SMU ECONOMICS & STATISTICS
WORKING PAPER SERIES

Two-sample estimation of poverty rates for disabled people: an application to Tanzania

Tomoki Fujii

January 2008

Paper No. 02-2008

Two-sample estimation of poverty rates for disabled people: an application to Tanzania*

Tomoki Fujii

School of Economics

Singapore Management University

90 Stamford Road, Singapore 178903

Phone: +65-6828-0279

Fax: +65-6828-0833

E-mail: tfujii@smu.edu.sg

*Initial stage of this research was prepared as a report for the World Bank. I thank Chris Elbers, Hans Hoogeveen, Wietze Lindeboom and Roy van der Weide for their comments. This study was supported by SMU Research Grant(07-C208-SMU-007). Usual caveats apply.

Two-sample estimation of poverty rates for disabled people: an application to Tanzania

Abstract

Estimating poverty measures for disabled people in developing countries is difficult, partly because relevant data are not available. We develop two methods to estimate poverty by the disability status of the household head. We extend the small-area estimation proposed by Elbers, Lanjouw and Lanjouw (2002, 2003) so that we can run a regression on head's disability status even when such information is unavailable in the survey. We do so by aggregation and by moment adjusted two sample instrumental variable estimation. Our results from Tanzania show that both methods work well, and that disability is indeed associated with poverty.

JEL classification code: C20, I10, I32

Key words: poverty, disability, Tanzania, aggregation, two-sample instrumental variable estimation

1 Introduction

Poverty and disability are fundamentally interlinked. The direction of causality between poverty and disability can go both ways. On the one hand, disability may cause poverty for at least three reasons. First, disabled people may not be able to earn income as much as they would without disability. Such loss of income can arise from loss in productivity as well as from social exclusion. Second, disabled people may incur additional costs such as medical expenses, equipment, adaptations to housing and specialized services (Elwan, 1999). Households with disabled members may also incur additional cost in the form of foregone incomes of other family members giving care to the disabled member. Third, disabled people may have limited access to services compared with non-disabled people. For example, social safety nets may not be accessible for disabled people when physical or social barriers exist (Mitra, 2004).

On the other hand, poverty may cause disability as well. Poor people are more likely to suffer not only from the lack of adequate food and water but from the lack of adequate and timely health care. They may also have to accept more hazardous working conditions and less safe living environment. For example, in Cambodia, it is not uncommon to see people living near a mine field near the Thai border in the north. According to an unpublished report by Action on Disability and Development (1997) cited in Yeo and Moore (2003), it is estimated that over 70 percent of land mine survivors had been farming or foraging with the full knowledge that they were doing so in areas infested with land mines.

As Yeo and Moore (2003) point out, despite an increasing awareness within development field that disabled people are among the poorest, studies on the relationship between poverty and disability in developing countries have been very limited. In particular, quantitative studies on the relationship between poverty and disability are scarce.

The lack of quantitative studies stems at least in part from the lack of relevant data on poverty and disability. Poverty analysis often relies on socioeconomic surveys representative at a highly aggregated level. Such survey data typically contain, if any, only a very limited number of people with disability because the proportion of disabled people to the total

population is often as low as a few percent. This makes it difficult to calculate reliable measures of poverty for disabled people.

Hence, the goal of this study is to develop methods to derive robust measures of poverty for disabled people. Our methods extend the small-area estimation developed by Elbers et al. (2002, 2003, hereafter ELL), which combine a survey and a census. As we shall discuss later, the ELL small-area estimation requires us to have all the regressors used in the regression model in both the census and survey. Hence, unless the naïve assumption that one single consumption regression model applies to both disabled and non-disabled groups holds, their methodology does not work in the absence of disability information in the survey.

In this study, we offer two methods of two-sample estimation to overcome this problem. First method is by aggregation, which is somewhat similar to Feige and Watts (1972). Aggregation is useful because disability information from another source can be often merged into the survey at an aggregated level. In the second method, we use a variant of the two-sample instrumental variable regression pioneered by Angrist and Krueger (1992). By choosing an appropriate instrument, we can estimate the consumption model for both disabled and non-disabled groups even when we do not observe the disability information in the survey. The empirical results from Tanzania indicate that both methods work reasonably well.

This paper is organized as follows. In Section 2, we review relevant literature on poverty and disability. In Section 3, we describe the data and present some summary statistics that motivate this study. In Section 4, we develop the methodology. In Section 5 presents the empirical results followed by discussions and conclusions in Section 6.

2 Relevant Studies on Poverty and Disability

Studies on the relationship between poverty and disability in developing countries are surprisingly limited (See Haveman and Wolfe (2000) for a good survey of economic studies on disability in developed countries). Elwan (1999) summarizes existing literature in developing countries, but a large part of the body of the literature is qualitative. One of the few quantitative studies is Masset and White (2004). They look at, among other things, the

relationship between consumption poverty and disability using data from Andhra Pradesh in India, Bulgaria, Ghana and Nicaragua. They find that disabled people are indeed poorer in all the places they studies except for Nicaragua. However, they do not carry out statistical test, and how much disabled people are poorer is not clear.

Filmer (2005) looks at the relationship between disability, poverty and schooling using 11 household surveys. He calculated the concentration index of disability against household economic status, as defined by consumption or asset index. He found that all but three surveys had negative concentration index, which indicates disabled people are disproportionately concentrated among the poor. However, only one survey had a value significantly different from zero at a five percent level.

Our study and Hoogeveen (2005) share the same motivation. Small target groups, such as disabled people, has a very small number of observations in surveys, and reliable estimate of poverty is difficult to derive. Thus, Hoogeveen (2005) applies the ELL small area estimation in order to estimate the poverty rates for the people headed by a disabled person in Uganda. He introduces interaction terms between several household-level regressors and the fraction of disabled people in the census enumeration area in the regression model. While he acknowledges the introduction of the interaction terms does not eliminate the bias in the model, his results support the qualitative evidence on the correlation between poverty and disability.

Lindeboom (2005) applied the same approach in Tanzania. He estimated poverty rates for households headed by a disabled and non-disabled person. While the difference in his study is statistically significant, the result must be interpreted with great caution as both the standard errors and the point estimates are likely to be biased as we shall argue in length later. Our study tries to overcome this problem by explicitly allowing the regression coefficient to be different across disabled and non-disabled groups.

We are not aware of any other quantitative studies on poverty and disability in Tanzania, there are a few studies that look at the relationship of disability to other dimensions of welfare. Taylor et al. (1991) show that the risk of death for visually impaired people over the

age of 40 have 3.33 times higher than normally sighted individuals after controlling for age, sex and village.

There is mixed evidence of social exclusion of disabled people. Kisanji (1995a,b) find that disabled people are not marginalized in the communities he studied. UNICEF (1999) makes similar remarks for children. On the other hand, Matuja and Rwiza (1994) showed that negative attitudes towards epilepsy among rural secondary school students are widespread.

These studies suggest that households with disabled people may face a reality that is quite different from households without disabled members. This, in turn, means that the relationship between consumption and other characteristics may be significantly different between those households with and without disabled people. Thus, we must allow for the possible differences between them when we run a consumption regression for the ELL small-area estimation.

3 Data and Measurement

As with the ELL small-area estimation, we combine a survey and a census. For the census, we use the Population and Housing Census for 2002. We use the long-form questionnaire of the census, which includes questions on the age, sex, relation to the household head, marital status, disability status, education and economic activity of the individual as well as the housing conditions and asset holdings of the household. The long-form questionnaire was used for about 1.2 million households out of about 6.8 million households in mainland Tanzania. We excluded Zanzibar from the analysis, because it is not covered in the survey. More detailed information on the census is given in National Bureau of Statistics (2003).

For the survey, we used the 2000/01 Household Budget Survey (HBS). It is representative at the level of twenty regions in mainland Tanzania. The survey data cover a wide range of household and individual characteristics, including many of the variables included in the census along with detailed information on consumption expenditure. The survey does not include a question on disability itself. However, the question on economic activities lists disability as one of the reasons for not being economically active. Hence, we have limited

information on disability in the survey.

The survey data set contain 22,178 households from 1,158 enumeration areas. It comes with the sampling weight for each observation. After eliminating observations with missing values for the variables used in this study, we were left with 21,608 observations from 1,148 enumeration areas. National Bureau of Statistics (2002) offers further information on the HBS data, including a range of summary statistics.

We followed the consumption-based definition of poverty given by National Bureau of Statistics (2002). National Bureau of Statistics (2002) first calculates the total household consumption expenditure for a 28-day period in the (male) adult equivalence scale, which accommodates different needs for different age and sex groups. One working-age male adult has a unit weight, while young children, elderly people and working-age female have a smaller weight. The household is considered poor when the household consumption expenditure per adult equivalent is less than the poverty line. The basic needs poverty line we adopted covers the cost for satisfying the minimum adult caloric requirement and some non-food consumption expenditure. After adjusting for regional price differences, the poverty line for mainland Tanzania is 7253 Tanzanian schillings. This is equal to 16.09 US dollars using the purchasing power parity conversion factor for 2002 reported in the World Development Indicator. According to this definition, 35.7 percent of the people in mainland Tanzania are poor.

It is difficult to define disability objectively. Most of the survey data with disability information in developing countries, including Tanzania, are self-reported. Thus, the perception of disability influences the reported status of disability. In Tanzania, some disabilities may go unrecognized until children go to school when learning difficulties as well as visual and hearing impairment are brought to notice (UNICEF, 1999). Our study is, therefore, constrained by the subjective nature of disability information.

We focus on the disability status of the household head. There are two reasons for this. First, as UNICEF (1999) implies, children's disability status may be less reliable than adults' disability status. Second, the level of the welfare of a household depends most heavily on

the household head. Hence, the impact of disability is likely most acute when the household head is disabled. We shall get back to this point later in this section.

Hereafter, we shall call a household headed by a disabled person a "disabled household." Table 1 provides us with the summary statistics by the disability status of the household head. About 2.4 percent of the population live in a disabled household. On average, disabled households are more likely to use firewood for cooking and lighting, less likely to use get piped water and have toilet. Furthermore, disabled households use weaker construction materials for their houses. The education of the head of disabled households is also lower than that of non-disabled households. The disabled households possess less assets, except for the hoe. These observations suggest that disabled households are indeed worse off. Yet, it is not clear how much they are worse off, and we cannot conclude anything about the differences in poverty rates for disabled households and non-disabled households.

The survey data also allow us to get a sense about the relationship between disability and poverty. While the survey doesn't ask about the individual disability status itself, we can identify those individuals whose primary or secondary economic activity is "not active" due to disability. Thus, we can identify in the survey the "economically disabled" households, which are headed by a person who is not economically active due to disability.

As can be seen from Table 2, the poverty rate for those living in economically disabled households is substantially higher than that for non-disable households. Yet, the difference is not statistically significant because the standard error associated with the poverty rate for disabled households is as high as 13.3 percent. The large standard error for disabled households is as a result of the small number of observations for the disabled households in the survey.

Using the survey data, we can also see the disability status of the household head is more strongly correlated with poverty than that of any other members. The poverty rate for households with at least one economically disabled people is 40.7 percent. This is still considerably higher than the national average of 35.7 percent, but much lower than 52.6 percent for the economically disabled households. While the standard errors are too large to

Table 1: Characteristics of non-disabled and disabled households. The numbers for housing conditions do not add up to hundred percent due to missing values. Author's calculation based on the census data.

Variable		Not Disabled	Disabled
Housing Conditions (numbers in percentage)			
Cooking	Electricity or Gas	1.0	0.4
	Paraffin Oil	3.6	1.6
	Firewood	72.3	84.9
	Charcoal	22.5	12.2
Lighting	Electricity/Solar	10.8	4.7
	Lamp	84.5	87.0
	Firewood	4.2	7.4
Water	Piped Water	36.5	31.0
	Protected Well/Spring/Rain Water	20.4	20.8
	Unprotected Well	24.9	27.5
	Unprotected Spring	4.2	5.0
	River Stream/Pond/Lake	12.4	14.7
Toilet	No Toilet	7.5	10.6
	Flush Toilet	3.2	1.4
	Pit Latrine	87.2	86.8
	Ventilated improved pit latrine	2.1	1.0
Floor	Mud	69.3	81.7
	Cement/Timber/Tile/Other	30.6	18.1
Wall	Timber	0.6	0.6
	Poles and Mud	33.8	42.4
	Sun dried Bricks	32.9	33.4
	Baked Bricks	17.4	13.3
	Stones/Cement Bricks	13.8	8.0
Roof	Grass /Bamboo	37.6	46.5
	Mud and Grass	9.9	12.7
	Concrete and Cement/Asbestos/Tiles	1.0	0.7
	Metal Sheets	51.3	39.6
Electricity		9.8	4.2
Asset Holding (numbers in percentage)			
Radio		53.1	38.0
Phone		3.3	1.3
Bike		34.2	24.3
Hoe		74.3	83.4
Barrow		15.1	13.0
Iron		4.7	3.3
Household size		4.50	4.20
Schooling of household head in years		4.95	3.26
Age of household head		41.4	51.7
% of male headed households		66.2	70.3
Number of households		1150667	30617

Table 2: Poverty rates for economically disabled and non-disabled households. The standard errors take into account clustering. All the numbers are due to author's calculation based on the survey data.

	Poverty Rate		Obs	Share
Economically Non-Disabled	35.7	(1.0)	21955	99.7
Economically Disabled	52.6	(13.3)	55	0.3
Tanzania	35.7	(1.0)	22010	100.0

draw any statistical conclusions, this observation warrants our focus on the disability status of the household head.

It should be also pointed out that economic disability likely represents severe forms of disability as mildly disabled people, especially household heads, would choose to work if they can. Our census and survey observations are consistent with this. As the last column of 2 shows, the share of people in economically disabled households is substantially smaller than the corresponding figure of 2.4 percent in the census. This point also indicates that census-based disability does not exactly correspond to economic disability in the survey. Hence, the poverty rates for disabled households we derive in subsequent sections are not directly comparable to Table 2. The numbers here should be taken only as a motivation for this study.

4 Methodology

The methodology of this study is partly built upon the ELL small-area estimation, which combines a census and a survey. The ELL small-area estimation is most popularly used for poverty mapping, in which poverty estimates for small areas are plotted on a map. It has also been used, among other things, to analyze geographic targeting (Elbers et al., 2007; Fujii, 2008), inequality (Elbers et al., 2004; Demombynes and Özler, 2005) and regression analysis at aggregated levels (Elbers et al., 2005).

Tarozzi and Deaton (2007) recently proposed a non-parametric version of small-area estimation. Their method is more robust to misspecification. However, we do not adopt this for two reasons. First, it is not readily applicable to this study, because our approach is derived

from the parametric specification of the ELL approach. Second, to the best of the author's knowledge, there is no empirical evidence that clearly rejects the ELL specification.

Let us first introduce the ELL small-area estimation in a simplest form and describe our estimation strategy. Suppose that we have C clusters in the sample, and let us denote the set of clusters by $\mathcal{C} = \{1, \cdots, C\}$. We denote the non-empty set of households in cluster $c \in \mathcal{C}$ by $\mathcal{H}_c$. Without loss of generality, we can let $\mathcal{H}_c = \{(c, 1), \cdots, (c, N_c)\}$ where $N_c \equiv \#\{\mathcal{H}_c\}$ is the number of households in cluster c.

We let the set of all households be $\mathcal{H} \equiv \cup_{c \in \mathcal{C}} \mathcal{H}_c$. The cluster membership function $\kappa : \mathcal{H} \to \mathcal{C}$ maps each household to the cluster it belongs to, so that $\kappa(h) = c \Leftrightarrow h \in \mathcal{H}_c$. Each household $h (\in \mathcal{H})$ has a weight w_h, which is typically the population expansion factor.

The ELL small-area estimation allows us to find an aggregate welfare measure $P_J \equiv P(\{y_h\}_{h \in J}, \{w_h\}_{h \in J})$ for a set J of households, where y_h measures the standards of living for household h. The aggregate welfare measures conventionally used include FGT poverty measures (Foster et al., 1984)) and inequality measures such as the Gini index. We take y_h to be per adult equivalent logarithmic consumption. The set J often, but not necessarily, represents a small geographic unit.

The ELL small-area estimation combines a survey and a census through a regression of y_h on a row L-vector of household-level variables $x_h (= (x_{h,1}, \cdots, x_{h,L}))$. These household-level variables are common between the census and the survey. They typically include demographic characteristics, housing conditions, education and characteristics of the community in which the household is located. Let us first decompose y_h into the conditional expectation and the error term u_h as follows:

$$\begin{aligned} y_h &= E[y_h | x_h] + u_h \\ &= x_h \beta + \eta_{\kappa(h)} + \epsilon_h \end{aligned} \tag{1}$$

Now, suppose that the conditional expectation $E[y_h|x_h]$ is a linear combination of x_h so that $E[y_h|x_h] = x_h\beta$, and that u_h can be expressed as the sum of a cluster-specific random effect $\eta_{\kappa(h)}$ and a household-specific random effect ϵ_h. Then, we have Eq(1).

We assume $E[\eta_{\kappa(h)}] = E[\epsilon_h] = 0$ for $\forall h$. Further, we assume η_c and ϵ_h are independently distributed across c and h respectively. They are also independent with each other and with x_h and w_h. Typically, $\eta_{k(h)}$ is assumed to be homoskedastic, but the heteroskedasticity of ϵ_h is allowed for. Note that Eq(1) describes conditional expectation and should not be interpreted as a causal model.

In a standard ELL small-area estimation, the regression coefficient, its associated variance-covariance matrix as well as the distributional parameters of η and ϵ are estimated using the survey. Then, y_h is repeatedly imputed to each census household in a Monte-Carlo simulation.

Now, let us denote by R the number of simulations. The regression parameter $\tilde{\beta}^{(r)}$ for the r-the round of the simulation for $r \in \{1, \cdots, R\}$ is randomly drawn from the estimated distribution of the parameter estimate $\hat{\beta}$. Furthermore, the cluster-specific random effect $\tilde{\eta}_c^{(r)}$ and the household-specific random effect $\tilde{\epsilon}_h^{(r)}$ are also drawn for each cluster c and each household h from their estimated distributions. Hence, in the r-th round of the simulation, the imputed welfare measure for household h is $\tilde{y}_h^{(r)} = x_h\tilde{\beta}^{(r)} + \tilde{\eta}_{\kappa(h)}^{(r)} + \tilde{\epsilon}_h^{(r)}$. The aggregate welfare measure for the r-the round is estimated at $\tilde{P}_J^{(r)} = P(\{\tilde{y}_h^{(r)}\}_{h\in J}, \{w_h^{(r)}\}_{h\in J})$. The point estimate and standard error are derived by taking the average and standard deviation of $\tilde{P}_J^{(r)}$ over r.

In our model, we assume that there are some groups that have different regression coefficients for at least some of the regressors. The groups we consider in our empirical application are disabled and non-disabled households, but they could be defined by other characteristics such as the race and religion of the household head. Suppose that there are G groups, and let a_{hg} be the group membership dummy, which takes one if household h belongs to group $g \in \{1, \cdots, G\}$ and zero otherwise. Each household belongs to exactly one group. We denote the G-row vector of group dummies for household h by $a_h = (a_{h1}, \cdots, a_{hg})$.

In this setup, we can write $x_h \equiv (x_h^e \quad a_h \otimes x_h^d)$ where a row L^e-vector of household-level characteristics x_h^e has the same regression coefficient across groups and a row L^d-vector of household characteristics x_h^d has a different regression coefficient across groups. The regression coefficient for x_h is a column L-vector of coefficients $\beta^e, \beta_1^d, \cdots, \beta_G^d$ stacked in one vector,

where $L = L^e + L^d G$. Hence, if x_h^e, a_h and x_h^d are observed both in the census and the survey, we observe x_h in both samples and thus the standard ELL small-area estimation applies.

However, it is not the case when the group information a_{hg} is missing in the survey. One quick fix is to simply assume that $\beta_1^d = \cdots = \beta_G^d$. This naïve assumption is clearly problematic when the coefficients are indeed different among different groups. The estimates of β^d under this assumption only reflect the relationship between consumption and x^d averaged over different groups. Hoogeveen (2005) and Lindeboom (2005) took this naïve approach. They, however, attempted to reduce the bias by including a number of cross-terms between community-level prevalence of disability and household level characteristics. This is in effect the same as replacing a_h with $\bar{a}_{\kappa(h)}$, where $\bar{a}_{\kappa(h)}$ is the community average of a_h in $\kappa(h)$. This approach would capture some of the variations across groups, but the bias in the imputed welfare measure remains. The size of the bias depends on how much $a_h \otimes x_h^d$ contributes to explaining the total variations of y_h.

Even if the bias is small, problems remain. Under the correct specification, both the point estimates and the standard errors for the coefficients on x_h^d depend on the group. However, the naïve model ignores the differences in standard errors across types, so that the standard errors for aggregate welfare measures are also incorrect. This is particularly problematic when our goal is to compare the aggregate welfare measures across groups.

Further, the naïve model cannot capture the heteroskedasticity in the unobserved household effects that may exist across groups. This is potentially an important problem. In our application, the presence of supporters for disabled households may be a source of such heterogeneity. Because the presence of such supporters is not observable, we are forced to treat it as a household-specific random effect. Because this random effect clearly does not affect non-disabled households, we have group-dependent heteroskedasticity in this case.

In this paper, we propose two different approaches to address these issues. The first approach is by way of aggregation, in which we merge the census-based disability information into the survey so that we can run regressions at an aggregated level. The second approach is by way of instrumental variables, where they must be constructed from variables common

between the census and the survey.

We make three additional requirements for these methods to work. First, we can merge aggregate group information into the survey. That is, we have the average of $a_h \otimes x_h^d$ in each survey cluster. This is often possible, so long as the survey and census (or any other tertiary data source) use harmonized administrative codes. This requirement is critical for the aggregation method.

Second, we allow for the heteroskedasticity across groups but not within each group. That is, the variance of ϵ depends only on the household group. This specification allows us to take into account household-level unobserved effects that are specific to the group, such as the presence of local supporters for disabled people. We let the variance of ϵ_h in group g be $\sigma^2_{\epsilon,g}$, and that of η_h be σ^2_η.

Finally, we assume that η and ϵ come from a one-parameter mean-zero distribution with reproductive property, where the parameter is the variance. We require this assumption because, unlike standard ELL applications, we cannot decompose the error into cluster-specific and household-specific random effects in our method.

Before proceeding, let us introduce some notations here. We let Y and U be an n-vector of y_h and u_h respectively, where $n \equiv \#\{\mathcal{H}\}$ is the total number of households. We let z_h be a row L-vector of instruments for x_h. We let X and Z be a $n \times L$ matrix of entire observations of x_h and z_h respectively. Using these notations, Eq(1) can be written as $Y = X\beta + U$.

Further, we let the weighting matrix be $W_H \equiv diag(w_1, \cdots, w_H)$. We write the variance-covariance matrix of U as $\Omega \equiv E[UU']$. We require Z to be full-ranked. Finally, we let K be a $C \times n$ cluster membership matrix whose (c, h) element is $\text{Ind}(\kappa(h) = c)$, where $\text{Ind}(\cdot)$ is an indicator function.

Hereafter, we make the following assumptions: Each cluster c has a double (N_c, η_c), which is independently and identically distributed. $N_c (\in \mathbf{N})$ is the number of households in cluster c and is uniformly bounded. Cluster c also has a triple $(\{x_h\}_{h\in\mathcal{H}_c}, \{z_h\}_{h\in\mathcal{H}_c}, \{\epsilon_h\}_{h\in\mathcal{H}_c})$. We assume (x_h, z_h, ϵ_h) is independently and identically distributed within and between clusters. Further, we assume x_h and z_h are uniformly bounded.

Aggregation Method A straightforward way to estimate Eq(1) in the presence of aggregate group information is to run the regression at an aggregated level. Our aggregation method is most closely related to Feige and Watts (1972). They investigate the properties of the estimator using aggregate data and the information loss due to aggregation. Welsch and Kuh (1976) considers a related problem for a random coefficient model. Empirical application of related models includes Polinsky (1977).

Unlike Feige and Watts (1972), we allow for the cluster-specific random effect. Now, let the cluster weight be $\tilde{w}_c \equiv \sum_{h \in \mathcal{H}_c} w_h$. Taking the weighted average of Eq(1) over cluster c, we get the following cluster-level equation:

$$\bar{y}_c = \bar{x}_c \beta + \bar{u}_c, \tag{2}$$

where $\bar{y}_c \equiv \frac{1}{\tilde{w}_c} \sum_{h \in \mathcal{H}_c} w_h y_h$, $\bar{x}_c \equiv \frac{1}{\tilde{w}_c} \sum_{h \in \mathcal{H}_c} w_h x_h$ and $\bar{u}_c \equiv \frac{1}{\tilde{w}_c} \sum_{h \in \mathcal{H}_c} w_h u_h$ $(= \eta_c + \frac{1}{\tilde{w}_c} \sum_{h \in \mathcal{H}_c} w_h \epsilon_h)$. It is straightforward to show that $E[\bar{u}_c] = 0$ and $\sigma^2_{u,c} \equiv Var[\bar{u}_c] = \sigma^2_\eta + \sum_{g=1}^{G} \bar{A}_{cg} \sigma^2_{\epsilon,g}$, where $\bar{A}_{cg} \equiv \frac{1}{\tilde{w}_c^2} \sum_{h \in \mathcal{H}_c} w_h^2 a_{hg}$.

Eq(2) can be estimated by a two-step feasible GLS regression because we can combine the survey and census at the cluster level so that we have both $\bar{y}_c$ and $\bar{x}_c$ in the combined data set. Now let $\bar{Y}(= KWY)$ and $\bar{U}(= KWU)$ be the C-vectors of $\bar{y}_c$ and $\bar{u}_c$ for all clusters. Similarly, let $\bar{X}$ be the $C \times L$ matrix be $\bar{x}_c(= KWX)$ for all clusters. We denote the cluster-level weighting matrix by $\bar{W} \equiv diag(\tilde{w}_1, \cdots, \tilde{w}_C)(= KWK^T)$. Further, we denote the variance-covariance matrix for the error terms at the cluster level by $\Omega_A \equiv E[\bar{U}\bar{U}'] = diag(\sigma^2_{u,1}, \cdots, \sigma^2_{u,C})$.

Using the matrix notation, we can write Eq(2) as $\bar{Y} = \bar{X}\beta + \bar{U}$. Premultiplying this by $\Omega_A^{-1/2}$ and running a weighted least squares regression, we would obtain the GLS estimator. This, of course, involves an unknown parameter Ω_A.

To estimate Ω_A, we first take the OLS residual $\hat{u}_c$ for Eq(2). Looking at the relationship between $\sigma^2_{u,c}$, σ^2_η and $\sigma^2_{\epsilon,g}$, we can see that an OLS regression of $\hat{u}_c^2$ on a constant and $\bar{A}_{c1}, \cdots, \bar{A}_{cG}$ gives consistent estimates of σ^2_η and $\sigma^2_{\epsilon,g}$ as the coefficients on the constant term and $\bar{A}_{cg}$ respectively. This in turn allows us to find a consistent estimate of Ω_A, which in

turn allows us to find a consistent estimate of β. Hence, under the regularity conditions for the standard feasible GLS regression model and letting $D \equiv \hat{\Omega}_A^{-T/2}\bar{W}\hat{\Omega}_A^{-1/2}$, we have the following aggregation method estimator and its variance estimator as follows:

$$\begin{aligned} \hat{\beta}_{AGG} &= (\bar{X}^T D\bar{X})^{-1}\bar{X}^T D\bar{Y}, \quad \text{with} \\ \widehat{Var}[\hat{\beta}_{AGG}] &= (\bar{X}^T D\bar{X})^{-1}\bar{X}^T\hat{\Omega}_A^{-T/2}\bar{W}^2\hat{\Omega}_A^{-1/2}\bar{X}(\bar{X}^T D\bar{X})^{-1}. \end{aligned} \tag{3}$$

Instrumental Variable Method While the aggregation method provides us with consistent estimates, it throws away a considerable amount of information, because we use the cluster averages $\bar{x}_c$ even when we can observe x^h. Hence, the goal in this section is to develop a method to run a regression at the household level.

To this end, we first find proxy variables $\tilde{a}_h$ for the group dummies a_h, which can be found or calculated for both census and survey. We can then find an L-vector of instrumental variables $z_h \equiv (x_h^e \quad \tilde{a}_h \otimes x_h^d)$ for x_h. We require z_h to be also uncorrelated with u_h.

This method works because the instrumental variable estimator consists of two sample moments. We can take the moment that involves y_h from the survey and the other moment that involves x_h from the census. Note that z_h has the same dimension as x_h by construction. This idea of combining two sample moments via instrumental variables originates from Angrist and Krueger (1992), in which they used the two-sample instrumental variable (TSIV) estimation to investigate the relationship between the age at the school entry and ultimate educational attainment by combining two US census data sets. Other empirical studies using a two-sample approach include Lusardi (1996), Bjorklund and Jantti (1997), Currie and Yelowitz (2000), and Dee and Evans (2003).

In a one-sample case, it is well known that the instrumental variable regression is identical to the two-stage least squares regression when the estimation equation is exactly identified. However, their two-sample analogues, the TSIV and the two-sample two-stage least squares (TS2SLS) estimators are not equivalent. Inoue and Solon (2006) argue many empirical researchers may have been (inadvertently) using two-sample two-stage least squares (TS2SLS) estimator instead of TSIV estimator.

Obvious question here is whether we should use TSIV or TS2SLS. Inoue and Solon (2006) showed that TS2SLS is asymptotically more efficient than TSIV. Fujii and van der Weide (2007) showed that TS2SLS is close to the mean-squared-error minimizing estimator, and argue that TS2SLS tend to behaves much better than TSIV in finite samples as well. The better property of TS2SLS stems from the moment adjustment implicit in TS2SLS. We shall highlight this point below.

These desirable properties of TS2SLS don't come without a price. In the standard framework of TS2SLS, we need an assumption of linear relationship between x_h and z_h, which is invalid in our application as we argue later. Hence, we propose variants of TSIV in this paper, in which we adjust the differences in moments for x between the two samples. We shall call them the moment-adjusted two-sample instrumental variable (MATSIV) estimators. We shall argue later that TS2SLS can be regarded as a special case of a MATSIV estimator and does not necessarily require the assumption of the first-stage regression.

Let us start with the one-sample standard IV estimator, which is $\hat{\beta}_{IV} = (Z^T W X)^{-1}(Z^T W Y) = (n^{-1} Z^T W X)^{-1}(n^{-1} Z^T W Y)$. This estimator does not work in our application because we do not have Z, X and Y in one sample. However, if we can find a suitable instrument in the two samples, we can take these two moments from the two different samples. Formally, we can use the following TSIV estimator:

$$\hat{\beta}_{TSIV} = (n_1^{-1} Z_1^T W_1 X_1)^{-1}(n_2^{-1} Z_2^T W_2 Y_2) \quad \text{with} \tag{4}$$

$$Var[\hat{\beta}_{TSIV}] = (n_1^{-1} Z_1^T W_1 X_1)^{-1}(n_2^{-1} Z_2^T W_2 \Omega_2 W_2 Z_2 n_2^{-1})(n_1^{-1} X_1^T W_1 Z_1)^{-1}, \tag{5}$$

where the subscripts 1 and 2 are used to denote Sample 1 (census) and Sample 2 (survey). We assume that $n_1 n_2^{-1}$ is fixed and the following limiting conditions hold for each of sample $i \in \{1, 2\}$,

$$\begin{cases} n_i^{-1} Z_i^T W_i X_i & \xrightarrow{p} & Q^0 \\ n_i^{-1} Z_i^T W_i \Omega_i W_i Z_i & \xrightarrow{p} & Q^1 \quad \text{as} \quad C \to \infty \\ n_i^{-1} Z_i^T W_i U_i & \xrightarrow{p} & O_{n_i \times 1} \end{cases} \tag{6}$$

where Q^0 is a positive definite matrix, and Q^1 is a symmetric positive definite matrix. Then,

noting that n_1 (and n_2) are approximately proportionate to C as $C \to \infty$, $\sqrt{C}$-asymptotically normality of the TSIV estimator follows from a variant of the standard argument (White, 1984, Chap. 5).

Six remarks are in order. First, unlike the one-sample case, TSIV estimator is biased. It is straightforward to show $E[\beta_{TSIV}] = (n_1^{-1} Z_1^T W_1 X_1)^{-1} (n_2^{-1} Z_2^T W_2 X_2)\beta \neq \beta$ in general. This bias is not observable precisely because X_1 is not observable. However, since Z contains some information on X, we may be able to correct this bias.

Second, this bias is related to the good properties of TS2SLS. In the standard framework of the two-stage least squares regression, we assume that there is a first-stage regression equation, $X = Z\Gamma + \Delta$, where Γ is an $L \times L$ matrix and Δ is an $n \times L$ matrix of error terms. It is straightforward to show that the TS2SLS estimator is $\hat{\beta}_{TS2SLS} = (n_1^{-1} Z_1^T W_1 X_1)^{-1} C (n_2^{-1} Z_2^T W_2 Y_s)$, where $C \equiv (n_1^{-1} Z_1^T W_1 Z_1)(n_2^{-1} Z_2^T W_2 Z_2)^{-1}$. Notice that the matrix C distinguishes between TSIV and TS2SLS. One can also intuitively see that this correction term is likely to work as the moment between X and Z can be reasonably well approximated by the moment between Z and Z.

Third, the assumption of the first-stage equation may be too restrictive for certain applications. In our application, we have constructed Z using a proxy of the disability status, which is unlikely to satisfy $X = Z\Gamma + \Delta$. This is not a drawback of the TS2SLS estimator. As we shall argue later, it does not necessarily require the first-stage equation since it is a valid estimator under alternative assumptions. On the other hand, if we are willing to assume the first-stage equation, the standard ELL technique could be used because OLS of z on x does no obvious harm for the purpose of imputation.

Fourth, Ω_2 is an unknown parameter and is not readily computable in general. An alternative way to estimate $\widehat{Var}[\hat{\beta}]$ is to independently bootstrap the two samples and repeatedly estimate $\hat{\beta}$. While this is a useful way to find an estimate of $Var[\hat{\beta}]$, it does not work well for the ELL small-area estimation because we still need estimates of $\sigma_{\epsilon,g}$ and σ_η for imputing y_h.

Hence, we use the $\hat{\sigma}_{\epsilon,g}$ and $\hat{\sigma}_\eta$ taken from the Aggregation Method. Let $B \equiv (a_{hg})$

be an $n \times G$ household group matrix. Then, we can get a consistent estimate $\hat{\Omega}$ using census as follows: $diag(B \cdot (\hat{\sigma}^2_{\epsilon,1}, \cdots, \hat{\sigma}^2_{\epsilon,G})^T) + \hat{\sigma}^2_\eta K^T K$. Replacing $n_2^{-1} Z_2^T W_2 \Omega_2 W_2 Z_2$ by $n_1^{-1} Z_1^T W_1 \hat{\Omega}_1 W_1 Z_1$ in Eq(5), we have $\widehat{Var}[\hat{\beta}_{TSIV}]$. Note that we need to use the census (Sample 1) here because only the census contains B.

Fifth, Aggregation Method and Instrumental Variable Method use the same estimates of $\sigma_{\epsilon,g}$ and σ_η. Therefore, the differences in poverty estimates only result from the differences in the estimates of $\hat{\beta}$ and $\widehat{Var}[\hat{\beta}]$.

Sixth, the consistency result rests on the sample moments from the two samples converging to the same moment. Therefore, the TSIV estimator is vulnerable to sample-specific shocks. Suppose for example that, instead of z_h, we observe $\tilde{z}_h \equiv z_h + f_h$ in Sample 2, where f_h is a non-stochastic L-row vector. Clearly, TSIV is no longer consistent in this case, because $n_2^{-1} \tilde{Z}_2^T W_2 X_2$ does not in general converge to Q^0 as $C \to \infty$. Note that the bias induced by the sample-specific shock may not be negligible, even if f_h is small relative to z_h so that $f_{h,l} \ll z_{h,l}$ for $\forall l \in \{1, \cdots, L\}$,

For practical applications, such sample-specific shocks may be important. It is often the case that two samples are taken at different points in time so that the observed variables are referenced to different points in time. Age is a typical example that shifts over time. When we have two large samples, the existence of f_h poses little problem as we can adjust the first moment easily. That is, we can simply subtract the sample mean of $\tilde{z}_h$ in Sample 2 and add the sample mean of z_h in Sample 1. Hence, replacing Z_2 by the moment adjusted observations ζ_2 in Eq(4), we have the first-moment-adjusted TSIV estimator (MATSIV1) estimator.

Adjustment of this sort could be carried out for higher order moments. For example, we can calculated a TSIV estimator adjust for the mean and variance of z_h, which we call MATSIV2. However, as we shall discuss in the next section, MATSIV2 does not perform better than MATSIV1. This is because higher-order sample moments are generally less stable.

Let us now write the MATSIV estimator formally. Let ι_i be an n_i-column vector of ones for sample $i \in \{1, 2\}$. We define the mean vector $\bar{Z}_i \equiv (\iota_i^T W_i \iota_i)^{-1} (\iota_i^T W_i Z_i)$, and the variance

matrix $\Sigma_{Z_i} \equiv (Z_i - \iota_i \bar{Z}_i)^T W_i (Z_i - \iota_i \bar{Z}_i)$. Then, TSIV, TS2SLS, MATSIV1 and MATSIV2 can all be expressed in the following generic form:

$$\begin{aligned} \hat{\beta}_{MATSIV} &= (n_1^{-1} Z_1^T W_1 X_1)^{-1} (n_2^{-1} \zeta_2^T W_2 Y_2) \quad \text{with} \\ \widehat{Var}[\hat{\beta}_{MATSIV}] &= n_2^{-1} (n_1^{-1} Z_1^T W_1 X_1)^{-1} (n_1^{-1} \zeta_1^T W_1 \hat{\Omega}_1 W_1 \zeta_1)(n_1^{-1} X_1^T W_1 Z_1)^{-1}, \end{aligned} \tag{7}$$

where, for $i \in \{1, 2\}$, $\zeta_i = Z_i$ for TSIV, $\zeta_i = Z_i C^T$ for TS2SLS, $\zeta_i = Z_i + \iota_i(\bar{Z}_1 - \bar{Z}_i)$ for MATSIV1, and $(Z_i - \iota_i \bar{Z}_i)\Sigma_{Z_i}^{-1/2} \Sigma_{Z_1}^{1/2} + \iota_i \bar{Z}_1$ for MATSIV2.

We have four points to make. First, it is straightforward to verify that $\zeta_1 = Z_1$ holds for TSIV, MATSIV1 and MATSIV2. Similarly, by construction, we have $\bar{Z}_1 = \bar{\zeta}_2$ for MATSIV1 and MATSIV2, and $\Sigma_{Z_1} = \Sigma_{\zeta_2}$ for MATSIV2. This implies that MATSIV1 and MATSIV2 are robust to the "shifting" of z_h. In addition, MATSIV2 is robust to the "scaling" of z_h.

Second, the TS2SLS estimator can be interpreted as a valid MATSIV estimator by modifying the usual assumptions. That is, instead of the first-stage regression equation for TS2SLS, we can assume that $Z_i^T W_i Z_i \xrightarrow{p} Q^2$ for each $i \in \{1, 2\}$ as $C \to \infty$ where Q^2 is a positive definite symmetric matrix.

Third, MATSIV1 and MATSIV2 are robust with respect to sample-specific shocks that only affect the measurement of z_h. However, they are not robust to shocks that affect the measurement of both z_h and y_h. Suppose, for example, that we observe $\tilde{z}_h \equiv k z_h$ and $\tilde{y}_h \equiv k y_h$ instead of z_h and y_h, where k is a positive constant. Then, TSIV, MATSIV and MATSIV are not consistent in the presence of the sample-specific multiplicative shock. On the other hand, TS2SLS is still consistent because $\tilde{\zeta}_2^T W_2 Y_2 == \tilde{Z}_2 (n_2^{-1} \tilde{Z}_2^T W_2 \tilde{Z}_2)^{-1} (n_1^{-1} Z_1^T W_1 Z_1) W_2 \tilde{Y} = Z_2 (n_2^{-1} Z_2^T W_2 Z_2)^{-1} (n_1^{-1} Z_1^T W_1 Z_1) W_2 Y = \zeta_2^T W_2 Y_2$. Therefore, MATSIV1, MATSIV2 and TS2SLS possess different types of robustness to sample-specific shocks.

Fourth, the calculation of Eq(7) is conditional on an unobservable variable X_2 (along with observable variables X_1, Z_1, and Z_2). For the purpose of imputation, this is reasonable because we indeed have observations for both x_h and z_h in the census. However, Eq(7) does not give the correct variance in the usual TS2SLS framework (See, Fujii and van der Weide (2007)), in which the first-stage regression equation is assumed, because X_2 must be treated

as a random variable.

So far, we have said nothing about the choice of proxy variable $\tilde{a}_h$. We consider two alternative strategies that have some general applicability. First alternative is the conditional probability of being in each group. For example, we can assume that the probability of household being in group g conditional on v_h is $Prob_g(v_h;\theta)$ where v_h is a vector of household characteristics and θ is the parameter of the model. v_h must be common between survey and census, so that θ can be estimated using the census data. We can impute the probability in the survey by $\hat{a}_{hg} \equiv Prob_t(v_h;\hat{\theta})$.

An alternative is the average of the group dummies in the cluster. That is, we can let $\hat{a}_{hg} = \bar{a}_{\kappa(h)g}$, where $\bar{a}_{cg} \equiv \frac{1}{\bar{w}_c}\sum_{h\in\mathcal{H}_c} w_h a_{hg}$. This is straightforward to implement as we have disability information in the census. This has another advantage for the comparison between aggregation method and instrumental variable method, because we use exactly the same set of information.

Once we have $\hat{\beta}$ and $\widehat{Var}[\hat{\beta}]$, we can carry out the Monte-Carlo simulation. The procedure is almost the same as the ELL small-area estimation. We draw $\tilde{\beta}^{(r)}$ for the r-the round of simulation from a normal with mean $\hat{\beta}$ and variance $\widehat{Var}[\hat{\beta}]$. The variances of the error terms, $\tilde{\sigma}_\eta^{2,(r)}$ and $\tilde{\sigma}_\epsilon^{2,(r)}$, are jointly drawn from the residual regression estimates, while ensuring both are non-negative. Cluster-specific random effect and household-specific random effect $\tilde{\epsilon}_{hg}^{(r)}$ are drawn from the (standardized) empirical distribution and augmented by $\tilde{\sigma}_\eta^{(r)}$ and $\tilde{\sigma}_\epsilon^{(r)}$ respectively. Once we have drawn these parameters, we can calculate $\tilde{y}_h^{(r)}$ for each census record. The remaining steps are the same as the ELL small-area estimation.

The assumptions we made in this section are fairly general. While the empirical focus of this study is on disabled households, both the Aggregation Method and the Instrumental Variable Method are potentially applicable to many other issues, in which the group information is not available but possibly important and can be approximated.

5 Empirical Results

We have first used the Aggregation Method. This method requires us to merge the survey and census at the cluster-level. Because the census and survey administration codes at the level of enumeration areas are not fully unified, we take a ward as a cluster instead. The census and survey data used in this study have 2,457 and 801 wards respectively. As a result, the standard errors in this study may be slightly upward biased, because cluster-level random effect would have cancelled out at the level of enumeration area. However, this effect is relatively small because most of the survey wards have only one enumeration area.

In principle, we could construct different consumption models for different zones or regions. However, we constructed a single consumption model for the mainland Tanzania, because we had relatively small number of wards in the survey. This is, however, not a major drawback as we are interested in the poverty statistics at the level of mainland Tanzania. The regression results for the Aggregation Method (*i.e.* Eq(3)) are given in Column (1) of Table 3.

The estimates of σ^2_η and $\sigma^2_{\epsilon,g}$ are given in Table 4, which are used for both the Aggregate Method and the Instrumental Variable Method. There are two points to note here. First, the variance of the cluster-specific random effect σ^2_η is statistically significant, but it is much smaller than the magnitude of the individual-specific effects. Second, while the difference between $\sigma^2_{\epsilon,\text{non-disabled}}$ and $\sigma^2_{\epsilon,\text{disabled}}$ is not statistically significant at a conventional level, $\sigma^2_{\epsilon,\text{disabled}}$ is much higher than $\sigma^2_{\epsilon,\text{non-disabled}}$. This is consistent with our earlier conjecture that there may exist household-specific random effects that are only relevant to disabled households.

We used an identical set of regressors for the Instrumental Variable Method to allow for direct comparisons. For the Instrumental Variable Method, we have tried two different instruments. First, using the census, we ran a logit regression of the disability status a_h on a variety of household characteristics v_h that are common between the census and survey. Then, we calculated $\hat{a}_h$ and z_h for both samples. The results for the logit regression are shown in Table 6 in the Appendix. This study has only two groups, but we can use a multinomial logit

Table 3: Regression estimation results for Aggregation Method and Instrumental Variable Method (A).

Variable	(1) AGG		(2) TSIV-A		(3) MATSIV1-A		(4) MATSIV2-A		(5) TS2SLS-A	
	Coef	S.E.	Coef	S.E.	Coef	S.E.	Coef	S.E.	Coef	S.E.
Constant	10.363 ***	(0.364)	-10.697 ***	(0.208)	9.048 ***	(0.208)	10.779 ***	(0.208)	9.583 ***	(0.453)
Head's age	-0.045 ***	(0.014)	0.891 ***	(0.004)	-0.009 **	(0.004)	-0.120 ***	(0.004)	-0.022	(0.014)
Head's age sq. /1000	0.361 ***	(0.130)	-7.295 ***	(0.044)	0.072	(0.044)	1.048 ***	(0.044)	0.177	(0.113)
Light from paraffin	-0.121	(0.111)	-2.512 ***	(0.087)	0.105	(0.087)	0.271 ***	(0.087)	0.017	(0.092)
Light from electricity	-0.354	(0.313)	-8.308 ***	(0.149)	0.116	(0.149)	0.924 ***	(0.149)	0.127	(0.168)
Electricity available	0.133	(0.289)	7.799 ***	(0.119)	0.111	(0.119)	-0.243 **	(0.119)	-0.032	(0.113)
No toilet	-0.090	(0.119)	-1.182 ***	(0.091)	-0.008	(0.091)	-0.627 ***	(0.091)	-0.109	(0.121)
Earth floor	-0.211 **	(0.095)	0.450 ***	(0.089)	-0.142	(0.089)	-1.171 ***	(0.089)	-0.167 *	(0.097)
Piped water	-0.007	(0.046)	2.904 ***	(0.112)	0.065	(0.112)	0.628 ***	(0.112)	0.091	(0.103)
River/dam/lake water	0.032	(0.058)	-0.015	(0.106)	0.132	(0.106)	0.938 ***	(0.106)	0.124	(0.115)
Rain/protected spring	-0.080	(0.064)	-1.138 ***	(0.089)	0.026	(0.089)	0.212 **	(0.089)	0.032	(0.104)
Bamboo & grass roofs	-0.136 **	(0.069)	2.121 ***	(0.115)	-0.128	(0.115)	-0.658 ***	(0.115)	-0.077	(0.127)
Mud & grass roofs	0.120	(0.090)	-0.953 ***	(0.113)	0.087	(0.113)	0.911 ***	(0.113)	0.132	(0.129)
Concrete roofs	0.518 **	(0.209)	-1.788 ***	(0.295)	0.282	(0.295)	1.939 ***	(0.295)	0.287	(0.347)
Total number of rooms	0.024	(0.026)	-0.850 ***	(0.013)	-0.006	(0.013)	-0.124 ***	(0.013)	-0.035	(0.029)
Have telephone	0.614 **	(0.271)	-9.011 ***	(0.129)	-0.159	(0.129)	1.450 ***	(0.129)	0.353	(0.452)
Have radio	0.225 ***	(0.070)	2.558 ***	(0.035)	0.121 ***	(0.035)	0.904 ***	(0.035)	0.117 ***	(0.042)
Head's education	0.013	(0.011)	0.152 ***	(0.005)	0.029 ***	(0.005)	0.165 ***	(0.005)	0.023 ***	(0.007)
Tanga region	0.099 *	(0.057)	-0.976 ***	(0.226)	0.030	(0.226)	0.671 ***	(0.226)	0.117	(0.239)
Mbeya region	0.117 **	(0.059)	-0.754 ***	(0.224)	0.351	(0.224)	1.948 ***	(0.224)	0.246	(0.234)
Singida region	-0.103	(0.075)	3.119 ***	(0.288)	-0.297	(0.288)	-1.593 ***	(0.288)	-0.188	(0.246)
Tabora region	0.152 **	(0.059)	-0.299	(0.235)	0.270	(0.235)	1.707 ***	(0.235)	0.221	(0.236)
Rukwa region	0.207 ***	(0.076)	-0.405	(0.343)	0.149	(0.343)	0.798 **	(0.343)	0.103	(0.346)
Disabled household	0.066	(0.725)	-67.438 ***	(1.134)	0.637	(1.134)	2.658 **	(1.134)	0.586	(3.178)

* Significant at a 10% level
** Significant at a 5% level
*** Significant at a 1% level

Table 4: Residual regression results. Standard errors are heteroskedasticity-consistent.

Variable	Coef	S.E.
σ^2_η	0.090 ***	(0.008)
$\sigma^2_{\epsilon,\text{non-disabled}}$	0.835	(2.183)
$\sigma^2_{\epsilon,\text{disabled}}$	28.047	(42.843)

*** Significant at a 1% level.

model when we have more than two groups. The regression results for the TSIV, MATSIV1, MATSIV2 and TS2SLS estimators are given in Columns (2)-(5) in Table 3. The standard errors for TSIV, MATSIV1 and MATSIV2 are identical by construction. We added the suffix "-A" to clarify that the instrumental variables are based on the logit regression model.

We use the suffix "-B" for the alternative instrumental variable that is constructed from the cluster-average of the group dummies $\bar{a}_{\kappa(h)g}$. Since this instrument is weak and the results seem unreliable, we prefer the results that use the logit regression. However, this method uses the same proxy variable as the Aggregate Method. Hence, the results for the alternative instrumental variable allows us to make a fair comparison between the Aggregate Method and the Instrumental Variable Method. The regression results based on the alternative instrumental variable are provided in Table 7 in the Appendix.

While we cannot decisively conclude from Tables 3 and 7 which estimator we should choose, they provide us with some insights into which estimators are likely to perform better in practice. There are four points worth making here.

First, the TSIV estimator does not seem to give us reasonable results, whichever instrumental variables are used. Almost all coefficients are highly significant, but their absolute values seem unreasonably large. Further, comparison between TSIV-A and TSIV-B suggest that the TSIV estimator is very unstable. Given that the TSIV estimator is vulnerable to various kinds of sample-specific shocks, this is not a surprising result. Compared with TSIV estimator, MATSIV1, MATSIV2 and TS2SLS estimators all are more robust. In particular, the results for TS2SLS-A and TS2SLS-B are remarkably similar.

Second, the signs of the coefficients for MATSIV2 and TS2SLS estimators are identical for both instrumental variables. MATSIV1 also has a similar pattern of signs. However, if we take into account the absolute value of the estimated coefficients, the estimates for TS2SLS is closer to MATSIV1 than MATSIV2.

Third, one yardstick for judging the performance of different estimators is whether the significant coefficients have the "right" sign. In this criterion, AGG, MATSIV1-A, TS2SLS-A, MATSIV1-B, and TS2SLS-B pass this test.

Table 5: Poverty estimates based on different methods. All the numbers are in percentage.

	AGG		TSIV-A		MATSIV1-A		MATSIV2-A		TS2SLS-A		TSIV-B		MATSIV1-B		MATSIV2-B		TS2SLS-B	
	Est.	S.E.	Est.	S.E.	Est.	S.E.	Est.	S.E.	Est.	S.E.	Est.	S.E.	Est.	S.E.	Est.	S.E.	Est.	S.E.
Non-Disabled	40.4	(6.4)	32.6	(1.6)	42.8	(6.2)	50.0	(1.4)	42.0	(6.8)	60.3	(1.4)	44.3	(6.3)	51.0	(1.8)	42.2	(7.4)
Disabled	50.2	(5.2)	93.8	(7.6)	48.7	(7.0)	46.6	(6.8)	48.5	(10.8)	7.8	(8.6)	45.6	(10.8)	41.3	(11.5)	48.6	(10.0)
Total	40.7	(6.2)	34.3	(1.6)	43.0	(6.0)	49.9	(1.4)	42.2	(6.5)	58.9	(1.4)	44.4	(6.0)	50.7	(1.7)	42.4	(7.1)

Fourth, it is worth noting that the coefficient on disabled household is positive for most models. This, of course, does not mean that disabled households are better off. This may be simply because of statistical error. It is possible that, conditional on housing conditions and other household characteristics, the disabled households actually are better off. This is consistent with poor living conditions for disabled households we found in Table 1.

Using the estimates given in Tables 3, 4, and 7 as well as the variance-covariance matrix associated with them, we carried out Monte-Carlo simulation as described in the previous section for 400 times. The point estimates and the standard errors for the poverty rate are summarized in Table 5. We should note five points here.

First, the results for TSIV-A, MATSIV2-A, TSIV-B, and MATSIV2-B are significantly different from the survey-only estimate of 35.7 percent and hence not consistent with the survey-only estimates. Hence, AGG, MATSIV1-A, TS2SLS-A, MATSIV1-B and TS2SLS-B are more preferable in this regard.

Second, the difference in poverty rates between non-disabled and disabled households is significant at 5 percent level for the Aggregation Method and TSIV-A. However, the difference is not statistically significant for all the other methods. Hence, among our preferred estimates, AGG allows us to compare the two groups in the sharpest manner.

Third, among the Instrumental Variable Method estimators, TS2SLS is most robust. MATSIV1, on the other hand, is not as robust as TS2SLS. The results for TS2SLS-B are encouraging as TS2SLS works even with a weak instrument such as the cluster-average of the group dummies. This finding is consistent with Inoue and Solon (2006) and Fujii and van der Weide (2007) that show the advantages of TS2SLS.

Fourth, under our assumptions, the method proposed by Hoogeveen (2005) does not produce correct point estimates or standard errors. Hence, the poverty estimates reported by Lindeboom (2005) are also biased. However, the bias for the point estimates is small compared with the magnitude of standard errors in our study. His estimates of 40.4 percent and 33.9 percent for disabled and non-disabled households are not significantly different from our AGG, TS2SLS-A and TS2SLS-B estimates. Further, the difference in poverty rates

between the two groups is also similar between this study and his study. Both studies imply a difference of about 6 to 10 percentage points.

However, these observations do not necessarily warrant the results of Lindeboom (2005). The standard errors for the poverty rates reported in Lindeboom (2005) are 0.9 percent and 0.7 percent for disabled and non-disabled households. These numbers seem to be misleadingly small, given that we are unable to observe the disability status in the survey.

Finally, our AGG, TS2SLS-A and TS2SLS-B estimates are closer to the numbers we would expect from Table 2 than the estimate by Lindeboom (2005) is. The poverty rate for disabled households is likely smaller than that for economically disabled households, because the heads in the latter households would be suffering from severer disability. Hence, we would expect the point estimate to be slightly smaller than 52.7 percent, but not much smaller. Given this, our AGG, TS2SLS-A and TS2SLS-B estimates seem to be in the right range. The balance of evidence suggests that these results are quite plausible.

6 Discussion and Conclusion

This paper developed methods to estimate the poverty rates for a small group such as disabled households. Our methods are built upon the ELL small-area estimation. However, the ELL estimation has the limitation that we need to have all the regressors in both of the two samples. This limitation can be problematic when we want to compare different types of groups that may be systematically different.

We overcome this issue in two different ways. One way is by aggregation and the other way is by applying the two-sample instrumental variable regression. Since all the methods we considered are derived from the same equation, Eq(1), we can make a fair comparison between them. We applied our methods to Tanzania using a census and a survey. Close inspection of our results indicates some of our methods work better than others.

Most notably, the standard TSIV estimator does not produce reasonable results, because it is vulnerable to sample-specific shocks. MATSIV1 and MATSIV2 estimators are more robust with respect to the choice of instrumental variables than the TSIV estimator. However,

MATSIV2 also did not produce reasonable results, because the second-order moment is not as reliable as the first-order moment. MATSIV1 produced plausible results, but it does not work well with weak instruments.

The AGG and the TS2SLS estimators seem to work best among all the methods we considered. Further, both AGG and TS2SLS yielded results that are consistent with each other and with what we would expect from the survey-only estimation. Using the aggregation method, we can conclude the poverty rate for disabled households is significantly higher than that for non-disabled households at a 5 percent level. Given these, our preferred results in this paper is the aggregation method.

Both the regression estimates and poverty estimates indicate that TS2SLS estimator is very insensitive to the choice of instrumental variables, compared with TSIV, MATSIV1 and MATSIV2. While our results are consistent with the recent findings about the superiority of the TS2SLS estimator, it is not necessarily the case that TS2SLS performs better than MATSIV1 or MATSIV2. They are robust to different types of sample-specific shocks, and, under certain circumstances, MATSIV1 or MATSIV2 may perform better than TS2SLS. One may always want to try TS2SLS to start with, but it is also useful to examine whether there may be sample-specific shocks that MATSIV1 and MATSIV2 are more robust to.

While our empirical results show that the aggregation method produces slightly better results than TS2SLS, this is not necessarily the case. One way to judge whether the aggregation method works better than the instrumental variable method is to consider where the variations lie. It is obvious that, when there is no variations in the cluster-level prevalence of disability $\bar{a}_{cg}$, the aggregation method will not work. On the other hand, the instrumental variable method works well even when there is no variations in $\bar{a}_{cg}$ across clusters, so long as there are variations within each cluster and a good proxy variable $\tilde{a}_h$.

Our finding that the disabled households are indeed poorer than the non-disabled households is not particularly surprising. However, it is worth reiterating two points. First, our results are consistent with the earlier study by Lindeboom (2005), which uses the methodology proposed by Hoogeveen (2005). However, the standard errors reported in Lindeboom

(2005) does not take into account the possibility of systematic differences across the two groups, and are likely severely biased downward.

Second, the point estimate for disabled households reported in Lindeboom (2005) seems too small given the numbers in Table 2. The balance of evidence suggests that our results for the aggregation method are more reliable.

Our empirical results indicate that both the aggregation method and instrumental variable method work well and can be useful. They allow us to make comparison of poverty rates between disabled and non-disabled households with consistent standard errors. Our method is applicable in a variety of situations as the issue of missing group information is common.

References

Action on Disability and Development (1997) 'Tour report of visit to Cambodia.' Unpublished manuscript, Action on Disability and Development

Angrist, J., and A. Krueger (1992) 'The effect of age at school entry on educational attainment: An application of instrumental variables with moments from two samples.' *Journal of the American Statistical Association* 87(418), 328–336

Bjorklund, A., and M. Jantti (1997) 'Intergenerational income mobility in Sweden compared to the United States.' *American Economic Review* 87(5), 1009–1018

Currie, J., and A. Yelowitz (2000) 'Are public housing projects good for kids?' *Journal of Public Economics.* 75, 99–124

Dee, T.S., and W.N. Evans (2003) 'Teen drinking and educational attainment: Evidence from two-sample instrumental variable estimates.' *Journal of Labor Economics* 21(1), 178–209

Demombynes, G., and B. Özler (2005) 'Crime and local inequality in South Africa.' *Journal of Development Economics* 76(2), 265–292

Elbers, C., J.O. Lanjouw, and P. Lanjouw (2002) 'Micro-level estimation of welfare.' Policy Research Department Working Paper 2911, The World Bank

_ (2003) 'Micro-level estimation of poverty and inequality.' *Econometrica* 71(1), 355–364

_ (2005) 'Imputed welfare estimates in regression analysis.' *Journal of Economic Geography* 5(1), 101–118

Elbers, C., P. Lanjouw, A. Mistiaen, B. Özler, and K. Simler (2004) 'On the unequal inequality of poor communities.' *World Bank Economic Review* 18(3), 401–421

Elbers, C., T. Fujii, P. Lanjouw, B. Özler, and W. Yin (2007) 'Poverty alleviation through geographic targeting.' *Journal of Development Economics* 83(1), 198–213

Elwan, Ann (1999) 'Poverty and disability: A survey of the literature.' SP Discussion Paper 9932, The World Bank

Feige, E.L., and H.W. Watts (1972) 'An investigation of the consequences of partial aggregation of micro-economic data.' *Econometrica* 40(2), 343–360

Filmer, D. (2005) 'Disability, poverty and schooling in developing countries: Results from 11 household surveys.' SP Discussion Paper 0539, The World Bank

Foster, J., J. Greer, and E. Thorbecke (1984) 'A class of decomposable poverty measures.' *Econometrica* 52(3), 761–766

Fujii, T. (2008) 'How well can we target aid with rapidly collected data? empirical results for poverty mapping from cambodia.' *World Development*

Fujii, T., and R. van der Weide (2007) 'Optimal linear instrumental variable estimator.' mimeo, Singapore Management University and the World Bank

Haveman, R., and B. Wolfe (2000) 'The economics of disability and disability policy.' In *Handbook of Health Economics: Volume 1B,* ed. A.J. Culyer and J.P. Newhouse pp. 995–1051

Hoogeveen, J. (2005) 'Measuring welfare for small but vulnerable groups poverty and disability in Uganda.' *Journal of African Economies* 14(4), 603–631

Inoue, A., and G. Solon (2006) 'Two-sample instrumental variable estimators.' NBER Technical Working Paper 0311, National Bureau of Economic Research

Kisanji, J. (1995a) 'Interface between culture and disability in the Tanzanian context: Part i.' *International Journal of Disability, Development and Education* 42(2), 93–108

_ (1995b) 'Interface between culture and disability in the Tanzanian context: Part ii.' *International Journal of Disability, Development and Education* 42(2), 109–124

Lindeboom, W. (2005) 'Disability and poverty in Tanzania.' mimeo, World Bank

Lusardi, A. (1996) 'Permanent income, current income, and consumption: Evidence from two panel data sets.' *Journal of Business & Economic Statistics* 14(1), 81–90

Masset, E., and H. White (2004) 'Are chronically poor people being left out of progress towards the millennium development goals? a quantitative analysis of older people, disabled people and orphans.' *Journal of Human Development* 5(2), 279–297

Matuja, W.B.P., and H.T. Rwiza (1994) 'Knowledge, attitude and practice (KAP) towards epilepsy in secodary students in Tanzania.' *Central African Journal of Medicine*

Mitra, S. (2004) 'Disability and social safety nets in developing countries.' SP Discussion Paper 0509, The World Bank

National Bureau of Statistics (2002) 'Household budget survey 2000/01.' Technical Report, National Bureau of Statistics, President's Office, Planning and Privatization, United Republic of Tanzania

_ (2003) '2002 population and housing census: general report.' Technical Report, National Bureau of Statistics, President's Office, Planning and Privatization, United Republic of Tanzania

Polinsky, A.M. (1977) 'The demand for housing: A study in specification and grouping.' *Econometrica* 45(2), 447–461

Tarozzi, A., and A. Deaton (2007) 'Using census and survey data to estimate poverty and inequality for small areas.' Working Paper 998, Industrial Relations Section, Department of Economics, Princeton University

Taylor, H.R., S. Katala, B. Muñoz, and V. Turner (1991) 'Increase in mortality associated with blindness in rural Africa.' *Bulletin of the World Health Organization* 69(3), 335–338

UNICEF (1999) *Children in Need of Special Protection Measures* (Dar es Salaam, Tanzania)

Welsch, R.E., and E. Kuh (1976) 'The variances of regression coefficient estimates using aggregate data.' *Econometrica* 44(2), 353–363

White, H. (1984) *Asymptotic Theory for Econometricians* (Academic Press)

Yeo, R., and K. Moore (2003) 'Including disabled people in poverty reduction work: "nothing about us, without us".' *World Development* 31(3), 571–590

Appendix: Additional Estimation Results

Table 6: Logit Regression for the Instrumental Variable Method (A).

Variable	Coef	S.E.
Constant	4.587***	(0.081)
Have radio	0.169***	(0.014)
Have bike	0.202***	(0.015)
Have iron	0.166***	(0.037)
Cooking paraffin	0.343***	(0.048)
Cooking firewood	-0.242***	(0.026)
Light from electricity	0.865***	(0.068)
Light from paraffin	0.551***	(0.059)
Light from firewood	0.319***	(0.062)
Rain/protected spring	-0.042***	(0.015)
River/dam/lake water	-0.046***	(0.016)
No toilet	-0.071***	(0.019)
Concrete/cement/tile/timber floor	0.140***	(0.023)
Mud walls	-0.100***	(0.013)
Baked brick walls	-0.037*	(0.020)
Mud & grass roofs	-0.106***	(0.017)
Metal sheets roofs	0.100***	(0.016)
Household size	0.023***	(0.004)
Household size sq/1000	-0.630***	(0.000)
Head's age	-0.019***	(0.001)
Head is male	-0.657***	(0.015)
Head is married	0.466***	(0.015)
Head's educ in yrs	0.081***	(0.006)
Head's educ sq/1000	-2.990***	(0.000)
Average age in hh	-0.009***	(0.001)
Fraction illiterate	-0.322***	(0.033)
Dependent ratio	0.029***	(0.001)
Fraction w. any educ	-0.414***	(0.037)
Fraction retired/unemployed	-1.361***	(0.023)
Average educ in cluster	-0.015**	(0.007)

* Significant at a 10% level
** Significant at a 5% level
*** Significant at a 1% level

Table 7: Regression estimation results for the Instrumental Variable Method (B).

Variable	TSIV-B		MATSIV1-B		MATSIV2-B		TS2SLS-B	
	Coef	S.E.	Coef	S.E.	Coef	S.E.	Coef	S.E.
Intercept	-14.118***	(0.231)	9.010***	(0.231)	10.701***	(0.231)	9.573***	(0.506)
Head's age	0.863***	(0.004)	-0.009**	(0.004)	-0.121***	(0.004)	-0.022*	(0.013)
Head's age squared/1000	-0.008***	(0.048)	0.000	(0.048)	0.001***	(0.048)	0.000	(0.121)
Light from paraffin	-0.835***	(0.093)	0.123	(0.093)	0.311***	(0.093)	0.021	(0.097)
Light from electricity	-6.407***	(0.153)	0.137	(0.153)	0.970***	(0.153)	0.132	(0.164)
Electricity available	7.913***	(0.119)	0.112	(0.119)	-0.241**	(0.119)	-0.032	(0.114)
No toilet	-1.469***	(0.092)	-0.011	(0.092)	-0.633***	(0.092)	-0.108	(0.120)
Earth floor	0.160*	(0.087)	-0.145*	(0.087)	-1.178***	(0.087)	-0.168*	(0.093)
Piped water	2.818***	(0.112)	0.064	(0.112)	0.625***	(0.112)	0.088	(0.105)
Water from River/Dam/Lake	-0.148	(0.106)	0.130	(0.106)	0.934***	(0.106)	0.123	(0.115)
Water from rain/protected spring	-1.247***	(0.089)	0.025	(0.089)	0.208**	(0.089)	0.029	(0.107)
Grass leaves & bamboo roofs	1.914***	(0.115)	-0.131	(0.115)	-0.663***	(0.115)	-0.078	(0.127)
Mud & grass roofs	-0.492***	(0.112)	0.092	(0.112)	0.921***	(0.112)	0.131	(0.127)
Concrete roofs	-1.593***	(0.294)	0.284	(0.294)	1.945***	(0.294)	0.288	(0.347)
Total number of rooms	-0.713***	(0.014)	-0.004	(0.014)	-0.121***	(0.014)	-0.034	(0.029)
Have telephone	-8.661***	(0.128)	-0.155	(0.128)	1.457***	(0.128)	0.353	(0.451)
Have radio	3.464***	(0.040)	0.131***	(0.040)	0.924***	(0.040)	0.117***	(0.045)
Head's education	0.252***	(0.006)	0.030***	(0.006)	0.167***	(0.006)	0.023***	(0.008)
Tanga region	0.272	(0.229)	0.044	(0.229)	0.697***	(0.229)	0.114	(0.234)
Mbeya region	0.913***	(0.229)	0.370	(0.229)	1.985***	(0.229)	0.246	(0.237)
Singida region	2.604***	(0.288)	-0.302	(0.288)	-1.599***	(0.288)	-0.180	(0.243)
Tabora region	1.173***	(0.238)	0.287	(0.238)	1.738***	(0.238)	0.219	(0.235)
Rukwa region	1.605***	(0.347)	0.171	(0.347)	0.845**	(0.347)	0.107	(0.354)
Disabled household	57.386***	(2.959)	2.025	(2.959)	5.441*	(2.959)	0.623	(2.661)

* Significant at a 10% level
** Significant at a 5% level
*** Significant at a 1% level