.469	0.473	0.472	0.470	0.472	0.658	0.473	0.473	0.469	0.473	0.473
	p	0.002	0.002	0.002	0.002	0.002	0.002	0.002	0.002	0.002	0.002	0.004	0.002	0.002	0.002	0.002	0.002
	q	0.001	0.002	0.002	0.002	0.002	0.002	0.002	0.002	0.002	0.002	0.002	0.002	0.002	0.002	0.002	0.002
K = 50	a	0.047	0.047	0.051	0.048	0.047	0.047	0.048	0.048	0.047	0.047	0.054	0.048	0.047	0.047	0.047	0.048
	b	0.512	0.520	0.537	0.534	0.527	0.523	0.530	0.514	0.528	0.519	0.561	0.533	0.528	0.522	0.524	0.529
	p	0.002	0.002	0.002	0.002	0.002	0.002	0.002	0.002	0.002	0.002	0.003	0.002	0.002	0.002	0.002	0.002
	q	0.001	0.001	0.002	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.002	0.001	0.001	0.001	0.001	0.001
n = 100,000																	
K = 10	a	0.009	0.011	0.011	0.011	0.011	0.011	0.011	0.011	0.011	0.011	0.027	0.011	0.011	0.011	0.011	0.011
	b	0.100	0.109	0.109	0.109	0.109	0.109	0.109	0.109	0.109	0.109	0.187	0.109	0.109	0.109	0.109	0.109
	p	0.000	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001
	q	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.001	0.000	0.000	0.000	0.000	0.000
K = 20	a	0.008	0.009	0.011	0.009	0.010	0.009	0.010	0.009	0.009	0.009	0.014	0.009	0.010	0.009	0.010	0.010
	b	0.091	0.093	0.099	0.093	0.093	0.093	0.093	0.092	0.092	0.092	0.127	0.092	0.093	0.093	0.093	0.093
	p	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.001	0.000	0.000	0.000	0.000	0.000
	q	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
K = 50	a	0.009	0.010	0.010	0.010	0.010	0.010	0.010	0.010	0.010	0.010	0.011	0.010	0.010	0.010	0.010	0.010
	b	0.089	0.090	0.092	0.091	0.091	0.091	0.091	0.090	0.091	0.090	0.102	0.091	0.090	0.091	0.091	0.091
	p	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.001	0.000	0.000	0.000	0.000	0.000
	q	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000

Notes: The results are obtained for 500 Monte Carlo replications and the parameter setting: $a = 4.4$, $b = 69$, $P = 0.7$, $q = 0.6$

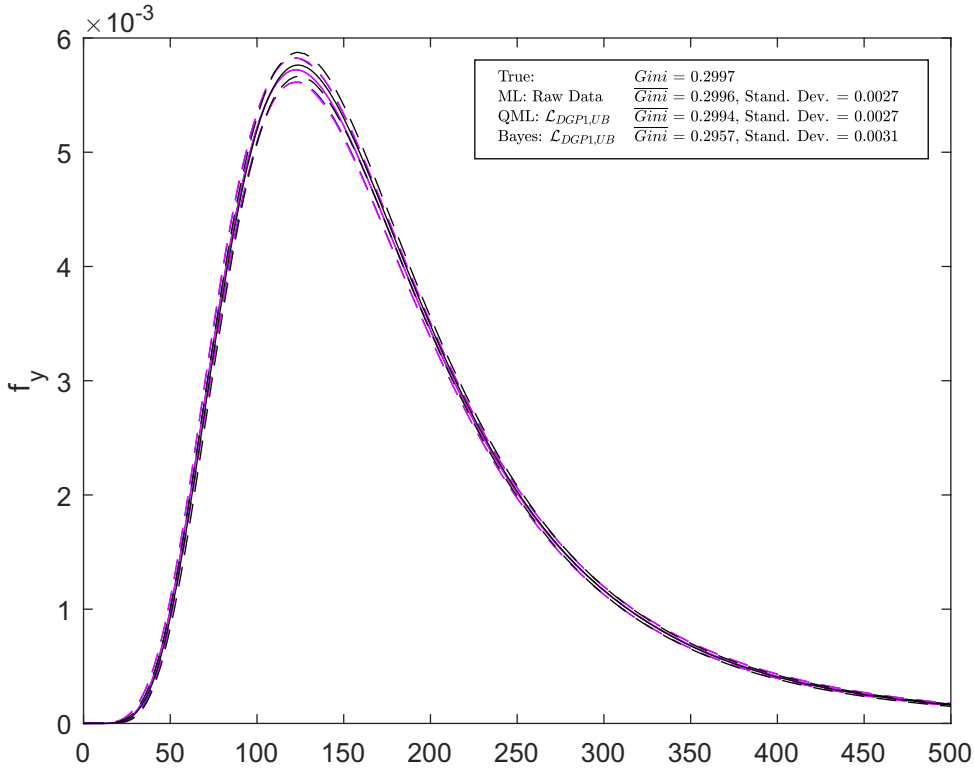

Figure 3. Average estimated income densities under parameter setting (i) for both, grouped data with unknown boundaries (K = 10 income groups; black line: estimated via Quasi Maximum Likelihood (QML); magenta line: estimated via Bayes) and raw data (blue line, estimated via ML), along with corresponding 95% pointwise confidence intervals under DGP1 with $n = 10,000$ (dashed lines). The average densities and confidence intervals are computed using the 500 estimated income distributions from the Monte Carlo experiment of section V. The figure also reports average estimates of the Gini coefficient and according finite sample standard errors which are computed as the sample standard deviation over the 500 Gini estimates

not available in practice. The same holds for DGP1 vs. DGP2: The data generating process itself does not have a strong effect on the MSEs. The worst MSE results are obtained for *ML Multi*, irrespective of the parameter setting, sample size and number of income groups – a finding which is explained by the method's limited use of the available data information. We also do not find clear evidence of decreasing MSEs with increasing number of income groups K, although the results provide some indication that this might be the case for low sample sizes n.

Taken all together, the results indicate a sound and stable performance of the QML- and Bayesian estimators under DGP1 and DGP2 and known-/unknown group boundaries. This performance appears to be robust against varying DGPs, parameterizations, sample sizes and numbers of income groups.

We now focus on parametrization (i), which is characterized by rather high differences in the MSEs across models and settings, and analyse the effect of the parameter uncertainty on estimates of the income distribution itself. Figure 3 depicts average estimated income distributions for both grouped data under DGP1 with unknown boundaries (estimated

via QML and Bayes) and raw data, along with corresponding 95% pointwise confidence intervals, which are computed using the 500 estimated income distributions from the simulation experiment. The figure also reports average estimates of the Gini coefficient and corresponding standard errors. The differences in the estimates under grouping and raw data appear minor. We conclude that the grouping itself generates only moderate losses in estimation uncertainty regarding the income distribution itself and derived measures like the Gini coefficient, even for unknown group boundaries. This is an important finding, since international income data are usually provided in grouped form and one might reasonably expect severe statistical limitations by this data format compared to raw data. Our results imply that this is actually not the case.

VI. Empirical application

We now apply our QML- and Bayesian estimation schemes to grouped household income data from the World Bank website *PovcalNet* provided for the year 2013 (income is measured in purchasing power parity Dollar rates, see *PovcalNet* for details). We consider a selection of four countries: India Rural, Peru, Ethiopia and Iraq. The data consist of group-specific mean incomes $\bar{y}_i$ and population shares c_i for 10 income groups, where the grouping mechanism corresponds to DGP1 with unknown boundaries (constant population shares $c_i = 0.1\ \forall i$). The complete data set is given in Table 6.

Table 7 reports the QML- and Bayesian parameter estimates under the GB2 distribution along with estimates of the Gini coefficient and the headcount ratio (HC) with according standard errors. For a given poverty line x, the headcount ratio is the proportion of population with income less than x. Hence

$$HC = F(x; \hat{\theta}),$$

where we set $x = 57.79$ as provided by the World Bank. The Gini coefficient is obtained as

$$\text{Gini} = -1 + \frac{2}{E[y]} \int_0^\infty y\, F(y;\theta)\, f(y;\theta)\, dy,$$

and the integral is evaluated numerically. We also consider forecasts of observed income shares s_i for $i = 1,\ldots,10$ as a criterion for the goodness of fit of the income distribution (see also Hajargasht *et al.*, 2012). Predicted cumulative income shares η_i are obtained by the first-moment distribution function, $\hat{\eta}_i = F_1(z_i; \hat{\theta})$. Predicted income shares are then computed as $\hat{s}_i = \hat{\eta}_i - \hat{\eta}_{i-1}$. Table 7 reports predictions of the income shares for the first and the last group. Accurate predictions for the first group are of special importance for poverty measurement, while predictions for the last group suffer from the thick right tail of typical income data.

The reported parameter estimates are obtained under the log-likelihood $\mathcal{L}_{\text{DGP1, UB}}$ in equation (14). For the Bayesian estimates we also report numerical standard errors computed by a Parzen-based spectral estimator (see e.g. Kim, Shephard and Chib, 1998) which addresses the numerical uncertainty of the estimates induced by the simulation-based MCMC approach. The last line in Table 7 reports the average MH acceptance ratio, which takes a minimum of 0.55 for Iraq and a maximum of 0.92 for Ethiopia, implying a high degree of simulation efficiency as reflected by a fast mixing of the generated Markov chains.

TABLE 6

Household income data for the empirical application of section VI obtained from the World Bank website PovcalNet for the year 2013 measured in purchasing power parity Dollar rates.

	n	$\bar{y}_1$	$\bar{y}_2$	$\bar{y}_3$	$\bar{y}_4$	$\bar{y}_5$	$\bar{y}_6$	$\bar{y}_7$	$\bar{y}_8$	$\bar{y}_9$	$\bar{y}_{10}$
India Rural	28799	40.0996	52.6935	61.4572	69.5890	77.5614	87.3438	99.1566	115.2944	141.8047	270.7199
Peru	28099	60.6701	114.9490	162.6194	209.5336	259.5349	317.4554	383.6402	472.5938	629.3494	1267.8513
Ethiopia	27755	28.8701	43.9975	53.2265	61.2832	69.7703	78.6416	89.2979	104.5388	129.5908	248.7848
Iraq	24923	68.4795	95.2616	113.4749	130.8265	148.9821	168.7104	192.7976	225.8518	275.5820	441.1694

Notes: See *PovcalNet* for details

TABLE 7

Bayesian and Quasi Maximum Likelihood (QML) estimates of model parameters, poverty and inequality measures and income shares $\hat{s}_1$ and $\hat{s}_{10}$ obtained under the GB2 distribution for the PovcalNet data

	India Rural		Peru		Ethiopia		Iraq	
	Bayes	QML	Bayes	QML	Bayes	QML	Bayes	QML
a	3.0809	3.0842	1.6403	1.6421	4.4050	4.3979	1.5137	1.5179
	(0.1340)	(0.1585)	(0.0785)	(0.0804)	(0.1815)	(0.1999)	(0.1058)	(0.1417)
	[0.0008]	[0.0003]		[0.0006]		[0.0012]		
b	55.5631	55.6948	386.2183	385.9497	69.1348	69.1478	105.7670	106.2234
	(1.5815)	(1.6864)	(9.6912)	(11.2499)	(0.5848)	(0.6178)	(6.3412)	(8.5476)
	[0.0106]		[0.0376]		[0.0020]		[0.0696]	
p	2.3017	2.2769	1.2221	1.2141	0.7106	0.7099	5.2060	5.0863
	(0.2207)	(0.2429)	(0.0885)	(0.0878)	(0.0398)	(0.0435)	(0.7813)	(1.0372)
	[0.0015]		[0.0004]		[0.0001]		[0.0090]	
q	0.8938	0.8903	1.8283	1.8165	0.5939	0.5938	2.9642	2.9191
	(0.0504)	(0.0609)	(0.1436)	(0.1517)	(0.0315)	(0.0353)	(0.3351)	(0.4265)
	[0.0003]		[0.0006]		[0.0001]		[0.0037]	
Gini	0.3078	0.3074	0.4377	0.4375	0.3303	0.3299	0.2947	0.2954
	(0.0021)	(0.0025)	(0.0023)	(0.0026)	(0.0025)	(0.0027)	(0.0016)	(0.0017)
	[< 0.0001]		[< 0.0001]		[< 0.0001]		[< 0.0001]	
HC	0.2051	0.2051	0.0418	0.0418	0.3045	0.3045	0.0195	0.0195
	(0.0020)	(0.0022)	(0.0010)	(0.0010)	(0.0024)	(0.0025)	(0.0007)	(0.0008)
	[< 0.0001]		[< 0.0001]		[< 0.0001]		[< 0.0001]	
s_1 emp	0.0395	0.0395	0.0156	0.0156	0.0318	0.0318	0.0368	0.0368
$\hat{s}_1$	0.0398	0.0396	0.0157	0.0156	0.0320	0.0319	0.0371	0.0368
s_{10} emp	0.2665	0.2665	0.3269	0.3269	0.2740	0.2740	0.2370	0.2370
$\hat{s}_{10}$	0.2614	0.2620	0.3231	0.3241	0.2713	0.2717	0.2357	0.3371
acc-ratio	0.8114		0.8787		0.9224		0.5469	

Notes: HC: Headcount ratio; s_i emp: observed income share for the i'th group. Bayesian estimates: posterior standard errors in parentheses; MC standard errors computed by a Parzen based spectral estimator in square brackets; All estimates are computed using 120,000 MH draws and a burnin of the first 20,000 draws; Lognormal priors with standard deviation 100 centered at the QML estimates; acc-ratio: average MH acceptance ratio α computed over the 100,000 MG draws after burnin. QML estimates: asymptotic standard errors in parentheses

Figure 4 finally depicts Bayesian estimates of the income distributions (posterior means of the income pdf as a function of the GB2 parameters) along with 95% pointwise posterior high-density regions.

The Bayesian- and QML parameter estimates as well as the associated standard errors appear very similar in value, as expected under uninformative prior distributions. All standard errors reported in Table 7 are low and indicate a high level of estimation precision, in particular for the Gini coefficient and the headcount ratio, which are of special interest in applied economic research. The estimation precision is also reflected by the Bayesian density estimates in Figure 4 with very tight 95% posterior high-density intervals. The numerical standard errors under the MCMC estimation scheme are low and in most cases

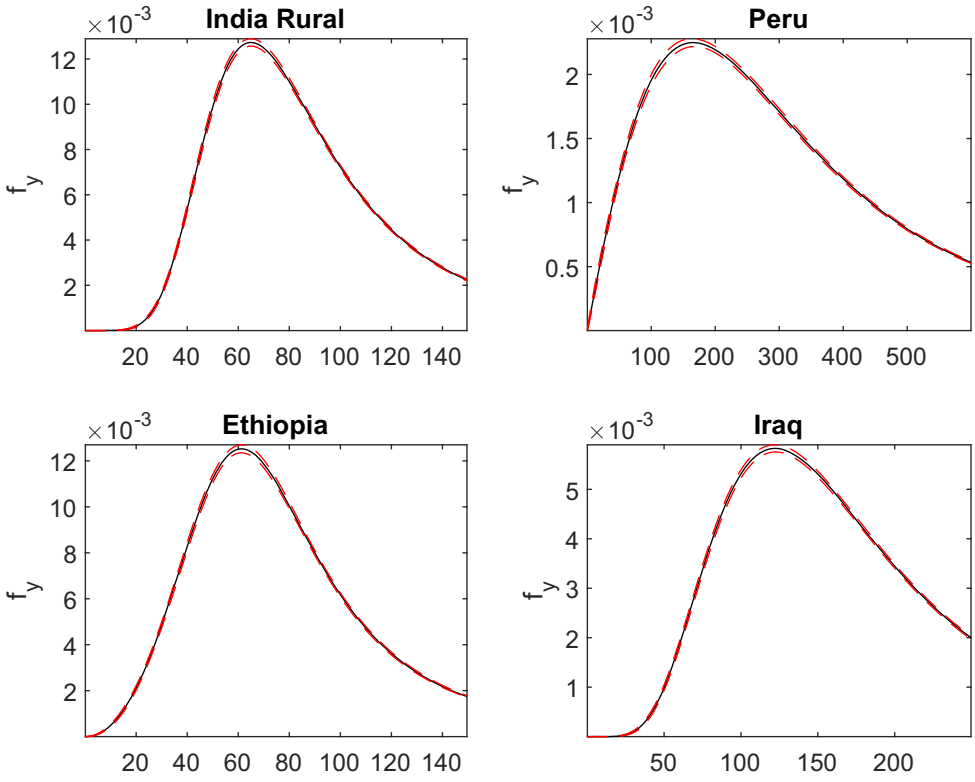

Figure 4. Bayesian estimates of the GB2 income densities (posterior means of the income pdf as a function of the GB2 parameters) along with 95% pointwise posterior high-density regions for India Rural, Peru, Ethiopia and Iraq

less than 1 per cent of the associated posterior standard deviations. Also the income share predictions for the first and the last income group appear very accurate. The highest absolute prediction errors for the income shares are obtained for the last income group, where the heavy right tail of the income distribution makes accurate predictions rather hard to obtain. The nested B2 ($a = 1$), Singh–Maddala ($P = 1$) and Dagum distributions ($q = 1$) are rejected by t-tests at the 5% significance level and not included in the according posterior high-density regions. India Rural builds the only exception since a standard t-test does not reject the Dagum distribution with $q = 1$ at the 5% significance level.

In order to assess the relative benefit of the GB2 for modelling grouped income data we compare Root Mean Squared Errors (RMSEs) for the forecasted income shares under the nested B2, Singh–Maddala and Dagum distributions. The results are reported in Table 8. Note that we do not assess the significance of RMSE differences, since each RMSE is based on 10 observations only, and we only report results for the Bayesian estimates, since the QML results are very similar. The GB2 performs best in all cases except for Peru, where the lowest RMSE is obtained under the Singh–Maddala distribution. We conclude that our findings overall support the adequacy of the GB2 for modelling international income data. However note that the obtained RMSEs are very low for all considered income distributions.

TABLE 8

Root mean squared errors (RMSEs) for income share predictions (parameter estimates: Bayesian)

Distribution	India Rural	Peru	Ethiopia	Iraq
GB2	**0.0018**	0.0015	**0.0009**	**0.0005**
B2	0.0070	0.0032	0.0062	0.0007
Singh-Maddala	0.0027	**0.0006**	0.0026	0.0082
Dagum	0.0024	0.0067	0.0047	0.0065

Notes: Bold numbers indicate the lowest RMSEs. The RMSEs are obtained as $RMSE = \sqrt{K^{-1}\sum_{i=1}^{K}(\hat{s}_i - s_i)^2}$

VII. Conclusion

In this paper we develop a general framework for QML- and Bayesian estimation of parametric income distributions for grouped data with potentially unknown group boundaries. Our approach accounts for two data generating processes of practical relevance and incorporates the information of group mean incomes into the likelihood. The method thereby generalizes the ML and QML frameworks of McDonald (1984), Nishino and Kakamu (2011), and Hitomi *et al.* (2008) which either neglect the informational content of the group mean incomes (McDonald and Nishino and Kakamu) and/or do not provide statistically sound inference in the presence of empirically realistic data generating processes and unknown group boundaries.

A Monte Carlo simulation experiment shows a good and stable performance of the proposed QML- and Bayesian estimation schemes under DGP1 and DGP2 and known-/unknown group boundaries. This performance appears to be robust against varying DGPs, parameterizations, sample sizes and numbers of income groups. Our results also indicate significant improvements over the conventional multinomial ML approach and we obtain accurate parameter estimates which come close to those obtained for individual income data. The results also indicate an overall comparable estimation precision under known and unknown group boundaries for both, DGP1 and DGP2.

We finally apply the QML approach to World Bank data for four countries and find evidence for the GB2 distribution relative to its nested competitors such as the Beta2, Singh–Maddala and the Dagum distribution. The obtained estimates of inequality and poverty measures as well as predictions of income shares show a high degree of accuracy.

Conflict of Interest Statement

The authors declare that there is no conflict of interest.

Funding

This research did not receive any specific grant from funding agencies in the public, commercial or not-for-profit sectors.

Appendix A: Regularity conditions for consistency and asymptotic normality

Let the QML objective function $\mathcal{L}_\cdot(\theta;\cdot) = n \cdot Q_n(\theta)$. Also denote $\tilde{\mu}_i(\theta) = E(y \mid z_{i-1}, z_i; \theta)$ and $\tilde{\sigma}_i^2(\theta) = Var(y \mid z_{i-1}, z_i; \theta)$. For identification we assume that $\theta = \theta_0$ and no other $\theta \in \Theta$ satisfies $\tilde{\mu}_i(\theta) = \tilde{\mu}_i(\theta_0)$, $\tilde{\sigma}_i^2(\theta) = \tilde{\sigma}_i^2(\theta_0)$ and $f_y(z_i;\theta) = f_y(z_i;\theta_0)$.

Note that the unique maximizer of $Q_n(\theta)$, denoted by $\hat{\theta}$, represents an Extremum Estimator (c.f. Hayashi, 2000). According to the standard requirements for the consistency and asymptotic normality of Extremum Estimators we employ the following two sets of assumptions (see e.g. Hayashi, 2000, p. 456 ff.).

Assumption 1. (Consistency) Assume that

(i) the parameter space Θ is a compact subset of $\mathcal{R}^p$;
(ii) $Q_n(\theta)$ is a continuous measurable function in θ;
(iii) there exists a function $Q_0(\theta)$ such that
  (a) (identification) $Q_0(\theta)$ is uniquely maximized on Θ at $\theta_0 \in \Theta$,
  (b) (uniform convergence) $Q_n(\cdot)$ converges uniformly in probability to $Q_0(\cdot)$.

Assumption 2. (Asymptotic normality) Assume that

(i) θ_0 is in the interior of Θ;
(ii) $\dfrac{\partial}{\partial\theta}\displaystyle\int f_y(z;\theta)dz = \int \frac{\partial}{\partial\theta} f_y(z;\theta)\, dz$;
(iii) $\dfrac{\partial Q_n(\theta)}{\partial\theta\partial\theta'}$ exists and is continuous in θ
(iv) $\dfrac{\partial Q_n(\theta^*)}{\partial\theta\partial\theta'} \xrightarrow{p} H(\theta_0)$ is a non-singular matrix for any consistent estimator θ^*;
(v) $\sqrt{n}\dfrac{\partial Q_n(\theta_0)}{\partial\theta} \xrightarrow{d} N(0, M(\theta_0))$, where $M(\theta_0)$ is a symmetric p.d. matrix.

Appendix B: Consistency and asymptotic normality of the QML estimator under DGP1

We start with the situation, where the group boundaries are known. Recognizing that $\mu_i = \frac{n_i - 1}{n_i}\tilde{\mu}_i + z_i/n_i$ and $\sigma_i^2 = \frac{n_i - 1}{n_i^2}\tilde{\sigma}_i^2$ we obtain by equation (13) the QML objective

$$
\begin{aligned}
n \cdot Q_n(\theta) = \Omega - 0.5[&\ln \tilde{\sigma}_K^2(\theta) - \ln n_K + \tilde{\sigma}_K^{-2}(\theta) n_K (\bar{y}_K - \tilde{\mu}_K(\theta))^2] \\
&+ (n - n_{K-1}^c)\ln[1 - F_y(z_{K-1};\theta)] \\
&+ \sum_{i=1}^{K-1} \Bigg\{ -0.5 \Bigg[\ln \tilde{\sigma}_i^2(\theta) - 2\ln(n_i) + \ln(n_i - 1) \\
&+ \tilde{\sigma}_i^2(\theta)^{-2} \frac{n_i^2}{n_i - 1} \left(\bar{y}_i - \frac{n_i - 1}{n_i}\tilde{\mu}_i(\theta) - \frac{z_i}{n_i} \right)^2 \Bigg] \\
&+ (n_i^c - n_{i-1}^c - 1)\ln[F_y(z_i;\theta) - F_y(z_{i-1};\theta)] + \ln f_y(z_i;\theta) \Bigg\}.
\end{aligned}
$$

Note that under DGP1 the proportions $c_i = n_i/n$ are fixed. According to point (iii) of Assumption 1 (see Appendix A) we observe that Ω is $o(n)$ and by the WLLN we have uniform convergence $\bar{y}_i \xrightarrow{p} \tilde{\mu}_i(\theta_0)$, and $z_i = y_{[n_i^c]} \xrightarrow{p} F_y^{-1}(c_i^c;\theta_0) \widehat{=} q_y(c_i^c;\theta_0)$, where $c_i^c = \sum_{\ell=1}^{i} c_\ell$. We then obtain

$$\begin{aligned} Q_n(\theta) \xrightarrow{p} & -0.5[\tilde{\sigma}_K^{-2}(\theta) c_K(\tilde{\mu}_K(\theta_0) - \tilde{\mu}_K(\theta))^2] \\ & + c_K \ln[1 - F_y(q_y(c_{K-1}^c;\theta_0);\theta)] \\ & + \sum_{i=1}^{K-1} -0.5[\tilde{\sigma}_i^2(\theta)^{-2} c_i(\tilde{\mu}_i(\theta_0) - \tilde{\mu}_i(\theta))^2] \\ & + c_i \ln[F_y(q_y(c_i^c,\theta_0);\theta) - F_y(q_y(c_{i-1}^c,\theta_0);\theta)], \end{aligned}$$

which is uniquely maximized for $\theta = \theta_0$. It follows that $\hat{\theta} \xrightarrow{p} \theta_0$ (see e.g. Hayashi, 2000).

For the derivation of the asymptotic distribution of $\hat{\theta}$ we employ Assumption 2 of Appendix A and note that by CLT arguments analogous to our asymptotic approximation in equation (8), involving the convergence of the characteristic function of the sum of iid random variables, Hitomi *et al.* (2008) show that for QML functions based on the Gaussian likelihood approximation in equation (8) it holds that $H(\theta_0) = -M(\theta_0)^{-1}$ (Information Matrix Equality, IME). By standard results, IME still holds if the likelihood is enriched by the joint density of the group boundaries, as long as we assume sufficient smoothness of the income distribution $F_y(\cdot)$, such that the order of integration and differentiation can be interchanged, see point (ii) of Assumption 2 (c.f. Amemiya, 1985, p. 17). In this case the QML estimator $\hat{\theta}$ is asymptotically efficient with asymptotic covariance matrix $ACOV(\hat{\theta}) = -H(\theta_0)^{-1}/n$ (see Hitomi *et al.*, 2008).

We obtain

$$\begin{aligned} \frac{\partial Q_n(\theta)}{\partial\theta\partial\theta'} = -0.5 \Bigg\{ & \left[\frac{1}{\tilde{\sigma}_K^2(\theta)n} - \frac{n_K}{\tilde{\sigma}_K^4(\theta)n}(\bar{y}_K - \tilde{\mu}_K(\theta))^2\right] \frac{\partial^2\tilde{\sigma}_K^2(\theta)}{\partial\theta\partial\theta'} \\ & + \left[-\frac{1}{\tilde{\sigma}_K^4(\theta)n} + \frac{2n_K}{\tilde{\sigma}_i^6(\theta)n}(\bar{y}_K - \tilde{\mu}_K(\theta))^2\right] \frac{\partial\tilde{\sigma}_K^2(\theta)}{\partial\theta}\frac{\partial\tilde{\sigma}_K^2(\theta)}{\partial\theta'} \\ & + \left[\frac{-2n_K}{\tilde{\sigma}_K^2(\theta)n}(\bar{y}_K - \tilde{\mu}_K(\theta))\right] \frac{\partial^2\tilde{\mu}_K(\theta)}{\partial\theta\partial\theta'} + \left[\frac{2n_K}{\tilde{\sigma}_K^2(\theta)n}\right] \frac{\partial\tilde{\mu}_K(\theta)}{\partial\theta}\frac{\partial\tilde{\mu}_K(\theta)}{\partial\theta'} \\ & + \left[\frac{2n_K}{\tilde{\sigma}_K^4(\theta)n}(\bar{y}_K - \tilde{\mu}_K(\theta))\right] \frac{\partial\tilde{\mu}_K(\theta)}{\partial\theta}\frac{\partial\tilde{\sigma}_K^2(\theta)}{\partial\theta'} \\ & + \left[\frac{2n_K}{\tilde{\sigma}_K^4(\theta)n}(\bar{y}_K - \tilde{\mu}_K(\theta))\right] \frac{\partial\tilde{\sigma}_K^2(\theta)}{\partial\theta}\frac{\partial\tilde{\mu}_K(\theta)}{\partial\theta'} - \left[\frac{n - n_{K-1}^c}{n(1 - F_y(z_{K-1};\theta))}\right] \frac{\partial^2 F_y(z_{K-1};\theta)}{\partial\theta\partial\theta'} \\ & - \left[\frac{n - n_{K-1}^c}{n(1 - F_y(z_{K-1};\theta))^2}\right] \frac{\partial F_y(z_{K-1};\theta)}{\partial\theta}\frac{\partial F_y(z_{K-1};\theta)}{\partial\theta'} \Bigg\} \\ & - \sum_{i=1}^{K-1} 0.5 \Bigg\{ \left[\frac{1}{\tilde{\sigma}_i^2(\theta)n} - \frac{n_i^2}{\tilde{\sigma}_i^4(\theta)(n_i-1)n}(\bar{y}_i - (n_i-1)\tilde{\mu}_i(\theta)/n_i - z_i/n_i)^2\right] \frac{\partial^2\tilde{\sigma}_i^2(\theta)}{\partial\theta\partial\theta'} \end{aligned}$$

$$
\begin{aligned}
&+\left[-\frac{1}{\tilde{\sigma}_i^4(\theta)n}+\frac{2n_i^2}{\tilde{\sigma}_i^6(\theta)(n_i-1)n}(\bar{y}_i-(n_i-1)\tilde{\mu}_i(\theta)/n_i-z_i/n_i)^2\right]\frac{\partial\tilde{\sigma}_i^2(\theta)}{\partial\theta}\frac{\partial\tilde{\sigma}_i^2(\theta)}{\partial\theta'}\\
&+\left[\frac{-2n_i}{\tilde{\sigma}_i^2(\theta)n}(\bar{y}_i-(n_i-1)\tilde{\mu}_i(\theta)/n_i-z_i/n_i)\right]\frac{\partial^2\tilde{\mu}_i(\theta)}{\partial\theta\partial\theta'}+\left[\frac{2(n_i-1)}{\tilde{\sigma}_i^2(\theta)n}\right]\frac{\partial\tilde{\mu}_i(\theta)}{\partial\theta}\frac{\partial\tilde{\mu}_i(\theta)}{\partial\theta'}\\
&+\left[\frac{2n_i}{\tilde{\sigma}_i^4(\theta)n}(\bar{y}_i-(n_i-1)\tilde{\mu}_i(\theta)/n_i-z_i/n_i)\right]\frac{\partial\tilde{\mu}_i(\theta)}{\partial\theta}\frac{\partial\tilde{\sigma}_i^2(\theta)}{\partial\theta'}\\
&+\left[\frac{2n_i}{\tilde{\sigma}_i^4(\theta)n}(\bar{y}_i-(n_i-1)\tilde{\mu}_i(\theta)/n_i-z_i/n_i)\right]\frac{\partial\tilde{\sigma}_i^2(\theta)}{\partial\theta}\frac{\partial\mu_i(\theta)}{\partial\theta'}\Bigg\}\\
&-\frac{n_i^c-n_{i-1}^c-1}{n(F_y(z_i;\theta)-F_y(z_{i-1};\theta))^2}\left[\frac{\partial F_y(z_i;\theta)}{\partial\theta}-\frac{\partial F_y(z_{i-1};\theta)}{\partial\theta}\right]\left[\frac{\partial F_y(z_i;\theta)}{\partial\theta}-\frac{\partial F_y(z_{i-1};\theta)}{\partial\theta}\right]'\\
&+\frac{n_i^c-n_{i-1}^c-1}{n(F_y(z_i;\theta)-F_y(z_{i-1};\theta))}\left[\frac{\partial^2F_y(z_i;\theta)}{\partial\theta\partial\theta'}-\frac{\partial^2F_y(z_{i-1};\theta)}{\partial\theta\partial\theta'}\right]\\
&-\frac{1}{f_y(z_i;\theta)^2n}\frac{\partial f_y(z_i;\theta)}{\partial\theta}\frac{\partial f_y(z_i;\theta)}{\partial\theta'}+\frac{1}{f_y(z_i;\theta)n}\frac{\partial^2f_y(z_i;\theta)}{\partial\theta\partial\theta'},
\end{aligned}
$$

which implies

$$
\begin{aligned}
H(\theta_0)&=\operatorname{plim}\left(\frac{\partial Q_n(\theta_0)}{\partial\theta\partial\theta'}\right)\\
&=-\left[\frac{\partial\tilde{\mu}_K(\theta_0)}{\partial\theta}\frac{\partial\tilde{\mu}_K(\theta_0)}{\partial\theta'}\right]\frac{c_K}{\tilde{\sigma}_K^2(\theta_0)}-\frac{1}{c_K}\left[\frac{\partial F_y(q_y(c_{K-1}^c,\theta_0);\theta_0)}{\partial\theta}\frac{\partial F_y(q_y(c_{K-1}^c,\theta_0);\theta_0)}{\partial\theta'}\right]\\
&\quad+\sum_{i=1}^{K-1}-\left[\frac{\partial\tilde{\mu}_i(\theta_0)}{\partial\theta}\frac{\partial\tilde{\mu}_i(\theta_0)}{\partial\theta'}\right]\frac{c_i}{\tilde{\sigma}_i^2(\theta_0)}-\frac{1}{c_i}\left[\frac{\partial F_y(q_y(c_i^c,\theta_0);\theta_0)}{\partial\theta}-\frac{\partial F_y(q_y(c_{i-1}^c,\theta_0);\theta_0)}{\partial\theta}\right]\\
&\quad\times\left[\frac{\partial F_y(q_y(c_i^c,\theta_0);\theta_0)}{\partial\theta}-\frac{\partial F_y(q_y(c_{i-1}^c,\theta_0);\theta_0)}{\partial\theta}\right]'\\
&=\sum_{i=1}^{K}-\left[\frac{\partial\tilde{\mu}_i(\theta_0)}{\partial\theta}\frac{\partial\tilde{\mu}_i(\theta_0)}{\partial\theta'}\right]\frac{c_i}{\tilde{\sigma}_i^2(\theta_0)}-\frac{1}{c_i}\left[\frac{\partial F_y(q_y(c_i^c;\theta_0);\theta_0)}{\partial\theta}-\frac{\partial F_y(q_y(c_{i-1}^c;\theta_0);\theta_0)}{\partial\theta}\right]\\
&\quad\times\left[\frac{\partial F_y(q_y(c_i^c;\theta_0);\theta_0)}{\partial\theta}-\frac{\partial F_y(q_y(c_{i-1}^c;\theta_0);\theta_0)}{\partial\theta}\right]',
\end{aligned}
$$

where $\theta_0=\operatorname{plim}(\hat{\theta})$ denotes the true value of θ, $q_y(\cdot;\theta_0)=F_y^{-1}(\cdot;\theta_0)$ denotes the quantile function of y, $c_i^c=\sum_{\ell=1}^{i}c_\ell$ and by definition of the first and the last income group $\partial F_y(q_y(c_0^c,\theta_0);\theta_0)/\partial\theta\,\widehat{=}\,0$ and $\partial F_y(q_y(c_K^c,\theta_0);\theta_0)/\partial\theta\,\widehat{=}\,0$.

For the case of unknown group boundaries we use results of Beach and Davidson (1983) leading to

$$
\sqrt{n}(\tilde{y}-L(q_y(c^c)))\xrightarrow{d}N(0,\ \Omega), \tag{21}
$$

where $\tilde{y}=(\tilde{y}_1,\ldots,\tilde{y}_K)$ with $\tilde{y}_i=\sum_{j=1}^{i}c_j\bar{y}_j$, and $L(q_y(c^c))=(L_1,\ldots,L_K)'$ with $L_i=L(q_y(c_i^c))=\int_0^{q_y(c_i^c)}yf_y(y)dy$.

The limiting covariance $\Omega = (\omega_{ij})$ obtains as

$$\omega_{ij} = m_i + (c_i^c q_y(c_i^c) - L_i)(q_y(c_j^c) - c_j^c q_y(c_j^c) + L_j) - q_y(c_i^c)L_i \quad \text{for } i \leqslant j, \tag{22}$$

where

$$m_i = \int_0^{q_y(c_i^c)} y^2 f_y(y)\, dy. \tag{23}$$

From these results we obtain

$$\underline{\bar{y}} \overset{\text{appr:}}{\sim} N\left(\mu^*,\ \frac{1}{n}\Psi\right), \tag{24}$$

where $\mu^* = (\mu_1^*, \ldots, \mu_K^*)'$, with $\mu_i^* = (1/c_i)\int_{q_y(c_{i-1}^c;\theta)}^{q_y(c_i^c;\theta)} y f_y(y) dy$, and $\Psi = DB\Omega B'D'$, with $D = \text{diag}(c_1^{-1}, \ldots, c_K^{-1})$ and

$$B = \begin{bmatrix} 10 & 0 & \ldots & 0 & 0 \\ -11 & 0 & \ldots & 0 & 0 \\ \vdots\vdots & \vdots & & \vdots & \vdots \\ 00 & 0 & \ldots & -1 & 1 \end{bmatrix}. \tag{25}$$

From (24) we obtain the QML objective $nQ_n(\theta) = \mathcal{L}_{\text{DGP1, UB}}(\theta; \underline{\bar{y}})$ in equation (14). Then since $\underline{\bar{y}} \xrightarrow{p} \mu^*(\theta_0)$

$$Q_n(\theta) \xrightarrow{p} (\mu^*(\theta_0) - \mu^*(\theta))'\Psi^{-1}(\mu^*(\theta_0) - \mu^*(\theta)), \tag{26}$$

which is uniquely maximized for $\theta = \theta_0$.

Let $H = \Psi/n$ and $x = (\underline{\bar{y}} - \mu^*)$. Then $\dot{H}_i = \frac{\partial H}{\partial \theta_i}$ and $\ddot{H}_{ij} = \frac{\partial^2 H}{\partial \theta_i \partial \theta_j}$, while $\dot{x}_i$ and $\ddot{x}_{ij}$ are defined accordingly. We obtain

$$\begin{aligned} \frac{\partial^2 Q_n(\theta)}{\partial \theta_i \partial \theta_j} = -\frac{1}{2n} tr[&\ddot{H}_{ij}H^{-1} - \dot{H}_i H^{-1}\dot{H}_j H^{-1} \\ &+ x'H^{-1}\ddot{x}_{ij} - x'H^{-1}\dot{H}_j H^{-1}\dot{x}_i + \dot{x}_j'H^{-1}\dot{x}_i \\ &+ \dot{x}_i'H^{-1}\dot{x}_j - \dot{x}_i'H^{-1}\dot{H}_j H^{-1}x + \ddot{x}_{ij}'H^{-1}x \\ &- x'H^{-1}\dot{H}_i H^{-1}\dot{x}_j - \dot{x}_j'H^{-1}\dot{H}_i H^{-1}x \\ &- x'H^{-1}[\ddot{H}_{ij} - \dot{H}_i H^{-1}\dot{H}_j - \dot{H}_j H^{-1}\dot{H}_i]H^{-1}x], \end{aligned} \tag{27}$$

implying

$$\frac{\partial^2 Q_n(\theta_0)}{\partial \theta_i \partial \theta_j} \xrightarrow{p} -\frac{1}{2}\left[\frac{\partial \mu^*(\theta_0)'}{\partial \theta_i}\Psi^{-1}\frac{\partial \mu^*(\theta_0)}{\partial \theta_j} + \frac{\partial \mu^*(\theta_0)'}{\partial \theta_j}\Psi^{-1}\frac{\partial \mu^*(\theta_0)}{\partial \theta_i}\right] \tag{28}$$

and

$$H(\theta_0) = \text{plim}\left(\frac{\partial^2 Q_n(\theta)}{\partial \theta \partial \theta'}\right) = -\frac{\partial \mu^*(\theta_0)'}{\partial \theta}\Psi^{-1}\frac{\partial \mu^*(\theta_0)}{\partial \theta}. \tag{29}$$

Appendix C: Consistency and asymptotic normality of the QML estimator under DGP2

We start with the situation, where the group boundaries are known. We then obtain by equation (17) the Quasi-Log-Likelihood

$$n \cdot Q_n(\theta) = \Omega + \sum_{i=1}^{K} \{-0.5[\ln \tilde{\sigma}_i^2(\theta) - \ln(n_i) + \tilde{\sigma}_i^{-2}(\theta) n_i (\bar{y}_i - \tilde{\mu}_i(\theta))^2] + n_i \ln \pi_i(\theta)\}.$$

According to condition (iii) for the consistency of Extremum Estimators (see Assumption 1 of Appendix A) we observe that Ω is $o(n)$ and by the WLLN we have uniform convergence $\bar{y}_i \xrightarrow{p} \tilde{\mu}_i(\theta_0)$ and $n_i/n \xrightarrow{p} \pi_i(\theta_0)$. We then obtain

$$Q_n(\theta) \xrightarrow{p} \sum_{i=1}^{K} \{-0.5[\tilde{\sigma}_i^{-2}(\theta)\pi_i(\theta_0)(\tilde{\mu}_i(\theta_0) - \tilde{\mu}_i(\theta))^2] + \pi_i(\theta_0) \ln(\pi_i(\theta))\},$$

which is uniquely maximized for $\theta = \theta_0$.

The asymptotic covariance matrix of $\hat{\theta}$ is obtained under Assumption 2 (see Appendix A) via

$$\begin{aligned}
n\frac{\partial Q_n(\theta)}{\partial\theta\partial\theta'} = -0.5\Bigg\{ & \sum_{i=1}^{K} \left[\frac{1}{\tilde{\sigma}_i^2(\theta)} - \frac{n_i}{\tilde{\sigma}_i^4(\theta)}(\bar{y}_i - \tilde{\mu}_i(\theta))^2\right] \frac{\partial^2 \tilde{\sigma}_i^2(\theta)}{\partial\theta\partial\theta'} \\
& + \left[-\frac{1}{\tilde{\sigma}_i^4(\theta)} + \frac{2n_i}{\tilde{\sigma}_i^6(\theta)}(\bar{y}_i - \tilde{\mu}_i(\theta))^2\right] \frac{\partial\tilde{\sigma}_i^2(\theta)}{\partial\theta} \frac{\partial\tilde{\sigma}_i^2(\theta)}{\partial\theta'} \\
& + \left[\frac{-2n_i}{\tilde{\sigma}_i^2(\theta)}(\bar{y}_i - \tilde{\mu}_i(\theta))\right] \frac{\partial^2\tilde{\mu}_i(\theta)}{\partial\theta\partial\theta'} \\
& + \left[\frac{2n_i}{\tilde{\sigma}_i^2(\theta)}\right] \frac{\partial\tilde{\mu}_i(\theta)}{\partial\theta} \frac{\partial\tilde{\mu}_i(\theta)}{\partial\theta'} \\
& + \left[\frac{2n_i}{\tilde{\sigma}_i^4(\theta)}(\bar{y}_i - \tilde{\mu}_i(\theta))\right] \frac{\partial\tilde{\mu}_i(\theta)}{\partial\theta} \frac{\partial\tilde{\sigma}_i^2(\theta)}{\partial\theta'} \\
& + \left[\frac{2n_i}{\tilde{\sigma}_i^4(\theta)}(\bar{y}_i - \tilde{\mu}_i(\theta))\right] \frac{\partial\tilde{\sigma}_i^2(\theta)}{\partial\theta} \frac{\partial\tilde{\mu}_i(\theta)}{\partial\theta'} - 2n_i \frac{\partial^2 \ln \pi_i(\theta)}{\partial\theta\partial\theta'}\Bigg\},
\end{aligned}$$

where since $\bar{y}_i \xrightarrow{p} \tilde{\mu}_i(\theta_0)$ and $n_i/n \xrightarrow{p} \pi_i(\theta_0)$

$$H(\theta_0) = \text{plim}\left(\frac{\partial Q_n(\theta_0)}{\partial\theta\partial\theta'}\right) = \sum_{i=1}^{K} \pi_i(\theta_0) \frac{\partial^2 \ln \pi_i(\theta_0)}{\partial\theta\partial\theta'} - \left[\frac{\partial\tilde{\mu}_i(\theta_0)}{\partial\theta} \frac{\partial\tilde{\mu}_i(\theta_0)}{\partial\theta'}\right] \frac{\pi_i(\theta_0)}{\tilde{\sigma}_i^2(\theta_0)}. \tag{30}$$

Note that the asymptotic covariance matrix of $\hat{\theta}$, $ACOV_{\text{DGP2}}(\hat{\theta}) = -\frac{1}{n}H(\theta_0)^{-1}$, corresponds to the one obtained under the QML approach of Hitomi *et al.* (2008) and the GMM estimators of Hajargasht *et al.* (2012) and Griffiths and Hajargasht (2015). Also note that the same asymptotic covariance applies if the group boundaries are unknown (with the parameter vector θ augmented by the set of group boundaries).

Final Manuscript Received: June 2020

References

Amemiya, T. (1985). *Advanced Econometrics*, Basil Blackwell, Oxford.

Bandourian, R., McDonald, J. B. and Turley, R. S. (2003). 'Income distributions: an inter-temporal comparison over countries', *Estadistica*, Vol. 55, pp. 135–152.

Beach, C. and Davidson, R. (1983). 'Distribution-free statistical inference with lorenz curves and income shares', *Review of Economic Studies*, Vol. 50, pp. 723–735.

Chen, Y. (2018). 'A unified approach to estimating and testing income distributions with grouped data', *Journal of Business and Economic Statistics*, Vol. 36, pp. 438–455.

Chotikapanich, D. (ed.) (2008). *Modeling Income Distributions and Lorenz Curves*, Springer, New York.

Chotikapanich, D. and Griffiths, W. E. (2000). 'Posterior distributions for the gini coefficient using grouped data', *Australian & New Zealand Journal of Statistics*, Vol. 42, pp. 383–392.

Chotikapanich, D., Griffiths, W. E. and Rao, D. S. P. (2007). 'Estimating and combining national income distributions using limited data', *Journal of Business and Economic Statistics*, Vol. 25, pp. 97–109.

Chotikapanich, D., Griffiths, W. E., Rao, D. S. P. and Valencia, V. (2012) 'Global income distributions and inequality, 1993 and 2000: incorporating country-level inequality modelled with beta distributions', *The Review of Economics and Statistics*, Vol. 94, 52–73.

Cowell, F. A. (1991). 'Grouping bounds for inequality measures under alternative informational assumptions', *Journal of Econometrics*, Vol. 48, pp. 1–14.

David, H. A. and Nagaraja, H. N. (2003). *Order Statistics*, Wiley, Hoboken, New Jersey.

Griffiths, W. E. and Hajargasht, G. (2012). *GMM Estimation of Mixtures from Grouped Data: an Application to Income Distributions*, Working Paper.

Griffiths, W. E. and Hajargasht, G. (2015). 'On GMM estimation of distributions from grouped data', *Economics Letters*, Vol. 126, pp. 122–126.

Hajargasht, G. and Griffiths, W. E. (2020). 'Minimum distance estimation of parametric lorenz curves based on grouped data', *Econometric Reviews*, Vol. 39, pp. 344–361.

Hajargasht, G., Griffiths, W. E., Brice, J., Rao, D. S. P. and Chotikapanich, D. (2012). 'Inference for income distributions using grouped data', *Journal of Business & Economic Statistics*, Vol. 30, pp. 563–575.

Hayashi, F. (2000). *Econometrics*, Princeton University Press, Princeton, NJ.

Hitomi, K., Liu, Q.-F., Nishiyama, Y. and Sueishi, N. (2008). 'Efficient estimation and model selection for grouped data with local moments', *Journal of the Japan Statistical Society*, Vol. 38, pp. 131–143.

Kakamu, K. (2016). 'Simulation studies comparing dagum and singh–maddala income distributions', *Computational Economics*, Vol. 48, pp. 593–605.

Kakamu, K. and Nishino, H. (2019). 'Bayesian estimation of beta-type distribution parameters based on grouped data', *Computational Economics*, Vol. 54, pp. 625–645.

Kim, S., Shephard, N. and Chib, S. (1998). 'Stochastic volatility: likelihood inference and comparison with arch models', *The Review of Economic Studies*, Vol. 65, pp. 361–393.

Kleiber, C. and Kotz, S. (2003). *Statistical Size Distributions in Economics and Actuarial Sciences*, Wiley, New York.

McDonald, J. B. (1984). 'Some generalized functions for the size distribution of income', *Econometrica*, Vol. 52, pp. 647–663.

Nadarajah, S. (2005). 'Sums, products and ratios of generalized beta variables', *Statistical Papers*, Vol. 47, pp. 69–90.

Nishino, H. and Kakamu, K. (2011). 'Grouped data estimation and testing of gini coefficients using lognormal distributions', *Sankhya B*, Vol. 73, pp. 193–210.

Parker, S. C. (1999). 'The generalized beta as a model for the distribution of earnings', *Economics Letters*, Vol. 62, pp. 197–200.

Stoye, J. (2010). 'Partial identification of spread parameters', *Quantitative Economics*, Vol. 1, pp. 323–357.

Wu, X. (2006). *Inference and Density Estimation with Interval Statistics*, Working Paper.

Wu, X. and Perloff, J. (2005). 'China's income distribution, 1985–2001', *The Review of Economics and Statistics*, Vol. 87, pp. 763–775.