

**On Statistical Inference for Inequality Measures
Calculated from Complex Survey Data**

**Judith A. Clarke
Nilanjana Roy**

*Department of Economics, University of Victoria
Victoria, B.C., Canada V8W 2Y2*

June 2009

Abstract

We examine inference for Generalized Entropy and Atkinson inequality measures with complex survey data, using Wald statistics with variance-covariance matrices estimated from a linearization approximation rather than the δ -method. Testing the equivalence of two or more inequality measures, including sub-group decomposition indices and group shares, are covered. We illustrate with Indian data from three surveys, examining children's height-for-age, an anthropometric measure that can indicate long-term malnutrition. Sampling involved an urban/rural stratification with clustering before selection of households. We compare the linearization complex survey outcomes with those from an incorrect iid assumption and a bootstrap that accounts for the survey design. For our samples, the results from the easy to implement linearization method and more computationally burdensome bootstrap are in close agreement.

Keywords: complex survey; inequality; Generalized Entropy; Atkinson; decomposition; linearization

JEL Classifications: C12, C42, D31

Author Contact:

Nilanjana Roy, Dept. of Economics, University of Victoria, P.O. Box 1700, STN CSC, Victoria, B.C., Canada V8W 2Y2; e-mail: nroy@uvic.ca; Phone: (250) 472-4417; FAX: (250) 721-6214

I. Introduction

The study of inequality, especially income inequality or income distribution, has been regaining attention among economists and other social scientists since the late 1980s, after a period of relative neglect throughout the rest of the last century. As Atkinson (1997, p. 297) points out “*For much of this century, it (the subject of income distribution) has been very much out in the cold*”. The revival of interest has resulted in a growing literature on the theory of measurement of inequality as well as formal analysis of the statistical properties of such measures. However, gaps still remain between theoretical developments and empirical applications.

A key manifestation of this gap is that much of the applied inequality research does not undertake statistical inference. Most empirical studies¹ use inequality measures to make inter-temporal or inter-regional comparisons of inequality and/or for studying policy impacts (e.g., to examine the effect of a tax policy) but the conclusions are usually based on comparisons of the numerical estimates rather than on formal statistical testing.² One argument (see Maasoumi, 1997) used by applied investigators to support such practices has been that their large samples do not warrant concern about precision but in reality, large standard errors can still arise with such data sets. Furthermore, as the majority of the statistical theory in this area is based on large sample or asymptotic approximations, the use of large samples in most income inequality studies actually makes it more meaningful to report standard errors and undertake statistical tests. So, why the common lack of statistical inference? We believe two factors are perhaps at play – applicability of current theoretical results and ease of use of relevant theory.

On the first factor, many papers provide theoretical contributions to the topic of inference with inequality measures using asymptotic approximations, including, to name only a few, Cowell (1989), Binder and Kovačević (1995), Van de gaer et al. (1999), Schluter and Trede (2002), Biewen and Jenkins (2006), Davidson and Flachaire (2007), Bhattacharya (2007) and Davidson (2009). Most of these studies

¹ See, e.g., Maasoumi (1997) for a survey of the theoretical and empirical literature on income inequality.

² See, e.g., Ram (2006).

focus on an identically and independently distributed (iid) framework, whereas data commonly used by applied researchers come from complex surveys. Although some of these iid papers consider weights, these take on a different role than in a complex survey. For instance, Cowell (1989) examines inference for decomposable inequality measures with random household weights that convert the observed household distribution into a personal distribution. Schluter and Trede (2002) allow for contemporaneous dependences within households, but assume that households are iid. Correlation is also introduced by Van de gaer et al. (1999) but it is temporal dependence rather than correlations arising from the survey design. However, most inequality measures are calculated from data obtained from a complex survey design, leading to (asymptotic) variances and covariances that can be quite different from those generated under simple random sampling (SRS) or the iid with weights framework. Consequently, to date, much of the theoretical work on asymptotic inference does not pertain to the data often used by empirical researchers.

Data obtained under a complex survey design typically involves both stratification and clustering, undertaken to ensure adequate representation of groups of interest, in addition to minimizing financial and administrative costs of surveying a population. Stratification, which can substantially reduce survey costs for a given level of precision, results in the breakdown of the “identical” part of an iid assumption – even when members are independent within a stratum, they are unlikely to come from the same distribution across strata. Moreover, sample observations are likely correlated when the survey design involves clustering, so violating the “independent” part of an iid assumption. Clustering, such as interviewing several households on the same block or from the same village, likely introduces a common unobserved cluster-specific effect, which needs to be accounted for when undertaking inference.

Theoretical contributions that do incorporate such effects include Binder and Kovačević (1995) and, more recently, Biewen and Jenkins (2006) and Bhattacharya (2007). Each of these provides explicit ways to obtain (at least) asymptotic variances for various inequality measures when data are obtained from complex surveys. Binder and Kovačević (1995) use linearization methods based on so-called estimating equations to obtain variance estimators for a few of the inequality measures (Gini

coefficient, coefficient of variation, an “exponential measure” and Lorenz curve ordinates) allowing for complex survey data. Asymptotic inference for the Lorenz curve and the Gini coefficient, assuming complex survey data, is developed by, for example, Bhattacharya (2007). The paper most akin to our work is Biewen and Jenkins (2006), who use a commonly applied survey “linearization method” to obtain asymptotic variances for the Atkinson (1970) and Generalized Entropy measures with complex survey data. Like the common δ -method estimator, the linearization estimator is obtained from a Taylor series approximation.

One ideal feature of the expressions provided by Biewen and Jenkins (2006) is that they address the second factor we identified above for the possible lack of statistical inference in applied research – ease of use. This factor seems important, as it has been identified by others (see, e.g., Giles, 2004 and Davidson, 2009), in explaining the lack of use of asymptotic variance formulae and subsequent hypothesis testing in applied inequality research. Asymptotic variances obtained via the linearization approach avoid the calculation of a potentially complicated variance-covariance matrix that arises with the usual δ -method. Indeed, the variances can be obtained with minimal coding in standard software packages. One element not covered by Biewen and Jenkins, as well as the other cited references that account for complex sampling, is that they do not examine the elements arising from a decomposition analysis, such as the “between” and “within” components or any share measures that may be generated from these parts. Decomposing inequality measures is standard applied practice. In addition, Biewen and Jenkins do not indicate how to extend the approach to test hypotheses involving two or more inequality measures.³

Our goal is to provide these missing pieces. For the Atkinson and Generalized Entropy families, we give linearization expressions for sub-group decomposition measures, the between and within components and for sub-group shares of overall inequality, describing how to use these in standard software packages to generate estimates of asymptotic variances with only a few lines of code. In addition, we show explicitly how to easily extend the method to calculate asymptotic variance-

³ For their empirical application, Biewen and Jenkins (2006) simply provide values of statistics for significance testing and a test statistic for a difference hypothesis.

covariance matrices so that Wald statistics can be formed to test hypotheses involving two or more inequality measures. In particular, we cover testing the equivalence of inequality measures that may be simple inequality indices, sub-group decomposition indices, between and within measures and group shares of overall inequality. Modified statistics that account for small sampling design degrees of freedom are also given. As in Biewen and Jenkins (2006), our expressions are applicable to the study of inequality of not just income but many other well-being variables such as wages, years of schooling, height-for-age etc. Consequently, the results should be of interest to a wide range of empirical researchers in various fields.

We illustrate using Indian data from three National Family and Health Surveys (1992/93 (NFHS-1), 1998/99 (NFHS-2) and 2005/06 (NFHS-3)), examining children's height-for-age, an anthropometric measure that can indicate growth retardation and cumulative growth deficits, suggestive of long-term malnutrition. The sampling design involved an urban/rural stratification with one or two stages of clustering prior to the selection of households. In addition to providing variance estimates for simple inequality indices, sub-group decomposition measures and sub-group shares of overall inequality, based on the urban/rural split, we undertake tests of equality of these measures across the three surveys. A brief examination of gender differences in inequality is also provided. India has been experiencing rapid economic growth since the 1990s along with poverty reduction. However, this has been accompanied by rising economic inequality within urban areas and also between urban and rural sectors. But as Deaton and Drèze (2002, p.3744) rightly ask "*What about other types of social inequality, involving other dimensions of well-being.....?*". Clearly health is an important dimension of well-being and health inequality among children is a worthy issue to explore. Testing across surveys allows us to answer what has been happening to health inequality among children over time.

Health inequality has been studied by others, who convincingly argue that measures of health inequality are important in their own right and not just because of the possible correlation between income and health. Ram (2006) estimates cross-country inequality in life expectancy over the period 1980-2000 using three different inequality measures. Gini coefficients for health and education using Latin American

children's height data and data on years of schooling for women aged 22-30 are calculated by Sahn and Younger (2006). Height inequality among adults in Sub-Saharan Africa is the subject of Moradi and Baten's (2005) work. Pradhan et al. (2003) decompose world health inequality, as measured by height inequality among pre-school children, into within-country and between-country inequality using one of Theil's (1967) measures. Neither Ram (2006), Moradi and Baten (2005) nor Pradhan et al. (2003) report standard errors associated with their inequality measures and despite concluding differences in the numerical estimates, they do not undertake formal hypothesis testing. While Sahn and Younger (2006) report standard errors and undertake significance testing, they do not mention how they obtain their standard errors. Our work is directly relevant to such studies, as data are typically obtained from Demographic and Health Surveys, which employ stratified multi-stage cluster sampling.

As bootstrapping offers a viable alternative to our considered method, albeit less computationally friendly, we also provide standard error estimates and hypothesis test p-values from a bootstrap experiment that allows for the complex survey design.⁴ Undertaking inference for inequality measures via bootstrapping was first proposed by Mills and Zandvakili (1997), who examine the Gini coefficient along with the two Theil (1967) measures, and extended to all the Generalized Entropy and Atkinson indices by Biewen (2002). Since the publication of Mills and Zandavakili (1997), applied researchers in this area have begun using resampling methods to undertake statistical inference, such as for standard error estimates and p-values; e.g., Barrett et al. (2000), Gray et al. (2003a), Mills and Zandvakili (2004) and Davidson (2009). Some of these studies use complex survey data that is not accounted for when undertaking their bootstrap experiments, leading to inappropriate standard errors and p-values. Our bootstrap samples, on the other hand, are drawn allowing for the complex survey. Finally, we use the linearization method to calculate variance-covariance matrices (and subsequent test statistics) under a false iid assumption but allowing for sampling weights. Although this last case misinterprets the role weights

⁴ A novel feature of our work compared with that of Biewen and Jenkins (2006).

play under an iid assumption, it is a useful illustration of an error that might inadvertently arise in applied research.

The paper is organized as follows. Section 2 reviews the inequality measures considered in our work. Our results on the estimators of the variance-covariance matrices and subsequent inference are presented in section 3. Section 4 describes our setup for the bootstrap experiment for the application. The setting, data and results from the empirical illustration are detailed in section 5 and section 6 concludes. Relevant formulae are provided in an appendix.

II. Inequality measures

Many measures, or indices, of inequality can be obtained from a population, each with a different sensitivity to inequality in the upper or lower tail of the distribution. We examine indices that belong to the Generalized Entropy (GE) class, I_{GE}^α , and Atkinson (A) class, I_A^ϵ . Theil's (1967) two information indices are special cases: specifically, the Theil-1 index, I_{T1} , arises when $\alpha \rightarrow 1$, whereas the Theil-2⁵ index, I_{T2} , results by letting $\alpha \rightarrow 0$. In addition, setting $\alpha = 2$ gives half of the coefficient of variation squared. Accordingly, the parameter α determines the sensitivity of the index to inequality; changes in the underlying distribution's upper tail are more important for larger positive α while a greater response to inequality in the lower tail occurs when α becomes more negative. The parameter $\epsilon (\geq 0)$ for the Atkinson indices is often called the inequality aversion parameter (or preference for equality parameter), as larger values lead to greater sensitivity to inequality in the lower tail (or more aversion to inequality). Each member of the Atkinson family of inequality measures has an ordinally equivalent member of the GE family (but not *vice versa*).

For consistency, we adopt the basic setup of Biewen and Jenkins (2006), where each inequality measure is written in terms of population totals of the variable of interest (denoted as y) that captures some aspect of well-being

⁵ Also often termed the mean logarithmic deviation.

$$U_{\theta} = \sum_{h=1}^L \sum_{i=1}^{N_h} \sum_{j=1}^{M_i} (y_{hij})^{\theta}, \quad (1)$$

$$T_{\theta} = \sum_{h=1}^L \sum_{i=1}^{N_h} \sum_{j=1}^{M_i} (y_{hij})^{\theta} (\log y_{hij}) \quad (2)$$

summed over the stages of the complex survey sampling design, assumed to involve $h=1, \dots, L$ strata, $i=1, \dots, N_h$ clusters in stratum h and $j=1, \dots, M_i$ individuals in cluster i . The parameter θ is predetermined by which particular index is adopted, being either 0 or 1 for the T totals and 0, 1, α or $(1-\varepsilon)$ for the U totals. Note that U_0 is then the finite population size. It will not matter whether there are more stages of sampling beyond these, as the nonparametric variance estimator is computed from the quantities formed from the M_i ultimate clusters.⁶

Given (1) and (2), the population indices we examine⁷ are

$$I_{GE}^{\alpha} = (\alpha^2 - \alpha)^{-1} \left(U_0^{\alpha-1} U_1^{-\alpha} U_{\alpha} - 1 \right), \quad \alpha \in \mathfrak{R} \setminus \{0, 1\} \quad (3)$$

$$I_{T1} = T_1 U_1^{-1} - \log(U_1 U_0^{-1}), \quad \alpha \rightarrow 1 \quad (4)$$

$$I_{T2} = -T_0 U_0^{-1} + \log(U_1 U_0^{-1}), \quad \alpha \rightarrow 0 \quad (5)$$

$$I_A^{\varepsilon} = 1 - U_0^{-\varepsilon/(1-\varepsilon)} U_1^{-1} U_{1-\varepsilon}^{1/(1-\varepsilon)}, \quad \varepsilon \geq 0, \varepsilon \neq 1 \quad (6)$$

$$I_A^1 = 1 - U_0 U_1^{-1} \exp(T_0 U_0^{-1}), \quad \varepsilon \rightarrow 1. \quad (7)$$

Estimators of these indices, \hat{I}_{GE}^{α} , \hat{I}_{T1} , \hat{I}_{T2} , \hat{I}_A^{ε} and \hat{I}_A^1 , are generated by using the complex survey sample totals

$$\hat{U}_{\theta} = \sum_{h=1}^L \sum_{i=1}^{n_h} \sum_{j=1}^{m_i} w_{hij} (y_{hij})^{\theta}, \quad (8)$$

$$\hat{T}_{\theta} = \sum_{h=1}^L \sum_{i=1}^{n_h} \sum_{j=1}^{m_i} w_{hij} (y_{hij})^{\theta} (\log y_{hij}) \quad (9)$$

where n_h is the number of sampled first stage clusters and m_i is the number of sampled units in cluster i . As the complex survey design results in units with (usually) different probabilities of being sampled, the weight, w_{hij} , is included to

⁶ See, for example, Cochran (1977) and Skinner et al. (1989).

⁷ See, for instance, Cowell (1989).

account for such differential sampling rates, in addition to any adjustments for nonresponse and inadequate frame coverage.

Each of these inequality measures is decomposable, in that total inequality of the population can be subdivided into a weighted average of the inequality within chosen sub-groups of the population (the “within” component) and the inequality between these sub-groups (the “between” component). Sub-groups are typically based on grouping together units that share a common feature, for instance: place of residence; gender; race; educational attainment of the household head.. Specifically, we suppose the population comprises G mutually exclusive and exhaustive sub-groups ($g=1, \dots, G$) with sub-group population totals

$${}_gU_\theta = \sum_{h=1}^L \sum_{i=1}^{N_h} \sum_{j=1}^{M_i} {}_gD_{hij}(y_{hij})^\theta, \quad (10)$$

$${}_gT_\theta = \sum_{h=1}^L \sum_{i=1}^{N_h} \sum_{j=1}^{M_i} {}_gD_{hij}(y_{hij})^\theta (\log y_{hij}) \quad (11)$$

where ${}_gD_{hij}$ is a dummy variable that is 1 when unit hij belongs to sub-group g , 0 otherwise. Then, the sub-group inequality measures are

$${}_gI_{GE}^\alpha = (\alpha^2 - \alpha)^{-1} \left(({}_gU_0^{\alpha-1})({}_gU_1^{-\alpha})({}_gU_\alpha) - 1 \right), \quad \alpha \in \mathfrak{R} \setminus \{0, 1\} \quad (12)$$

$${}_gI_{T1} = ({}_gT_1)({}_gU_1^{-1}) - \log \left(({}_gU_1)({}_gU_0^{-1}) \right), \quad \alpha \rightarrow 1 \quad (13)$$

$${}_gI_{T2} = -({}_gT_0)({}_gU_0^{-1}) + \log \left(({}_gU_1)({}_gU_0^{-1}) \right), \quad \alpha \rightarrow 0 \quad (14)$$

$${}_gI_A^\varepsilon = 1 - ({}_gU_0^{-\varepsilon/(1-\varepsilon)})({}_gU_1^{-1})({}_gU_{1-\varepsilon}^{1/(1-\varepsilon)}), \quad \varepsilon \geq 0, \varepsilon \neq 1 \quad (15)$$

$${}_gI_A^1 = 1 - ({}_gU_0)({}_gU_1^{-1}) \exp \left(({}_gT_0)({}_gU_0^{-1}) \right), \quad \varepsilon \rightarrow 1. \quad (16)$$

The decomposition is additive when the index (say I) satisfies the constraint

$$I = \sum_{g=1}^G \omega_g I_g + B = W + B \quad (17)$$

where for the g 'th sub-group ($g=1, \dots, G$): I_g is the inequality index and ω_g is the weight function that depends only on sub-group means and population sizes. The terms B and W are, respectively, the “between” and “within” components, with the sub-group inequality indices only entering via W . This decomposition is empirically

useful as then total inequality is the sum of a term, W , which considers the inequality within each population sub-group plus the term, B , which is the inequality arising when there is no inequality within each group. Studying inequality in such a manner provides information on possible causes of inequality that may prove useful to policy makers. GE indices are additively decomposable in this way, whereas A indices cannot be decomposed additively into inter- and intra-group components. However, defining B and W as before, for an arbitrary A index, I , we have the decomposition

$$(1 - I) = (1-B)(1-W) \quad (18)$$

or, equivalently, $I = W + B - W \times B$. Given these definitions, Table 1 provides the weights for the within component and the between measure for each of our considered inequality indices. Here on, we denote the within components for the indices as W_{GE}^α , W_{T1} , W_{T2} , W_A^ε and W_A^1 , and the between components as B_{GE}^α , B_{T1} , B_{T2} , B_A^ε and B_A^1 . Point estimates are obtained by replacing the population totals with their sample counterparts – we denote the estimators as \hat{W}_{GE}^α , \hat{W}_{T1} , \hat{W}_{T2} , \hat{W}_A^ε , \hat{W}_A^1 , \hat{B}_{GE}^α , \hat{B}_{T1} , \hat{B}_{T2} , \hat{B}_A^ε and \hat{B}_A^1 .

Empirically, interest also lies with the shares of these components to total inequality. Specifically: the contribution of the between component to overall inequality ($S_{GE,B}^\alpha$, $S_{T1,B}$, $S_{T2,B}$, $S_{A,B}^\varepsilon$ and $S_{A,B}^1$) where, generically, $S_B = B/I$; the share of the within component to overall inequality ($S_{GE,W}^\alpha$, $S_{T1,W}$, $S_{T2,W}$, $S_{A,W}^\varepsilon$ and $S_{A,W}^1$) where, generically, $S_W = W/I$; and the proportion of total inequality taken by the within-group component of sub-group g (${}_gS_{GE,W}^\alpha$, ${}_gS_{T1,W}$, ${}_gS_{T2,W}$, ${}_gS_{A,W}^\varepsilon$ and ${}_gS_{A,W}^1$) where, generically, ${}_gS_W = {}_gW/I$ with $W = \sum_{g=1}^G {}_gW$. Estimators of

these shares, denoted with a circumflex, are formed using the relevant sample counterparts. Note that $S_B + S_W = 1$ for the GE indices, but not for the A measures. Despite this shortcoming for the latter indices, the share information still provides guidance on how inequality is changing from, for example, one survey to another.

TABLE 1

GE and A indices: between component and weights for the within component

<i>Index</i>	<i>Weight</i>	<i>Between</i>
	ω_g	B
I_{GE}^α	$\left(\frac{g U_0}{U_0}\right)^{1-\alpha} \left(\frac{g U_1}{U_1}\right)^\alpha$	$(\alpha^2 - \alpha)^{-1} \left[\sum_{g=1}^G \left(\frac{g U_0}{U_0}\right)^{1-\alpha} \left(\frac{g U_1}{U_1}\right)^\alpha - 1 \right]$
I_{T1}	$\left(\frac{g U_1}{U_1}\right)$	$\sum_{g=1}^G \left(\frac{g U_1}{U_1}\right) \log\left(\frac{g U_1 / U_1}{g U_0 / U_0}\right)$
I_{T2}	$\left(\frac{g U_0}{U_0}\right)$	$\sum_{g=1}^G \left(\frac{g U_0}{U_0}\right) \log\left(\frac{g U_0 / U_0}{g U_1 / U_1}\right)$
I_A^ε	$\left(\frac{g U_1}{U_1}\right)$	$1 - \left[\sum_{g=1}^G \left(\frac{\left(\frac{g U_{1-\varepsilon}}{U_{1-\varepsilon}}\right)^{1/(1-\varepsilon)}}{\left(\frac{g U_0}{U_0}\right)^{\varepsilon/(1-\varepsilon)}} \right) \right]^{-1}$
I_A^1	$\left(\frac{g U_1}{U_1}\right)$	$1 - (\exp(T_0 / U_0)) \left[\sum_{g=1}^G \left(\frac{g U_0}{U_0}\right) \exp(g T_0 / g U_0) \right]^{-1}$

Note: The table provides the weights used to form the within component for each index; i.e.,

$$W = \sum_{g=1}^G \omega_g I_g, \text{ where } I_g \text{ is the } g\text{'th sub-group's inequality index (} g=1, \dots, G\text{).}$$

We now turn to obtaining nonparametric estimators of the variances of the inequality statistics, along with an estimator of the variance-covariance matrix of any linear combination of two or more of these statistics, needed for hypothesis tests.

III. Inference using linearization estimators of variance-covariance matrices

Having estimated the sample inequality measures, we now turn to estimating sampling variability along with undertaking hypothesis tests involving two or more inequality measures, taking account of the complex survey design. Questions might include: Are sub-group inequality indices equal? Has inequality changed from one survey to another? Do the sub-groups have equal within shares? Are the shares equal across two or more surveys? To address such questions, we propose use of a linearization estimator, formed via a first-order Taylor series approximation, of the

relevant variance-covariance matrix when forming an appropriate Wald statistic. This method, which straightforwardly accommodates the complex survey design, avoids complicated covariance calculations that often arise with the δ -method, formed from the same Taylor series approximation (see, e.g., Cowell, 1989; Schluter and Trede, 2002; Van de gaer et al., 1999; Bhattacharya, 2007).

To be general, let $\Omega = [\vartheta_1, \dots, \vartheta_K]'$ be a K -dimensional vector of inequality quantities; e.g., Ω might contain sub-group inequality measures for a survey or consist of inequality indices for several surveys. We consider testing the null hypothesis $H_0: R\Omega = r$ against a two-sided alternative hypothesis using a Wald statistic, where R is a nonstochastic $q \times K$ matrix and r is a nonstochastic q -dimensional vector. Let $\hat{\Omega} = [\hat{\vartheta}_1, \dots, \hat{\vartheta}_K]'$ be the estimator of Ω formed using the relevant inequality estimators defined in section 2. Then, the usual Wald statistic is

$$WT = (R\hat{\Omega} - r)' [V\hat{a}r(R\hat{\Omega} - r)]^{-1} (R\hat{\Omega} - r), \quad (19)$$

where $V\hat{a}r(R\hat{\Omega} - r) = RV\hat{a}r(\hat{\Omega})R'$ is an estimator of the asymptotic variance-covariance matrix of $(R\hat{\Omega} - r)$. Obtaining the so-called linearization rule for $V\hat{a}r(\hat{\Omega})$ is our focus in this section; we denote this estimator as $V\hat{a}r_L(\hat{\Omega})$.⁸ Each inequality measure in Ω is a nonlinear function of relevant population totals U_θ , T_θ , ${}_gU_\theta$ and ${}_gT_\theta$ - we suppose there are P distinct population totals used in forming Ω , placed in a vector $\Phi = [\phi_1, \dots, \phi_P]'$; i.e., $\Omega = f(\Phi) = [f_1(\Phi), \dots, f_K(\Phi)]'$. To make this notation concrete, as an illustration, suppose we wish to test the equality of three sub-group Theil-1 indices; i.e., $H_0: {}_1I_{T1} = {}_2I_{T1} = {}_3I_{T1}$. Then, $K=3$, $q=2$,

$$P=9, \Omega = [{}_1I_{T1} \quad {}_2I_{T1} \quad {}_3I_{T1}]', \hat{\Omega} = [{}_1\hat{I}_{T1} \quad {}_2\hat{I}_{T1} \quad {}_3\hat{I}_{T1}]', R = \begin{bmatrix} 1 & -1 & 0 \\ 1 & 0 & -1 \end{bmatrix}, r = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \text{ and}$$

$$\Phi = [{}_1T_1 \quad {}_2T_1 \quad {}_3T_1 \quad {}_1U_1 \quad {}_2U_1 \quad {}_3U_1 \quad {}_1U_0 \quad {}_2U_0 \quad {}_3U_0]'$$

Let $\hat{\Phi}$ be a consistent estimator of Φ and assume $f(\cdot)$ is appropriately differentiable. A first-order Taylor series approximation is then:

⁸ Our approach extends that of Biewen and Jenkins (2006) to the vector case. It is based on Woodruff (1971), who extended some results due to Keyfitz (1957). The survey literature usually terms this the "linearization method" for variance estimation; see, e.g., Skinner et al., (1989, pp50-51 and p54).

$$\hat{\Omega} = f(\hat{\Phi}) \cong f(\Phi) + \sum_{p=1}^P f^P(\Phi)(\hat{\phi}_p - \phi_p),$$

where $f^P(\Phi) = \partial f(\Phi) / \partial \phi_p$ is a K -dimensional vector of partial derivatives. The asymptotic variance-covariance matrix of $\hat{\Omega}$ is then approximated by

$$\text{Var}\left(\sum_{p=1}^P f^P(\Phi)\hat{\phi}_p\right) = \text{Var}(F(\Phi)\hat{\Phi}) \quad (20)$$

where $F(\Phi)$ is the $K \times P$ derivative matrix. Noting that $\text{Var}(F(\Phi)\hat{\Phi}) = F(\Phi)\text{Var}(\hat{\Phi})F(\Phi)'$ leads to the usual δ -method estimator that requires evaluation of the $P \times P$ variance-covariance matrix of $\hat{\Phi}$, which can be difficult, especially with the complex survey design. The “linearization method” avoids generating this matrix by rewriting (20) in an alternative manner. Specifically, as $\Phi = \sum_{h=1}^L \sum_{i=1}^{N_h} \sum_{j=1}^{M_i} t_{hij}$, where t_{hij} is

a P -dimensional vector with p 'th element $t_{p,hij}$, we have⁹

$$\begin{aligned} \text{Var}\left(\sum_{p=1}^P f^P(\Phi)\hat{\phi}_p\right) &= \text{Var}\left(\sum_{p=1}^P f^P(\Phi) \sum_{h=1}^L \sum_{i=1}^{n_h} \sum_{j=1}^{m_i} w_{hij} t_{p,hij}\right) \\ &= \text{Var}\left(\sum_{h=1}^L \sum_{i=1}^{n_h} \sum_{j=1}^{m_i} w_{hij} \left[\sum_{p=1}^P f^P(\Phi) t_{p,hij}\right]\right) = \text{Var}\left(\sum_{h=1}^L \sum_{i=1}^{n_h} \left[\sum_{j=1}^{m_i} w_{hij} \gamma_{hij}\right]\right) \end{aligned} \quad (21)$$

where $\gamma_{hij} = \sum_{p=1}^P f^P(\Phi) t_{p,hij}$. Assuming¹⁰: (i) the ultimate clusters selected within

strata are independent; (ii) the n_h sampled units within each stratum h are selected with replacement;¹¹ and (iii) $n_h \geq 2$, we obtain the linearization estimator

⁹ For our illustration, $t_{hij} = [{}_1D_{hij}(y_{hij}) {}_2D_{hij}(y_{hij})(\log y_{hij}) {}_3D_{hij}(y_{hij})(\log y_{hij}) {}_1D_{hij}(y_{hij}) {}_2D_{hij}(y_{hij}) {}_1D_{hij}(y_{hij}) {}_1D_{hij} {}_2D_{hij} {}_3D_{hij}]'$.

¹⁰ See, for example, Skinner et al. (1989, p47).

¹¹ This assumption is usually always violated with surveys, in which case the formula generally leads to overestimation. An alternative assumption is that the n_h ultimate clusters within stratum h form a simple random sample without replacement from the stratum, $h=1, \dots, L$; see, e.g., Kalton (1977). This results in adding a finite population correction involving the factor (n_h/N_h) . The correction adds little when this factor is small, as is often the case.

$$\begin{aligned}
\text{V}\hat{\text{a}}\text{r}_L(\hat{\Omega}) &= \text{V}\hat{\text{a}}\text{r}\left(\sum_{h=1}^L \sum_{i=1}^{n_h} \left[\sum_{j=1}^{m_i} w_{hij} \hat{\gamma}_{hij} \right]\right) \\
&= \sum_{h=1}^L n_h \text{V}\hat{\text{a}}\text{r}\left(\sum_{j=1}^{m_i} w_{hij} \hat{\gamma}_{hij}\right) \tag{22}
\end{aligned}$$

where $\hat{\gamma}_{hij}$ is γ_{hij} with $\hat{\Phi}$ replacing Φ in the derivatives; we denote the k 'th element of $\hat{\gamma}_{hij}$ as $\hat{\gamma}_{k,hij}$, $k=1, \dots, K$. Switching the summation order has reduced the problem to one of obtaining a variance-covariance matrix for a total; e.g., Cochran (1977) and Skinner et al. (1989). Applying standard formulae, the linearization method estimator of the variance-covariance matrix is then

$$\begin{aligned}
\text{V}\hat{\text{a}}\text{r}_L(\hat{\Omega}) &= \\
&\sum_{h=1}^L \frac{n_h}{(n_h - 1)} \sum_{i=1}^{n_h} \left[\left(\sum_{j=1}^{m_i} w_{hij} \hat{\gamma}_{hij} - \frac{\sum_{i=1}^{n_h} \sum_{j=1}^{m_i} w_{hij} \hat{\gamma}_{hij}}{n_h} \right) \left(\sum_{j=1}^{m_i} w_{hij} \hat{\gamma}_{hij} - \frac{\sum_{i=1}^{n_h} \sum_{j=1}^{m_i} w_{hij} \hat{\gamma}_{hij}}{n_h} \right)' \right] \tag{23}
\end{aligned}$$

The elements of this matrix provide variance estimators for each individual inequality measure on the diagonal and covariance estimators on the off-diagonals. Standard software packages can easily generate this estimated variance-covariance matrix. For instance, with Stata (StataCorp., 2005) after defining the survey's strata, ultimate cluster and weight variables, along with generating each element of the vector $\hat{\gamma}_{hij}$, use of the *svy: total* command generates the matrix (23). To enable each of the inequality measures reported in section 2 to be part of any hypothesis test, in addition to providing a linearization method standard error, we need formulae to generate a corresponding $\hat{\gamma}_{k,hij}$. We provide these in the appendix.¹² Note that we do not provide explicit formulae for generating the linearization method variances for the

¹² In the appendix, for consistency we also report the equations for the full sample inequality indices obtained by Biewens and Jenkins (2006). The rest, to our knowledge, are new.

estimated between component shares for members of the GE family because they are the same as for the within components. This is not so for Atkinson indices.

A substantial body of research exists on the asymptotic properties of complex survey estimators of totals, smooth/nonsmooth linear/nonlinear functions of totals, and the corresponding linearization variance/covariance estimators: e.g., Krewski and Rao (1981) and Rao and Wu (1988). Most assume with replacement selection of the primary sampling units, in which case Krewski and Rao's (1981) analysis establishes consistency of the total estimators and smooth nonlinear functions of them, as used for our inequality measures. Asymptotic normality is also shown, along with consistency of variance estimators using the linearization method. Key is that there are no isolated, influential, values in the clusters. Applying these results to without replacement sampling typically follows directly when the sampling fraction $f_h = n_h/N_h$ is small, as is usual with the surveys used to generate inequality measures. Further, when the analysis assumes that it is the total number of clusters that goes to infinity (with fixed and finite secondary stage units and fixed number of strata), sampling with or without replacement has no effect on the asymptotic results; e.g., Bhattacharya (2005, 2007). Accordingly, we assume that our inequality estimators are consistent and asymptotically normal with the linearization method variance estimator also consistent. So, WT using $\text{Var}_L(\hat{\Omega})$, denoted as WT_L , is approximately χ_q^2 under its null hypothesis.

Although unlikely to arise with most socio-economic datasets, inverting the variance-covariance matrix $\text{Var}_L(\hat{\Omega})$ may be an issue when the number of sampled clusters relative to the number of strata (i.e., $\sum n_h - L$) is small compared to K , the dimension of $\hat{\Omega}$.¹³ One way to proceed is to use the estimated variance-covariance matrix under simple random sampling, then adjust for the design effects of complex sampling (see, e.g., Skinner et al., 1989, pp 90-103). The adjusted Wald statistic is

$$\text{WT}_{L,\text{adj}} = (\mathbf{R}\hat{\Omega} - \mathbf{r}) \left[\mathbf{R} \text{Var}_{\text{SRS}}(\hat{\Omega}) \mathbf{R}' \right]^{-1} (\mathbf{R}\hat{\Omega} - \mathbf{r}) / (\text{tr}(\hat{\Sigma})/e), \quad (24)$$

¹³ The effective number of degrees of freedom (under some assumptions) lies between the smallest of the (n_h-1) 's and $\sum n_h - L$, but is usually assumed to be the latter for the survey.

where $\hat{V}\hat{a}_{\text{SRS}}(\hat{\Omega})$ is the estimated variance-covariance matrix under a simple random sampling assumption, $e = (\text{tr}(\hat{\Sigma}))^2 / \text{tr}(\hat{\Sigma}^2)$ and

$$\hat{\Sigma} = (\text{R}\hat{V}\hat{a}_{\text{SRS}}(\hat{\Omega})\text{R}')^{-1}(\text{R}\hat{V}\hat{a}_{\text{L}}(\hat{\Omega})\text{R}') \quad (25)$$

is the design effect measure. The asymptotic null distribution of $\text{WT}_{\text{L,adj}}$ is approximated by that of a χ_e^2 variate.

To end this section, we comment on the applicability of our results to two other sampling schemes: simple random sampling (SRS), where every unit has an identical chance of being selected into the sample; and the frequently studied iid framework where, for the i 'th unit, the variable of well-being, y_i , and weight, w_i , are viewed as iid draws from a population (y, w) . Nicely, our results are easily modified to handle both cases. Under SRS, a self-weighting design where n units are selected (with replacement) from a finite population of N , the formulae in the appendix apply with the sampling weights either ignored or simply set to 1 for all units, and the summation is over $i=1, \dots, n$ rather than over the various stages of the complex survey design. The iid framework regards the inequality measure as a function of population moments rather than population totals, with the treatment of the weights being different than under a complex survey design.¹⁴ Despite these disparities, the numerical estimates of linearization variance-covariance matrices can be obtained using the formulae in the appendix with the summation being over $i=1, \dots, n$ rather than over the stages of the complex survey design.

IV. Inference using bootstrapping

Maintaining the general notation from the previous section, we first outline how we obtained a bootstrap variance-covariance estimator for the nonlinear estimator $\hat{\Omega} = f(\hat{\Phi})$; we denote this estimator as $\hat{V}\hat{a}_{\text{BT}}(\hat{\Omega})$.¹⁵ The method, which involves the following steps for each bootstrap sample, guarantees that the replicate sample has the same sampling design as the parent sample. The steps, V1 through V5, are:

¹⁴ For example, Cowell (1989), Van de gaer et al. (1999), Biewen and Jenkins (2006).

¹⁵ Our outlined approach is commonly termed the rescaling bootstrap (e.g., Rao and Wu, 1988).

V1. Draw a simple random sample of n_h clusters with replacement from the clusters within stratum h independently for each stratum ($h=1, \dots, L$).¹⁶

V2. When a cluster is selected into the bootstrap replicate, all secondary and successive units from the selected cluster are retained, along with their corresponding sampling weights.

V3. Let r_{hi}^b ($0 \leq r_{hi}^b \leq n_h$) be the number of times that cluster i from stratum j is included in bootstrap replicate b . The bootstrap sampling weight is then

$$w_{hij}^b = w_{hij} r_{hi}^b \quad (26)$$

so that $w_{hij}^b = 0$ if cluster i is not selected in the b 'th bootstrap sample.¹⁷

V4. Let $\hat{\Phi}^b$ be the estimated Φ using the p 'th bootstrap totals $\hat{\phi}_p^b$ formed from replicate b . Specifically, with $\theta = 0, 1$ or α , depending on the term of interest, $\hat{\phi}_p^b$ will be one of the following totals:

$$\hat{U}_\theta^b = \sum_{h=1}^L \sum_{i=1}^{n_h} \sum_{j=1}^{m_i} w_{hij}^b (y_{hij})^\theta,$$

$$\hat{T}_\theta^b = \sum_{h=1}^L \sum_{i=1}^{n_h} \sum_{j=1}^{m_i} w_{hij}^b (y_{hij})^\theta (\log y_{hij}),$$

$${}_g \hat{U}_\theta^b = \sum_{h=1}^L \sum_{i=1}^{n_h} \sum_{j=1}^{m_i} w_{hij}^b ({}_g D_{hij})(y_{hij})^\theta,$$

$${}_g \hat{T}_\theta^b = \sum_{h=1}^L \sum_{i=1}^{n_h} \sum_{j=1}^{m_i} w_{hij}^b ({}_g D_{hij})(y_{hij})^\theta (\log y_{hij}).$$

Then form the bootstrap copy of Ω , $\hat{\Omega}^b = f(\hat{\Phi}^b)$.

V5. Repeat steps V1-V4 B_V times to give B_V bootstrap estimators of Ω , $\hat{\Omega}^1, \dots, \hat{\Omega}^{B_V}$, and compute the bootstrap estimator of the variance-covariance matrix:

¹⁶ Undertaking the resampling with replacement simplifies the procedure and should not be an issue with most surveys used to generate inequality measures. Although the number of clusters to be resampled is often chosen to be (n_h-1) to ensure unbiased estimation (at least asymptotically), it is computationally easier with Stata to select n_h clusters from each strata. The effect of this is minimal in our case given the large number of clusters in the surveys.

¹⁷ Modifying the sampling weight by $(n_h/(n_h-1))$ occurs when (n_h-1) clusters are drawn – see footnote 16.

$$\text{V}\hat{\text{a}}\text{r}_{\text{BT}}(\hat{\Omega}) = \frac{1}{B_V - 1} \sum_{b=1}^{B_V} (\hat{\Omega}^b - \hat{\Omega})(\hat{\Omega}^b - \hat{\Omega})' . \quad (27)$$

We set $B_V = 200$. Aside from providing standard errors, this rule is used to form a bootstrapped sample value of any Wald statistics for inference; specifically, we form $\text{WT}_{\text{BT}} = (\mathbf{R}\hat{\Omega} - \mathbf{r})' [\mathbf{R}\text{V}\hat{\text{a}}\text{r}_{\text{BT}}(\hat{\Omega})\mathbf{R}']^{-1} (\mathbf{R}\hat{\Omega} - \mathbf{r})$. Turning to bootstrapping p-values, we undertake a double bootstrap with the following steps.

W1.Undertake steps V1 to V4 above to give the first ($b=1$) bootstrap sample with its estimator of Ω , $\hat{\Omega}^1 = f(\hat{\Phi}^1)$.

W2.Treating this first bootstrap replicate as the parent sample, repeat steps V1 to V4 again to give B_V estimates of Ω , $\hat{\Omega}_1^1, \dots, \hat{\Omega}_{B_V}^1$. Form the associated estimator of the variance-covariance matrix for this first bootstrap replicate sample:

$$\text{V}\hat{\text{a}}\text{r}_{\text{BT}}(\hat{\Omega}^1) = \frac{1}{B_V - 1} \sum_{b=1}^{B_V} (\hat{\Omega}_b^1 - \hat{\Omega}^1)(\hat{\Omega}_b^1 - \hat{\Omega}^1)' . \quad (28)$$

W3.Form the associated bootstrap Wald statistic:

$$\text{WT}_{\text{BT}}^1 = (\hat{\Omega}^1 - \hat{\Omega})' \mathbf{R}' [\mathbf{R}\text{V}\hat{\text{a}}\text{r}_{\text{BT}}(\hat{\Omega}^1)\mathbf{R}']^{-1} \mathbf{R}(\hat{\Omega}^1 - \hat{\Omega}), \quad (29)$$

where, recall, $\hat{\Omega}$ is the original sample's estimate of Ω and is used to so-call centre the statistic because our data may not have been drawn from a population that satisfies H_0 ; see, e.g., Hall and Wilson (1991).

W4.Repeat steps W1 through W3 B_W times to obtain B_W values of

WT: $\text{WT}_{\text{BT}}^1, \dots, \text{WT}_{\text{BT}}^{B_W}$. We choose $B_W = 99$ to correspond with a nominal 10% or

5% level.¹⁸ The bootstrapped p-value is $p = \left(\sum_{b=1}^{B_W} \mathbf{I}(\text{WT}_{\text{BT}}^b > \text{WT}_{\text{BT}}) + 1 \right) / (B_W + 1)$.

¹⁸ Given a nominal level for the test of α_w , a choice of B_W that leads to $\alpha_w(B_W+1)$ being an integer results in an exact Monte Carlo test when the statistic is pivotal; Dufour and Kiviet (1998). For a nonpivotal statistic (as is ours) it is not necessary to choose B_W in such a way, but, as advocated by (for example) Davidson and MacKinnon (2000), it would seem reasonable to still follow such a practice.

V. Empirical example: height inequality among Indian children

Here we apply some of our results to study health inequality among Indian children, based on the anthropometric measure height. As a detailed analysis of health inequality among Indian children is beyond our scope, our example illustrates, using a policy-relevant problem, the usefulness of statistical testing involving the simple inequality indices or their decompositions, as well as the effectiveness of our proposed methods compared to the computationally burdensome bootstrapping procedure. For space reasons, we only report results using the Theil-1 measures.¹⁹

We provide standard errors for simple inequality indices and sub-group decomposition measures based on an urban/rural split, as well as undertake tests for equality of these measures across two or three surveys, details of which are provided in the next subsection. Whether rural and urban regions differ in health inequality, as represented by children's height inequality, is of interest given the strong evidence of varying economic inequality across these regions (e.g., Deaton and Drèze, 2002). We also examine gender differences in height inequality, of concern given the debate on whether girls and boys are equally well cared for due to the preference for sons, particularly in rural districts. This predilection, formed from social, cultural, economic and religious desires and norms, suggests that there may be health and nutritional discrimination against girls,²⁰ which may show up in health inequality measures. In particular, as argued by, for instance, Moradi and Baten (2005), poorer households may reduce resources allocated to girls while maintaining those for boys in leaner times, which may show up as gender differences in health inequality.

For each case, we compare the linearization outcomes with those from two other scenarios: (i) assuming (incorrectly) that the standardized heights and the sample weights are iid draws from a common population; and (ii) from using the bootstrap procedure described in section 4, designed to account for the complex survey design.

¹⁹ Results for other measures are available on request.

²⁰ For instance, Kadi et al. (1996) and Tarozzi and Mahajan (2007) report that girls are more nutritionally deprived compared to boys. In contrast, Griffiths et al. (2002) and Marcoux (2002), among others, find little evidence of gender differentials in food consumption.

Survey design and data characteristics

Our data come from the three Indian National Family Health Surveys (NFHS), conducted under the agency of the International Institute for Population Sciences (IIPS): NFHS-1 (1992/93), NFHS-2 (1998/99) and (2005/06). All states and the National Capital Territory of Delhi²¹ are included, providing indicators on family welfare, maternal and child health, and nutrition.²² Due to coverage differences between surveys, we only include children (i) whose mothers were interviewed with the Women's Questionnaire, (ii) who are less than three years of age and (iii) who lived in states other than Sikkim, Andhra Pradesh, Himachal Pradesh, Madhya Pradesh, Tamil Nadu and West Bengal. This resulted in 20,410 children from NFHS-1, 18,520 children from NFHS-2 and 18,146 children from NFHS-3.

Stratified multi-stage cluster sampling generated the data with the design being roughly similar for each survey. We sketch out the key stages from NFHS-3; see IIPS (2007a,b; 2000; 1995) for full details. Each state was sampled separately with urban and rural areas forming the first stage strata. Two phases of cluster sampling came next for rural regions: random selection of villages followed by households. A three-stage procedure was adopted for urban areas: selection of wards followed by census enumeration blocks followed by households. On average, 30 households were targeted for interviewing from each village or census enumeration block. This sampling scheme results in the number of clusters far exceeding the number of strata so that testing should not suffer from a shortage of sampling design degrees of freedom.²³ The survey method also ensured self-weighting at the domain level (i.e., the urban and rural areas of each state) so that each child in the same domain has a common sampling weight (the inverse of the probability of selection).

Prior to estimating height inequalities, we account for natural/biological median height differences of children across gender and age by converting the individual heights into percent-of-median, which is simply the height of an individual child relative to the median height of comparable children in a reference population,

²¹ Covering more than 99 percent of India's population; other union territories were not surveyed.

²² See International Institute for Population Sciences (2007a, b; 2000; 1995) for full details.

²³ Specifically, there are 559 clusters for NFHS-1, 549 for NFHS-2 and 2719 for NFHS-3.

expressed as a percentage.²⁴ The current reference population group is the World Health Organization's (WHO) Child Growth Standards (WHO Multicentre Growth Reference Study Group, 2006; WHO, 2006), formed from a multiethnic sample of healthy children; we refer to this as the WHO-MGRS standard. Specifically, let h_{iga} be the height of the i th Indian child of gender g and aged month a ($g=1$ when the child is a boy, 2 when the child is a girl; $a=0, \dots, 35$). Let Md_{ga} be the median from the WHO-MGRS standard for a child of gender g and age a in months. The percent-of-median is $P_{iga} = 100(h_{iga}/Md_{ga})$; when gender and age is not of issue, we denote this as P_i . For example, if a 6-month old boy's height is 65.3cms then his $P_i=96.6\%$ as the WHO-MGRS median height is 67.6236cms for a 6-month old boy.

Using percent-of-median does not account for the natural variability in height, normal for healthy children, around the median height, that differs across and within age and gender groups. We believe this is likely not an issue when examining total inequality over surveys, as it seems reasonable to assume that the natural inequality in our Indian children is fairly stable over the time frame of our study and across regions. However, care is needed when considering gender differences in inequality, as natural inequality varies significantly across boys and girls, dependent on age. For this case, we estimate the natural inequality by taking appropriate draws of children's height from the WHO-MGRS reference growth curves. We now turn to our results.

Overall inequality

Table 2 provides the estimated inequality indices, I_{T1} , along with standard errors for the three surveys. For each case, we report three standard error estimates based on: our proposed linearization method, a (false) iid assumption with weights and the bootstrap approach that accounts for the complex survey sample design; these are denoted as se_L , se_{IID} and se_{BT} respectively. Irrespective of index, we see that inequality has declined over the three surveys, with the change between NFHS-2 and NFHS-3 being far more than the decline that occurred between NFHS-2 and NFHS-1. That the standard errors allowing for the complex survey design are larger than those

²⁴ See, e.g., Gershwin et al. (2000, pp. 7-8). Studies that use percent-of-median include Prudhon et al. (1996) and Zainah et al. (2001).

under the false iid assumption (by, approximately, 9% to 24%) highlights the importance of taking account of the design when estimating standard errors, the heterogeneity between and within clusters increasing the variance from the iid case. Typically, the linearization and bootstrap standard errors are in close agreement. Given the ease of forming the linearization standard errors compared to coding the bootstrap, these outcomes support using the former approach.

How do these inequality outcomes compare with those from a sample of “healthy” children? To examine this, we drew random samples of children’s height from the distributions that generated the WHO MGRS growth curves, with samples constructed to match the age/gender structure of each NFHS survey. Drawing samples in this way ensures that we allow for the differences in natural inequality across age and gender. The corresponding NFHS-3 Theil-1 index is 6.169E-04 with that for the other surveys being minimally different. The inequality for our Indian children is over four times that of the natural inequality of healthy children, which provides some indication of the disparities in health of Indian children and the relevance of examining whether inequality has changed.

TABLE 2

Estimates of overall height inequality using Theil 1 and standard errors

	<i>NFHS-1</i> <i>(1992/93)</i>	<i>NFHS-2</i> <i>(1998/99)</i>	<i>NFHS-3</i> <i>(2005/06)</i>	<i>% change</i> <i>NFHS-2</i> <i>/NFHS-1</i>	<i>% change</i> <i>NFHS-3</i> <i>/NFHS-1</i>	<i>% change</i> <i>NFHS-3</i> <i>/NFHS-2</i>
\hat{I}_{TI}	2.655E-03	2.609E-03	2.263E-03	-1.73%	-14.76%	-13.26%
se _L	4.635E-05	3.919E-05	3.734E-05			
se _{IID}	3.543E-05	3.405E-05	3.414E-05			
se _{BT}	4.772E-05	3.644E-05	3.944E-05			

Note: The standard error obtained via the linearization method (accounting for the complex survey) is denoted by se_L, that from the linearization approach assuming (falsely) that sampling is iid with weights as se_{IID} and that from the bootstrap method (allowing for the complex survey) by se_{BT}.

Returning back to the results in Table 2, having standard error estimates enables us to ask whether the changes in the indices are statistically significant. Outcomes from hypothesis tests to address this question are given in Table 3. We provide Wald

statistics, associated χ^2 and bootstrapped p-values from four tests: the first three test equality of indices across two surveys while the fourth test is for equality of the indices across the three surveys. Results are reported using the three different approaches to estimating variances. We denote Wald statistics by WT_L , W_{IID} and WT_{BT} and associated p-values by p_L , p_{IID} and p_{BT} . We assume that the samples across surveys are independent, allowing the variance of the difference in inequality indices to be the sum of the variances from each individual survey. This is a reasonable assumption, as the clusters are sampled independently from one survey to another.

TABLE 3

Testing whether overall inequality has changed across surveys using Theil 1

	<i>Hypothesis test</i>			
	<i>NFHS-1=NFHS-2</i>	<i>NFHS-2=NFHS-3</i>	<i>NFHS-1=NFHS-3</i>	<i>NFHS-1=NFHS-2=NFHS-3</i>
$WT_L (p_L)$	0.586 (0.444)	40.956 (0.000)	43.573 (0.000)	59.019 (0.000)
$W_{IID} (p_{IID})$	0.894 (0.344)	51.617 (0.000)	63.764 (0.000)	77.785 (0.000)
$WT_{BT} (p_{BT})$	0.599 (0.500)	41.625 (0.010)	40.281 (0.010)	55.803 (0.010)

Notes: The table reports Wald statistics and associated p-values for equality of indices. The subscripts are: L = complex survey linearization; IID = iid with weights linearization; BT= complex survey bootstrap.

Turning to the test outcomes, the change between NFHS-1 and NFHS-2 is not statistically significant, while that between NFHS-2 and NFHS-3, and NFHS-1 and NFHS-3 are statistically significant. The bootstrap and linearization methods are in close agreement, again supporting use of the linearization approach over the more computationally intensive bootstrap. Using the iid outcomes does not qualitatively change the results. Our findings perhaps suggest that the high income growth observed in India over the last 15 years or so has taken time to impact health inequality of children, as there is no significant change between NFHS-1 and NFHS-2 but there are strong declines in overall health inequality between NFHS-2 and NFHS-3. One possibility for the delayed impact could be habit persistence in food

consumption, as it may take time for any income increase to lead to consumption of more nutritious food.

Inequality by place of residence and gender

In contrast to the previous subsection that examined only overall inequality measures, this subsection contains results based on various sub-groups (sector- or gender-specific) and on the decomposition of the overall inequality by place of residence (rural or urban). Turning first to the urban sector, Table 4 provides estimated Theil 1 indices, their percentage changes and standard errors from the three methods, and Table 5 reports on tests across surveys.

TABLE 4

Urban sector: estimates of height inequality and standard errors for Theil 1

	<i>NFHS-1</i> <i>(1992/93)</i>	<i>NFHS-2</i> <i>(1998/99)</i>	<i>NFHS-3</i> <i>(2005/06)</i>	<i>% change</i> <i>NFHS-2</i> <i>/NFHS-1</i>	<i>% change</i> <i>NFHS-3</i> <i>/NFHS-1</i>	<i>% change</i> <i>NFHS-3</i> <i>/NFHS-2</i>
\hat{U}_{TI}	2.445E-03	2.140E-03	2.160E-03	-12.47%	-11.66%	0.93%
se_L	8.297E-05	6.444E-05	6.830E-05			
se_{IID}	6.458E-05	5.721E-05	6.268E-05			
se_{BT}	8.953E-05	6.6290E-05	6.430E-05			

Notes: The standard error obtained via the linearization method (accounting for the complex survey) is denoted by se_L , that from the linearization approach assuming (falsely) that sampling is iid with weights as se_{IID} and that from the bootstrap method (allowing for the complex survey) by se_{BT} . The subscript U in \hat{U}_{TI} indicates index estimates for the urban sector.

The results in Table 4 highlight the importance of sector-specific analysis, as the observed decline in health inequality of children occurred between NFHS-1 and NFHS-2 for this sector, with minimal change (indeed a nominal increase) in inequality between NFHS-2 and NFHS-3. This contrasts with the findings observed for overall inequality. In terms of standard errors, those under the false iid assumption are again smaller than those that account for the complex survey design, more so for NFHS-1 than for the other two surveys. We again observe that the bootstrap and linearization standard errors are in close agreement.

The hypothesis test outcomes provided in Table 5 support the statistical significance of the inequality change between the first two surveys but that there is no significant change in inequality between NFHS-2 and NFHS-3. The outcomes are qualitatively consistent across the methods used to generate variances.

TABLE 5

Testing whether height inequality of urban children has changed using Theil 1

	<i>Hypothesis test</i>			
	<i>NFHS-1= NFHS-2</i>	<i>NFHS-2= NFHS-3</i>	<i>NFHS-1= NFHS-3</i>	<i>NFHS-1= NFHS-2= NFHS-3</i>
WT _L (p _L)	8.428 (0.004)	0.047 (0.828)	7.011 (0.008)	9.654 (0.008)
WT _{IID} (p _{IID})	12.496 (0.000)	0.058 (0.810)	9.997 (0.002)	14.738 (0.001)
WT _{BT} (p _{BT})	7.495 (0.040)	0.049 (0.800)	6.664 (0.020)	8.594 (0.030)

Notes: The table reports Wald statistics and associated p-values for equality of urban indices. The subscripts are: L = complex survey linearization; IID = iid with weights linearization; BT= complex survey bootstrap.

Our findings on urban health inequality contrast with those on urban income inequality (at least for the 1990s); e.g., Deaton and Drèze (2002). This highlights the importance of exploring the impact of the economic reforms on not just income or per capita consumption expenditure inequality but also on other social inequality measures such as children’s health inequality. Aside from examining raw decomposition measures by region of residence, it is also of interest to ascertain the contribution of the between component to total inequality and how this share has changed across surveys. We report this information in Tables 6 and 7; Table 6 provides estimates of shares along with standard errors using the three variance methods and Table 7 details outcomes from hypothesis tests.

Around 1% of total health inequality arises from inequality between the rural and urban sectors, implying that the majority of inequality arises from within each sector. The standard error estimates indicate, however, that this small between component is significantly different from zero, even allowing for the lower precision from the complex survey design. This result contrasts with most income inequality findings,

where the between component often contributes more to the overall inequality than the within component.²⁵ It is useful to note, however, that our result is similar to that of Pradhan et al. (2003), who, using children’s height data, find that the within country health inequality component dominates world health inequality. Here, we find that this holds even within regions for an intra-country study. Whether this is a common finding for health inequality remains to be seen. At least at the 5% level, there is no evidence to suggest different between group shares across surveys when using the complex survey linearization or bootstrap variance estimates. However, rejection of this null of equality occurs under the false iid assumption in two of the four cases, highlighting, again, the importance of accounting for the survey design.

TABLE 6

Urban-rural between component shares and standard errors (%’s)

	<i>NFHS-1</i> <i>(1992/93)</i>	<i>NFHS-2</i> <i>(1998/99)</i>	<i>NFHS-3</i> <i>(2005/06)</i>
$\hat{S}_{T1,B}$	0.50%	1.04%	0.75%
se _L	0.19%	0.22%	0.18%
se _{IID}	0.11%	0.16%	0.15%
se _{BT}	0.23%	0.22%	0.17%

Note: The standard error obtained via the linearization method (accounting for the complex survey) is denoted by se_L, that from the linearization approach assuming (falsely) that sampling is iid with weights as se_{IID} and that from the bootstrap method (allowing for the complex survey) by se_{BT}.

The final set of results we provide are given in Table 8, based on gender-specific inequality. Health inequality indices for boys and girls are reported, along with outcomes from hypothesis tests that examine whether height inequality for girls is the same as that for boys. Examining the estimated values, inequality for girls is higher than for boys, with the two-sided hypothesis tests indicating that this difference is statistically significant, irrespective of method used to estimate the variances.²⁶

²⁵ An example of an exception is Gray et al. (2003b), who report a 1% between-group inequality share when comparing incomes of those born in Canada, immigrants who arrived before 1981, and immigrants who arrived after 1981.

²⁶ Interestingly, we see an example (NFHS-2 boys) where the linearization and bootstrap standard errors are marginally smaller than the iid standard error.

TABLE 7

Testing whether the Theil 1 between group (urban/rural) share has changed

	Hypothesis test			
	NFHS-1= NFHS-2	NFHS-2= NFHS-3	NFHS-1= NFHS-3	NFHS-1= NFHS-2= NFHS-3
	$S_{T1,B}$			
WT _L (p _L)	3.346 (0.067)	1.023 (0.312)	0.891 (0.345)	3.351 (0.187)
WT _{IID} (p _{IID})	7.328 (0.007)	1.688 (0.194)	1.691 (0.193)	7.465 (0.024)
WT _{BT} (p _{BT})	2.864 (0.100)	1.108 (0.290)	0.754 (0.320)	2.892 (0.240)

Note: The table reports Wald statistics and associated p-values for equality of urban/rural between shares. . The subscripts are: L = complex survey linearization; IID = iid with weights linearization; BT= complex survey bootstrap.

However, these results do not allow for the variation in natural inequality across gender. To estimate this, we drew random samples of healthy children's height from the distributions used to generate the WHO-MGRS growth curves, with the samples constructed to have the same age/gender structure as our NFHS samples. The boys and girls natural inequality Theil-1 estimates from these simulated samples of healthy children, denoted \hat{B}_{T1}^N and \hat{G}_{T1}^N respectively, are reported in the bottom part of Table 8, along with the differences $(\hat{B}_{T1} - \hat{B}_{T1}^N)$ and $(\hat{G}_{T1} - \hat{G}_{T1}^N)$, so-called adjusted inequalities, that estimate the inequality in children's height due to poor health and nutrition. The genetic natural inequality in the heights of healthy girls (as represented by percent-of-median) exceeds that of boys, mitigating much of the observed inequality differences between our Indian boys and girls when we do not account for natural inequality. Indeed, although not reported in the table, tests of the

TABLE 8

Boys and girls: height inequality and standard errors using Theil-1

	<i>NFHS-1</i> (1992/93)	<i>NFHS-2</i> (1998/99)	<i>NFHS-3</i> (2005/06)	% change <i>NFHS-2</i> / <i>NFHS-1</i>	% change <i>NFHS-3</i> / <i>NFHS-1</i>	% change <i>NFHS-3</i> / <i>NFHS-2</i>
$\hat{B}I_{T1}$	2.561E-03	2.472E-03	2.153E-03	-3.48%	-15.93%	-12.90%
se _L	5.655E-05	4.278E-05	4.595E-05			
se _{IID}	4.880E-05	4.326E-05	4.323E-05			
se _{BT}	5.696E-05	3.816E-05	4.784E-05			
$\hat{G}I_{T1}$	2.752E-03	2.758E-03	2.383E-03	0.22%	-13.41%	-13.60%
se _L	5.792E-05	5.956E-05	5.569E-05			
se _{IID}	5.140E-05	5.335E-05	5.363E-05			
se _{BT}	5.935E-05	5.867E-05	5.980E-05			
	boys=girls					
WT _L	8.061	17.984	11.152			
(p _L)	(0.005)	(0.000)	(0.001)			
WT _{IID}	7.311	17.333	11.168			
(p _{IID})	(0.007)	(0.000)	(0.001)			
WT _{BT}	8.248	18.911	9.883			
(p _{BT})	(0.010)	(0.010)	(0.010)			
	natural inequality					
$\hat{B}I_{T1}^N$	5.851E-04	5.788E-04	5.703E-04			
$\hat{G}I_{T1}^N$	6.673E-04	6.706E-04	6.673E-04			
	adjusted inequality					
$(\hat{B}I_{T1} - \hat{B}I_{T1}^N)$	1.976E-04	1.893E-04	1.583E-04			
$(\hat{G}I_{T1} - \hat{G}I_{T1}^N)$	2.085E-04	2.087E-04	1.716E-04			

Notes: The subscript L denotes use of the linearization method (accounting for the complex survey) to form variances, IID implies that the variances are obtained using the linearization approach under a false iid assumption with weights, and BT refers to the bootstrap method that accommodates the complex survey design. The subscript B and G in $\hat{B}I_{T1}$ and $\hat{G}I_{T1}$, and $\hat{B}I_{T1}^N$ and $\hat{G}I_{T1}^N$, indicate Theil-1 index estimates for boys and girls respectively.

hypothesis that $(\hat{I}_{T1}^B - \hat{I}_{T1}^N) = (\hat{I}_{T1}^G - \hat{I}_{T1}^N)$ confirm that we cannot reject equality of these adjusted inequality measures, suggesting that these samples do not support gender differences in health inequality, at least at the national level.

VI. Concluding comments

In this paper, we have considered undertaking inference on GE and Atkinson inequality indices when data are drawn under a complex survey design. Variance-covariance matrices are obtained via a linearization method rather than the δ -approach, so avoiding the calculation of often cumbersome covariance expressions. One of our key contributions is to obtain expressions that enable inference for the components of common decompositions of the inequality measures, including “between” and “within” elements and any subsequent share measures generated from these. A key benefit of using these expressions for inference is ease of coding in standard software packages (e.g., Stata), in contrast to the coding that must be undertaken to bootstrap variances and p-values.

Our illustrative application using height-for-age data on Indian children from three surveys highlights the importance of accounting for the stratified multi-stage cluster sampling design and the (typically) similar outcomes obtained using the linearization and bootstrap methods. This latter finding is particularly encouraging for applied researchers. Although it is unclear as to whether our findings can be broadly generalized, as the sampling design degrees of freedom for our data are relatively large, it is clear that the linearization approach to inference provides a user friendly way to undertake inference for inequality measures.

Acknowledgement

The authors thank Brennan Thompson and conference participants at the Canadian Economic Association meeting for helpful comments and suggestions. We are grateful to Macro International Inc. for supplying the data, and would like to extend special thanks to Glen Heller for his prompt and detailed responses to our data queries.

References

- Atkinson, A.B. (1970). 'On the measurement of inequality', *Journal of Economic Theory*, Vol. 2, pp. 244-263.
- Atkinson, A.B. (1997). 'Bringing income distribution in from the cold', *Economic Journal*, Vol. 107, pp. 297-321.
- Barrett, G.F., Crossley, T.F. and Worswick, C. (2000). 'Consumption and income inequality in Australia', *Economic Record*, Vol.76, pp. 116-138.
- Bhattacharya, D. (2005). 'Asymptotic inference from multi-stage samples', *Journal of Econometrics*, Vol. 126, pp. 145-171.
- Bhattacharya, D. (2007). 'Inference on inequality from household survey data', *Journal of Econometrics*, Vol. 137, pp. 674-707.
- Biewen, M. (2002). 'Bootstrap inference for inequality, mobility and poverty measurement', *Journal of Econometrics*, Vol. 108, pp. 317-342.
- Biewen, M. and Jenkins, S.P. (2006). 'Variance estimation for generalized entropy and Atkinson inequality indices: the complex survey data case', *Oxford Bulletin of Economics and Statistics*, Vol. 68, pp. 371-383.
- Binder, D.A. and Kovačević, M.S. (1995). 'Estimating some measures of income inequality from survey data: an application of the estimating equations approach', *Survey Methodology*, Vol. 21, pp. 137-145.
- Cochran, W.G. (1977). *Sampling Techniques*, 3rd Edition., Wiley, New York.
- Cowell, F.A. (1989). 'Sampling variance and decomposable inequality measures', *Journal of Econometrics*, Vol. 42, pp. 27-41.
- Davidson, R. (2009). 'Reliable inference for the Gini index', *Journal of Econometrics*, Vol. 150, pp. 30-40.
- Davidson, R. and Flachaire, E. (2007). 'Asymptotic and bootstrap inference for inequality and poverty measures', *Journal of Econometrics*, Vol. 141, pp. 141-166.
- Davidson, R. and MacKinnon, J.G. (2000). 'Bootstrap tests: how many bootstraps', *Econometric Reviews*, Vol. 19, pp. 55-68.
- Deaton, A. and Drèze, J. (2002). 'Poverty and inequality in India: a re-examination', *Economic and Political Weekly*, September, pp. 3729-3748.

- Dufour, J.-M. and Kiviet, J.F. (1998). 'Exact inference methods for first-order autoregressive distributed lag models', *Econometrica*, Vol. 66, pp. 79-104.
- Gershwin, M.E., German, J.B. and Keen, C.L. (2000). *Nutrition and Immunology: Principles and Practice*, Humana Press, New York.
- Giles, D.E.A. (2004). 'Calculating a standard error for the Gini Coefficient: some further results', *Oxford Bulletin of Economics and Statistics*, Vol. 66, pp. 425-433.
- Gray, D., Mills, J.A. and Zandvakili, S. (2003a). 'Statistical analysis of inequality with decompositions: the Canadian experience', *Empirical Economics*, Vol. 28, pp. 291-302.
- Gray, D., Mills, J.A. and Zandvakili, S. (2003b). 'Immigration, assimilation and inequality of income distribution in Canada', Working Paper, Department of Economics, University of Ottawa.
- Griffiths, P., Matthews, Z. and Hinde, A. (2002). 'Gender, family and the nutritional status of children in three contrasting states of India', *Social Science and Medicine*, Vol. 55, pp. 775-790.
- Hall, P. and Wilson, S.R. (1991). 'Two guidelines for bootstrap hypothesis testing', *Biometrics*, Vol. 47, pp. 757-762.
- International Institute for Population Sciences (IIPS). (1995). *National Family Health Survey (MCH and Family Planning), India 1992-93*. Bombay: IIPS.
- International Institute for Population Sciences (IIPS) and ORC Macro. (2000). *National Family Health Survey (NFHS-2), 1998-99: India*. Mumbai: IIPS.
- International Institute for Population Sciences (IIPS) and Macro International. (2007a). *National Family Health Survey (NFHS-3), 2005-06: India*, Vol. I. Mumbai: IIPS.
- International Institute for Population Sciences (IIPS) and Macro International. (2007b). *National Family Health Survey (NFHS-3), 2005-06: India*, Vol. II. Mumbai: IIPS.
- Kadi, P.B., Khateeb, J. and Patil, M.S. (1996). 'Development of rural children – a bias', *Indian Journal of Maternal and Child Health*, Vol. 7, pp. 24-27.
- Kalton, G. (1977). 'Practical methods for estimating survey sampling errors', *Bulletin of the International Statistical Institute*, Vol. 47, pp. 495-514.

- Keyfitz, N. (1957). 'Estimates of sampling variance where two units are selected from each stratum', *Journal of the American Statistical Association*, Vol. 52, pp. 503-510.
- Krewski, D. and Rao, J.N.K. (1981). 'Inference from stratified samples: properties of the linearization, jackknife and balanced repeated replication methods', *Annals of Statistics*, Vol. 9, pp. 1010-1019.
- Maasoumi, E. (1997). 'Empirical analysis of inequality and welfare', in H. Pesaran and P. Schmidt (eds.), *Handbook of Applied Econometrics: Microeconomics*, Vol. II, pp. 202-245, Blackwell, Malden, Mass.
- Marcoux, A. (2002). 'Sex differentials in undernutrition: a look at survey evidence', *Population and Development Review*, Vol. 28, pp. 275-284.
- Mills, J.A. and Zandvakili, S. (1997). 'Statistical inference via bootstrapping for measures of inequality', *Journal of Applied Econometrics*, Vol. 12, pp. 133-150.
- Mills, J.A. and Zandvakili, S. (2004). Analysis of gender based family income inequality in Canada, *Applied Economics Letters*, Vol. 11, pp. 469-472.
- Moradi, A. and Baten, J. (2005). 'Inequality in Sub-Saharan Africa: new data and new insights from anthropometric estimates', *World Development*, Vol. 33, pp. 1233-1265.
- Pradhan, M., Sahn, D.E. and Younger, S.D. (2003). 'Decomposing world health inequality', *Journal of Health Economics*, Vol. 22, pp. 271-293.
- Prudhon, C., Briend, A., Laurier, D., Golden, M.H. and Mary, J.Y. (1996). 'Comparison of weight- and height-based indices for assessing the risk of death in severely malnourished children', *American Journal of Epidemiology*, Vol. 144, pp. 116-123.
- Ram, R. (2006). 'State of the "life span revolution" between 1980 and 2000', *Journal of Development Economics*, Vol. 80, pp. 518 – 526.
- Rao, J.N.K. and Wu, C.F.J. (1988). 'Resampling inference with complex survey data', *Journal of the American Statistical Association*, Vol. 83, pp. 231-241.
- Sahn, D.E. and Younger, S.D. (2006). 'Changes in inequality and poverty in Latin America: looking beyond income to health and education', *Journal of Applied Economics*, Vol. IX, pp. 215-234.
- Schluter, C. and Trede, M. (2002). 'Statistical inference for inequality and poverty measurement with dependent data', *International Economic Review*, Vol. 43, pp. 493-508.

- Skinner, C.J., Holt, D. and Smith, T.M.F. (1989). *Analysis of Complex Surveys*, Wiley, Chichester.
- StataCorp. (2005). *Stata Statistical Software: Release 9.0.*, Stata Corporation, College Station, TX.
- Tarozzi, A. and Mahajan, A. (2007). 'Child nutrition in India in the nineties', *Economic Development and Cultural Change*, Vol. 55, pp. 441-486.
- Theil, H. (1967). *Economics and Information Theory*, North-Holland, Amsterdam.
- Van de gaer, D., Funnell, N. and McCarthy, T. (1999). 'Statistical inference for two measures of inequality when incomes are correlated', *Economics Letters*, Vol. 64, pp. 295-300.
- WHO, (2006). *WHO Child Growth Standards: Length/height-for-age, Weight-for-age, Weight-for-length, Weight-for-height and Body Mass Index-for-age: Methods & Development*, Department of Nutrition for Health and Development, World Health Organization.
- WHO Multicentre Growth Reference Study Group, (2006). 'WHO child growth standards based on length/height, weight and age', *Acta Paediatrica*, Vol. 95 (s450), pp. 76-85.
- Woodruff, R.S. (1971). 'A simple method for approximating the variance of a complicated estimate', *Journal of the American Statistical Association*, Vol. 66, pp. 411-414.
- Zainah, S.H., Ong, L.C., Sofiah, A., Poh, B.K. and Hussain, I.H.M.I. (2001). 'Determinants of linear growth in Malaysian children with cerebral palsy', *Journal of Paediatrics & Child Health*, Vol. 37, pp. 376-381.

Appendix A: Formulae for $\hat{\gamma}_{k,hij}$ ²⁷

A.1 Sample total indices

$$\begin{aligned}\hat{I}_{GE}^\alpha : \hat{\gamma}_{k,hij} &= \alpha^{-1} \hat{U}_\alpha \hat{U}_1^{-\alpha} \hat{U}_0^{\alpha-2} - (\alpha-1)^{-1} \hat{U}_\alpha \hat{U}_0^{\alpha-1} \hat{U}_1^{-\alpha-1} y_{hij} + (\alpha^2 - \alpha)^{-1} \hat{U}_0^{\alpha-1} \hat{U}_1^{-\alpha} y_{hij}^\alpha; \\ \hat{I}_{T1} : \hat{\gamma}_{k,hij} &= \hat{U}_1^{-1} y_{hij} \log y_{hij} + \hat{U}_0^{-1} - \hat{U}_1^{-1} (\hat{T}_1 / \hat{U}_1 + 1) y_{hij}; \\ \hat{I}_{T2} : \hat{\gamma}_{k,hij} &= -\hat{U}_0^{-1} \log y_{hij} + \hat{U}_1^{-1} y_{hij} + \hat{U}_0^{-1} (\hat{T}_0 / \hat{U}_0 - 1); \\ \hat{I}_A^\varepsilon : \hat{\gamma}_{k,hij} &= (\varepsilon / (1 - \varepsilon)) \hat{U}_0^{-1/(1-\varepsilon)} \hat{U}_1^{-1} \hat{U}_{1-\varepsilon}^{1/(1-\varepsilon)} + \hat{U}_0^{-\varepsilon/(1-\varepsilon)} \hat{U}_1^{-2} \hat{U}_{1-\varepsilon}^{1/(1-\varepsilon)} y_{hij} \\ &\quad - (1/(1-\varepsilon)) \hat{U}_1^{-1} (\hat{U}_{1-\varepsilon} / \hat{U}_0)^{\varepsilon/(1-\varepsilon)} y_{hij}^{1-\varepsilon}; \\ \hat{I}_A^1 : \hat{\gamma}_{k,hij} &= \hat{U}_0^{-1} (\hat{I}_A^1 - 1) \log y_{hij} + \hat{U}_0^{-1} (\hat{I}_A^1 - 1) (1 - \hat{T}_0 / \hat{U}_0) + \hat{U}_1^{-1} (1 - \hat{I}_A^1) y_{hij}.\end{aligned}$$

A.2 Sample sub-group indices

$$\begin{aligned}{}_g \hat{I}_{GE}^\alpha : \hat{\gamma}_{k,hij} &= \alpha^{-1} ({}_g \hat{U}_\alpha) ({}_g \hat{U}_1^{-\alpha}) ({}_g \hat{U}_0^{\alpha-2}) ({}_g D_{hij}) - (\alpha-1)^{-1} ({}_g \hat{U}_\alpha) ({}_g \hat{U}_0^{\alpha-1}) ({}_g \hat{U}_1^{-\alpha-1}) \times \\ &\quad ({}_g D_{hij}) y_{hij} + (\alpha^2 - \alpha)^{-1} ({}_g \hat{U}_0^{\alpha-1}) ({}_g \hat{U}_1^{-\alpha}) ({}_g D_{hij}) y_{hij}^\alpha; \\ {}_g \hat{I}_{T1} : \hat{\gamma}_{k,hij} &= ({}_g \hat{U}_1^{-1}) ({}_g D_{hij}) y_{hij} \log y_{hij} + ({}_g \hat{U}_0^{-1}) ({}_g D_{hij}) - ({}_g \hat{U}_1^{-1}) ({}_g \hat{T}_1 / {}_g \hat{U}_1 + 1) \times \\ &\quad ({}_g D_{hij}) y_{hij}; \\ {}_g \hat{I}_{T2} : \hat{\gamma}_{k,hij} &= -({}_g \hat{U}_0^{-1}) ({}_g D_{hij}) \log y_{hij} + ({}_g \hat{U}_1^{-1}) ({}_g D_{hij}) y_{hij} + ({}_g \hat{U}_0^{-1}) ({}_g \hat{T}_0 / {}_g \hat{U}_0 - 1) \times \\ &\quad ({}_g D_{hij}); \\ {}_g \hat{I}_A^\varepsilon : \hat{\gamma}_{k,hij} &= (\varepsilon / (1 - \varepsilon)) ({}_g \hat{U}_1^{-1}) ({}_g \hat{U}_{1-\varepsilon}) / ({}_g \hat{U}_0)^{1/(1-\varepsilon)} ({}_g D_{hij}) + ({}_g \hat{U}_0^{-\varepsilon/(1-\varepsilon)}) ({}_g \hat{U}_1^{-2}) \times \\ &\quad ({}_g \hat{U}_{1-\varepsilon}^{1/(1-\varepsilon)}) ({}_g D_{hij}) y_{hij} - (1/(1-\varepsilon)) ({}_g \hat{U}_1^{-1}) ({}_g \hat{U}_{1-\varepsilon}) / ({}_g \hat{U}_0)^{\varepsilon/(1-\varepsilon)} ({}_g D_{hij}) y_{hij}^{1-\varepsilon}; \\ {}_g \hat{I}_A^1 : \hat{\gamma}_{k,hij} &= ({}_g \hat{U}_0^{-1}) ({}_g \hat{I}_A^1 - 1) ({}_g D_{hij}) \log y_{hij} + ({}_g \hat{U}_0^{-1}) ({}_g \hat{I}_A^1 - 1) (1 - {}_g \hat{T}_0 / {}_g \hat{U}_0) ({}_g D_{hij}) \\ &\quad + ({}_g \hat{U}_1^{-1}) (1 - {}_g \hat{I}_A^1) ({}_g D_{hij}) y_{hij}.\end{aligned}$$

A.3 Sample within indices

$$\begin{aligned}\hat{W}_{GE}^\alpha : \hat{\gamma}_{k,hij} &= (\alpha^2 - \alpha)^{-1} \hat{U}_0^{\alpha-2} \hat{U}_1^{-\alpha} \left[\sum_{g=1}^G \left\{ (\alpha-1) ({}_g \hat{U}_1 / {}_g \hat{U}_0)^\alpha - \alpha ({}_g \hat{U}_1 / {}_g \hat{U}_0)^{\alpha-1} y_{hij} \right. \right. \\ &\quad \left. \left. + y_{hij}^\alpha \right] \hat{U}_0 ({}_g D_{hij}) + ({}_g \hat{U}_\alpha - ({}_g \hat{U}_0^{1-\alpha}) ({}_g \hat{U}_1^\alpha) \right] \left[\alpha (1 - (\hat{U}_0 / \hat{U}_1) y_{hij}) - 1 \right]; \\ \hat{W}_{T1} : \hat{\gamma}_{k,hij} &= -\hat{U}_1^{-2} y_{hij} \sum_{g=1}^G [{}_g \hat{T}_1 - ({}_g \hat{U}_1) \log ({}_g \hat{U}_1 / {}_g \hat{U}_0)] + \hat{U}_1^{-1} \sum_{g=1}^G ({}_g D_{hij}) [{}_g \hat{U}_1 / {}_g \hat{U}_0 \\ &\quad + y_{hij} \log y_{hij} - (1 + \log ({}_g \hat{U}_1 / {}_g \hat{U}_0)) y_{hij}];\end{aligned}$$

²⁷ The expressions given in subsections A.2 to A.7 are our contributions whereas those in A.1 are from Biewen and Jenkins (2006).

$$\begin{aligned}\hat{W}_{T2} : \hat{\gamma}_{k,hij} &= \hat{U}_0^{-2} \sum_{g=1}^G \left\{ \hat{U}_0 ({}_g D_{hij}) \left[({}_g \hat{U}_0 / {}_g \hat{U}_1) y_{hij} + \log({}_g \hat{U}_1 / {}_g \hat{U}_0) - 1 - \log y_{hij} \right] \right. \\ &\quad \left. + ({}_g \hat{T}_0) - ({}_g \hat{U}_0) \log({}_g \hat{U}_1 / {}_g \hat{U}_0) \right\}; \\ \hat{W}_A^\varepsilon : \hat{\gamma}_{k,hij} &= \hat{U}_1^{-1} \sum_{g=1}^G ({}_g \hat{U}_0^{-\varepsilon/(1-\varepsilon)}) ({}_g \hat{U}_{1-\varepsilon}^{1/(1-\varepsilon)}) \left[\hat{U}_1^{-1} y_{hij} + (\varepsilon/(1-\varepsilon)) ({}_g \hat{U}_0^{-1}) ({}_g D_{hij}) \right. \\ &\quad \left. - (1/(1-\varepsilon)) ({}_g \hat{U}_{1-\varepsilon}^{-1}) ({}_g D_{hij}) y_{hij}^{1-\varepsilon} \right]; \\ \hat{W}_A^1 : \hat{\gamma}_{k,hij} &= \hat{U}_1^{-1} \sum_{g=1}^G \exp({}_g \hat{T}_0 / ({}_g \hat{U}_0)) \left[({}_g \hat{U}_0 / \hat{U}_1) y_{hij} - ({}_g D_{hij}) (1 - {}_g \hat{T}_0 / {}_g \hat{U}_0 + \log y_{hij}) \right].\end{aligned}$$

A.4 Sample between indices

$$\begin{aligned}\hat{B}_{GE}^\alpha : \hat{\gamma}_{k,hij} &= (\alpha^2 - \alpha)^{-1} \hat{U}_0^{\alpha-2} \hat{U}_1^{-\alpha-1} \left\{ \sum_{g=1}^G \left\{ \hat{U}_0 \hat{U}_1 ({}_g D_{hij}) \left[\alpha ({}_g \hat{U}_0 / {}_g \hat{U}_1)^{1-\alpha} y_{hij} \right. \right. \right. \\ &\quad \left. \left. - (\alpha - 1) ({}_g \hat{U}_0 / {}_g \hat{U}_1)^{-\alpha} \right] + ({}_g \hat{U}_0^{1-\alpha}) ({}_g \hat{U}_1^\alpha) \left[(\alpha - 1) \hat{U}_1 - \alpha \hat{U}_0 \right] \right\} \}; \\ \hat{B}_{T1} : \hat{\gamma}_{k,hij} &= \hat{U}_0^{-1} + \hat{U}_1^{-1} \left\{ -y_{hij} + \sum_{g=1}^G ({}_g D_{hij}) \left[(1 + \log({}_g \hat{U}_1 / {}_g \hat{U}_0)) y_{hij} - {}_g \hat{U}_1 / {}_g \hat{U}_0 \right] \right\} \\ &\quad - \hat{U}_1^{-2} \sum_{g=1}^G ({}_g \hat{U}_1) \log({}_g \hat{U}_1 / {}_g \hat{U}_0) y_{hij}; \\ \hat{B}_{T2} : \hat{\gamma}_{k,hij} &= -\hat{U}_0^{-1} \sum_{g=1}^G ({}_g \hat{U}_0 / {}_g \hat{U}_1) ({}_g D_{hij}) y_{hij} + \hat{U}_0^{-1} \sum_{g=1}^G (1 + \log({}_g \hat{U}_0 / {}_g \hat{U}_1)) ({}_g D_{hij}) \\ &\quad + \hat{U}_1^{-1} y_{hij} - \hat{U}_0^{-1} \left(1 + \hat{U}_0^{-1} \sum_{g=1}^G ({}_g \hat{U}_0) \log({}_g \hat{U}_0 / {}_g \hat{U}_1) \right); \\ \hat{B}_A^\varepsilon : \hat{\gamma}_{k,hij} &= (1/(1-\varepsilon)) (\hat{U}_{1-\varepsilon} / \hat{U}_0)^{\varepsilon/(1-\varepsilon)} \left[\sum_{g=1}^G ({}_g \hat{U}_0^{-\varepsilon/(1-\varepsilon)}) ({}_g \hat{U}_{1-\varepsilon}^{1/(1-\varepsilon)}) \right]^{-2} \left\{ \left(\sum_{g=1}^G ({}_g \hat{U}_0^{-\varepsilon/(1-\varepsilon)}) \times \right. \right. \\ &\quad \left. \left. ({}_g \hat{U}_{1-\varepsilon}^{1/(1-\varepsilon)}) \left(\varepsilon (\hat{U}_{1-\varepsilon} / \hat{U}_0) - y_{hij}^{1-\varepsilon} \right) + \hat{U}_{1-\varepsilon} \left[\sum_{g=1}^G ({}_g D_{hij}) ({}_g \hat{U}_{1-\varepsilon}) / ({}_g \hat{U}_0) \right]^{\varepsilon/(1-\varepsilon)} \left[y_{hij}^{1-\varepsilon} \right. \right. \right. \\ &\quad \left. \left. - \varepsilon ({}_g \hat{U}_{1-\varepsilon} / {}_g \hat{U}_0) \right] \right\}; \\ \hat{B}_A^1 : \hat{\gamma}_{k,hij} &= \exp(\hat{T}_0 / \hat{U}_0) \left[\sum_{g=1}^G ({}_g \hat{U}_0) \exp({}_g \hat{T}_0 / {}_g \hat{U}_0) \right]^{-1} \left\{ -\log y_{hij} + (\hat{T}_0 / \hat{U}_0 - 1) \right. \\ &\quad \left. + \hat{U}_0^{-1} \left[\sum_{g=1}^G ({}_g \hat{U}_0) \exp({}_g \hat{T}_0 / {}_g \hat{U}_1) \right]^{-1} \left[\sum_{g=1}^G (\exp({}_g \hat{T}_0 / {}_g \hat{U}_0)) ({}_g D_{hij}) \times \right. \right. \\ &\quad \left. \left. (1 - {}_g \hat{T}_0 / {}_g \hat{U}_0 + \log y_{hij}) \right] \right\}.\end{aligned}$$

A.5 Sample within shares

$$\begin{aligned} \hat{S}_{GE,W}^{\alpha} : \hat{\gamma}_{k,hij} = & \left(\hat{U}_{\alpha} - \hat{U}_0^{1-\alpha} \hat{U}_1^{\alpha} \right)^{-1} \sum_{g=1}^G ({}_g D_{hij}) \left[-\alpha \left({}_g \hat{U}_1 / {}_g \hat{U}_0 \right)^{\alpha-1} y_{hij} + (\alpha-1) \left({}_g \hat{U}_1 / {}_g \hat{U}_0 \right)^{\alpha} \right. \\ & \left. + y_{hij}^{\alpha} \right] + \left(\hat{U}_{\alpha} - \hat{U}_0^{1-\alpha} \hat{U}_1^{\alpha} \right)^{-2} \left[(1-\alpha) \left(\hat{U}_1 / \hat{U}_0 \right)^{\alpha} + \alpha \left(\hat{U}_1 / \hat{U}_0 \right)^{\alpha-1} y_{hij} - y_{hij}^{\alpha} \right] \times \\ & \sum_{g=1}^G \left({}_g \hat{U}_{\alpha} - \left({}_g \hat{U}_0^{1-\alpha} \right) \left({}_g \hat{U}_1^{\alpha} \right) \right); \end{aligned}$$

$$\begin{aligned} \hat{S}_{T1,W} : \hat{\gamma}_{k,hij} = & \left(\hat{T}_1 - \hat{U}_1 \log(\hat{U}_1 / \hat{U}_0) \right)^{-2} \left[\left(1 + \log(\hat{U}_1 / \hat{U}_0) \right) y_{hij} - y_{hij} \log y_{hij} - \hat{U}_1 / \hat{U}_0 \right] \times \\ & \left[\sum_{g=1}^G \left({}_g \hat{T}_1 - \left({}_g \hat{U}_1 \right) \log \left({}_g \hat{U}_1 / {}_g \hat{U}_0 \right) \right) \right] + \left(\hat{T}_1 - \hat{U}_1 \log(\hat{U}_1 / \hat{U}_0) \right)^{-1} \times \\ & \sum_{g=1}^G ({}_g D_{hij}) \left[{}_g \hat{U}_1 / {}_g \hat{U}_0 - \left(1 + \log \left({}_g \hat{U}_1 / {}_g \hat{U}_0 \right) \right) y_{hij} + y_{hij} \log y_{hij} \right]; \end{aligned}$$

$$\begin{aligned} \hat{S}_{T2,W} : \hat{\gamma}_{k,hij} = & \left(-\hat{T}_0 + \hat{U}_0 \log(\hat{U}_1 / \hat{U}_0) \right)^{-2} \left[\log y_{hij} + 1 - \log(\hat{U}_1 / \hat{U}_0) - \left(\hat{U}_0 / \hat{U}_1 \right) y_{hij} \right] \times \\ & \left[\sum_{g=1}^G \left(-\left({}_g \hat{T}_0 \right) + \left({}_g \hat{U}_0 \right) \log \left({}_g \hat{U}_1 / {}_g \hat{U}_0 \right) \right) \right] + \left(-\hat{T}_0 + \hat{U}_0 \log(\hat{U}_1 / \hat{U}_0) \right)^{-1} \times \\ & \sum_{g=1}^G ({}_g D_{hij}) \left[-\log y_{hij} + \log \left({}_g \hat{U}_1 / {}_g \hat{U}_0 \right) - 1 + \left({}_g \hat{U}_0 / {}_g \hat{U}_1 \right) y_{hij} \right]; \end{aligned}$$

$$\begin{aligned} \hat{S}_{A,W}^{\varepsilon} : \hat{\gamma}_{k,hij} = & \left(\hat{U}_1 - \hat{U}_0^{-\varepsilon/(1-\varepsilon)} \hat{U}_1^{1/(1-\varepsilon)} \right)^{-1} \left\{ \left[1 - \left(\hat{U}_1 - \hat{U}_0^{-\varepsilon/(1-\varepsilon)} \hat{U}_1^{1/(1-\varepsilon)} \right)^{-1} \times \right. \right. \\ & \left. \left(\hat{U}_1 - \sum_{g=1}^G \left({}_g \hat{U}_0^{-\varepsilon/(1-\varepsilon)} \right) \left({}_g \hat{U}_1^{1/(1-\varepsilon)} \right) \right) y_{hij} \right] - \hat{U}_0^{-\varepsilon/(1-\varepsilon)} \hat{U}_1^{1/(1-\varepsilon)} \times \\ & \left(\hat{U}_1 - \hat{U}_0^{-\varepsilon/(1-\varepsilon)} \hat{U}_1^{1/(1-\varepsilon)} \right)^{-1} \left(\hat{U}_1 - \sum_{g=1}^G \left({}_g \hat{U}_0^{-\varepsilon/(1-\varepsilon)} \right) \left({}_g \hat{U}_1^{1/(1-\varepsilon)} \right) \right) \times \\ & \left[\varepsilon/(1-\varepsilon) \hat{U}_0^{-1} - (1/(1-\varepsilon)) \hat{U}_1^{-1} y_{hij}^{1-\varepsilon} \right] + \sum_{g=1}^G \left({}_g \hat{U}_0^{-\varepsilon/(1-\varepsilon)} \right) \left({}_g \hat{U}_1^{1/(1-\varepsilon)} \right) \times \\ & \left. \left. \left({}_g D_{hij} \right) \left[\left(\varepsilon/(1-\varepsilon) \right) \left({}_g \hat{U}_0^{-1} \right) - \left(1/(1-\varepsilon) \right) \left({}_g \hat{U}_1^{-1} \right) y_{hij}^{1-\varepsilon} \right] \right\}; \end{aligned}$$

$$\begin{aligned} \hat{S}_{A,W}^1 : \hat{\gamma}_{k,hij} = & \left(\hat{U}_1 - \hat{U}_0 \exp(\hat{T}_0 / \hat{U}_0) \right)^{-1} \left\{ y_{hij} - \sum_{g=1}^G ({}_g D_{hij}) \exp \left({}_g \hat{T}_0 / {}_g \hat{U}_0 \right) \times \right. \\ & \left. \left(\log y_{hij} + 1 - {}_g \hat{T}_0 / {}_g \hat{U}_0 \right) \right\} + \left(\hat{U}_1 - \hat{U}_0 \exp(\hat{T}_0 / \hat{U}_0) \right)^{-2} \times \\ & \left[\hat{U}_1 - \sum_{g=1}^G \left({}_g \hat{U}_0 \right) \exp \left({}_g \hat{T}_0 / {}_g \hat{U}_0 \right) \right] \left[-y_{hij} + \exp(\hat{T}_0 / \hat{U}_0) \left(1 - \hat{T}_0 / \hat{U}_0 + \log y_{hij} \right) \right]. \end{aligned}$$

A.6 Sample between shares

$\hat{S}_{GE,B}^\alpha$: $\hat{\gamma}_{k,hij}$ is the same as for $\hat{S}_{GE,W}^\alpha$;

$\hat{S}_{T1,B}$: $\hat{\gamma}_{k,hij}$ is the same as for $\hat{S}_{T1,W}$;

$\hat{S}_{T2,B}$: $\hat{\gamma}_{k,hij}$ is the same as for $\hat{S}_{T2,W}$;

$$\begin{aligned} \hat{S}_{A,B}^\varepsilon : \hat{\gamma}_{k,hij} = & (1/(1-\varepsilon))\hat{U}_0^{-\varepsilon/(1-\varepsilon)}\hat{U}_1^{1/(1-\varepsilon)}\left(1-\hat{U}_0^{-\varepsilon/(1-\varepsilon)}\hat{U}_1^{-1}\hat{U}_1^{1/(1-\varepsilon)}\right)^{-1}\times \\ & \left\{-\varepsilon\hat{U}_0^{-1}\hat{U}_1^{-1}\left(1-\hat{U}_0^{-\varepsilon/(1-\varepsilon)}\hat{U}_1^{-1}\hat{U}_1^{1/(1-\varepsilon)}\right)^{-1}+\left(\sum_{g=1}^G({}_g\hat{U}_0^{-\varepsilon/(1-\varepsilon)})({}_g\hat{U}_1^{1/(1-\varepsilon)})\right)^{-1}\right\}\times \\ & \left[1+\hat{U}_0^{-\varepsilon/(1-\varepsilon)}\hat{U}_1^{-1}\hat{U}_1^{1/(1-\varepsilon)}\left(1-\hat{U}_0^{-\varepsilon/(1-\varepsilon)}\hat{U}_1^{-1}\hat{U}_1^{1/(1-\varepsilon)}\right)^{-1}\right]\times \\ & \left(\varepsilon\hat{U}_0^{-1}+\hat{U}_{1-\varepsilon}^{-1}y_{hij}^{1-\varepsilon}\right)+(1-\varepsilon)\hat{U}_1^{-2}\left(1-\hat{U}_0^{-\varepsilon/(1-\varepsilon)}\hat{U}_1^{-1}\hat{U}_1^{1/(1-\varepsilon)}\right)^{-1}\times \\ & \left[-1+\hat{U}_0^{-\varepsilon/(1-\varepsilon)}\hat{U}_1^{1/(1-\varepsilon)}\left(\sum_{g=1}^G({}_g\hat{U}_0^{-\varepsilon/(1-\varepsilon)})({}_g\hat{U}_1^{1/(1-\varepsilon)})\right)^{-1}\right]y_{hij}+ \\ & \hat{U}_{1-\varepsilon}^{-1}\hat{U}_1^{-1}\left(1-\hat{U}_0^{-\varepsilon/(1-\varepsilon)}\hat{U}_1^{-1}\hat{U}_1^{1/(1-\varepsilon)}\right)^{-1}y_{hij}^{1-\varepsilon}-\left(\sum_{g=1}^G({}_g\hat{U}_0^{-\varepsilon/(1-\varepsilon)})({}_g\hat{U}_1^{1/(1-\varepsilon)})\right)^{-2}\times \\ & \left.\sum_{g=1}^G({}_gD_{hij})({}_g\hat{U}_{1-\varepsilon}/{}_g\hat{U}_0)^{\varepsilon/(1-\varepsilon)}\left(\varepsilon({}_g\hat{U}_{1-\varepsilon}/{}_g\hat{U}_0)-y_{hij}^{1-\varepsilon}\right)\right\}. \end{aligned}$$

$$\begin{aligned} \hat{S}_{A,B}^1 : \hat{\gamma}_{k,hij} = & \hat{U}_1^{-1}\exp(\hat{T}_0/\hat{U}_0)\left(1-(\hat{U}_0/\hat{U}_1)\exp(\hat{T}_0/\hat{U}_0)\right)^{-2}\log y_{hij}-\exp(\hat{T}_0/\hat{U}_0)\times \\ & \left(1-(\hat{U}_0/\hat{U}_1)\exp(\hat{T}_0/\hat{U}_0)\right)^{-1}\left[\sum_{g=1}^G({}_g\hat{U}_0)\exp({}_g\hat{T}_0/{}_g\hat{U}_0)\right]^{-1}\times \\ & \left[(\hat{U}_0/\hat{U}_1)\exp(\hat{T}_0/\hat{U}_0)\left(1-(\hat{U}_0/\hat{U}_1)\exp(\hat{T}_0/\hat{U}_0)\right)^{-1}+1\right]\log y_{hij} \\ & -\left(1-(\hat{U}_0/\hat{U}_1)\exp(\hat{T}_0/\hat{U}_0)\right)^{-2}\hat{U}_1^{-1}\exp(\hat{T}_0/\hat{U}_0)\left(-1+(\hat{T}_0/\hat{U}_0)\exp(\hat{T}_0/\hat{U}_0)\right)\times \\ & \left[1-\hat{U}_0\exp(\hat{T}_0/\hat{U}_0)\left(\sum_{g=1}^G({}_g\hat{U}_0)\exp({}_g\hat{T}_0/{}_g\hat{U}_0)\right)^{-1}\right]+\left(1-(\hat{U}_0/\hat{U}_1)\exp(\hat{T}_0/\hat{U}_0)\right)^{-1} \\ & \times\exp(\hat{T}_0/\hat{U}_0)\left(\sum_{g=1}^G({}_g\hat{U}_0)\exp({}_g\hat{T}_0/{}_g\hat{U}_0)\right)^{-1}\left(\hat{T}_0/\hat{U}_0-1\right)-\hat{U}_0\hat{U}_1^{-2}\times \\ & \exp(\hat{T}_0/\hat{U}_0)\left(1-(\hat{U}_0/\hat{U}_1)\exp(\hat{T}_0/\hat{U}_0)\right)^{-2}\left[1-\hat{U}_0\exp(\hat{T}_0/\hat{U}_0)\times \right. \\ & \left.\left(\sum_{g=1}^G({}_g\hat{U}_0)\exp({}_g\hat{T}_0/{}_g\hat{U}_0)\right)^{-1}\right]y_{hij}+\hat{U}_0\exp(\hat{T}_0/\hat{U}_0)\left(1-(\hat{U}_0/\hat{U}_1)\times \right. \end{aligned}$$

$$\exp(\hat{T}_0 / \hat{U}_0) \left[\sum_{g=1}^G (\hat{U}_0) \exp(\hat{T}_0 / \hat{U}_0) \right]^{-2} \left[\sum_{g=1}^G \exp(\hat{T}_0 / \hat{U}_0) (1 - \hat{T}_0 / \hat{U}_0) (\hat{D}_{hij}) + \hat{U}_0 \exp(\hat{T}_0 / \hat{U}_0) (1 - (\hat{U}_0 / \hat{U}_1) \exp(\hat{T}_0 / \hat{U}_0))^{-1} \times \left[\sum_{g=1}^G (\hat{U}_0) \exp(\hat{T}_0 / \hat{U}_0) \right]^{-2} \left[\sum_{g=1}^G \exp(\hat{T}_0 / \hat{U}_0) (\hat{D}_{hij}) y_{hij} \right] \right].$$

A.7 Sample sub-group within shares

$$\begin{aligned} \hat{S}_{GE,W}^\alpha : \hat{\gamma}_{k,hij} &= -(\hat{U}_\alpha - \hat{U}_0^{1-\alpha} \hat{U}_1^\alpha)^{-1} (\hat{D}_{hij}) \left[\alpha (\hat{U}_1 / \hat{U}_0)^{\alpha-1} y_{hij} + \right. \\ &\quad \left. (1-\alpha) (\hat{U}_1 / \hat{U}_0)^\alpha - y_{hij}^\alpha \right] + (\hat{U}_\alpha - \hat{U}_0^{1-\alpha} \hat{U}_1^\alpha)^{-2} (\hat{U}_\alpha - (\hat{U}_0^{1-\alpha} \hat{U}_1^\alpha)) (\hat{U}_1 / \hat{U}_0)^\alpha - y_{hij}^\alpha \times \\ &\quad \left[\alpha (\hat{U}_1 / \hat{U}_0)^{\alpha-1} y_{hij} + (1-\alpha) (\hat{U}_1 / \hat{U}_0)^\alpha - y_{hij}^\alpha \right]; \\ \hat{S}_{T1,W} : \hat{\gamma}_{k,hij} &= (\hat{T}_1 - \hat{U}_1 \log(\hat{U}_1 / \hat{U}_0))^{-2} (\hat{T}_1 - \hat{U}_1 \log(\hat{U}_1 / \hat{U}_0)) \times \\ &\quad \left[-y_{hij} \log y_{hij} - \hat{U}_1 / \hat{U}_0 + (1 + \log(\hat{U}_1 / \hat{U}_0)) y_{hij} \right] + (\hat{T}_1 - \hat{U}_1 \log(\hat{U}_1 / \hat{U}_0))^{-1} \times \\ &\quad (\hat{D}_{hij}) \left[(\hat{U}_1 / \hat{U}_0) - (1 + \log(\hat{U}_1 / \hat{U}_0)) y_{hij} + y_{hij} \log y_{hij} \right]; \\ \hat{S}_{T2,W} : \hat{\gamma}_{k,hij} &= (-\hat{T}_0 + \hat{U}_0 \log(\hat{U}_1 / \hat{U}_0))^{-2} (\log y_{hij} - \log(\hat{U}_1 / \hat{U}_0) + 1 - (\hat{U}_0 / \hat{U}_1) y_{hij}) \times \\ &\quad (-\hat{T}_0 + (\hat{U}_0) \log(\hat{U}_1 / \hat{U}_0)) - (-\hat{T}_0 + \hat{U}_0 \log(\hat{U}_1 / \hat{U}_0))^{-1} \times \\ &\quad (\hat{D}_{hij}) (\log y_{hij} - \log(\hat{U}_1 / \hat{U}_0) + 1 - (\hat{U}_0 / \hat{U}_1) y_{hij}); \\ \hat{S}_{A,W}^\varepsilon : \hat{\gamma}_{k,hij} &= (\hat{U}_1 - \hat{U}_0^{-\varepsilon/(1-\varepsilon)} \hat{U}_1^{1/(1-\varepsilon)})^{-2} \left\{ (\hat{U}_1 - \hat{U}_0^{-\varepsilon/(1-\varepsilon)} \hat{U}_1^{1/(1-\varepsilon)}) \times \right. \\ &\quad \left. [y_{hij} + (1/(1-\varepsilon)) \hat{U}_0^{-\varepsilon/(1-\varepsilon)} \hat{U}_1^{1/(1-\varepsilon)} (\varepsilon \hat{U}_0^{-1} - \hat{U}_1^{-1} y_{hij}^\varepsilon)] + \right. \\ &\quad \left. (\hat{U}_1 - \hat{U}_0^{-\varepsilon/(1-\varepsilon)} \hat{U}_1^{1/(1-\varepsilon)}) (\hat{D}_{hij}) [y_{hij} + (1/(1-\varepsilon)) \hat{U}_0^{-\varepsilon/(1-\varepsilon)} \hat{U}_1^{1/(1-\varepsilon)} \times \right. \\ &\quad \left. (\varepsilon \hat{U}_0^{-1}) - (\hat{U}_1^{-1} y_{hij}^\varepsilon)] \right\}; \\ \hat{S}_{A,W}^1 : \hat{\gamma}_{k,hij} &= (\hat{U}_1 - \hat{U}_0 \exp(\hat{T}_0 / \hat{U}_0))^{-1} \left\{ \hat{U}_0^{-2} (\hat{U}_0) (\hat{U}_1 - \hat{U}_0 y_{hij}) + (\hat{D}_{hij}) [\hat{U}_1 / \hat{U}_0 - \right. \\ &\quad \left. \exp(\hat{T}_0 / \hat{U}_0) (1 - (\hat{T}_0 / \hat{U}_0) + \log y_{hij}) \right\} + \hat{U}_0 (\hat{U}_1 - \hat{U}_0 \exp(\hat{T}_0 / \hat{U}_0))^{-2} \times \\ &\quad (\hat{U}_1 / \hat{U}_0 - \exp(\hat{T}_0 / \hat{U}_0)) \left[-y_{hij} + \exp(\hat{T}_0 / \hat{U}_0) (1 - \hat{T}_0 / \hat{U}_0 + \log y_{hij}) \right]. \end{aligned}$$