

Fuzzy chronic poverty: A proposed response to Measurement Error for Intertemporal Poverty Measurement

Catherine Porter *
Heriot-Watt University

Gaston Yalonetzky
University of Leeds

June 4, 2017

Abstract

A number of chronic poverty measures are now empirically applied to quantify the prevalence and intensity of chronic poverty, vis-a-vis transient experiences, using panel data. Welfare trajectories over time are assessed in order to identify the chronically poor and distinguish them from the non-poor, or the transiently poor, and assess the extent and intensity of intertemporal poverty. We examine the implications of measurement error in the welfare outcome for some popular discontinuous chronic poverty measures, and propose corrections to these measures that seeks to minimize the consequences of measurement error. The approach is based on a novel criterion for the identification of chronic poverty that draws on fuzzy set theory. We illustrate the empirical relevance of the approach with a panel dataset from rural Ethiopia and some simulations.

JEL Codes: I32, D63

Keywords: Intertemporal poverty, poverty measurement, measurement error, fuzzy sets theory.

*Corresponding author. Catherine.porter@hw.ac.uk. We would like to thank two anonymous referees, Roberto Zelli, Florent Bresson, Natalie Quinn and seminar participants at the IARIW 32nd General Conference, Boston, August, 2012, for very helpful comments and suggestions. Any errors remain our own.

1 Introduction

Chronic poverty measurement has grown in academic and policy interest over the past ten years in particular, not least due to the increasing availability of panel datasets, especially in developing countries. Several intertemporal poverty measures have now been proposed, and are in use in empirical applications around the world. While no one measure has yet become the standard, several extensions of the Foster-Greer-Thorbecke set of static measures (Foster et al., 1984) are currently proposed. Porter and Quinn (2013) review the intertemporal poverty measurement literature, and show that some of the well-established properties of static poverty measurement are not easily extended to the intertemporal context. We do not review all of them here, but we note that several options are available to the poverty analyst. Intertemporal poverty measures have been proposed, inter alia, by Jalan and Ravallion (2000); Porter and Quinn (2008); Hoy and Zheng (2011); Calvo and Dercon (2009); Foster (2009); Bossert et al. (2012); Gradin et al. (2012); Foster and Santos (2013); Dutta et al. (2013).

In particular, there has been a policy interest in quantifying duration of poverty, as well as in identifying those who may be said to be “chronically” poor as opposed to transiently poor. In parallel, chronicity is a concept which many authors in the intertemporal poverty literature have sought to incorporate. This is an appealing concept in the intertemporal context: all other things equal, the length of time spent in poverty may have a more than one-for-one impact on the underlying wellbeing of a person.¹ There is also a direct analogy to the unemployment literature, which shows that spending longer time in unemployment may also decrease the chances of exiting unemployment. The same may be posited regarding poverty. Meanwhile there is a clear consensus on the substantial measurement error present in almost all survey data in developing and developed countries (Deaton and Grosh, 2000; Bound et al., 2001).

The literature has not yet managed to design a *continuous* measure of intertemporal poverty that incorporates an appropriate concept of *duration sensitivity*, even though these

¹See for example several qualitative research papers and summaries from the Chronic Poverty Research Centre, www.chronicpoverty.org.

two properties are not incompatible in theory. Continuity is an important and desirable property of any poverty measure, given that any discontinuity would render the measure highly sensitive to small changes in the wellbeing indicator being used. Of particular concern, the measure would also be sensitive to measurement error generating spurious fluctuations around poverty lines; in turn leading to misclassifications of people as either non (chronic) poor, or (chronic) poor. In the static and multidimensional context, similar concerns have motivated the incorporation of insights from fuzzy set theory (see e.g. Lemmi and Betti, 2006), in order to better identify the poor, and to avoid the problem of setting a poverty line that classifies people as poor or non-poor, with nothing in between (Pritchett, 2006). More recently, Marano et al. (2015) proposed a method to mitigate the potential impact of measurement error on the estimation of persistence and transition probabilities into and out of poverty. Their method combines a fuzzy approach to poverty identification with latent class Markov models. Verma et al. (2015) adopt a fuzzy approach with a continuous fuzzy identification function, applied to EU data.

Our focus in this paper is somewhat pragmatic, building on these insights for application with widely used monetary measures of chronic poverty that speak to the policy literature. We create a “thick” poverty line enabling us to mitigate the potentially excessive sensitivity of discontinuous intertemporal poverty measures to spurious transitions across the poverty line. Specifically we propose a generalization of two popular intertemporal poverty measures: the measure of Foster (2009) and the more recent measure of Gradin et al. (2012). Hence we are not concentrating on the impact of measurement error on transition probabilities, as some of the recent literature. Rather, we are proposing amendments which reduce the influence of measurement error on the quantification of duration-sensitive poverty. The two new proposals are characterized by a lower sensitivity to transitions around the poverty lines. In accordance with fuzzy set theory applied to poverty measurement, our measures allow *some* people to have a fuzzy poverty status, somewhere between being poor and non-poor. We dedicate special attention to monetary measures of wellbeing and to measurement error which may be negatively correlated with the true value of wellbeing indicators (e.g. income or consumption), since the scant

available evidence points to this phenomenon (Gibson et al., 2015). Hence, our concern is that existing intertemporal poverty measures might *underestimate* the extent of chronic poverty, and/or the intensity of duration-sensitive poverty experiences, in the face of this form of measurement error.

While we expect our proposed adjustments to be helpful in the case of several classes of intertemporal poverty measures, we illustrate with one class of measures that aims to identify the chronically poor (Foster, 2009), and with another class which does not distinguish between the chronically poor and the rest (never poor and transiently poor), but rather renders the individual poverty function dependent on the number of contiguous poverty periods (Gradin et al., 2012).

We explore these measures' empirical implications with some simulated data, as well as the Ethiopian Panel Household Survey. In the Ethiopian case-study, we find that a crisp-poverty-line application of the Gradin et al. measures underestimates intertemporal poverty vis-a-vis the fuzzy-poverty-line application. Meanwhile, whether the Foster measure reports higher poverty statistics for the crisp or the fuzzy poverty line depends crucially on the choice of chronic-poverty duration cut-off. The simulations show that in the case of classical or mean-reverting measurement error, the fuzzy measures are closer to the "true" (error free) measures, and that it is possible to minimise exclusion errors by fuzzifying only above the poverty line. Our simulations therefore provide suggestive evidence that the two measures (Foster and Gradin et al.) may *underestimate* chronic poverty and intertemporal poverty, respectively, in the presence of both classic and mean-reverting measurement error.

The rest of the paper proceeds as follows. Firstly, we briefly introduce a few ideas about intertemporal poverty measurement, followed by a basic notion of poverty identification with fuzzy sets. Then we dedicate two sections, respectively, for the new proposals generalizing the measures of Foster (2009) and Gradin et al. (2012). Then we proceed to the empirical section. There, we first provide a discussion of empirical issues in measurement error, followed by the simulations, and then by the application to the Ethiopian dataset. Finally, the paper ends with some concluding remarks.

2 Intertemporal Poverty Measurement

Most intertemporal poverty measures build on the class of measures known as p-alpha, or FGT, which were introduced by Foster et al. (1984) in order to quantify poverty at the individual and social levels in one specific time period. Let $x_n (\in \mathbb{R}_+)$ be the attainment (e.g. income) of household n (out of N households) and z be the poverty line. A person is deemed poor if $x_n < z$. Then the FGT class is given by:²

$$P_\alpha(X) = \frac{1}{N} \sum_{n=1}^N \left(1 - \frac{x_n}{z}\right)^\alpha \mathbb{I}(x_n \leq z), \quad (1)$$

for $\alpha \geq 0$; which satisfy the axioms of FOCUS, ANONYMITY, WEAK MONOTONICITY, WEAK TRANSFER, SUBSET CONSISTENCY and POPULATION SIZE NEUTRALITY. Additionally, the measures satisfy STRICT MONOTONICITY and CONTINUITY when $\alpha > 0$, and STRICT TRANSFER when $\alpha > 1$. They have become very well-known and widely applied.

Now consider a matrix X , whose N rows have information on the wellbeing attainments of N individuals across a time span of T periods. Each column, therefore, hosts the attainment distribution across the population in a specific time period:

$$X = \begin{pmatrix} x_{11} & x_{12} & \cdots & x_{1T} \\ x_{21} & x_{22} & \cdots & x_{2T} \\ \vdots & \vdots & \ddots & \vdots \\ x_{N1} & x_{N2} & \cdots & x_{NT} \end{pmatrix}. \quad (2)$$

A typical attainment element of the matrix is: $x_{nt} (\in \mathbb{R}_+)$, that is, the attainment of individual n in period t . The poverty lines, specific to each period, are denoted by z_t (from a vector of poverty lines, $Z : (z_1, \dots, z_t, \dots, z_T)$), and a person is deemed poor in period t if: $x_{nt} < z_t$. When conceptualising poverty over time, it is useful to think about the **trajectory** of wellbeing attainments experienced by an individual n (Porter and Quinn, 2008), that is, the n th row of the data matrix $\mathbf{x}_n = (x_{n1}, x_{n2}, \dots, x_{nT})$.

² $\mathbb{I}(\cdot)$ is an indicator function equal to 1 whenever the content in parenthesis is true. Otherwise it is equal to 0.

Most intertemporal poverty measures are constructed in two steps. In the first step, an individual measure is devised as a function of the individual trajectory vector \mathbf{x}_n . In the second step, all the individual (intertemporal) poverty measures are aggregated in order to produce a social poverty measure.

Some earlier measures developed in the literature (e.g. Jalan and Ravallion, 1998 which is an extension of Rodgers and Rodgers, 1993) addressed that first step, i.e. the challenge of transforming the trajectory vector into a measure of individual intertemporal poverty, by averaging of the income stream over time, meaning to capture the notion of permanent income. The method is straightforward, and intuitively appealing: if a household's average income lies below the poverty line, then they can be deemed chronically poor. And their poverty gap (along with square poverty gap and further poverty measures) can be calculated also using this average, applying the FGT class of measures in (1).

This approach has been criticised partly because it allows a period of high income to compensate for a period in severe poverty. In response, several authors (Calvo and Dercon, 2009; Foster, 2009) proposed alternative extensions of the static FOCUS axiom (Foster and Shorrocks, 1991) to the intertemporal context. They propose that the principle of STRONG FOCUS should apply to any chronic poverty measure; that is, the poverty measure should not be sensitive to changes in wellbeing, *in any time period when wellbeing is above the poverty line*.³

The concept of STRONG FOCUS is not sufficient to distinguish between the group of static poor and intertemporal poor, which led Foster (2009) to introduce the concept of DURATION SENSITIVITY, which is at the heart of the identification strategy in his measure: only those people who are poor for at least a certain proportion of time qualify as chronic poor. Porter and Quinn (2008) show that this property is incompatible with another property, that penalises depth of poverty and allows a non-zero elasticity of substitution of wellbeing between periods (INTERTEMPORAL TRANSFER). Which of these properties one wishes to incorporate in the analysis is a normative choice, and depends on the policy

³For an alternative view see Dutta et al. (2013). See also Foster and Santos (2013) for an intermediate approach between the perfect compensation implied by the framework of Jalan and Ravallion (1998) and the null compensation implied by the strong focus axiom.

context and the data under consideration.

Two properties that capture the specific concept of chronicity, or length of time spent under the poverty line, have been proposed so far in the literature: the first relates to the total number of time periods spent in poverty, regardless of their order in time. This has been termed DURATION SENSITIVITY by Porter and Quinn (2013) and TIME MONOTONICITY by Foster (2009). The second is contiguity of poverty (introduced by Bossert et al., 2012, whose measure is generalised by Gradin et al., 2012, and recently further developed by Dutta et al., 2013). CONTIGUITY refers to the concept that *consecutive* spells of poverty without any recovery time in between may be more damaging to wellbeing than when there is some recovery time in between. So, for example in a three-period panel, a sequence [poor, poor, non-poor] would be ranked as worse off than [poor, non-poor, poor] for a poverty measure satisfying CONTIGUITY. Both these are appealing normative properties.

However, another property that is highly desirable is CONTINUITY, which is motivated by the notion that an infinitesimal change in wellbeing in any period should lead to no more than an infinitesimal change in the value of the individual trajectory measure, i.e. the evaluation of intertemporal poverty (Porter and Quinn, 2013). If the trajectory ordering is not continuous then we may find trajectories which are ordered in a perverse way (see Quinn (2014) for a further illustration). For empirical applications this is also extremely important: a discontinuous measure would be excessively sensitive to measurement error, at any point of discontinuity.

3 Duration-sensitive Poverty Measures

3.1 The chronic poverty measures of Foster (2009)

We first tackle one of the most popular measures, that proposed by Foster (2009), which has increasingly been adopted in policy applications (e.g. Perez-Mayo, 2009; Nunez Velasquez, 2009), in addition to being the inspiration behind the class of multidimensional poverty indices proposed by Alkire and Foster (2011). Foster proposes a property of

Time Monotonicity, whereby an additional period of poverty experienced by an already chronically poor person should lead to an increase in the poverty measure.

Foster’s measure includes a ‘double cutoff’: (1) A poverty line indicating material deprivation in one time period, and (2) a duration cutoff indicating the minimum proportion of periods in poverty required for a household, or individual, to be deemed chronically poor. This measure satisfies a focus axiom stating its insensitivity to any period-specific deprivation from people who are not identified as chronically poor.

Foster’s chronic poverty measure is based on a deprivation count which has a very simple and intuitive understanding: a person is chronically poor (as opposed to transiently poor, or non-poor) if they experience poverty during a minimum proportion of time periods. A count of deprivation periods is computed weighting each deprivation period with weights v_t , from a vector of weights $V : (v_1, \dots, v_t, \dots, v_T)$, such that: $v_t \in]0, 1[\subset \mathbb{R}_{++}$ | $\sum_{t=1}^T v_t = 1$. Hence the weighted number of deprivation periods suffered by individual n is: $c_n \equiv \sum_{t=1}^T v_t \mathbb{I}(z_t > x_{nt})$.⁴

Foster (2009) identifies the chronically poor as those whose weighted deprivation count is at least as high as the duration cutoff, $\tau \in [0, 1] \subset \mathbb{R}$. The poverty identification function is thus:

$$\varphi(c_n; \tau) \equiv \mathbb{I}(c_n \geq \tau). \quad (3)$$

Then, for an individual poverty function, Foster (2009) proposes a weighted sum of the powered and censored normalized poverty gaps in every period, i.e. the FGT metric (Foster et al., 1984):

$$p(\mathbf{x}_n; Z, V, \tau, \alpha) \equiv \varphi(c_n; \tau) \sum_{t=1}^T v_t \left(1 - \frac{x_{nt}}{z_t}\right)^\alpha \mathbb{I}(z_t > x_{nt}), \quad \alpha \geq 0. \quad (4)$$

Note that, in this approach to chronic poverty measurement, the sequence and timing of poverty spells does not impact the individual measure. Foster (2009) calls this property ‘time anonymity’.

⁴For simplicity, we keep the weights at one for our empirical illustrations, though these could be altered.

Finally, the social poverty measure, P , has a functional form satisfying desirable properties like individual anonymity, population principle, and additive decomposability:

$$P \equiv \frac{1}{N} \sum_{n=1}^N p(\mathbf{x}_n; Z, V, \tau, \alpha). \quad (5)$$

3.2 The intertemporal poverty measures of Gradin et al. (2012)

The second measure under consideration is the one by Gradin et al. (2012) (GDC), which is a generalisation of Bossert et al. (2012). This measure features a different duration-sensitivity property, whereby consecutive spells of poverty are weighted more heavily. Bossert et al. (2012, p. 1) observe that: “[t]he negative effects of being in poverty are cumulative, hence a two-period poverty spell is much harder to handle than two one-period spells that are interrupted by one (or more) period(s) out of poverty.”

Gradin et al. (2012) take a similar approach to that of Foster (2009), in that the measure is an intertemporal sum of FGT per-period poverty measures. However, they do not incorporate the duration cutoff for identification. This means that anyone with any period in poverty at all is included in the set of the intertemporally poor (in the poverty identification literature this would be deemed a *union* identification approach). Put it differently, this measure of intertemporal poverty is not meant to differentiate between chronic and transient poverty.

In order to penalise contiguous periods of poverty Gradin et al. (2012) introduce a weight multiplying the FGT normalized poverty gap. This weight, w_{nt} , depends on the length of a contiguous poverty spell, denoted by s_{nt} . Thereby the same poverty shortfall gets weighted more heavily if it belongs in a longer experience of uninterrupted poverty:

$$p_G(\mathbf{x}_n; Z, S, \alpha) = \left[\frac{1}{T} \sum_{t=1}^T \left(1 - \frac{x_{nt}}{z}\right)^\gamma \mathbb{I}(z > x_{nt}) w_{nt} \right]^\alpha \quad \alpha \geq 0, \quad \gamma \geq 0; \quad (6)$$

where

$$w_{nt} = \left(\frac{s_{nt}}{T} \right)^\beta, \quad \beta > 0. \quad (7)$$

and S is the vector of poverty spells, s_{nt} . So, for example, a single period in poverty en-

ters with a weight of $(1/T)^\beta$; whereas both periods in a two-period spell would be weighted by $(2/T)^\beta$ as in (7). As noted by Porter and Quinn (2013), the Gradin *et al.* measure satisfies WEAK IDENTIFICATION, GENERAL FOCUS, WEAK MONOTONICITY, STRONG FOCUS, RESTRICTED STRICT MONOTONICITY (if $\gamma > 0$) and CONTIGUITY; but not CONTINUITY, NON-DECREASING COMPENSATION or TIME SYMMETRY. Its discontinuities mean that it does not satisfy INTERTEMPORAL TRANSFER or DURATION SENSITIVITY although it does satisfy each of these for certain poverty trajectories. Gradin et al. (2012) note that the measures proposed by Foster (2009) are a special case of their measure if $\beta = 0$ and $\alpha = 1$. However, this is strictly true only if we set all time-period weights equal to $\frac{1}{T}$ and $\tau = 0$ in the framework proposed by Foster (2009).

Finally, the social poverty measure, P , can be constructed by inserting (6) into the general form (5). Gradin et al. (2012) note that the measure by Bossert et al. (2012) is a specific subclass characterized by $\beta = \alpha = 1$.

4 Poverty identification with fuzzy sets

In order to compensate for the potential effects of measurement error on duration-sensitive chronic poverty measures, we propose a generalization of the two measures outlined above, building on the fuzzy set literature. Fuzzy set theory has been used extensively in the social sciences (e.g. see Ragin, 2000; Smithson and Verkuilen, 2006). In the poverty literature, fuzzy set theory was introduced as an alternative identification criterion by researchers who were unhappy with the blunt dichotomy posed by traditional poverty lines for the identification of the poor. Instead, they opted for the membership functions used in fuzzy set theory (see e.g. Lemmi and Betti, 2006). While we do not intend to contest the practice of setting a poverty line for identification purposes, we do worry about the consequences of using a traditional poverty line in chronic poverty assessments based on duration-sensitive measures, when transitions across the line may be taking place spuriously due to measurement error. We discuss the empirical considerations around the likely type of measurement error after this section.

Since traditional measurement error corrections are usually not readily available (for

a comprehensive treatment, see Bound et al., 2001), we propose a fuzzy-style adjustment to the period-specific poverty lines, and then to the identification criteria of both the time-specific poor and the chronically poor. This adjustment smooths out the impact of (potentially spurious) transitions that take place across, and in close proximity to, the poverty lines. In that way, we generalize some of the proposed duration-sensitive measures of chronic poverty.

An illustration of our proposed identification adjustment is in Figure 1 where a traditional poverty line, z , is compared against a ‘thick’ poverty line bounded by z_1 and z_2 , such that $z_1 < z < z_2$. This is the general class of fuzzy poverty lines introduced by Dombi (1990). In a traditional identification approach, a person is deemed poor if his/her income is below z , and non-poor otherwise. Under a fuzzy approach, poverty status ceases to be dichotomic if a person’s income is in the interval $[z_1, z_2]$; for example, in the proposal by Dombi (1990) the membership function in that interval is given by:

$$\pi_{nt} = \frac{(z_2 - x_{nt})^2}{(x_{nt} - z_1)^2 + (z_2 - x_{nt})^2}.$$

Two important features of our application of a fuzzy identification approach to chronic poverty measurement stand out: i) transitions across the vicinity of z do not generate abrupt changes in poverty status when the ‘thick’ poverty line is used. For big changes in poverty status to happen, the magnitude of the transition has to be large enough to cross from below z_1 to above z_2 (or the other way around). In those cases, we assume that the transition is less likely to be spurious (e.g. driven by measurement error). ii) our fuzzy identification approach can be fine-tuned, thus tailored to different concerns around measurement error, by either changing the values of $[z_1, z_2]$ or by changing the parameters that control the shape of the membership function.

As it is clear from definition (3), a change in x_{nt} that modifies the deprivation status in period t , i.e. a transit across z_t , increases or decreases c_n in the amount v_t . In turn such a perturbation may or may not change $\varphi(c_n; \tau)$ from 1 to 0 (or viceversa), in the case of measures like Foster (2009)’s. As long as there is transit across z_t , a change in individual poverty status is possible, *irrespective of the magnitude of the change in x_{nt} that caused the transit.*

However we do not want small, and potentially spurious, changes around z_t to have a significant effect on chronic poverty status. In order to reduce the likelihood of such occurrence, we propose an alternative poverty identification function, which follows definition (3), however, deprivation in any individual period is determined by the fuzzy poverty line introduced by Dombi (1990):

$$\pi_{nt} = \left\{ \begin{array}{lll} 1 & \text{if} & x_{nt} < z_{1t} \\ \frac{(z_{2t}-x_{nt})^2}{(x_{nt}-z_{1t})^2+(z_{2t}-x_{nt})^2} & \text{if} & z_{1t} \leq x_{nt} \leq z_{2t} \\ 0 & \text{if} & x_{nt} > z_{2t} \end{array} \right\}, \quad (8)$$

where $z_{1t} \leq z_t \leq z_{2t}$, i.e. there is now a ‘thick’ poverty line above and below the original. Also, we note here that, if we are particularly concerned with errors of exclusion (i.e. underestimation of the chronically poor), rather than those of inclusion (overestimation), we may wish to set the lower bound of the thick poverty line at z , and an upper bound somewhere above it. The choice of the bounds for the ‘thick’ poverty line is discussed further in the empirical section below.

We could also consider alternative membership functions instead of (8). For instance the class of membership functions proposed by Chakravarty (2006), in which $\pi_{nt} = \left[\frac{z_2 - x_{nt}}{z_2 - z_1} \right]^\theta$ if $z_{1t} \leq x_{nt} \leq z_{2t}$. An illustration of this membership function can be found in figure 3 in annex 6, though we note that these are not twice-differentiable.

PLACE FIGURE 1 HERE

4.1 Poverty identification with fuzzy sets: the case of the measures by Foster (2009)

Drawing on the preceding section, the next step in our proposal to amend the Foster (2009) measures is to redefine the intertemporal deprivation count: $c_n^\pi \equiv \sum_{t=1}^T v_t \pi_{nt}$. Then the new individual poverty function is:

$$p_\pi(\mathbf{x}_n; Z_\pi, V, \tau, \alpha) \equiv \varphi(c_n^\pi; \tau) \sum_{t=1}^T v_t \pi_{nt} \left[1 - \frac{x_{nt}}{z_{t2}}\right]^\alpha \mathbb{I}(z_{t2} > x_{nt}), \alpha \geq 0, \quad (9)$$

where $\varphi(c_n^\pi; \tau) = \mathbb{I}(c_n^\pi \geq \tau)$, and the vector Z_π is now made of trios of poverty lines, one per time period, $Z_\pi := \{z_{11}, z_1, z_{21}; \dots; z_{1t}, z_t, z_{2t}; \dots, z_{1T}, z_T, z_{2T}\}$.

Finally, the new social poverty function is:

$$P_\pi \equiv \frac{1}{N} \sum_{n=1}^N p_\pi(x_n; Z_\pi, V, \tau, \alpha). \quad (10)$$

For the rest of the paper, especially in the empirical application, we mainly focus on the measure with $\alpha = 0$, though we also illustrate the case of $\alpha > 0$ in table 7. Two interesting differences between the class of measures in (9) and the original one by Foster (2009) are worth highlighting. Firstly, our proposal fulfills the original properties of the Foster (2009) measures, in addition, now, to continuity in each period.⁵ Hence, for instance, a transit across z_t is less likely to change c_n^π by a full amount of v_t . The change, Δc_n^π depends now on the magnitude of the change in x_{nt} , Δx_{nt} :

$$\Delta c_n^\pi = [\pi_{nt}(x_{nt} - \Delta x_{nt}) - \pi_{nt}(x_{nt})]v_t. \quad (11)$$

The lower sensitivity of c_n^π to the same change in x_{nt} , as reflected in (11), is the main feature rendering P_π better protected from drastic changes in period-specific deprivation status, and chronic poverty status, due to small and potentially spurious transits across z_t .

However this new specification has other consequences. A second, expectable, difference is that the baseline number of chronically poor people according to P_π in (10) need not coincide with that according to P in (5). For example, in the case of deprived people in period t , the following condition, for continuous variables, establishes the circumstances under which P_π *overstates* the proportion of deprived people in period t :

⁵Though note that the presence of τ retains a discontinuity that we could also in principle remove by fuzzifying, though we do not at this point.

$$\int_{z_{1t}}^{z_t} [1 - \pi(x)] dF(x) < \int_{z_t}^{z_{2t}} \pi(x) dF(x), \quad (12)$$

where $F(x)$ is the cumulative distribution function of x and $\pi(x)$ is the membership function. The left-hand side of (12) measures the ‘loss’ in full deprivation status experienced by those who still have partial deprivation status, i.e. individuals for whom $z_{1t} \leq x_{nt} \leq z_t$. The right-hand side measures the acquired partial deprivation status among individuals who, otherwise, would not be considered deprived in period t , i.e. people for whom $z_t \leq x_{nt} \leq z_{2t}$. Whenever the latter is greater than the former *for every* t , the social poverty headcount is never lower according to P_π .

The case of $\alpha = 0$

When $\alpha = 0$, the individual poverty function in (9) reduces to:

$$p_\pi(\mathbf{x}_n; Z_\pi, V, \tau, 0) = \varphi(c_n^\pi; \tau) \sum_{t=1}^T v_t \pi_{nt}, \quad (13)$$

then, following Foster (2009), the social poverty function can be expressed as the product of the chronic poverty headcount times the average proportion of poverty periods among the chronically poor (hence why it is also known as a *duration-adjusted headcount ratio*):

$$P_{\pi;}(0) \equiv \frac{1}{N} \sum_{n=1}^N p_\pi(x_n; Z_\pi, V, \tau, 0) = H_\pi D_\pi, \quad (14)$$

where:

$$H_\pi \equiv \frac{1}{N} \sum_{n=1}^N \varphi(c_n^\pi; \tau), \quad (15)$$

and:

$$D_\pi \equiv \frac{1}{HN} \sum_{n=1}^N \varphi(c_n^\pi; \tau) c_n^\pi = \frac{P_{\pi;}(0)}{H}. \quad (16)$$

Given that P in definition (14) can be expressed in terms of a chronic poverty headcount (H) multiplied by an average proportion of poverty periods among the chronically poor

(D), then it is clear, from (14), (15), and (16), that a fuzzy identification function can change not only the chronic poverty headcount (H), but also the average duration statistic (D).

Comparison with alternative tools to deal with measurement error: the case of the Foster (2009) measures

Even if ‘thickening’ the poverty line in every period does help mitigate the effects of measurement error in the case of Foster (2009) measures, there are conceivable alternatives for the same purpose, including performing robustness checks with alternative poverty lines (i.e. changing z) and/or alternative duration cut-offs (i.e. changing τ).⁶ These are all legitimate and potentially useful alternatives, in fact some could even be implemented in tandem (e.g. combinations of duration cutoff with fuzzy poverty lines). We illustrate this in our empirical section, by varying τ . As a note of caution, here we would like to illustrate the potentially different effects of some of these possible adjustments on poverty measurement with Foster (2009) measures, vis-a-vis using fuzzy poverty lines.

We will focus the illustration on the case of negative correlation between measurement error and the true value of the wellbeing indicator, which has gained notoriety due to its recent empirical detection in a developing country Gibson et al. (2015), and is connected to the concern for so-called errors of exclusion (i.e. failing to classify poor households as poor). Consider, first, a situation with $z = 100$, and a hypothetical household A observed over $T = 4$ with the following achievement vector: $\mathbf{x}_A = (110, 120, 105, 110)$. Now imagine that the vector of true unobserved values (i.e. purged of measurement error) is: $\mathbf{x}_A^* = (80, 120, 70, 80)$. That is, if we could observe \mathbf{x}_A^* we would classify A as poor with any $\tau \leq 0.75$ (assuming time periods weighted evenly). Clearly, with $z = 100$ and a ‘crisp’ application of the Foster (2009) measure, household A would be ‘off the radar’ for any τ . That is, reducing τ would not be helpful in this circumstance. By contrast, consider a simple fuzzy poverty identification function $\pi_{nt} = \frac{120-x_{nt}}{120-100}$ if $100 \leq x_{nt} \leq 120$, otherwise $\pi_{nt} = 1$ if $x_{nt} < 100$ and $\pi_{nt} = 0$ if $x_{nt} > 120$. In this case, we would get the following

⁶We would like to thank an anonymous referee for highlighting this point.

poverty identification vector (π_{At} for every period): (0.5,0,0.75,0.5). If every time period is weighted equally, this yields a total deprivation score of 0.4375. Therefore, household A would be identified as chronically poor with any $\tau \leq 0.4375$ in a ‘fuzzy’ version of the Foster (2009) measure.

Now consider household B with observed achievement vector: $\mathbf{x}_B = (80, 120, 105, 110)$, and a vector of true unobserved achievements: $\mathbf{x}_B = (70, 120, 90, 95)$. If we could observe \mathbf{x}_B^* we would not classify B as poor with $\tau > 0.75$ (assuming time periods weighted evenly). Now imagine we raise z from 100 to 110. Then we classify B as poor for any $\tau \leq 0.75$. However, if we raise z from 100 to 120, then B would be deemed poor for any τ . By contrast, consider the poverty identification function of the previous example. We would get the following poverty identification vector for B (π_{Bt}): (1,0,0.75,0.5); which yields a total deprivation count of 0.5575. Therefore, household B would be identified as chronically poor with any $\tau \leq 0.5575$. Here raising the poverty line evenly would be more successful in reducing potential error of exclusion, but at the expense of higher error of inclusion. More importantly, if we believe that higher incomes are less likely to be overestimated by measurement error, then a fuzzy approach would enable us to reflect that judgment by attributing a lower membership function value to higher incomes within the fuzzy domain. By contrast, even increases in z would treat all incomes below it in the same manner for poverty identification purposes.

There are certainly many more household situations that could be considered. The point of the above illustration is to highlight the differential impacts of alternative robustness adjustments to the measure, and how a fuzzy approach to the period-specific poverty line can be helpful toward mitigating the impact of measurement error on chronic poverty identification.

4.2 Poverty identification with fuzzy sets: the case of the measures by Gradin et al. (2012)

In the case of the measures by Gradin et al. (2012), the concern with measurement error generating a transit across z_t , in turn changing the deprivation status in period t , is not

that the individual chronic poverty status may be affected, since these measures do not intend to identify the chronically poor from the rest (put it differently, they rely on a union approach to identification, i.e. $\tau = 0$ implicitly). However, as is clear from the weight equation (7), a small disturbance can produce significant changes in the spell variables, i.e. s_{nt} , which in turn affect the weights. This becomes apparent by examining the formula for s_{nt} :

$$s_{nt} = \left[\sum_{i=t-m}^{t+n} \mathbb{I}(z_i > x_{ni}) \right] \left[\prod_{i=t-m}^{t+n} \mathbb{I}(z_i > x_{ni}) \right] \mathbb{I}(z_i \leq x_{n,t-m-1}) \mathbb{I}(z_i \leq x_{n,t+n+1}). \quad (17)$$

As is clear from equation (17), changes in period poverty status, both within $t - m$ and $t + n$, as well as in the immediately adjacent periods ($t - m - 1$, $t + n + 1$), can generate discontinuous changes in s_{nt} . Our proposal seeks to reduce this sensitivity to small changes in x_{nt} generating transit across z_t , by introducing π_{nt} , from equation (8), into expression (17), thereby ‘thickening’ the poverty lines. This yields the following spell value function:

$$s_{nt}^f = \left[\sum_{i=t-m}^{t+n} \pi_{ni} \right] \left[\prod_{i=t-m}^{t+n} \pi_{ni} \right] \mathbb{I}(z_{2i} \leq x_{n,t-m-1}) \mathbb{I}(z_{2i} \leq x_{n,t+n+1}). \quad (18)$$

An illustration

Here we provide one illustration of the impact of ‘thickening’ the poverty lines in the context of the Gradin et al. (2012) poverty measures. The four panels of Figure 2 show the income profiles of an individual over three periods. According to the top left panel, the individual is poor in periods 1 and 3 if poverty line z is used. On the top right panel, the individual’s income in period 2 is lower enough to render him/her poor. Comparing the poverty spells of the two top panels it turns out that: $S^{tl} := (1, 0, 1)$, while $S^{tr} := (3, 3, 3)$ (where ‘tl’ and ‘tr’ denote, respectively, the top left and the top right panels). Let $\alpha = 1$ in equation (6) and g_t^γ be the (FGT) normalized poverty gap in period t . Then $p_G^{tl} = (g_1^\gamma + g_3^\gamma)(\frac{1}{3})^\beta$, $p_G^{tr} = (\sum_{t=1}^3 g_t^\gamma)(\frac{3}{3})^\beta$; and the difference between the two is:

$$\Delta p_G^{top} \equiv p_G^{tr} - p_G^{tl} = 3^{-\beta} [g_2^\gamma (3)^\beta + (g_1^\gamma + g_3^\gamma)(3^\beta - 1^\beta)]. \quad (19)$$

PLACE FIGURE 2 HERE

By contrast, the two bottom panels perform the same comparison but using a ‘thick’ poverty line, between z_1 and z_2 , for period poverty identification, and z for the normalized poverty gaps. Using π_{nt} with the membership function proposed by Chakravarty (2006), it turns out that: $s_t^{bl} = 2 + (\frac{z_2 - x_2}{z_2 - z_1})^\theta \forall t = 1, 2, 3$, while $s_t^{br} = 2 + (\frac{z_2 - x_2 + \epsilon}{z_2 - z_1})^\theta \forall t = 1, 2, 3$ (where “bl” and “br” denote, respectively, the bottom left and the bottom right panels, and ϵ represents the drop in income on the right-half panels). Then $p_G^{bl} = 3^{-\beta} \sum_{t=1}^3 g_t^\gamma (2 + [\frac{z_2 - x_2}{z_2 - z_1}]^\theta)^\beta$, $p_G^{br} = 3^{-\beta} \sum_{t=1}^3 g_t^\gamma (2 + [\frac{z_2 - x_2 + \epsilon}{z_2 - z_1}]^\theta)^\beta$; and the difference between the two is:

$$\Delta p_G^{bot} \equiv p_G^{br} - p_G^{bl} = 3^{-\beta} \sum_{t=1}^3 g_t^\gamma [(2 + [\frac{z_2 - x_2 + \epsilon}{z_2 - z_1}]^\theta)^\beta - (2 + [\frac{z_2 - x_2}{z_2 - z_1}]^\theta)^\beta]. \quad (20)$$

Comparing (19) against (20), it is clear that the impact of ϵ should be milder on Δp_G^{bot} than on Δp_G^{top} as long as: $\theta > 0$, $\beta \geq 1$, $x_2 - \epsilon > z_1$ and $z_2 > x_2$. For instance, when $\beta = \theta = 1$, as in (21):

$$\Delta p_G^{bot}(\beta = \theta = 1) = \frac{\epsilon}{3(z_2 - z_1)} \sum_{t=1}^3 g_t^\gamma < \frac{2}{3}(g_1^\gamma + g_3^\gamma) + g_2^\gamma = \Delta p_G^{top}(\beta = \theta = 1). \quad (21)$$

5 Empirical considerations

5.1 What kind of measurement error do we expect to encounter?

The existence of measurement error in income and consumption is of huge concern in many areas of applied economics (Bound et al., 2001), and the consequences of measurement error for econometric estimation are well documented. Many theoretical studies of the consequences of measurement error tend to assume for convenience that measurement error is ‘classical’, that is, with a mean of zero, and no correlation with the variable(s) of interest (Bound et al., 2001). However, Ravallion (1994) showed that this may have little

effect for static poverty measurement. Classical measurement error should not affect the distribution of the variable. Ravallion shows that if random measurement error is added to two distributions, one of which has higher poverty, then the measurement error should not affect the mean, and the poverty ranking of households should not change. Our analysis above shows that when multiple time periods are involved, then this may not be the case.

What form does measurement error take? In the US, several studies have validated earnings data, and found a high degree of measurement error; further, that the likely form is in fact non-classical: mean-reverting, and serially correlated (Bound and Krueger, 1991; Pischke, 1995). Poverty studies in developing countries (especially outside Latin America) have tended to focus on consumption as a measure of welfare. The rationale is that consumption is likely less subject to systematic measurement error than income or earnings, i.e. due to underreporting (Deaton and Grosh, 2000). Several studies in developing countries have shown the importance of correcting for measurement error when estimating e.g. mobility and poverty dynamics (Antman and McKenzie, 2007; Glewwe, 2007; Lee et al., 2016), however none to our knowledge have applied any methods of measurement-error correction to duration-based chronic poverty measurement.

Given these concerns, how should the analyst choose the bounds for the ‘thick’ poverty line? Traditional applications of the ‘totally fuzzy relative approach’ effectively give each and every individual a non-zero value for the fuzzy poverty measure (Cheli and Lemmi, 1995). We do not proceed in this way, as otherwise we would lose some of the interesting properties characterizing intertemporal poverty measurement (e.g. duration-sensitivity). In an earlier contribution, Cerioli and Zani (1990) propose that a fuzzy poverty measure (based on FGT) could have a subsistence poverty line, z , as the minimum bound, and mean income as the maximum. For our purpose, we are interested in errors of inclusion and exclusion brought about through measurement error. We therefore seek empirical evidence on what the extent of measurement error is likely to be in a consumption survey.

There is relatively little information on the actual parameters of measurement error in a developing country context - e.g. what on average do we expect the mean, variance

and correlation with welfare to be? Beegle et al. (2012) recently conducted a randomized control trial of consumption in the context of a household survey in Tanzania. The authors compare several methods to elicit recall of food consumption. The benchmark is a daily visit to the household with individual diary for each day. The method used by the ERHS survey, 7-day recall at the household level, is also included. The results show that the 7-day recall method is subject to underestimating the level of consumption, by approximately 20%. Using the same experiment, Gibson et al. (2015) also conclude through regression analysis that measurement error is thus mean-reverting, given a negative correlation between the true level of consumption and the amount of error, and substantial (as shown by a higher variance of consumption relative to the benchmark).

Gibson et al. (2015) estimated regressions on alternative consumption data in order to estimate parameters of the following equation:

$$x^* = \theta + \lambda x + v, \tag{22}$$

with x^* the error-ridden consumption, x true consumption. λ shows the correlation between true consumption and the error-ridden consumption. v is a random variable with mean zero and variance σ^2 . Gibson et al. (2015) find that λ lies between 0.4 and 0.6, and also that it is higher in rural than urban areas. The variance of the error-ridden consumption is therefore lower than that of true consumption.

Given this information, we provide results varying the bandwidth of the thick poverty line by 10, 20 and 30% of the original poverty line for a well-used Ethiopian household dataset. First we conduct some illustrative simulations based on “true” consumption data (artificially constructed) and then measurement error that is a) classical and b) mean-reverting.

5.2 Simulations on synthetic data

We create a synthetic dataset of a panel of 1500 individuals, and six time periods, to mirror that of the actual survey data used in the empirical section below. For convenience, we

call the underlying wellbeing measure “consumption” Our first simple example is created to fix ideas on movements around the poverty line. We examine a case where error-free consumption stays constant over time, and therefore all movements observed in the measured data are due to measurement error. Let households be allocated randomly to either “high” or “low” consumption, of either 90 or 110 units, which stays constant across all 6 periods. We draw the poverty line at 100. Given this setup, we can consider how to set the fuzzy poverty lines - for example 10% of the poverty line is one standard deviation of the error (by construction) and will therefore capture a large portion of the measurement error. We then add measurement error to create the error-ridden wellbeing measure. We first add classical measurement error such that $\theta = 0$, $\lambda = 1$ in equation (22) above, and $v \sim N(0, \sigma_v^2)$. We set $\sigma_v = 10$.

For the Foster measures we set $\tau = 4$ and $v_t = \frac{1}{T} \forall t$. In this construction we can easily see that exactly half of the households are chronic poor (since they are always poor) and half are not (never poor). This would also not change by varying the level of τ . In table 1, we show the “true” headcount measure (50%). Taking 1000 draws from the measurement error distribution, we calculate the average Foster headcount measures (crisp and fuzzy) in the face of measurement error. The headcount of the crisp measure falls to 37%. However, fuzzifying the line at 10% above and below yields 44%, and fuzzifying only (at 10%) above the line gives similar results to the “true” data. Looking at the aggregate rate may not give the full picture in terms of targeting - the two averages could be quite similar but comprise different individuals/households. Therefore in the second and third columns of table 1 we show the proportion of exclusion and inclusion errors over the 1000 simulation draws. The crisp measure has quite a high 12% of exclusion errors which may be worrying for policy. The fuzzy measure reduces this to 5.7%, and fuzzifying only at the upper bound almost eliminates exclusion errors completely. The price for this is 1% inclusion errors - it would be a normative choice whether this is worth the tradeoff in policy applications. Note that in this particular example the inclusion errors increase as the bound is increased upward, so it is important not to set the bounds higher than required. Adding mean reverting measurement error ($\theta = 40$, $\lambda = 0.6$) causes the crisp

measure to drop below half of that for the true data. The fuzzy measure again comes closer to the true measure, especially when fuzzifying only above the poverty line, though the errors of inclusion do increase slightly more in this case. We also calculate the GDC measure (with $\beta = 1$), and adding measurement error changes the crisp measure from 0.5 to 0.32. The fuzzy measures come closer to the true measure; especially when we fuzzy only above the poverty line, in which case, the measures are almost identical.

PLACE TABLE 1 HERE

Our second synthetic example is more realistic based on empirical observations and other studies in the literature. The “true” consumption measure x is lognormally distributed in period one, and in subsequent time periods welfare evolves through a mildly autoregressive (and therefore converging) process, with a random disturbance (e.g. Fields, 2008; Burger et al., 2016). We write the evolution of true consumption x over time as:

$$x_t = \beta x_{t-1} + u_{it}. \quad (23)$$

The mean of consumption in the first period is 150^7 and we again set the poverty line to 100 for convenience. Table 1 shows the results for the Foster measure (left hand columns). We calculate the Foster measure for the true consumption data as 0.198. Adding classical measurement error as above ($N(0,10)$) reduces the crisp measure to 0.14, with 5% exclusion errors. The fuzzy measures reduce this to 3% and 1% when fuzzifying above and below, and only above, respectively. The inclusion errors as before are low, but do increase to just under 3% when we only fuzzy above z .

We then add mean-reverting measurement error (such that $\lambda = 0.6$) as in the first example. This further reduces the Foster measure to 0.12, with a lower impact on the fuzzy measure, to 0.15. As in the first example too, fuzzifying only above the poverty line almost eliminates errors of exclusion though inclusion errors are again higher -almost 4%. For the GDC measures we also find that fuzzifying only above the poverty line brings the measure closer to that from the true distribution. The usefulness of these examples of course

⁷We set this higher than in the previous example as there is a greater variance in the consumption outcomes, and we avoid the issue of negative values when measurement error is introduced.

depends on how realistic the assumptions are around i) the underlying data generating process of the consumption data and ii) the type and parameters of the measurement error. We also present results below using real data to see how the results differ between the fuzzy and crisp measures (though of course we do not know the true consumption in this case).

5.3 Empirical application: Ethiopia

We explore the empirical implications of these generalizations using the Ethiopian Rural Household Survey (ERHS). The ERHS is a well-known panel dataset from a developing country that has been extensively used for poverty and mobility analysis (Baulch and Hoddinott, 2000; Dercon and Shapiro, 2007; Dercon et al., 2012). The ERHS contains data on just over 1100 households in 15 villages, observed at six points in time over a fifteen year period, 1994 – 2009. The timing of the rounds is not even, with fieldwork in 1994, 1995, 1997, 1999, 2004 and 2009.⁸ We use information on household consumption, that households were asked to recall for the week prior to the survey, including food that was home grown, bought at market, and received as a gift or benefit from government. In this way, we can assume that any consumption smoothing that the household intended, and was able to implement, would have been implemented. Below, we note the likely measurement error that this method may incorporate.

The poverty line is village-specific, and represents the amount needed to consume just over 2000 calories per day per adult equivalent, plus some very basic non-food items (such as firewood to cook). It is thus an extremely austere poverty line, around one-third of the commonly used “dollar a day” international poverty line. In each round we also deflated consumption and the poverty line by a village-specific food price index based on prices collected at the community level, and thus we construct a measure of consumption per adult equivalent. For more details on this survey and the calculation of the, by now, quite widely used consumption basket, see Dercon and Krishnan (1998). The poverty line is on average 43 Ethiopian Birr (1994 prices) per adult equivalent in the household.

⁸Two rounds were actually fielded in 1994, but only six months apart, so we drop the second one.

PLACE TABLE 2 HERE

Several authors have analysed wellbeing based on consumption measures in the ERHS, including most recently Baulch (2011) and Dercon et al. (2012). Table 2 shows that cross sectional, or ‘snapshot’, poverty fell in the study villages between 1994 and 2004, with the headcount (P_0) falling from just under 43% to just under 20%, but then the headcount rate increased between 2004 and 2009 back to 35%. The other two measures, the average poverty gap (P_1) and the poverty severity index (P_2), followed a similar trend.

Table 3 shows the tabulation of number of periods spent in poverty. Looking at households over time, there is a lot of movement in and out of poverty, and fewer than a third of all households have never experienced any poverty at all. However, only 2% recorded consumption below the poverty line in every visit over the ten-year period. Hence we are faced with exactly the kind of exercise that was outlined in the theory section above. Some households have longer periods in poverty, but do not fall very much below the poverty line; some have fewer episodes of poverty but some of those are very severe.

PLACE TABLE 3 HERE

We now calculate the “fuzzy” poverty measures outlined above, the chronic poverty measure of Foster (2009) and the intertemporal poverty measure of Gradin et al. (2012) by taking an upper and lower bound around the poverty line.

Given the discussion above, we begin by setting the upper bound for the fuzzy set (z_2) at 10% above the poverty line, and symmetrically with the lower bound (i.e. $z_1 = 0.9z$, $z_2 = 1.1z$). We note that this means that $\pi_{nt} = 1$ only if consumption is below the lower bound poverty line z_1 , and $\pi_{nt} = 0$ for consumption above the upper bound z_2 . Between z_1 and z_2 π_{nt} receives a value between zero and one. We choose the fuzzy membership function proposed by Dombi (1990); we show results for alternative specifications based on the proposal by Chakravarty (2006) (in Annex 6 below), and in fact the results do not

change substantially.

How many households have recorded consumption levels that are between the upper and lower bounds of the “thick” poverty lines as discussed above? To understand how much impact the fuzzying of poverty identification will have, we note that 545 observations (just under 8% of all observations) lie between z_1 and z_2 at $\pm 10\%$ fuzzy set. If we increase the set to $\pm 20\%$ the number doubles, and at $\pm 30\%$ just over a fifth of the observations (1373) are included. We conclude, then, that the fuzzy identification has the potential to affect poverty measures that incorporate a strict cutoff.

We begin our analysis with the Foster measure, and considering $v_t = \frac{1}{T} \forall t$. For the calculations, we note from table 4 that the measure depends crucially on the value of choice parameter τ , i.e. the duration cutoff. If $\tau = 0.66$, or $4/6$ periods (those poor in 4 or more of the 6 periods are classified as poor), then 16.5% of the sample will be classified as poor. If we increase the required number of periods in poverty to $5/6$ for chronic classification, then only 8% are defined as chronic poor. Recall that the other measures in the Foster class of indices are calculated based on this identification step (duration, poverty gap, squared poverty gap).

We also note that, for “*crisp*” poverty identification in our setting of discrete time periods (or rounds of survey data) defining those who are chronically poor if the deprivation score is higher than or equal to 0.66 is empirically equivalent to chronic poverty identification when the deprivation score is strictly above 0.5 (see table 4, “crisp” column). In other words, in our setting, the members of the chronic poor set are the same whether we defined them as having spent four or more periods in poverty *or* strictly more than three periods in poverty. This is to be expected with “crisp” (i.e. non-fuzzied) measures. However, as the results in table 4 show, the difference in the way τ is defined, can and does matter when calculating fuzzy poverty measures.

PLACE TABLE 4 HERE

We illustrate the effect of increasing the bandwidth of the fuzzy line on the headcount

(H and H_π) and the duration-adjusted headcount measure (P and P_π) in table 4. In the first row, the first column shows the headcount, or proportion of households classified as chronically poor, at 16.5%. Increasing the thick poverty line at 10% reduces this to 13.5%. At 20% bandwidth, 12% of households are considered chronic poor, and at 30% bandwidth 11% of households are chronically poor. The second row incorporates the fuzzyness into a duration cutoff based on a strict inequality, with $\tau_a = 0.5$ (i.e. the deprivation score has to be strictly higher than 0.5 to identify the household as chronically poor). In this case, the proportion deemed chronically poor is increasing in the bandwidth of the thick poverty line, classifying just over 20% of households as chronically poor for the highest of our chosen illustrative bandwidths. The next two columns use the identification choices mentioned above in order to calculate the corresponding duration-adjusted headcounts, which behave similarly. The poverty measures with $\alpha = 1$ and $\alpha = 2$ also decrease correspondingly, though to a lesser extent, since those gaps that are included through the fuzzifying are weighted lower than more extreme poverty.

We also discussed above that, under different scenarios, the Foster measure with “crisp” poverty lines may overestimate, or underestimate, chronic poverty (vis-a-vis “fuzzy” alternatives) in the presence of measurement error, when there are transitions of a small amount above and below the poverty line that may be spurious due to measurement error. The simulations also showed that fuzzifying with the poverty line as the lower bound brought us closer to the “true” measures in the cases that we presented. We may therefore wish to make a slightly different normative choice, which is to set the thick poverty line at z as a minimum, and allow periods in which consumption is just above the poverty line to still be considered poor. The assumption here would be that we care more about measurement error that misclassifies a household just above the poverty line, rather than just below it as we wish to penalise errors of exclusion more heavily than those of inclusion. Table 7 in the Annex below shows headcount measures for all of our different assumptions. The last two rows consider the poverty line z as the lower bound. By design, this would increase the poverty measures; e.g. the headcount from under 21% to just under 24%. The change may seem minor, but it could be important in terms of targeting. This

reflects also the results from the simulation, that if we care most about exclusion errors, then this would be an appropriate choice.

Next we show our calculation of the Gradin et al. “fuzzy measures”. In this case, there is no second duration cutoff, due to the union approach to poverty membership. However as outlined above, in calculating the individual poverty measure, each poverty episode is weighted by the length of the “spell” to which it belongs (see equation (17)). For simplicity, we again calculate only the headcount measure (i.e. $\gamma = 0$), and we use values $\alpha = \beta = 1$ to illustrate the change in the measure when we “thicken” the poverty line.

Table 5 shows the results. As the bandwidth widens, the measure increases, by approximately 1% for every 10% of poverty-line widening. This increase is due to the net lengthening in spell duration; e.g. the effect of periods being reclassified from 0 (in terms of poverty status) to a non-zero amount on spell lengthening overtaking the spell-shortening effect of periods being reclassified from 1 to a lesser amount (see equations (17) and (18)). For the 10% bandwidth this represents a change for 448 of the people-year observations, and for the 30% bandwidth the change affects 1369 people-year observations. (By contrast, in the case of Foster measure an increase in the headcount, or lack thereof, was conditional upon the choice of the duration cutoff).

PLACE TABLE 5 HERE

Our results show some material differences between the crisp and fuzzy measures as calculated. Does this matter in practice in terms of rankings of households in the distribution? In the annex, table 9 and 10 rank the 15 villages in the survey by poverty first with the crisp measure, then the fuzzy. Interestingly, but perhaps also expected, we found that rankings are altered in the middle of the rank distribution. This seems to make sense if we consider that the top and bottom ranks are made of villages hosting higher proportions of households which either spend most periods safely out of poverty, i.e. away from the poverty line, or spend most periods with very low consumption levels. In other words, households whose membership functions and poverty measures are less likely to be affected by the fuzzifying vis-a-vis households which spend more time nearer the fuzzy-

line region. We looked at large vs small households, given the findings of Caeyers et al. (2012), who found that measurement error was likely higher for larger households, when comparing paper surveys with computer-based surveys that corrected for inconsistencies. The ranking does not change, but the difference between the fuzzy and crisp measures is higher for the larger households as shown in table 6.

PLACE TABLE 6 HERE

The magnitude of change for the fuzzy measures is of course proportional to the bandwidth, and the choice of this we would see as a pragmatic issue, depending on the perceived level of measurement error. We replicated our results using alternative definitions of the fuzzy set, and they are not substantially different. Finally, one may be concerned with errors of exclusion, rather than inclusion in the case of poverty targeting; in which case, we would recommend setting the lower bound of the thick poverty line to the original poverty line, and creating a bandwidth above it.

6 Conclusions

This paper presents an empirical adjustment for some increasingly used chronic or intertemporal poverty measures which show desirable normative properties, yet may be excessively sensitive to measurement error, due to the discontinuities inherent in their calculation. The adjustment is fairly simple and empirically practical. Moreover, depending on the policy concern (e.g. errors of exclusion), it can bear conceptual and/or practical advantages vis-a-vis alternative possible adjustments (e.g. changing duration cutoffs). Drawing on fuzzy set theory, we construct a ‘thick’ poverty line that enters into the poverty identification step of the poverty measures in each and every time period. This thickening of the poverty line allows us to remove some of the discontinuities in the measures, thus rendering the measures less sensitive to measurement error, and without affecting any of their other properties.

The empirical section presents some simple simulations, and empirical results for data on rural Ethiopia, showing that, in this case, the choice of functional form for the fuzzy poverty identification method is less important than the size of the bandwidth (i.e. the

poverty line's 'thickness'). In our simulations, the fuzzy adjustment around the poverty line shows that both measures by Foster (2009) and Gradin et al. (2012) may underestimate intertemporal poverty (as defined by each measure respectively) in the presence of measurement error. It is possible that, with a longer time series, the potential for measurement error to affect the results would be greater. The poverty analyst should make choices on these sensitivity tests based on the appropriate objectives of the measurement exercise (e.g. if there is higher concern over exclusion errors vis-a-vis others). We hope that our proposed adjustment method adds to the toolkit suitable for these purposes.

The proposed adjustment focused on two classes of intertemporal poverty measures. We chose those two classes for two reasons, in addition to limited space: (1) they illustrate important traits of intertemporal poverty measurement with clarity (identification of the chronically poor in the case of Foster (2009), sensitivity to duration of contiguous spells in the case of Gradin et al., 2012), and (2) their discontinuities may render them highly sensitive to measurement error. However the adjustment could also be implemented in other equally insightful classes of measures. For example, some classes of measures axiomatically justify differential weighting of poverty spells according to their timing in households', or individuals', lifetimes (e.g. Hoy and Zheng, 2011; Hoy et al., 2012), e.g. prioritising spells experienced earlier in life. Likewise, classes of measures where spells outside poverty count toward mitigating the effects of poverty spells could also be considered (Dutta et al., 2013).⁹

⁹We thank an anonymous referee for this suggestion.

References

- Alkire, S. and J. Foster, “Counting and multidimensional poverty measurement,” *Journal of Public Economics*, 95(7-8), 476–487, 2011.
- Antman, F. and D. J. McKenzie, “Earnings mobility and measurement error: A pseudo-panel approach,” *Economic Development and Cultural Change*, 56(1), 125–161, 2007.
- Baulch, B., editor, *Why poverty persists? Poverty dynamics in Asia and Africa*, Edward Elgar, 2011.
- Baulch, B. and J. Hoddinott, “Economic mobility and poverty dynamics in developing countries,” *The Journal of Development Studies*, 36(6), 1–24, 2000.
- Beegle, K., J. De Weerd, J. Friedman, and J. Gibson, “Methods of household consumption measurement through surveys: Experimental results from Tanzania,” *Journal of Development Economics*, 98(1), 3–18, 2012.
- Bossert, W., S. Chakravarty, and C. D’Ambrosio, “Poverty and time,” *Journal of Economic Inequality*, 10(2), 145–62, 2012, doi 10.1007/s10888-011-9175-2.
- Bound, J., C. Brown, and N. Mathiowetz, “Measurement error in survey data,” in J. Heckman, editor, “*Handbook of Econometrics*,” 3707–3842, Elsevier, 2001.
- Bound, J. and A. B. Krueger, “The extent of measurement error in longitudinal earnings data: Do two wrongs make a right?” *Journal of Labor Economics*, 9(1), 1–24, 1991.
- Burger, R. P., S. Klasen, and A. Zoch, “Estimating income mobility when income is measured with error: The case of South Africa,” 2016, economic Research Southern Africa Working Paper 607.
- Caeyers, B., N. Chalmers, and J. De Weerd, “Improving consumption measurement and other survey data through CAPI: Evidence from a randomized experiment,” *Journal of Development Economics*, 98(1), 19–33, 2012.

- Calvo, C. and S. Dercon, “Chronic poverty and all that,” in T. Addison, D. Hulme, and R. Kanbur, editors, “Poverty Dynamics,” Oxford University Press, 2009.
- Cerioli, A. and S. Zani, “A fuzzy approach to the measurement of poverty,” in “Income and wealth distribution, inequality and poverty,” 272–284, Springer, 1990.
- Chakravarty, S., “An axiomatic approach to multidimensional poverty measurement via fuzzy set,” in A. Lemmi and G. Betti, editors, “Fuzzy Set Approach to Multidimensional Poverty Measurement,” Springer, 2006.
- Cheli, B. and A. Lemmi, “A “Totally” fuzzy and relative approach to the multidimensional analysis of poverty,” *Economic Notes*, 24(1), 115–134, 1995.
- Deaton, A. and M. Grosh, “Consumption,” in M. Grosh and P. Glewwe, editors, “Designing Household Survey Questionnaires for Developing Countries: Lessons from Fifteen Years of LSMS Experience,” Oxford University Press, Washington, DC, 2000.
- Dercon, S., J. Hoddinott, and T. Woldehanna, “Growth and chronic poverty: Evidence from rural communities in ethiopia,” *Journal of Development Studies*, 48(2), 238–53, 2012.
- Dercon, S. and P. Krishnan, “Changes in poverty in rural Ethiopia 1989-1995: measurement, robustness tests and decomposition,” 1998, CSAE WPS/98-7.
- Dercon, S. and J. Shapiro, “Moving on, staying behind, getting lost: lessons on poverty mobility from longitudinal data,” 2007, GPRG-WPS-075.
- Dombi, J., “Membership functions as an evaluation,” *Fuzzy Sets and Systems*, 35, 1–21, 1990.
- Dutta, I., L. Roope, and H. Zank, “On intertemporal poverty measures: the role of affluence and want,” *Social Choice and Welfare*, 41(4), 741–762, 2013.
- Fields, G. S., “A brief review of the literature on earnings mobility in developing countries,” 2008, ILR Working Paper 2-20-2008. Ithaca: Cornell University.

- Foster, J., “A class of chronic poverty measures,” in T. Addison, D. Hulme, and R. Kanbur, editors, “Poverty dynamics: interdisciplinary perspectives,” Oxford University Press, 2009.
- Foster, J., J. Greer, and E. Thorbecke, “A class of decomposable poverty measures,” *Econometrica*, 52(3), 761–6, 1984.
- Foster, J. and M. E. Santos, “Measuring chronic poverty,” in G. Betti and A. Lemmi, editors, “Poverty and Social Exclusion: New Methods of Analysis,” Routledge, 2013.
- Foster, J. and A. Shorrocks, “Subgroup consistent poverty indices,” *Econometrica*, 59(3), 687–709, 1991.
- Gibson, J., K. Beegle, J. De Weerd, and J. Friedman, “What does variation in survey design reveal about the nature of measurement errors in household consumption?” *Oxford Bulletin of Economics and Statistics*, 77(3), 466–474, 2015.
- Glewwe, P., “Measurement error bias in estimates of income and income growth among the poor: analytical results and a correction formula,” *Economic Development and Cultural Change*, 56(1), 163–189, 2007.
- Gradin, C., C. del Rio, and O. Canto, “Measuring poverty accounting for time,” *Review of income and wealth*, 58(2), 330–54, 2012.
- Hoy, M., B. Thompson, and B. Zheng, “Empirical issues in lifetime poverty measurement,” *Journal of Economic Inequality*, 10(2), 163–89, 2012.
- Hoy, M. and B. Zheng, “Measuring lifetime poverty,” *Journal of Economic Theory*, 146(6), 2544–62, 2011.
- Jalan, J. and M. Ravallion, “Determinants of transient and chronic poverty,” 1998, World Bank Policy Research Working Paper 1936.
- Jalan, J. and M. Ravallion, “Is transient poverty different? Evidence from Rural China,” *Journal of Development Studies*, 36(6), 2000.

- Lee, N., G. Ridder, and J. Strauss, “Estimation of poverty transition matrices with noisy data,” *Journal of Applied Econometrics*, 32(1), 37–55, 2016.
- Lemmi, A. and G. Betti, Fuzzy set approach to multidimensional poverty measurement, Springer, 2006.
- Marano, G., G. Betti, and F. Gagliardi, “Latent Class Markov Models for Measuring Longitudinal Fuzzy Poverty,” in M. Carpita, E. Brentari, and E. M. Qannari, editors, “Advances in Latent Variables,” *Studies in Theoretical and Applied Statistics*, 73–81, Springer International Publishing, 2015.
- Nunez Velasquez, J. J., “Current status and new approaches to the measurement of poverty,” *Estudios de Economia Aplicada*, 27(2), 325–46, 2009.
- Perez-Mayo, J., “A dynamic analysis of deprivation for Spain,” *Estudios de Economia Aplicada*, 27(2), 501–22, 2009.
- Pischke, J.-S., “Measurement error and earnings dynamics: Some estimates from the PSID validation study,” *Journal of Business & Economic Statistics*, 13(3), 305–314, 1995.
- Porter, C. and N. Quinn, “Intertemporal Poverty Measurement: tradeoffs and policy options,” 2008, CSAE Working Paper 2008-21.
- Porter, C. and N. Quinn, “Measuring intertemporal poverty. Policy options for the poverty analyst,” in G. Betti and A. Lemmi, editors, “Poverty and social exclusion: new methods of analysis,” 166–193, Routledge, 2013.
- Pritchett, L., “Who is not poor? Dreaming of a world truly free of poverty,” *The World Bank Research Observer*, 21(1), 1–23, 2006.
- Quinn, N., “Chronic and Transient Poverty in Rural Ethiopia: A New Decomposition,” 2014, paper presented at the 2014 IARIW Conference.
- Ragin, C., Fuzzy-set Social Science, University of Chicago Press, 2000.

- Ravallion, M., "Poverty rankings using noisy data on living standards," *Economics Letters*, 45(4), 481–485, 1994.
- Rodgers, J. and J. Rodgers, "Chronic poverty in the United States," *Journal of Human Resources*, 28(1), 25–54, 1993.
- Smithson, M. and J. Verkuilen, *Fuzzy Set Theory: Applications in the Social Sciences*, volume 147 of *Quantitative Applications in the Social Sciences*, Sage Publications, 2006.
- Verma, V., G. Betti, and F. Gagliardi, "Fuzzy measures of longitudinal poverty in a comparative perspective," *Social Indicators Research*, 130(2), 1–20, 2015.

Tables

Table 1: Simulations: Constant wellbeing and lognormal autoregressive consumption

	Foster-H0		Exclusion errors		Inclusion errors		GDC-0	
Crisp - true consumption	0.500	<i>0.198</i>	0.500	<i>0.187</i>
Classical M.E.								
Crisp	0.377	<i>0.148</i>	0.123	<i>0.055</i>	0.000	<i>0.005</i>	0.326	<i>0.145</i>
Fuzzy (0.9, 1.1)	0.443	<i>0.170</i>	0.057	<i>0.036</i>	0.000	<i>0.008</i>	0.429	<i>0.179</i>
Fuzzy (1, 1.1)	0.503	<i>0.211</i>	0.008	<i>0.013</i>	0.012	<i>0.026</i>	0.505	<i>0.204</i>
Mean-Reverting M.E.								
Crisp	0.241	<i>0.121</i>	0.263	<i>0.083</i>	0.003	<i>0.006</i>	0.249	<i>0.122</i>
Fuzzy (0.9, 1.1)	0.322	<i>0.149</i>	0.183	<i>0.058</i>	0.005	<i>0.010</i>	0.394	<i>0.171</i>
Fuzzy (1, 1.1)	0.521	<i>0.215</i>	0.044	<i>0.020</i>	0.065	<i>0.038</i>	0.506	<i>0.208</i>
N	1500							

Notes: Left hand columns refer to the “constant wellbeing” simulation results, right hand columns in *italics* refer to the lognormal autoregressive simulation results.

Table 2: **FGT Poverty, by year**

Year	P_0	P_1	P_2
1994	0.346	0.140	0.078
1995	0.378	0.151	0.081
1997	0.213	0.067	0.031
1999	0.232	0.073	0.033
2004	0.199	0.066	0.031
2009	0.347	0.127	0.064
N	1136		

Table 3: Number of periods in poverty

Item	Number	Per cent
Never Poor	343	30.19
Once	264	23.24
Twice	194	17.08
Three times	147	12.94
Four times	107	9.42
Five times	59	5.19
In every period	22	1.94
Total	1,136	100

Source: ERHS Data

Table 4: Fuzzy Foster measures

Measure	Crisp	Fuzzy 10	Fuzzy 20	Fuzzy 30
Headcount				
$\tau = 0.67$.165	.137	.118	.108
$\tau_a = 0.5$.165	.189	.196	.203
Duration-adjusted Headcount				
$\tau = 0.67$.125	.106	.093	.086
$\tau_a = 0.5$.125	.136	.139	.142
Poverty Gap				
$\tau = 0.67$.053	.054	.054	.055
$\tau_a = 0.5$.075	.076	.077	.079
Squared Poverty Gap				
$\tau = 0.67$.029	.03	.03	.03
$\tau_a = 0.5$.041	.041	.042	.042
N	1136			

Notes: Fuzzy poverty defined as the S-shaped membership function (Dombi, 1990), with fuzzifying above and below the poverty line at 10, 20, 30 percent respectively. The difference between the two τ is that in the first case, we set $\tau \geq 4$ periods. In the second case, $\tau > 3$. This does not change the crisp measure, but allows the fuzzifying to take effect above and below τ . See text for further discussion.

Table 5: GDC Headcount measures (s-convex)

Variable	$\beta = 1$	$\beta = 0.5$
Crisp	0.108	0.168
Fuzzy 10 percent	0.115	0.173
Fuzzy 20 percent	0.123	0.178
Fuzzy 30 percent	0.132	0.185
N	1136	

Gradin et al (2012) measures (GDC). The thick poverty line is defined as in table 4.

Table 6: Example: Measure differences by Household Size

	Small HH	Large HH
Foster Crisp	0.113	0.237
Foster Fuzzy	0.122	0.255
GDC Crisp	0.079	0.147
GDC Fuzzy	0.116	0.207

Notes: Fuzzy measures are calculated at 30% bandwidth, $\tau = 0.67$. Large household is defined as having greater than 5 members. Measures calculated as in table 4.

FIGURES

Figure 1: Fuzzy identification of deprivation status in period t (Dombi, 1990)

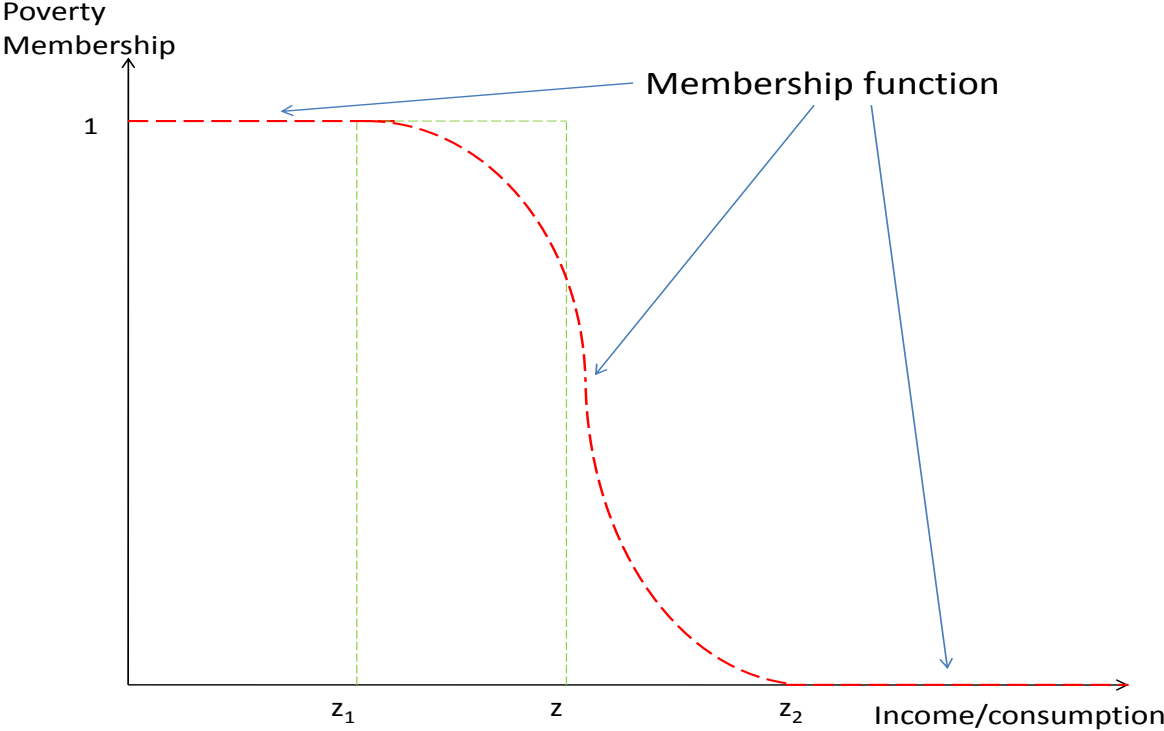
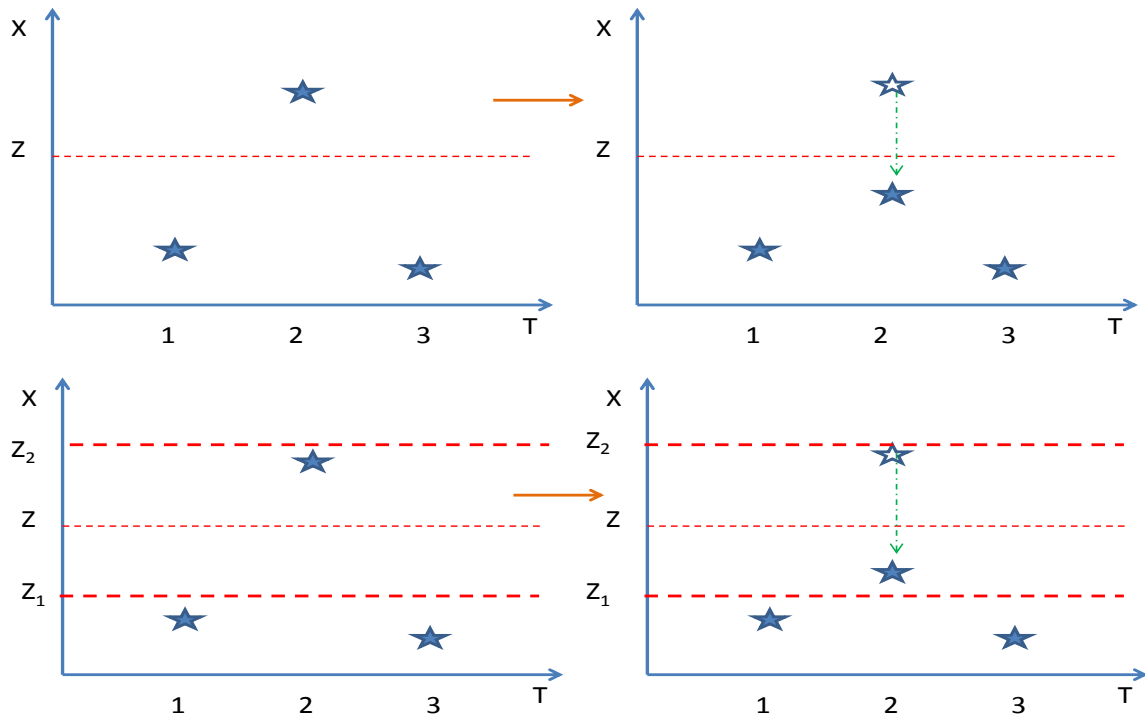


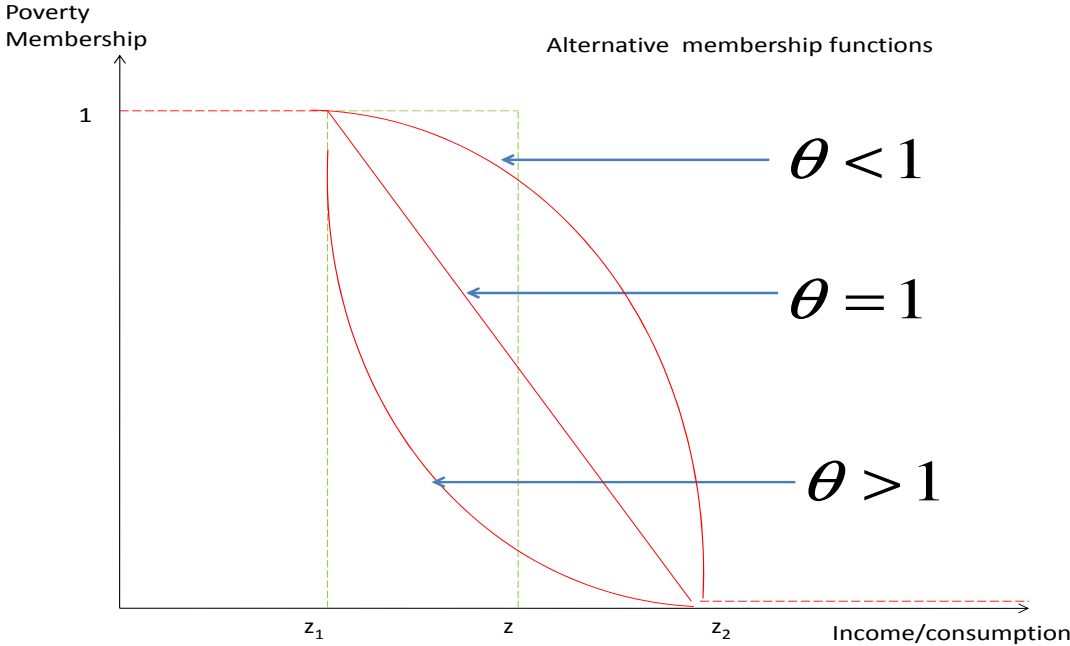
Figure 2: Fuzzy poverty spells



SUPPLEMENTARY INFORMATION

Annex 1: Alternative membership functions

Figure 3: Fuzzy identification of deprivation status in period t (Chakravarty, 2006)



Annex 2: Additional Tables based on alternative assumptions for the membership function

Table 7: Poverty headcounts, Foster measures, alternative membership function

Variable	Mean	Std. Dev.
Crisp	0.216	0.412
Fuzzy(10%), $\theta = 1$	0.181	0.385
Fuzzy(20%), $\theta = 1$	0.154	0.361
Fuzzy(30%), $\theta = 1$	0.142	0.349
Fuzzy(10%), $\theta = 0.5$	0.185	0.388
Fuzzy(20%), $\theta = 0.5$	0.168	0.374
Fuzzy(30%), $\theta = 0.5$	0.165	0.371
Fuzzy ($z=\min$, +20%), $\theta = 1$	0.229	0.42
Fuzzy ($z=\min$, +20%), $\theta = 0.5$	0.237	0.426
N		1136

Notes: As in the main tables, Fuzzy 10% means that the upper bound of the “thick” poverty line is at $1.1z$ and the lower bound at $0.9z$, similarly with 20 and 30% respectively. θ is the parameter referred to in figure 4, for the fuzzy membership function as proposed by Dombi (1990). The bottom two lines include the poverty line z as the lower bound of the thick poverty line, with the upper bound set at $1.2z$.

Table 8: GDC measures, alternative membership function

Variable	Mean	Std. Dev.
Fuzzy(10%), $\theta = 1$	0.115	0.184
Fuzzy(20%), $\theta = 1$	0.122	0.191
Fuzzy(30%), $\theta = 1$	0.131	0.199
Fuzzy(10%), $\theta = 0.5$	0.119	0.187
Fuzzy(20%), $\theta = 0.5$	0.131	0.198
Fuzzy(30%), $\theta = 0.5$	0.146	0.211
N		1136

Notes: As in the main tables, Fuzzy 10% means that the upper bound of the “thick” poverty line is at $1.1z$ and the lower bound at $0.9z$, similarly with 20 and 30% respectively. θ is the parameter referred to in figure 4, for the fuzzy membership function as proposed by Dombi (1990).

Table 9: ERHS Village rankings: Foster measure

	Foster	<i>RANK-F</i>	F-30	<i>RANK F-30</i>
Doma	0.694	1	0.600	1
Geblen	0.607	2	0.411	2
Dinki	0.327	3	0.250	3
Imdibir	0.246	4	0.158	4
Dinki	0.222	5	0.139	5
Adado	0.164	6	0.073	6
Gara Godo	0.156	7	0.052	8
Haresaw	0.130	8	0.072	7
Korodegaga	0.096	9	0.012	11
Trurufe Kechemene	0.063	10	0.025	9
Yetmen	0.021	11	0.000	12
Debre Birhan	0.019	12	0.000	12
Adele Keke	0.013	13	0.013	10
Shumsha	0.000	14	0.000	12
Sirbana Godeti	0.000	14	0.000	12

Notes: Foster measures calculated as in tables above. Rank refers to the village ranking in terms of the amount of poverty as per the measure. Fuzzy measure is calculated with 30% bandwidth.

Table 10: ERHS Village rankings: GDC measure

	GDC	<i>Rank-GDC</i>	GDC-30	<i>Rank GDC-30</i>
Gara Godo	0.383	1	0.443	1
Geblen	0.251	2	0.305	2
Doma	0.205	3	0.248	3
Dinki	0.157	4	0.197	5
Imdibir	0.151	5	0.217	4
Aze Deboa	0.111	6	0.123	8
Korodegaga	0.110	7	0.140	6
Adado	0.110	8	0.129	7
Haresaw	0.095	9	0.115	9
Adele Keke	0.049	10	0.063	11
Trurufe Ketchema	0.047	11	0.064	10
Debre Birhan	0.033	12	0.038	12
Yetmen	0.026	13	0.034	13
Shumsha	0.012	14	0.017	14
Sirbana Godeti	0.005	15	0.007	15

Notes: GDC measures calculated as in tables above. Rank refers to the village ranking in terms of the amount of poverty as per the measure. Fuzzy measure is calculated with 30% bandwidth.