WIDER Working Paper 2021/26

Estimating poverty transitions in Mozambique using synthetic panels

A validation exercise and an application to cross-sectional survey data

Vincenzo Salvucci and Finn Tarp*

February 2021

United Nations University World Institute for Development Economics Research
wider.unu.edu

Abstract: In this paper we first validate the use of the synthetic panels technique in the context of the 2014/15 intra-year panel survey data for Mozambique, and then apply the same technique to the 1996/97, 2002/03, 2008/09, and 2014/15 cross-sectional household budget surveys for the same country. We find that in most analyses poverty rates and poverty transitions estimated using synthetic panels provide results that are close to the true values obtained using the 2014/15 panel data. With respect to intra-year poverty dynamics, we find that Mozambique has a high intra-year variability in consumption and poverty, and a very high degree of intra-year poverty immobility, with a big portion of the population remaining either in poverty or out of poverty over the whole year, with smaller percentages of individuals moving upward or downward. With respect to the 1996/97, 2002/03, 2008/09, and 2014/15 cross-sectional surveys, our results suggest that in most year-to-year comparisons there is a greater proportion of people getting out of poverty than falling into poverty, consistent with the poverty-reduction process observed, but the percentage of people staying in poverty over time appears to be substantially higher, involving about one-third of the population in most years. Further analyses on the 2008/09 and 2014/15 surveys estimate that for an individual who was in the vulnerable group in 2008/09, there is a 60 per cent probability of remaining in the same group, whereas the probability of becoming non-vulnerable is lower than the probability of entering poverty. This constitutes the first attempt to provide an insight into poverty dynamics in Mozambique using all the available survey data.

Key words: poverty dynamics, poverty transitions, Mozambique, synthetic panels

JEL classification: C52, C53, I32

Acknowledgements: The authors would like to acknowledge the great help provided by Peter Lanjouw, Gerton Rongen, and David Garces-Urzainqui in the preparation of the technical work supporting the present analysis. We also acknowledge helpful comments by Eva-Maria Egger, Ines Ferreira, Kenneth Mdadila, Silvio Daidone, and Stevan Lee of Oxford Policy Management (OPM).

* University of Copenhagen, Copenhagen, Denmark; corresponding author: vincenzo.salvucci@gmail.com

This study has been prepared within the UNU-WIDER project Inclusive growth in Mozambique – scaling-up research and capacity implemented in collaboration between UNU-WIDER, University of Copenhagen, University Eduardo Mondlane, and the Mozambican Ministry of Economics and Finance. The project is financed through specific programme contributions by the governments of Denmark, Finland, and Norway.

Information and requests: publications@wider.unu.edu

ISSN 1798–7237 ISBN 978-92-9256-964-8

https://doi.org/10.35188/UNU-WIDER/2021/964-8

Typescript prepared by Gary Smith.

United Nations University World Institute for Development Economics Research provides economic analysis and policy advice with the aim of promoting sustainable and equitable development. The Institute began operations in 1985 in Helsinki, Finland, as the first research and training centre of the United Nations University. Today it is a unique blend of think tank, research institute, and UN agency—providing a range of services from policy advice to governments as well as freely available original research.

The Institute is funded through income from an endowment fund with additional contributions to its work programme from Finland, Sweden, and the United Kingdom as well as earmarked contributions for specific projects from a variety of donors.

Katajanokanlaituri 6 B, 00160 Helsinki, Finland

1 Introduction

Development economists have investigated movements out of and into poverty in several poor countries, relying on the availability of longitudinal data—that is, household survey data following the same households over many years—and collecting data at different points in time (e.g., Dercon and Krishnan 2000, 2003; Kumar and Quisumbing 2013). With such data, one can ask the question of which local, household, or individual characteristics are associated with positive or negative trajectories. Further, researchers can estimate the impact of specific policies or shocks, such as the recent cyclones hitting the central and northern regions of Mozambique. Another important question to ask is whether some households are trapped in poverty while others pursue a sustainable pathway towards higher levels of welfare and, if so, which factors determine these different trajectories. This question can only be analysed with data that allow comparing households to each other and also following them over time.

However, such household panel data spanning a period of a few years are generally not available for Mozambique. Nationally representative cross-sectional household budget survey data exist, but they are only collected every five or six years (*Inquérito aos Agregados Familiares Sobre Orçamento Familiar*; IAF or IOF). Therefore, most existing studies focusing on poverty and other welfare indicators and aggregates lack the dynamic dimension in their analyses.[1]

In this paper we validate the synthetic panels approach introduced by Dang et al. (2011, 2014b) and by Dang and Lanjouw (2013, 2014) using the most recent Mozambican data and apply it to previous cross-sectional household budget survey data. The synthetic panels methodology allows construction of synthetic panel data from repeated cross-sections and is based on an imputation procedure through which the values of the relevant welfare aggregate—income or consumption—for households observed at time 1 are estimated using households and community characteristics and welfare aggregates measured at time 0 (Dang and Lanjouw 2013, 2014; Dang et al. 2011, 2014b). This approach relies on imputation models and on the presence of time-invariant correlates of consumption in the survey. An extensive validation work based on actual panel data has been carried out by various researchers over recent years for different sets of countries. This work seems to suggest that the methodology is sufficiently reliable as an alternative to actual panels when it comes to estimating income/consumption dynamics (Bierbaum and Gassmann 2012; Bourguignon et al. 2015; Cruces et al. 2015; Dang and Lanjouw 2015; Dang and Lanjouw 2018; Dang et al. 2014a).

In the first part of the paper, we validate the use of the synthetic panels technique in the context of the 2014/15 intra-year panel survey data for Mozambique; in the second part of the paper, we apply the same technique to the 1996/97, 2002/03, 2008/09, and 2014/15 cross-sectional household budget surveys.

In the validation exercise, we describe the movements in and out of poverty between different quarters of the 2014/15 Mozambican household budget survey, analysing real and estimated poverty transitions and performing a series of tests to check the robustness of the results. In most of the analyses performed, the poverty rates and poverty transitions that we estimate using the synthetic panels approach provide results (bounds) that are close to the true values obtained using

[1] In their 2017 paper, Dang and Dabalen (2017) included Mozambique among the countries used to evaluate the chronic or transitioning poverty situation of African countries, but they limited their analysis to the household budget surveys for 2002/03 and 2008/09.

the panel data. Hence, the positive performance of the validation exercise in the 2014/15 survey case enables us to apply the synthetic panels technique to studying poverty transitions even in 'normal' survey years, when data are not collected as a panel. Thus, we also apply synthetic panels to the 1996/97, 2002/03, 2008/09, and 2014/15 cross-sectional household budget surveys. This implementation of the synthetic panels provides important insights into the poverty dynamics and trajectories in Mozambique, which have not previously been explored in detail due to the lack of longitudinal data.

This constitutes the first validation exercise of the synthetic panels method in an intra-year panel setting, and represents the first attempt to provide an insight into poverty dynamics in Mozambique using all the available survey data. The paper develops as follows: Section 2 presents the context; Section 3 describes the data and Section 4 presents the methodology. The results are discussed in Section 5, while Section 6 concludes.

2 Context

With about half of the population being considered poor according to the latest Poverty Assessment, Mozambique is still one of the poorest countries in the world, notwithstanding a decrease in the poverty rate of about 25 percentage points between 1996/97 and 2014/15 and years of sustained economic growth. Poverty is widespread in the country, but there are substantial differences between different provinces and between urban and rural areas (DEEF 2016). Looking at the four available household budget surveys implemented in 1996/97, 2002/03, 2008/09, and 2014/15, DEEF (2016) also reports that a fall in poverty occurred between 1996/97 and 2002/03, followed by a stagnation between 2002/03 and 2008/09, and a further decrease in the poverty rate between 2008/09 and 2014/15 (DEEF 2016). Poverty rates are presented in Table 1 for all years and at different levels of disaggregation (national level, rural/urban, regions, and provinces).

Table 1: Consumption poverty rates, 1996/97–2014/15 (%)

Area	1996/97	2002/03	2008/09	2014/15
National	69.7	52.8	51.7	46.1
Urban	61.8	48.2	46.8	37.4
Rural	71.8	55.0	53.8	50.1
North	67.3	51.9	45.1	55.1
Centre	74.1	49.2	57.0	46.2
South	65.5	59.9	51.2	32.8
Niassa	71.9	48.3	33.0	60.6
Cabo Delgado	59.1	60.3	39.0	44.8
Nampula	69.4	49.1	51.4	57.1
Zambezia	67.6	49.7	67.2	56.5
Tete	81.9	60.5	41.0	31.8
Manica	62.4	44.7	52.8	41.0
Sofala	87.8	41.3	54.4	44.2
Inhambane	83.0	78.1	54.6	48.6
Gaza	64.8	55.4	61.0	51.2
Maputo Province	65.6	59.0	55.9	18.9
Maputo City	47.1	42.9	29.9	11.6

Note: percentage of poor people over the total population for different areas and for all available household budget surveys.

Source: authors' compilation based on DEEF (2016).

The long-term trends in poverty rates have been analysed in detail in the various poverty assessments, but the country's poverty dynamics are less studied—mainly because household panel data are not available for Mozambique. This limitation was partially overcome when the National Statistics Office (INE) designed the 2014/15 household budget survey: it was designed as an intra-year household panel survey, with households interviewed three times over a 12-month period. Up to 2008/09, households were only interviewed once during the year, and even though the 1996/97, 2002/03, and 2008/09 surveys were designed so that each quarter was representative for the whole population, the sample in each quarter was too small to provide precise information about poverty dynamics (DEEF 2016).[2] The fact that the 2014/15 survey was designed as an intra-year household panel survey helped to take into account the intra-year variability in household consumption and poverty, which is significant in Mozambique due to both the high percentage of people working in subsistence agriculture and the recurrent natural shocks that hit the country (Arndt et al. 2012, 2016, 2018; DEEF 2016; DNEAP 2010; DNPO 1998, 2004; INE 2004, 2010, 2015). Given that 2014/15 is considered to be a normal year (DEEF 2016), this provides the opportunity to study in more depth the intra-year poverty dynamics. However, at the same time it provides the opportunity to validate the synthetic panels method used to estimate poverty and vulnerability transitions using repeated cross-sectional surveys in an intra-year panel setting, which to the best of our knowledge has not been pursued before.

Regarding intra-year variations, from an analysis of the 2014/15 household survey it emerges that poverty rates are clearly higher in the second survey quarter, which corresponds to the months mid-November 2014 to mid-February 2015. This is not unusual for Mozambique since these months represent the core of the rainy season for most areas and provinces, and they are often associated with scarce food reserves, high food prices, hunger, and higher poverty rates. Similar results have also been found in previous household surveys held in 1996/97, 2002/03, and 2008/09 (DEEF 2016; DNEAP 2010; DNPO 1998, 2004). Table 2 shows the results for the various quarters of 2014/15 as presented in DEEF (2016).

Table 2: Quarterly and aggregate poverty rates, 2014/15 (%)

	Q1 (Aug.–Nov. 2014)	Q2 (Nov. 2014–Feb. 2015)	Q4 (May 2015–Aug. 2015)	Entire sample (Aug. 2014–Aug. 2015)
National	43.9	55.0	43.2	46.1
Urban	34.8	42.8	36.9	37.4
Rural	48.1	60.6	46.1	50.1
Niassa	60.1	69.5	58.0	60.6
Cabo Delgado	45.5	54.4	40.8	44.8
Nampula	54.4	68.8	51.9	57.1
Zambezia	54.2	62.9	56.1	56.5
Tete	35.7	51.6	19.2	31.8
Manica	35.2	57.0	31.8	41.0
Sofala	38.1	47.7	48.8	44.2
Inhambane	43.8	53.6	49.4	48.6
Gaza	44.3	55.9	53.8	51.2
Maputo Province	17.3	21.7	19.5	18.9
Maputo City	13.7	11.7	11.7	11.6

Note: percentage of poor people in each quarter and for different geographic areas. The last column reports the poverty rate results for the entire sample (all quarters), already presented in Table 1.

Source: authors' calculations based on the household budget survey 2014/15 and on DEEF (2016).

[2] For an analysis of intra-year dynamics using the Mozambican household budget survey 2008/09 and exploiting the characteristic that these surveys are designed so that each quarter is representative for the whole population, see Arndt et al. (2016). Instead, Salvucci and Santos (2020) exploit the panel structure of the Mozambican household budget survey 2014/15 to assess the short-term impact on consumption and poverty of the 2015 flood in Mozambique.

3 Data

In this study, we use the quarterly data from the 2014/15 household budget survey (henceforth, IOF14) as our primary source of data for the validation exercise of intra-year poverty dynamics. For the application of synthetic panels to previous cross-sectional household surveys developed in the second part of the paper, we use the household budget surveys 1996/97, 2002/03, and 2008/09 (abbreviated as IAF96, IAF02, and IOF08, respectively). When applying the synthetic panels technique to all the available cross-sectional household surveys, we will also use the IOF14 survey and treat it as a standard cross-sectional household survey, without exploiting its panel structure.

All the IAFs/IOFs were designed and implemented by the Instituto Nacional de Estatística (INE), whereas the poverty analyses were performed by the Ministry of Economics and Finance with technical assistance from various partners, including IFPRI, UNU-WIDER, and the University of Copenhagen, depending on the survey year (DEEF 2016; DNEAP 2010; DNPO 1998, 2004; INE 2004, 2010, 2015). The various IAFs/IOFs are similar in many respects, despite some relatively minor differences in the structure of the questionnaires. They are representative of Mozambique as a whole, of rural and urban areas, and of each of the 11 provinces, including the capital, Maputo. Each family was interviewed at different times of the 12-month survey period, with questions about general characteristics, employment, education, access to basic services, daily expenses and household consumption from own production, possession of durable goods, housing conditions, receipts and transfers received and paid, income from various sources, as well as less frequent expenses. The consumption aggregate is computed using daily, monthly, and annual household expenditures; expenditures obtained from education-, health-, and tourism-specific modules; individual expenditures not captured in the household module; receipts in-kind from work and from other activities; imputed house rents; and imputed use value for durable goods.[3]

The IOF14 adds to the basic features of previous IAFs/IOFs the fact that it is a repeated interview (mini-panel) survey (DEEF 2016; INE 2015). It was carried out from mid-August 2014 to mid-August 2015, and the 12-month period was subdivided as follows: quarter 1, mid-August to mid-November 2014; quarter 2, mid-November 2014 to mid-February 2015; quarter 3, mid-February to mid-May 2015; and quarter 4, mid-May to mid-August 2015. Originally, it had been designed so that each household had to be interviewed four times over the four quarters of the year. However, for various reasons quarter 3 of the IOF14 survey ended up not being implemented, but fieldwork was reinstated in the fourth quarter.

Additional information on the four household budget surveys presented and on the poverty assessments that derived from the analysis of these data is found in a series of documents produced by both INE and the Ministry of Economics and Finance (DEEF 2016; DNEAP 2010; DNPO 1998, 2004; INE 2004, 2010, 2015).

[3] The fact that all the IAFs/IOFs are very similar in their scope and design is important, as the synthetic panels method can only be applied if data are comparable: the underlying population must be the same in all rounds of the survey, which makes it possible to use time-invariant household characteristics to predict household consumption. This implies, for example, that the sampling methodology is not modified over time. In the case of the IOF14 this is not problematic, as we work with different quarters from the same survey. However, for previous surveys as well, based on the technical household budget survey documents issued by INE, it seems that the sampling methodology was not changed across different survey rounds, even though relatively minor changes occurred over time: for example, non-essential survey modules were added or dropped depending on the survey year and the list of consumption items changed over time (INE 2004, 2010, 2015).

4 Methodology

In this section we describe the synthetic panels approach, which is first applied to the IOF14 survey quarters in order to validate the results from this approach against results from true panel data, and it is subsequently applied to previous household budget surveys as well, the IAF96, IAF02, and IOF08. Here, we greatly rely on and refer to Dang et al. (2011, 2014b) and Dang and Lanjouw (2013). The synthetic panels approach was first introduced by Dang et al. (2011, 2014b) and further developed in several other papers by the same authors, among which we highlight the contribution by Dang and Lanjouw (2013), which introduced some significant novelties in order to obtain point estimates for poverty and vulnerability transition.

Summarizing, the synthetic panels approach allows constructing synthetic panel data from repeated cross-sections. This method is based on an imputation procedure through which the values of the relevant welfare aggregate (income or consumption) for households observed at time 1 are estimated using households and community characteristics and welfare aggregates measured at time 0 (Dang et al. 2011, 2014b; Dang and Lanjouw 2013). Obviously, cross-sectional survey data do not provide information on household consumption for the same households over time, but under a series of assumptions it is possible to estimate the consumption that round 2 households would have had in round 1, specifying a consumption model for round 1 that is only based on time-invariant household characteristics (Dang et al. 2011).[4] Hence, round 1 consumption is first projected on time-invariant characteristics; subsequently, the OLS parameters that are estimated in this consumption model are applied to the same time-invariant household characteristics, but using the information collected in round 2. In this way, we can obtain an estimate of household consumption in round 1 for households interviewed in round 2. More formally, for the population as a whole, the linear projection of round 1 consumption or income, y_{i1}, onto x_{i1} is given by:

$$y_{i1} = \beta_1' x_{i1} + \varepsilon_{i1} \tag{1}$$

where x_{i1} is a vector of characteristics of household i in survey round 1 that are observed in both round 1 and round 2. This could include language, religion, and ethnicity, but also time-invariant characteristics of the household head such as sex, education, place of birth, parental education, and age.[5] Similarly, the linear projection of round 2 consumption or income y_{i2} onto x_{i2} is given by:

---

[4] The first assumption is that the underlying population must be the same in all rounds of the survey, which makes it possible to use time-invariant household characteristics to predict household consumption. This implies, for example, that the sampling methodology is not modified over time. Based on the technical household budget survey documents issued by INE, it derives that the sampling methodology was not changed across different quarters of the IOF14 (INE 2015). However, the underlying population might also change due to changes in the household composition (births, deaths, migration, etc.), but this difficulty can be overcome by restricting the sample, as explained in the rest of the section. The second assumption is that the correlation between the error terms of the consumption model in the two survey rounds should be non-negative. Dang et al. (2011) outline the reasons why this assumption is expected to be satisfied in most applications, and state that the two abovementioned assumptions are generally satisfied if the sample is restricted to households whose household head is 25–55 years old, as derived from the pseudo-panel literature (Dang et al. 2011). Therefore, we apply this restriction and limit the sample to households whose household head is aged 25–55 in the first survey round, while the age range is restricted accordingly in subsequent survey rounds. In the analysis of the IOF14, given that it is based on an intra-year panel, the age differences over different survey quarters are very small or non-existent.

[5] x_{i1} could also include time-varying characteristics of the household that can be recalled for round 1 in round 2 (Dang et al. 2011). For example, whether or not the household head is employed in round 1, and his or her occupation, their place of residence in round 1, etc.

$$y_{i2} = \beta_2^{'} x_{i2} + \varepsilon_{i2} \qquad (2)$$

where x_{i2} is the set of household characteristics in round 2 that are observed in both round 1 and round 2 surveys. The poverty line in periods 1 and 2 are indicated as z_1 and z_2, respectively. Now, in order to estimate the degree of mobility in and out of poverty we want to estimate, for example, what fraction of households in the population is poor in round 1 and non-poor in round 2,[6] or:

$$P\left(y_{i1} < z_1 \text{ and } y_{i2} > z_2 \right) \qquad (3)$$

However, we do not know y_{i1} and y_{i2} for the same households. Dang et al. (2011) provide and discuss the assumptions that need to be made in order to be able to estimate bounds for such quantities (Dang et al. 2011: 4–10). Moreover, they consider two approaches to estimate the bounds on mobility: a non-parametric approach, where no assumptions are made about the joint error distribution, and a parametric approach, where it is assumed that the joint error distribution is bivariate normal. In what follows, we apply both the non-parametric and the parametric approach to the analysis of poverty mobility.

Estimating the upper and lower bounds for poverty transitions requires a number of steps, which are explained in detail by Dang et al. (2011: 10–12). In general, also depending on the assumptions that are made regarding the joint distribution of the error terms in rounds 1 and 2, estimated mobility will be greater the less correlated are the error terms, as this implies that round 1 consumption is less correlated with consumption in round 2. When no correlation is assumed, the upper bounds for poverty mobility are obtained; when perfect correlation is assumed, we obtain the lower bounds for poverty mobility. If we indicate with ρ the correlation coefficient between the error terms in rounds 1 and 2, then the non-parametric estimates for the lower and upper bounds of poverty mobility correspond to assuming ρ is equal to either 1 or 0, respectively (Dang et al. 2011).[7]

The estimated bounds for poverty mobility, as will be observed, can be relatively wide. However, when they are too wide, they can provide little information on the underlying poverty transitions and be of little practical use. The width can be greatly reduced by improving the quality of the underlying consumption model. In the validation exercise based on the survey quarters of the IOF14, we use three different model specifications. In Model 1 we only use those variables that are strictly time invariant: household head gender, age, age squared, and education level. In Model 2 we also include rural and province dummies, and in Model 3 we add interaction terms between

---

[6] In the work of Dang et al. (2011), poverty mobility indicates that households have a different poverty status in the two survey rounds, whereas poverty immobility indicates that households have the same poverty status in the two survey rounds. Equation 3 represents the joint probability of being poor in round 1 and non-poor in round 2. Poverty dynamics has also been studied as a conditional probability—that is, the probability of being, for example, non-poor in round 2 given that the individual was poor in round 1.

[7] When zero correlation is assumed, Dang et al. (2011) propose to predict round 1 consumption/income using the equation $\hat{y}_{i1}^{2U} = \hat{\beta}_1^{'} x_{i1}^{2} + \hat{\varepsilon}_{i1}^{2}$, where $\hat{y}_{i1}^{2}$ is the round 1 predicted consumption/income for households in round 2 (the round is indicated by the superscript), obtained using the set of round 2 household characteristics, x_{i2}, and the first round OLS estimates of parameters, $\hat{\beta}_1$. The superscript U indicates that this is an upper bound. With respect to the error term, this is obtained for each household i by randomly drawing with replacement in round 2 from the empirical distribution of estimated residuals in round 1. The random drawing procedure is repeated R times and the average value obtained over the R repetitions is used. In our study, R = 50. When perfect correlation is assumed (ρ = 1) the estimates of the residuals obtained from round 2 can be directly used to predict round 1 consumption/income: $\hat{y}_{i1}^{2L} = \hat{\beta}_1^{'} x_{i1}^{2} + \hat{\varepsilon}_{i2}^{2}$, where the superscript L indicates that this is a lower bound.

gender and age, gender and education level, and rural and province dummies. Passing from Model 1 to Model 3, performance clearly improves. In the application of the synthetic panels to the IAF96, IAF02, IOF08, and IOF14 we only use one model, including gender and age of the household head, education level of the household head, and provincial and rural dummies.

As discussed in Section 1, extensive validation work based on actual panel data has been carried out by various researchers over recent years for different countries. This work seems to suggest that the methodology is sufficiently reliable as an alternative to actual panels when it comes to estimating income/consumption dynamics (Bierbaum and Gassmann 2012; Bourguignon et al. 2015; Cruces et al. 2015; Dang and Lanjouw 2015, 2018; Dang et al. 2014a).

The non-parametric method discussed until now only requires a few assumptions to estimate bounds for poverty mobility. However, if only a limited number of time-invariant characteristics can be used in the consumption model, the bounds obtained can be rather wide. If additional assumptions on the joint distribution of the error terms are introduced, then bounds can be sharpened and it is even possible to obtain point estimates for poverty mobility. Dang et al. (2011) present this parametric approach as a variant of their basic approach. This latter approach has the advantage of being applicable even in those cases in which the available number of time-invariant variables is very limited, which seems to be the case in most cross-sectional household surveys. In the parametric approach, the joint distribution of the error terms in rounds 1 and 2 is assumed to be bivariate normal and the correlation coefficient between them is indicated with ρ.[8] In order to sharpen the bounds' estimates, Dang et al. (2011) suggest, for example, to use 0.3 and 0.7, or 0.2 and 0.8, respectively as lower and upper values for ρ, which is found to significantly reduce the bounds. Dang et al. (2011) show that when the joint error distribution is assumed to be bivariate normal, then quantities of interest such as the fraction of the population that is poor in round 1 and non-poor in round 2 can be derived as follows:

$$\begin{aligned} P^E\left(y_{i1} < z_1 \text{ and } y_{i2} > z_2\right) &= P\left(\beta_1' x_{i1} + \varepsilon_{i1} < z_1 \text{ and } \beta_2' x_{i2} + \varepsilon_{i2} > z_2\right) \\ &= \Phi_2\left(\frac{z_1 - \beta_1' x_{i2}}{\sigma_{\varepsilon 1}}, \frac{z_2 - \beta_2' x_{i2}}{\sigma_{\varepsilon 2}}, -\rho\right) \end{aligned} \tag{4}$$

where Φ_2 indicates the bivariate normal cumulative distribution function and σ_ε represents the standard deviation of the error term, ε, in round 1 or 2. Expressions to derive the other quantities of interest are also provided by Dang et al. (2011). As previously discussed, it is clear from this equation that a lower value of ρ implies a higher probability of mobility between rounds 1 and 2. Hence, obtaining a better estimate for ρ rather than just using the boundaries $\rho = 0$ and $\rho = 1$ may greatly help in obtaining more precise estimates of poverty mobility.

In this respect, Dang and Lanjouw (2013) introduce a method to compute point estimates for ρ and in turn obtain point estimates for poverty and vulnerability transitions. They show that, under a number of assumptions, ρ should be bounded from above by the simple correlation coefficient between household consumption in rounds 1 and 2, which can be approximated by the synthetic panel cohort-level simple correlation coefficient, $\rho_{yi1\,yi2}$, and from below by the expression

[8] This assumption may hold in a number of cases; indeed, the distribution of income or consumption is often approximated using a log-normal distribution (Dang et al. 2011).

$\frac{\beta_1^{'} \operatorname{var}(x_i)\beta_2}{\sqrt{\operatorname{var}(y_{i1})\operatorname{var}(y_{i2})}}$, where β_1 and β_2 are the vectors of estimated coefficients from the consumption model and x_i represents the vector of household time-invariant characteristics. The partial correlation coefficient, ρ, can then be estimated using the following equation (Dang and Lanjouw 2013: 9–15):[9]

$$\rho = \frac{\rho_{y_{i1},y_{i2}}\sqrt{\operatorname{var}(y_{i1})\operatorname{var}(y_{i2})} - \beta_1^{'} \operatorname{var}(x_i)\beta_2}{\sigma_{\varepsilon 1}\sigma_{\varepsilon 2}} \tag{5}$$

Once ρ is estimated, the procedures to compute the point estimates for poverty mobility are also provided by Dang and Lanjouw (2013: 15–25) and Dang and Lanjouw (2014).

Moreover, following Dang and Lanjouw (2014), we also implement an analysis of vulnerability by identifying a group of vulnerable individuals within the non-poor group. The analysis of vulnerability permits defining a vulnerability line and in turn creating three groups: (1) the poor, defined as those individuals whose daily real consumption per capita lies below the poverty line; (2) the vulnerable, those individuals whose daily real consumption per capita lies between the poverty line and the vulnerability line; and (3) the non-vulnerable (alternatively defined by Dang and Lanjouw (2014) as 'middle-class', 'secure', or 'prosperous'), those individuals whose daily real consumption per capita lies above the vulnerability line. Dang and Lanjouw (2014) propose not to set the vulnerability line at a value that is an arbitrary scaling up of the poverty line, but deriving the vulnerability line from a specified index of vulnerability, which is defined either as the probability of becoming poor at time 2 conditional on being in the middle-class at time 1, or as the probability of becoming poor at time 2 conditional on being vulnerable at time 1. The former is indicated as P^1, and is then defined as the 'insecurity index'. Given two survey rounds, 1 and 2, and a specified insecurity index, P^1, the vulnerability line in round 1, v_1, should satisfy the equality: $P^1\left(y_2 \le z_2 \middle| y_1 > v_1\right)$. Conversely, when the index of vulnerability is defined as the probability of becoming poor at time 2 conditional on being vulnerable at time 1, then it is indicated as P^2 and is defined as the 'vulnerability index'. In this case, given two survey rounds, 1 and 2, and a specified vulnerability index, P^2, the vulnerability line in round 1, v_1, should satisfy the equality: $P^2\left(y_2 \le z_2 \mid z_1 < y_1 < v_1\right)$. For the properties of these indices, see Dang and Lanjouw (2014).

Thus, the procedure outlined by Dang and Lanjouw (2014) permits estimation of a vulnerability line and linking it directly with a vulnerability index, derived for example from budgetary planning, social welfare objectives or relative concepts of well-being. Once the value of the insecurity index, P^1, or the value of the vulnerability index, P^2, are selected, it is possible to derive the value for the vulnerability line.

If we assume, as described earlier in this section, that the error terms in survey rounds 1 and 2 have a bivariate normal distribution with correlation coefficient ρ, then quantities of interest such

[9] Dang and Lanjouw (2013) also propose an alternative approximation for ρ: $\rho = \frac{\rho_{y_{i1},y_{i2}} - \sqrt{R_1^2 R_2^2}}{\sqrt{1-R_1^2}\sqrt{1-R_2^2}}$, with the simple correlation coefficient as upper bound and the expression $\sqrt{R_1^2 R_2^2}$ as lower bound, where R_1^2 and R_2^2 represent the coefficients of determination obtained from estimating the consumption model in rounds 1 and 2.

as the fraction of the population that is poor in round 1 and vulnerable in round 2 can be derived as follows:

$$
\begin{aligned}
P^{E}\left(y_{i1} < z_1 \text{ and } z_2 < y_{i2} < v_2\right) = \Phi_2\left(\frac{z_1 - \beta_1^{'} x_{i2}}{\sigma_{\varepsilon 1}}, \frac{v_2 - \beta_2^{'} x_{i2}}{\sigma_{\varepsilon 2}}, \rho\right) \\
-\Phi_2\left(\frac{z_1 - \beta_1^{'} x_{i2}}{\sigma_{\varepsilon 1}}, \frac{z_2 - \beta_2^{'} x_{i2}}{\sigma_{\varepsilon 2}}, \rho\right)
\end{aligned} \tag{6}
$$

Expressions to derive the other quantities of interest are also provided by Dang and Lanjouw (2014). We use the same procedures described earlier in the section to obtain estimates for the correlation coefficient of the error terms in survey rounds 1 and 2, ρ.

In the present study, population weights and other survey design features specific to the surveys considered are applied. Moreover, we follow Dang et al. (2011), who suggest limiting the sample to households whose household head is 25–55 years old. In all the models presented, the dependent variable is the log of real household consumption per capita, which is obtained as described in DNPO (1998, 2004), DNEAP (2010), and DEEF (2016).

5 Results

In this section we first present the results with respect to the true rates of poverty and poverty transitions obtained using the IOF14 data for quarters 1, 2, and 4 (Section 5.1). Results are briefly described and then the poverty rates and poverty transitions results derived from the application of the synthetic panels approach are presented and compared to the true rates (Section 5.2). Extensions and robustness checks are presented in Section 5.3. Subsequently, we present the main results with respect to poverty dynamics using all the available surveys, from 1996/97 to 2014/15, presenting the consumption model and the upper and lower bounds estimated for poverty mobility and immobility (Section 5.3). In Section 5.4, a more detailed analysis of poverty dynamics is presented for the period 2008/09–2014/15, corresponding to the last two available surveys in Mozambique. The upper and lower bounds estimated for poverty mobility and immobility are presented, but the poverty dynamics results obtained using a parametric estimation approach are also described. Moreover, in Section 5.5 an analysis of vulnerability is provided, following Dang and Lanjouw (2014) and Dang and Dabalen (2017).

5.1 Poverty rates and poverty transitions, true rates obtained from the IOF14 panel data

The Fourth National Poverty Assessment for Mozambique, contained in DEEF (2016), presented the consumption poverty rates based on the IOF14 and computed them for the national level, urban, rural, and regional levels, and for various subpopulations (Castigo and Salvucci 2017; DEEF 2016). However, in that Poverty Assessment the panel dimension of the IOF14 was not exploited.[10] This choice derived from the fundamental need to make the poverty results comparable with those from previous years. Hence, to the best of our knowledge, this is the first analysis of the intra-year poverty dynamics for Mozambique using real panel data representative of the whole country. As outlined in Section 1, 2014/15 was considered a 'normal' year. Localized

[10] An estimate of quarterly poverty rates was presented in the Appendix of the Fourth National Poverty Assessment, but the results were not fully analysed or described.

shocks occurred, as described, for example, by Salvucci and Santos (2020), but at an aggregate level the country did not experience major shocks during the survey months (DEEF 2016). Nonetheless, when analysing the quarterly poverty rates, it is possible to notice a rather big increase in poverty in Q2 compared to the other two available quarters, Q1 and Q4 (Table 2). The poverty rates for Q1 and Q4 are similar and close to 43–44 per cent, whereas the poverty rate in Q2 is about 10 percentage points higher (about 55 per cent) (Table 2). The second survey quarter corresponds to the months mid-November 2014 to mid-February 2015, which corresponds to the core of the rainy season for most areas and provinces in Mozambique, and this period is often associated with scarce food reserves, high food prices, hunger, and higher poverty rates. However, this seems like a substantial increase over a period of just a few months. The increase observed is driven mainly by the increase in poverty in rural areas as a whole (+12.5 percentage points) and in the provinces of Nampula, Tete, and Manica, but most provinces present an increase between Q1 and Q2 that is close to or above 10 percentage points. Nonetheless, poverty rates also decrease by approximately the same amount between Q2 and Q4, leaving the poverty rate almost unchanged between Q1 and Q4. The poverty rates at the national level for each quarter are presented in Table 2, while summary statistics for the variables included and for all survey quarters, obtained without the restrictions imposed on the age of the household head, are found in Table A1 in the Appendix.

However, the poverty rates alone do not permit full understanding of the dynamics between the poverty and non-poverty statuses. In Table 3, the poverty transitions over the three available quarters of the IOF14 are presented. In panel A we present the fraction of the population that is in each of the four categories displayed—'Poor, poor', 'Poor, non-poor', 'Non-poor, poor', and 'Non-poor, non-poor'—over the three available quarters of the IOF14. For example, 'Poor, poor' indicates the fraction of the population that was poor at time 1 and poor at time 2. In panel B we show the fraction of the population in each of the four states displayed: 'Poor to poor', 'Poor to non-poor', 'Non-poor to poor', and 'Non-poor to non-poor' over the three available quarters of the IOF14. For example, 'Poor to poor' indicates the fraction of the poor population at time 1 that was also poor at time 2, (i.e. the probability of being poor at time 2 given that the individual was also poor at time 1).

Table 3: Poverty transitions between survey quarters of the 2014/15 household budget survey

	True rates, full sample	Q1–Q2	Q1–Q4	Q2–Q4
A: Unconditional probabilities	Poor, poor	0.364	0.299	0.351
	Poor, non-poor	0.072	0.137	0.189
	Non-poor, poor	0.185	0.132	0.076
	Non-poor, non-poor	0.379	0.432	0.368
B: Conditional probabilities	Poor to poor	0.835	0.687	0.639
	Poor to non-poor	0.165	0.313	0.344
	Non-poor to poor	0.328	0.235	0.169
	Non-poor to non-poor	0.672	0.765	0.817

Note: the probabilities presented are estimated using the national poverty lines provided in the household surveys. The 'unconditional probabilities' panel provides the fraction of the population in the selected age range that is in each of the four categories. For example, 'Poor, poor' under the column Q1–Q2 indicates the fraction of the population that was poor in quarter 1 and poor in quarter 2. The 'conditional probabilities' panel provides the probability of each of the four states. For example, 'Poor to poor' under the column Q1–Q2 indicates the probability of being poor in quarter 2, given that the individual was also poor in quarter 1.

Source: authors' calculations based on household survey data.

It emerges that a relatively high proportion of the population remains in its original poverty status, either poor or non-poor, in all the survey quarters. Indeed, the probability of remaining poor given that the individual was also poor in Q1 is above 80 per cent for Q2 and around 70 per cent for

Q4. Also, as expected from the poverty rates presented in Table 2, a higher proportion of the population is non-poor in Q1 but is found in poverty in Q2 compared to the proportion of the population that is poor in Q1 and non-poor in Q2 (18.5 versus 7.2 per cent). The probability of becoming poor in Q2 given that the individual was non-poor in Q1 is rather high at 33 per cent, twice as high as the probability of becoming non-poor in Q2 given that the individual was poor in Q1 (16.5 per cent). Conversely, at the end of the hunger season (Q2) more people seem to escape from poverty, so that the proportion of people in the status 'poor in Q2, non-poor in Q4' is higher than the proportion of people in the status 'non-poor in Q2, poor in Q4' (18.9 versus 7.6 per cent). In this case, the probability of becoming non-poor in Q4 given that the individual was poor in Q2 is much higher than the probability of becoming poor in Q4 given that the individual was non-poor in Q2 (34 versus 17 per cent). The transitions between Q1 and Q4 present a rather stable situation, with an almost equivalent proportion of people being poor in Q1 and non-poor in Q4 and people being non-poor in Q1 and poor in Q4. In this case, the probability of escaping poverty given that the individual was poor in Q1 is higher than the probability of falling into poverty given that the individual was non-poor in Q1, but they are not as far away as in the Q2–Q4 case (31 versus 23 per cent).

5.2 Validation

In this section we use the panel data available in the IOF14 to validate the synthetic panel approach as presented by Dang et al. (2011, 2014b) and Dang and Lanjouw (2013). The structure of this validation exercise greatly relies on the work by Cruces et al. (2015), who performed an extensive exercise for the case of Chile, Nicaragua, and Peru. We present the results for the three models outlined in the methodology section and for the three quarters of the IOF14. In particular, in Table 4 the poverty rates obtained using the synthetic panels methodology are presented for three different cases: (1) poverty rate in Q1 for households surveyed in Q2, obtained from Q2 data; (2) poverty rate in Q1 for households surveyed in Q4, obtained from Q4 data; and (3) poverty rate in Q2 for households surveyed in Q4, obtained from Q4 data.[11] These cases are of interest for a number of reasons: cases 1 and 2 show equivalent estimates of the poverty rate in Q1 using two different sets of subsequently collected data, Q2 and Q4. With respect to case 3, this provides an estimate of poverty in a 'negative' quarter (as outlined, Q2 corresponds with the rainy season and hunger period), using the household data from Q4, which instead is considered a rather positive/slightly prosperous quarter. The synthetic panel method provides very similar estimates in cases 1 and 2: the true estimates are contained in the bounds defined by the synthetic panels method in all the models considered when poverty rates are estimated using Q4 data; and the true estimates are contained in the bounds defined by the synthetic panels method in the more complex consumption estimation model (Model 3) when poverty rates are estimated using Q2 data. Concerning case 3, the true estimates always lie within the bounds defined by the synthetic panels method. Also, the bounds defined by our preferred consumption estimation model, Model 3, are always narrower than the 95 per cent confidence intervals (CIs) for the true estimates, and overlap with them in all cases. The true estimates presented here differ from those described in Section 5.1 because in this and the following subsections the estimation sample was restricted to households whose household head is 25–55 years old, as mentioned in the methodology section.

[11] The remaining cases (poverty rate in Q2 for households surveyed in Q1, poverty rate in Q4 for households surveyed in Q1, and poverty rate in Q4 for households surveyed in Q2) have also been estimated and are available on request.

Table 4: True and estimated poverty rates, IOF14

	Synthetic panel estimate upper bound			True rates			Synthetic panel estimate lower bound		
	[1]	[2]	[3]	Estimate	95 per cent CI		[3]	[2]	[1]
P^1 from Q2 data	0.447	0.45	0.453	0.450	0.422	0.479	0.409	0.403	0.384
N	7,114	7,114	7,114	7,247			7,114	7,114	7,114
P^1 from Q4 data	0.452	0.453	0.456	0.451	0.418	0.484	0.413	0.406	0.387
N	7,411	7,411	7,411	7,926			7,411	7,411	7,411
P^2 from Q4 data	0.561	0.565	0.565	0.554	0.515	0.592	0.495	0.487	0.464
N	7,411	7,411	7,411	7,247			7,411	7,411	7,411

Note: P^1, P^2, and P^4 indicate the poverty rate in Q1, Q2, and Q4, respectively. The true estimates presented here differ from those described in Table 3 because in this and the following subsections the estimation sample was restricted to households whose household head is 25–55 years old, as mentioned in the methodology section.

Source: authors' calculations based on household survey data.

We now turn to poverty transitions: in this case also poverty transition bounds are estimated using Models 1, 2, and 3, and they are presented in Table 5 for the transitions Q1–Q2, Q1–Q4, and Q2–Q4. As already noticed for the case of poverty rate bounds, estimation bounds get narrower as we pass from Model 1 to Model 2 and then to Model 3, the difference being more noticeable between Models 1 and 2 than between Models 2 and 3. In any case, results seem to be encouraging with respect to the application of the synthetic panels approach to situations in which panel data are not available: the true rates lie between the lower and upper bounds for all the transitions statuses (poor, poor; poor, non-poor; non-poor, poor; non-poor, non-poor), for the different quarter transitions (Q1–Q2, Q1–Q4, and Q2–Q4), and for all the underlying consumption models (1, 2, and 3). However, in this case the 95 per cent CIs are much narrower than the bounds estimated using the synthetic panels.

In the next section, we attempt to reduce these bounds by using a more complex consumption estimation model that also includes the enumeration areas among the regressors. Furthermore, we will also implement the method introduced by Dang and Lanjouw (2013) to derive point estimates for poverty transitions by computing reasonable values for the correlation coefficient between the error terms in rounds 1 and 2, ρ.

Table 5: Real and estimated poverty dynamics in Q1, Q2, and Q4 of the IOF14: non-parametric estimates of unconditional probabilities

Poverty dynamics Q1–Q2

	Synthetic panel estimate upper bound			True rates			Synthetic panel estimate lower bound		
Model	[1]	[2]	[3]	Estimate	95 per cent CI		[3]	[2]	[1]
Poor, poor	0.447	0.448	0.451	0.377	0.357	0.397	0.264	0.259	0.236
Poor, non-poor	0.000	0.002	0.002	0.074	0.065	0.083	0.145	0.144	0.148
Non-poor, poor	0.115	0.114	0.111	0.185	0.171	0.200	0.299	0.303	0.327
Non-poor, non-poor	0.438	0.436	0.436	0.364	0.345	0.384	0.293	0.294	0.290
N	7,114	7,114	7,114	7,114			7,114	7,114	7,114

Poverty dynamics Q1–Q4

Model	Synthetic panel estimate upper bound			True rates			Synthetic panel estimate lower bound		
	[1]	[2]	[3]	Estimate	95 per cent CI		[3]	[2]	[1]
Poor, poor	0.430	0.415	0.413	0.310	0.290	0.329	0.210	0.204	0.185
Poor, non-poor	0.022	0.038	0.043	0.142	0.128	0.155	0.203	0.202	0.202
Non-poor, poor	0.006	0.02	0.022	0.130	0.118	0.142	0.226	0.231	0.25
Non-poor, non-poor	0.542	0.527	0.522	0.418	0.399	0.438	0.362	0.363	0.363
N	7,411	7,411	7,411	7,411			7,411	7,411	7,411

Poverty dynamics Q2–Q4

Model	Synthetic panel estimate upper bound			True rates			Synthetic panel estimate lower bound		
	[1]	[2]	[3]	Estimate	95 per cent CI		[3]	[2]	[1]
Poor, poor	0.435	0.433	0.432	0.362	0.340	0.384	0.246	0.241	0.221
Poor, non-poor	0.126	0.132	0.133	0.192	0.175	0.209	0.249	0.246	0.243
Non-poor, poor	0.000	0.002	0.003	0.075	0.066	0.085	0.189	0.194	0.214
Non-poor, non-poor	0.439	0.433	0.432	0.356	0.337	0.376	0.316	0.319	0.322
N	7,411	7,411	7,411	7,411			7,411	7,411	7,411

Note: estimated using the national poverty lines provided with the survey data. Rows give the fraction of population in the selected age range that is in each of the four categories. For example, 'Poor, poor' in the upper panel indicates the fraction that was poor in Q1 and poor in Q2. The true estimates presented here differ from those described in Table 3 because in this and the following subsections the estimation sample was restricted to households whose household head is 25–55 years old, as mentioned in the methodology section.

Source: authors' calculations based on household survey data.

5.3 Extensions and robustness checks

In this section the validation of the synthetic panels method applied to the IOF14 quarterly data is extended and a series of robustness checks are performed. We first attempt to reduce the width of the estimated bounds by implementing a more complex consumption estimation model that also includes the enumeration areas among the regressors. We also implement a method introduced by Dang and Lanjouw (2013) to derive point estimates for poverty transitions, which are obtained by computing reasonable values for the correlation coefficient between the error terms in rounds 1 and 2, ρ.

As explained by Dang et al. (2011), the width of the estimated bounds for poverty mobility can be reduced by improving the quality of the underlying consumption model. In particular, we have observed in this study that our model improved when regional characteristics and variable interactions were included in the analysis. However, the variables included in the consumption model must retain the characteristic of being reasonably time invariant. Given that we are using as a basis for our analysis survey data collected over a relatively short period of time, it may be reasonable to consider the enumeration areas as time invariant, too. As expected, when dummies for the enumeration areas are included in the consumption model (we call this Model 4), the R-squared noticeably increases with respect to Models 1, 2, and 3. The R-squared is in the range 0.17–0.20 for Model 1, 0.25–0.28 for Model 2, and 0.27–0.29 for Model 3, whereas it increases to about 0.50 for Model 4. As a consequence, the estimated bounds for poverty rates and poverty mobility get narrower. In Table 6 we show the true poverty rates in each quarter, with their CIs, and compare them with the bounds obtained using Model 4. Differently from what was observed in Table 4, in this case the true rates lie outside the range defined by the lower and upper bounds in two out of three cases. Only the true rate for poverty in Q1 is contained within the interval defined by the estimated Q1 poverty rate, computed using Q4 data. However, poverty transitions are more precisely estimated using Model 4 than in previous cases. These are presented in Table 7. As observed for Models 1, 2, and 3 in Table 5, the true rates are always contained within the estimated bounds, but in this case the bounds are sensibly narrower.

Table 6: True and estimated poverty rates using an extended consumption model, IOF14

	Synthetic panel estimate upper bound	True rates			Synthetic panel estimate lower bound
	[4]	Estimate	95% CI		[4]
P^1 from Q2 data	0.447	0.450	0.422	0.479	0.403
N	7,114	7,247			7,114
P^1 from Q4 data	0.459	0.451	0.418	0.484	0.408
N	7,411	7,926			7,411
P^2 from Q4 data	0.529	0.554	0.515	0.592	0.471
N	7,411	7,247			7,411

Note: P^1, P^2, and P4 indicate the poverty rate in Q1, Q2, and Q4, respectively. The true estimates presented here differ from those described in Table 3 because in this and following subsections the estimation sample was restricted to households whose household head is 25–55 years old, as mentioned in the methodology section.

Source: authors' calculations based on household survey data.

Table 7: Real and estimated poverty dynamics in Q1, Q2, and Q4 of the IOF14. Non-parametric estimates of unconditional probabilities using an extended consumption model

Poverty dynamics Q1–Q2

	Synthetic panel estimate upper bound	True rates			Synthetic panel estimate lower bound
	[4]	Estimate	95 per cent CI		[4]
Poor, poor	0.405	0.377	0.357	0.397	0.287
Poor, non-poor	0.116	0.074	0.065	0.083	0.042
Non-poor, poor	0.275	0.185	0.171	0.200	0.157
Non-poor, non-poor	0.396	0.364	0.345	0.384	0.322
N	7,114	7,114			7,114

Poverty dynamics Q1–Q4

	Synthetic panel estimate upper bound	True rates			Synthetic panel estimate lower bound
	[4]	Estimate	95 per cent CI		[4]
Poor, poor	0.345	0.310	0.290	0.329	0.229
Poor, non-poor	0.179	0.142	0.128	0.155	0.114
Non-poor, poor	0.206	0.130	0.118	0.142	0.09
Non-poor, non-poor	0.451	0.418	0.399	0.438	0.386
N	7,411	7,411			7,411

Poverty dynamics Q2–Q4

	Synthetic panel estimate upper bound	True rates			Synthetic panel estimate lower bound
	[4]	Estimate	95 per cent CI		[4]
Poor, poor	0.366	0.362	0.340	0.384	0.257
Poor, non-poor	0.214	0.192	0.175	0.209	0.163
Non-poor, poor	0.178	0.075	0.066	0.085	0.069
Non-poor, non-poor	0.402	0.356	0.337	0.376	0.351
N	7,411	7,411			7,411

Note: estimated using the national poverty lines provided with the survey data. Rows give the fraction of the population in the selected age range that is in each of the four categories. For example, 'Poor, poor' in the upper panel indicates the fraction that was poor in Q1 and poor in Q2. The true estimates presented here differ from those described in Table 3 because in this and the following subsections the estimation sample was restricted to households whose household head is 25–55 years old, as mentioned in the methodology section.

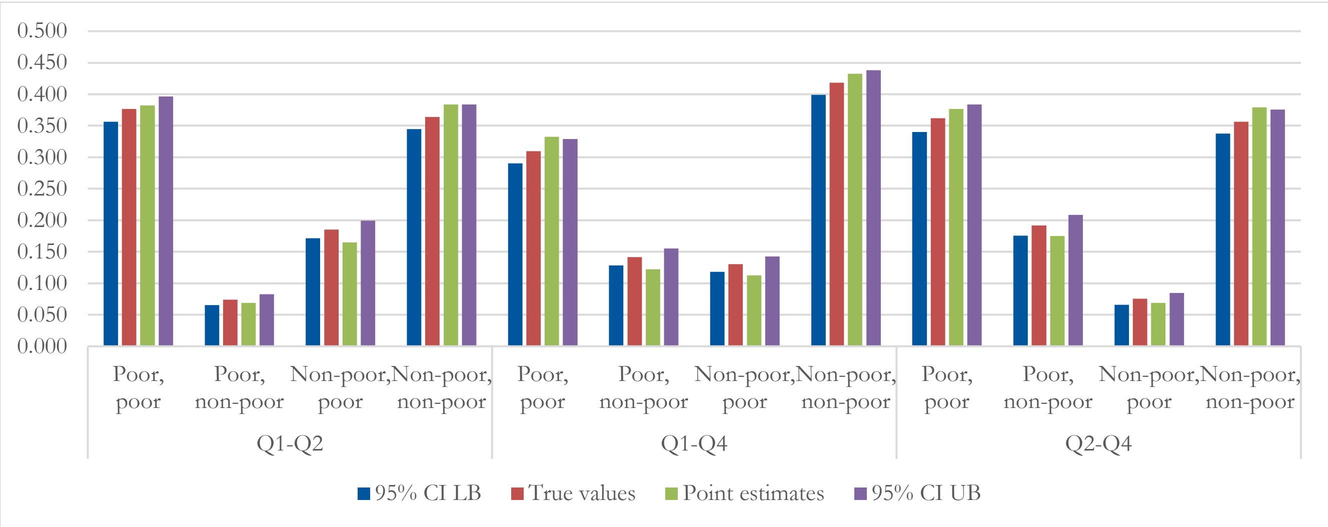
Source: authors' calculations based on household survey data.

As a final check, we also implement the method introduced by Dang and Lanjouw (2013) to derive point estimates for poverty transitions already described in the methodology section. Once ρ is estimated, the procedures to compute the point estimates for poverty mobility are also provided by Dang and Lanjouw (2013: 15–25) and Dang and Lanjouw (2014).

The values of ρ that we computed for the three transitions analysed, Q1–Q2, Q1–Q4, and Q2–Q4 are: 0.709, 0.688, and 0.709, respectively. They are all positive, as expected in most applications, and are within the range (0.2, 0.8) that Dang et al. (2011) consider as most reasonable based on previous validation studies and applications. In Figures 1 and 2 we present the true transition values with their 95 per cent CIs and the point estimates obtained using the values of ρ just presented. These are discussed for the three transitions analysed, Q1–Q2, Q1–Q4, and Q2–Q4. Moreover, we show the results for both unconditional and conditional probabilities.

With respect to all possible transitions, Q1–Q2, Q1–Q4, and Q2–Q4, it can be noticed that all the point estimates obtained using the values of ρ discussed above are close to the true transition values, and most of them lie within true values' 95 per cent CIs. In particular, our point estimates in the unconditional probabilities case seem to slightly underestimate poverty mobility (states 'poor, non-poor' and 'non-poor, poor') and slightly overestimate poverty immobility (states 'poor, poor' and 'non-poor, non-poor'). In the conditional probabilities case, there is not a clear pattern with respect to mobility/immobility, but all the point estimates lie within the true values' 95 per cent CIs, apart from the states 'poor, poor' and 'poor, non-poor' in the transition Q1–Q2. The latter point estimates differ from the true rates by no more than 3 percentage points, which seems like an extremely good performance for the synthetic panels method.

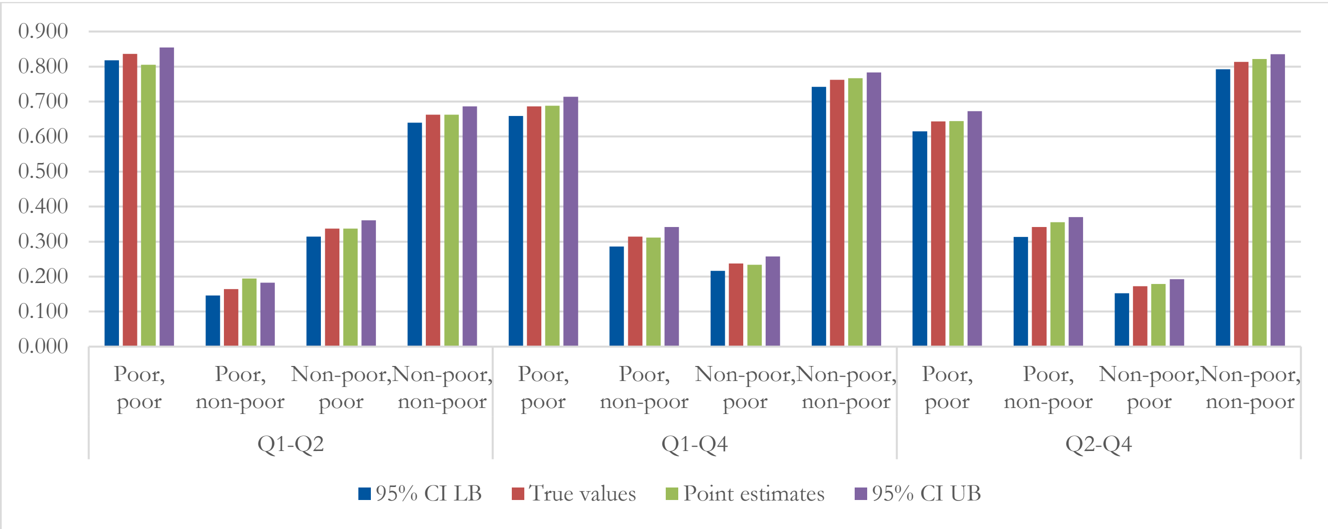
Figure 1: True transition rates (with CIs) and point estimates obtained using synthetic panels: unconditional probabilities



Note: the poverty rates are estimated using the national poverty lines provided with the survey data. Unconditional probabilities are presented for each of the four categories displayed. For example, 'Poor, poor' in the Q1–Q2 section indicates the fraction of the population that was poor in Q1 and poor in Q2. The true estimates presented here differ from those described in Table 2 because the estimation sample was restricted to households whose household head is 25–55 years old, as mentioned in the methodology section.

Source: authors' calculations based on household survey data.

Figure 2: True transition rates (with CIs) and point estimates obtained using synthetic panels: conditional probabilities



Note: the poverty rates are estimated using the national poverty lines provided with the survey data. Conditional probabilities are presented for each of the four categories displayed. For example, 'Poor to poor' in the Q1–Q2 section indicates the probability of being poor in quarter 2, given that the individual was also poor in quarter 1. The true estimates presented here differ from those described in Table 2 because the estimation sample was restricted to households whose household head is 25–55 years old, as mentioned in the methodology section.

Source: authors' calculations based on household survey data.

5.4 Poverty dynamics, 1996–2015

In order to analyse poverty dynamics in the period 1996/97–2014/15, we implemented a parsimonious consumption model regression, only including those covariates that were considered more likely to be time invariant: gender, age, and education level of the household head, but adding the provincial and rural dummies as well. It corresponds to Model 2 presented in Section 4.[12] As discussed in footnote 4, the sample is restricted to households whose household head is 25–55 years old in the first survey round under analysis, and the age range is restricted accordingly in subsequent survey rounds (see Dang et al. (2011) for details).

Summary statistics for the variables included and for all survey rounds, obtained without the restrictions imposed on the age of the household head, are presented in Table A2 in the Appendix. Results concerning the consumption model implemented for all available surveys are presented in Table 8. Based on the coefficients obtained from the consumption model, we computed the conditional and unconditional probabilities presented in Tables 9 and 10.[13]

[12] The coefficients for provincial and rural dummies are omitted in Table 8, but are displayed in Table A3 in the Appendix.

[13] It can be noticed that the boundaries of the estimated poverty dynamics are rather wide. As discussed in Section 4, the width of the estimated boundaries depends on the quality of the underlying consumption model—namely, the overall explanatory power and the statistical significance of the individual regressors. Since we only had relatively few time-invariant characteristics that could be reasonably included in the consumption model presented, we decided to use the urban/rural and the province dummies as well. Including these regional characteristics greatly increased the quality of the overall model and helped to take into account shocks occurring at the urban–rural/provincial level (Dang et al. 2011).

Table 8. Consumption model synthetic panel, Mozambique, 1996/97–2014/15: dependent variable is log of household consumption per capita

	1996/97	2002/03	2008/09	2014/15
Household head gender	–0.034	0.007	–0.013	–0.010
	(0.029)	(0.040)	(0.030)	(0.020)
Household head age	0.000	0.001	–0.001	0.001
	(0.001)	(0.001)	(0.002)	(0.001)
Primary school (first cycle, 5 years)	0.177	0.114	0.025	0.132
	(0.028)***	(0.033)***	(0.031)	(0.019)***
Primary school (second cycle, 7 years)	0.329	0.417	0.195	0.224
	(0.046)***	(0.055)***	(0.050)***	(0.024)***
Secondary school (first cycle, 10 years)	0.522	0.725	0.387	0.496
	(0.087)***	(0.064)***	(0.050)***	(0.030)***
Secondary school (second cycle, 12 years)	1.021	1.105	0.799	0.856
	(0.100)***	(0.079)***	(0.069)***	(0.044)***
Tertiary or higher (13+ years)	1.775	2.339	0.993	1.564
	(0.261)***	(0.296)***	(0.113)***	(0.053)***
Constant	3.256	3.242	3.599	3.769
	(0.074)***	(0.088)***	(0.084)***	(0.062)***
Adjusted R^2	0.163	0.191	0.126	0.251
N	5,757	6,130	7,765	22,578

Note: for education level dummies, the years in parentheses represent the corresponding education years. Provincial and rural dummies are omitted (shown in Table A2 in the Appendix). * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$.

Source: authors' calculations based on household survey data.

In Table 9 unconditional probabilities are presented. They give the fraction of population in the selected age range that is in each of the four categories displayed: 'poor, poor', 'poor, non-poor', 'non-poor, poor', and 'non-poor, non-poor'. For example, 'poor, poor' indicates the fraction of population that was poor in year 1 and poor in year 2. This information is important for the study of poverty mobility and immobility. Notwithstanding the relatively wide bounds, in the Mozambican case we can see that a big fraction of the population falls into the category 'poor, poor' in all years. This percentage is lowest when the years 2002/03 and 2014/15 are considered (22.5–35.6 per cent), but in general it appears that about one-third of the population is consistently in this category. We estimate much lower fractions of the population in the categories 'poor, non-poor' and 'non-poor, poor', whereas non-negligible fractions of the population are estimated in all years for the category 'non-poor, non-poor'. These results may point to a situation in which mobility between the different poverty states is not very likely, which is confirmed when analysing conditional probabilities. Conditional probabilities represent the probability of each of the four states displayed in Table 10: 'poor to poor', 'poor to non-poor', 'non-poor to poor', and 'non-poor to non-poor'. For example, 'poor to poor' indicates the probability of being poor in year 2, given that the individual was also poor in year 1. In this case, the estimated bounds are even wider than in the unconditional probability case, but the estimated conditional probabilities with respect to the two states 'poor to poor' and 'non-poor to non-poor' are on average much larger than those estimated for the two states 'poor to non-poor' and 'non-poor to poor'.

Table 9: Non-parametric estimates of unconditional probabilities of poverty mobility and poverty immobility, 1996/97–2014/15

Years	Category	Bounds	
1996/97–2002/03	Poor, poor	0.529	0.337
N = 5,969	Poor, non-poor	0.168	0.271
	Non-poor, poor	0.003	0.195
	Non-poor, non-poor	0.300	0.197
1996/97–2008/09	Poor, poor	0.498	0.303
N = 6,263	Poor, non-poor	0.115	0.243
	Non-poor, poor	0.035	0.230
	Non-poor, non-poor	0.352	0.224
1996/97–2014/15	Poor, poor	0.461	0.273
N = 16,889	Poor, non-poor	0.160	0.275
	Non-poor, poor	0.003	0.191
	Non-poor, non-poor	0.376	0.261
2002/03–2008/09	Poor, poor	0.409	0.230
N =6,716	Poor, non-poor	0.036	0.183
	Non-poor, poor	0.121	0.300
	Non-poor, non-poor	0.435	0.288
2002/03–2014/15	Poor, poor	0.366	0.205
N = 18,047	Poor, non-poor	0.083	0.206
	Non-poor, poor	0.102	0.263
	Non-poor, non-poor	0.449	0.326
2008/09–2014/15	Poor, poor	0.438	0.231
N = 20,781	Poor, non-poor	0.091	0.224
	Non-poor, poor	0.034	0.241
	Non-poor, non-poor	0.437	0.304

Note: the probabilities presented are estimated using the national poverty lines provided in the household surveys. The table provides the fraction of population in the selected age range that is in each of the four categories. For example, 'Poor, poor' indicates the fraction of population that was poor in year 1 and poor in year 2. The upper-bound estimates for poverty mobility (and lower-bound estimates for poverty immobility) are obtained by taking their average values over 50 repetitions (see Dang et al. 2011).

Source: authors' calculations based on household survey data.

Table 10: Non-parametric estimates of conditional probabilities of poverty mobility and poverty immobility, 1996/97–2014/15

Years	Category	Bounds	
1996/97–2002/03	Poor to poor	0.760	0.555
N = 5,969	Poor to non-poor	0.240	0.445
	Non-poor to poor	0.009	0.497
	Non-poor to non-poor	0.991	0.503
1996/97–2008/09	Poor to poor	0.813	0.555
N = 6,263	Poor to non-poor	0.187	0.445
	Non-poor to poor	0.090	0.506
	Non-poor to non-poor	0.910	0.494
1996/97–2014/15	Poor to poor	0.742	0.498
N = 16,889	Poor to non-poor	0.258	0.502
	Non-poor to poor	0.009	0.423
	Non-poor to non-poor	0.991	0.577
2002/03–2008/09	Poor to poor	0.920	0.557
N = 6,716	Poor to non-poor	0.080	0.443
	Non-poor to poor	0.217	0.510
	Non-poor to non-poor	0.783	0.490
2002/03–2014/15	Poor to poor	0.815	0.499
N = 18,047	Poor to non-poor	0.185	0.501
	Non-poor to poor	0.186	0.447
	Non-poor to non-poor	0.814	0.553
2008/09–2014/15	Poor to poor	0.828	0.508
N = 20,781	Poor to non-poor	0.172	0.492
	Non-poor to poor	0.072	0.442
	Non-poor to non-poor	0.928	0.558

Note: the probabilities presented are estimated using the national poverty lines provided in the household surveys. The table provides the probability of each of the four states. For example, 'poor to poor' indicates the probability of being poor in year 2, given that the individual was also poor in year 1. The upper-bound estimates for poverty mobility (and lower-bound estimates for poverty immobility) are obtained by taking their average values over 50 repetitions (see Dang et al. 2011).

Source: authors' calculations based on household survey data .

5.5 An analysis of poverty dynamics in the most recent surveys, 2008/09–2014/15

In this section a more detailed analysis of poverty dynamics is presented for the period 2008/09–2014/15. The household budget surveys IOF 2008/09 and IOF 2014/15 are the last two available surveys in Mozambique and are also the richest in terms of available information and number of samples. Consequently, the analysis of poverty mobility using these two survey rounds is probably the one that deserves more attention and that is more useful for policy-making purposes.

We report here also, for comparison, the upper and lower bound probabilities estimated for poverty mobility and immobility, as presented in Tables 9 and 10; however, in this case we also present the poverty dynamics results obtained using the parametric estimation approach briefly described in the methodology section and presented in detail by Dang et al. (2011, 2014b) and Dang and Lanjouw (2013), and applied, for example, by Dang and Dabalen (2017). In order to get the point estimates for poverty mobility, we need an estimate for the correlation coefficient between household consumption in the two survey rounds, ρ. Following Dang and Lanjouw (2013), we first approximate the simple correlation coefficient with the synthetic panel cohort-level simple correlation coefficient, and then compute the partial correlation coefficient, ρ, using the equations provided in Section 4. With an estimate for ρ in hand, we may then turn to obtaining the point estimates for poverty mobility.

In our case, using Equation 5 we estimate a value of ρ equal to 0.736. This is in line with theory that expects a value of ρ to be bounded in the interval [0, 1]. Plus, validation analyses of the synthetic panel approach implemented using surveys from other countries have found that ρ is generally found within the interval [0.2, 0.8] (Dang and Lanjouw 2013).[14] The point estimates for the unconditional and conditional poverty transition probabilities obtained using this estimate of ρ, together with the lower and upper bounds already shown in Tables 9 and 10, are presented in Table 11. This analysis confirms that there is a high percentage of the population that is poor in both the first and second periods (37 per cent), or that is non-poor in both the first and second periods (39 per cent). The percentage of the population that is estimated to be poor in 2008/09 and non-poor in 2014/15 is about 15 per cent, only 5 percentage points more than the proportion of the population that is non-poor in 2008/09 and poor in 2014/15 (about 10 per cent), meaning that overall poverty mobility is not very likely, but also that the movements between the poverty and non-poverty states occur in both directions and in comparable magnitudes.

Table 11: Unconditional and conditional probabilities for poverty mobility, point estimates and bounds, 2008/09 and 2014/15

	State	Bounds		Point estimate
Unconditional probabilities	Poor, poor	0.438	0.231	0.371
	Poor, non-poor	0.091	0.224	0.147
	Non-poor, poor	0.034	0.241	0.097
	Non-poor, non-poor	0.437	0.304	0.385
Conditional probabilities	Poor to poor	0.828	0.508	0.690
	Poor to non-poor	0.172	0.492	0.310
	Non-poor to poor	0.072	0.442	0.213
	Non-poor to non-poor	0.928	0.558	0.787

Note: the probabilities presented are estimated using the national poverty lines provided in the household surveys. The 'unconditional probabilities' panel provides the fraction of the population in the selected age range that is in each of the four categories. For example, 'Poor, poor' indicates the fraction of population that was poor in year 1 and poor in year 2. The 'conditional probabilities' panel provides the probability of each of the four states. For example, 'Poor to poor' indicates the probability of being poor in year 2, given that the individual was also poor in year 1. The upper-bound estimates for poverty mobility (and lower-bound estimates for poverty immobility) are obtained by taking their average values over 50 repetitions (see Dang et al. 2011).

Source: authors' calculations based on household survey data.

With respect to conditional probabilities, we estimate that the probability of a person being poor in 2014/15 given that he/she was poor in 2008/09 is substantially higher than the probability of becoming non-poor (69 versus 31 per cent). The high degree of poverty immobility is also reinforced by the high probability of remaining non-poor given that the person was non-poor in the previous period (79 per cent), as opposed to a probability of 21 per cent of becoming poor given that the individual was non-poor in 2008/09. Thus, conditional probabilities show that

[14] As mentioned, Dang and Lanjouw (2013) propose an alternative approximation for ρ: $\rho = \dfrac{\rho_{y_{i1}y_{i2}} - \sqrt{R_1^2 R_2^2}}{\sqrt{1-R_1^2}\sqrt{1-R_2^2}}$,

where ρ_{yi1yi2} is the simple correlation coefficient, and R_1^2 and R_2^2 represent the coefficients of determination obtained from estimating the consumption model in rounds 1 and 2. Using this alternative approximation, we obtain a value for ρ equal to 0.655.

upward mobility is still more likely than downward mobility (31 versus 21 per cent), but still less likely than remaining in the initial poverty or non-poverty state.

5.6 Poverty and vulnerability dynamics in most recent surveys, 2008/09–2014/15

In what follows, an analysis of vulnerability is also provided on top of the analysis of poverty dynamics, following the approach outlined by Dang and Lanjouw (2014). We already discussed in Section 4 that this is based on the identification of a group of vulnerable individuals out of the group of the non-poor. In particular, the vulnerable are those individuals whose daily real consumption per capita lies between the poverty line and the vulnerability line. Consequently, the non-vulnerable (also defined as 'middle-class', 'secure', or 'prosperous') are the individuals whose daily real consumption per capita lies above the vulnerability line. In this analysis, following Dang and Lanjouw (2014), we do not set the vulnerability line at a value that is an arbitrary scaling up of the poverty line, but we derive the vulnerability line from a specified vulnerability index,[15] which is defined either as the probability of becoming poor at time 2 conditional on being in the middle-class at time 1, or as the probability of becoming poor at time 2 conditional on being vulnerable at time 1. The former is indicated as P^1, and is then defined as the 'insecurity index'. Conversely, when the index of vulnerability is defined as the probability of becoming poor at time 2 conditional on being vulnerable at time 1, it is indicated as P^2 and is defined as the 'vulnerability index'. Once the value of the insecurity index, P^1, or the value of the vulnerability index, P^2, are selected—derived, for example, from budgetary planning, social welfare objectives, or relative concepts of well-being (Dang and Lanjouw 2014)—it is possible to compute the value for the vulnerability line and derive the transition probabilities for the three groups: poor, vulnerable, and middle-class. In the Mozambican case, we do not have clear guidance from economic and policy reports on these vulnerability-related targets, so in what follows we present a general analysis that can be subsequently tailored to the objectives of national policy-makers once they become available and/or official.

Computing the vulnerability line for different values of the insecurity and vulnerability indices, it appears that the probability of falling into poverty in 2014/15 given that the individual was vulnerable or even middle-class in 2008/09 is higher than in other contexts (see Dang and Dabalen 2017; Dang and Lanjouw 2014).[16] Hence, in the present analysis we start with a vulnerability index, P^2, of 25 per cent. This means fixing the probability of becoming poor in 2014/15 conditional on being vulnerable in 2008/09 at 25 per cent. This entails setting the vulnerability line at a value of 75.3 meticais, which in turn corresponds to scaling up the original poverty line by about 158 per cent and considering about 40 per cent of the population as 'vulnerable'.[17]

[15] Dang and Dabalen (2017) and Dang and Lanjouw (2016) present the advantages of their approach with respect to choosing the cut-off points identifying the different income groups using, for example, a range of fixed percentiles of the income distribution (as in Alesina and Perotti 1996) or some absolute cut-off thresholds (as in Banerjee and Duflo 2008).

[16] As an example, Dang and Dabalen (2017) in their study use a vulnerability index of 15 per cent, but in the present case if we implemented the same vulnerability index we would get a very high vulnerability line that would leave in the middle-class group only a tiny percentage of the population.

[17] With this vulnerability line, the insecurity index, P^1, corresponds to 3.4 per cent.

The proportion of the population in each group is shown in Table 12.[18] Poverty reduced in the 2008/09–2014/15 period by about 5.5 percentage points, whereas the vulnerable group increased only slightly in percentage terms, and the middle-class expanded.[19]

Table 12: Proportion of the population in the poor, vulnerable, and middle-class groups with an insecurity index, P^2, set at 25 per cent, 2008/09–2014/15

Year	Poor	Vulnerable	Middle-class
2008/09	0.527	0.397	0.076
2014/15	0.472	0.415	0.113

Note: the proportions presented are estimated using the national poverty lines provided in the household surveys and the newly computed vulnerability line.

Source: authors' calculations based on household surveys data.

The unconditional and conditional probabilities of mobility among the three groups identified by the poverty line and the newly computed vulnerability line are presented in Table 13.[20] With respect to unconditional probabilities, we have already noted that a big proportion of the population is estimated to be poor in both periods (37 per cent); nonetheless, a substantial proportion of the entire population, about 14 per cent, is found in the state 'poor, vulnerable', reflecting the reduction in poverty observed between the two surveys. At the same time, about 23 per cent of the population is found to be vulnerable in both periods, highlighting that even for households that are not in poverty, there is a relatively high chance to remain in the vulnerable group for relatively long periods. With respect to the middle-class, only a small proportion of the population is in this category, and the proportion of the population that is middle-class in both periods is even smaller (about 6 per cent). Overall, there seems to exist greater mobility between the poor and the vulnerable group, whereas much more limited mobility towards the middle-class is observed.

Concerning conditional probabilities, and given the selected vulnerability index and vulnerability line, it appears that there is a non-negligible probability for the poor in 2008/09 to become vulnerable in 2014/15 (about 28 per cent), but it seems extremely unlikely for the poor to become middle-class (about 2 per cent). Conversely, for an individual who was in the vulnerable group in 2008/09, there is a high probability of remaining in the same group (59 per cent), whereas the probability of becoming middle-class is lower than the probability of entering poverty (16 versus 25 per cent). At the same time, individuals who were middle-class in 2008/09 are only slightly more likely to remain in the same group than becoming vulnerable, which seems surprising and highlights the relatively high risk of a downward transition faced by even relatively well-off households.

[18] These numbers differ, even if not very substantially, from the official poverty estimates for Mozambique because of the restrictions in the age range adopted in the synthetic panel analysis.

[19] This scenario, in which the first group reduces and the second and third groups expand, is defined by Dang and Dabalen (2017) as a scenario of 'more positive pro-poor growth', which is the second possible best scenario in their classification.

[20] In what follows, only the point estimates for poverty transitions obtained using the parametric approach are presented.

Table 13: Point estimates for poverty–vulnerability transitions, 2008/09–2014/15

Unconditional probabilities	Point estimate	Conditional probabilities	Point estimate
Poor, poor	0.370	Poor to poor	0.692
Poor, vulnerable	0.137	Poor to vulnerable	0.284
Poor, middle-class	0.008	Poor to middle-class	0.024
Vulnerable, poor	0.096	Vulnerable to poor	0.250
Vulnerable, vulnerable	0.227	Vulnerable to vulnerable	0.592
Vulnerable, middle-class	0.063	Vulnerable to middle-class	0.158
Middle-class, poor	0.003	Middle-class to poor	0.034
Middle-class, vulnerable	0.037	Middle-class to vulnerable	0.446
Middle-class, middle-class	0.060	Middle-class to middle-class	0.520

Note: the probabilities presented are estimated using the national poverty lines provided in the household surveys and a vulnerability line obtained by setting the vulnerability index, P^2, at 25 per cent (see Section 4 and the Appendix). Only the point estimates for poverty transitions obtained using the parametric approach are presented. The 'unconditional probabilities' panel provides the fraction of the population in the selected age range that is in each of the nine categories. For example, 'poor, poor' indicates the fraction of population that was poor in year 1 and poor in year 2. The 'conditional probabilities' panel provides the probability of each of the nine states. For example, 'poor to poor' indicates the probability of being poor in year 2, given that the individual was also poor in year 1.

Source: authors' calculations based on household survey data.

6 Conclusions

Mozambique is one of the poorest countries in the world, with about half of the population still considered to be poor in 2014/15. However, the country has achieved rapid poverty reduction over the period 1996/97–2014/15, with poverty rates passing from about 70 per cent in 1996/97 to 46 per cent in 2014/15 (DEEF 2016). A number of scientific articles and reports has been produced documenting the trends and characteristics associated with poverty relative to the abovementioned period, but given the lack of longitudinal data, not much is known about poverty dynamics and trajectories. This is relevant in poverty analysis, especially for a country like Mozambique, where distinguishing between the chronic poor and those who only happen to be in poverty for a limited period of time is key to addressing the different facets of poverty and to designing effective policies. Indeed, policy-makers could consider addressing more transitory poverty, for example with safety net programmes that can be adjusted to the local or sectoral context, such as sustainable insurance schemes for farmers exposed to weather shocks, well-designed social protection systems, and employment insurance, among others. Conversely, households trapped in poverty will require different interventions that address the structural factors constraining these households from moving out of poverty, such as smallholder agriculture's reduced productivity, lack of decent jobs for the fast-growing young population, poor infrastructure, and access to public services, among others. It is a consolidated result in the literature that the longer people stay in poverty, the lower seems to be their chance of getting out of it.

In this paper we attempted to achieve two objectives: (1) validating the use of the synthetic panels technique in the context of the 2014/15 intra-year panel survey data for Mozambique, focusing on the movements in and out of poverty between different quarters and analysing real and estimated poverty transitions; and (2) applying the synthetic panels technique to the 1996/97, 2002/03, 2008/09, and 2014/15 cross-sectional household budget surveys—after having verified that this technique works reasonably well in the Mozambican context. This permitted us to shed some light on both the intra-year and over-time poverty dynamics in Mozambique.

With respect to the validation of the synthetic panels approach proposed by Dang et al. (2011, 2014b) and Dang and Lanjouw (2013, 2014), we find that poverty rates and poverty transitions estimated using the synthetic panels approach seem to provide results (bounds) that are close to the true values obtained using the real panel data. Moreover, the point estimates for poverty transitions are also close to the true transition rates, in both the unconditional and conditional probability cases, which indicates that this methodology performs well in this intra-year panel setting. We also tested the robustness of the estimations obtained using the synthetic panels approach for different specifications.

With respect to the intra-year poverty dynamics results, we show that the poverty rate is much higher in the months November–February and that Mozambique is characterized by a very high degree of poverty immobility, with a big portion of the population remaining either in poverty or out of poverty over the whole period analysed, with smaller percentages of individuals moving upward or downward. The percentage of individuals moving into poverty is higher between the first survey quarter (mid-August to mid-November 2014) and the second survey quarter (mid-November 2014 to mid-February 2015), corresponding to the dry and the rainy seasons, respectively; whereas the percentage of individuals moving out of poverty is higher between the second and the fourth survey quarters (mid-May to mid-August 2015), corresponding to the rainy and the subsequent dry seasons.

With respect to the application of the synthetic panels to the 1996/97, 2002/03, 2008/09, and 2014/15 cross-sectional household budget surveys, we found that an increasing proportion of people gets out of poverty over time, which is consistent with the poverty reduction process observed, but both the percentage of people staying in poverty and the percentage of people remaining out of poverty over time appear to be far higher, showing a sizeable degree of immobility with respect to the initial poverty status. The probability of being poor given that the individual was also poor in a previous survey round is on average much larger than the one estimated for the transition from poor to non-poor. Looking more closely at the most recent household surveys, we also find that about one-third of the population that was poor in 2008/09 is also poor in 2014/15, with a probability of a person being poor in 2014/15 given that he/she was poor in 2008/09 of about 70 per cent. Our estimates of conditional probabilities also show that upward mobility is still more likely than downward mobility (31 versus 21 per cent), but it is still less likely than remaining in the initial poverty or non-poverty state. When the non-poor are further divided into vulnerable and middle-class, we get a richer analysis with respect to poverty dynamics and trajectories: (1) we find that there seems to exist greater mobility between the poor and the vulnerable groups, whereas a much more limited mobility towards the middle-class is observed; (2) it emerges that in the Mozambican case about 23 per cent of the population is estimated to be vulnerable in both periods, highlighting that even for households that are not in poverty, there is a relatively high chance to remain in the vulnerable group for relatively long periods; (3) for an individual who was in the vulnerable group in 2008/09, there is a 60 per cent probability of remaining in the same group, whereas the probability of becoming middle-class is lower than the probability of entering poverty (16 versus 25 per cent); and (4) at the same time, individuals who were middle-class in 2008/09 are only slightly more likely to remain in the same group than to become vulnerable, which highlights the relatively high risk of a downward transition faced by even relatively well-off households.

This constitutes the first validation exercise of the synthetic panels method in an intra-year panel setting and represents the first attempt to provide an insight into poverty dynamics in Mozambique using all the available survey data. For policy-makers, this analysis provides important insights: given the high degree of poverty immobility and the high probability of remaining in the vulnerable group even if one manages to escape poverty, it seems reasonable to adopt a mix of temporary and chronic poverty approaches to tackle the poverty–vulnerability phenomenon in Mozambique.

However, the value added of this analysis lies exactly in this attempt to quantify the proportion of people transitioning from one state to the other across different rounds and provide an estimation of the probabilities of escaping poverty and/or vulnerability over time. Covering two decades of the post-conflict development, this study contributes to finding viable solutions to reach the Sustainable Development Goal 1 for Mozambique.

References

Alesina, A., and R. Perotti (1996). 'Income Distribution, Political Instability, and Investment'. *European Economic Review*, 40(6): 1203–28. https://doi.org/10.1016/0014-2921(95)00030-5

Arndt, C., M.A. Hussain, E.S. Jones, V. Nhate, F. Tarp, and J. Thurlow (2012). 'Explaining the Evolution of Poverty: The Case of Mozambique'. *American Journal of Agricultural Economics*, 94(4): 854–72. https://doi.org/10.1093/ajae/aas022

Arndt, C., E.S. Jones, and F. Tarp (2016). 'Mozambique: Off-Track or Temporarily Sidelined?' In C. Arndt, A. McKay, and F. Tarp (eds), Growth and Poverty in Sub-Saharan Africa. Oxford: Oxford University Press. https://doi.org/10.1093/acprof:oso/9780198744795.003.0009

Arndt, C., F. Castigo, M. Ibraimo, S. Jones, K. Mahrt, V. Salvucci, R. Santos, and F. Tarp (2018). 'Evolução da Pobreza e Bem-Estar em Moçambique, 1996/97-2014/15'. In S. Forquilha (ed.), *Desafios para Moçambique 2018*. Maputo: Instituto de Estudos Sociais e Económicos (IESE).

Banerjee, A., and E. Duflo. (2008). 'What Is Middle Class about the Middle Classes around the World?' *Journal of Economic Perspectives*, 22(2): 3–28. https://doi.org/10.1257/jep.22.2.3

Bierbaum, M., and F. Gassmann (2012). 'Chronic and Transitory Poverty in the Kyrgyz Republic: What Can Synthetic Panels Tell Us?' MERIT Working Paper 064. Maastricht: United Nations University–Maastricht Economic and Social Research Institute on Innovation and Technology (MERIT).

Bourguignon, F., H. Moreno, and H.A. Dang (2015). 'On the Construction of Synthetic Panels'. Presented at the North East Universities Development Consortium annual conference, Brown University, Providence RI.

Castigo, F., and V. Salvucci (2017). 'Estimativas e Perfil da Pobreza em Moçambique'. Inclusive Growth Mozambique Working Paper 2017. Available at: http://igmozambique.wider.unu.edu/pt/article/estimativas-e-perfil-da-pobreza-em-mo%C3%A7ambique-uma-analise-baseada-no-inqu%C3%A9rito-sobre (accessed 8 October 2019).

Cruces, G., P. Lanjouw, L. Lucchetti, E. Perova, R. Vakis, and M. Viollaz (2015). 'Estimating Poverty Transitions Repeated Cross-Sections: A Three-Country Validation Exercise'. *Journal of Economic Inequality*, 13: 161–179. https://doi.org/10.1007/s10888-014-9284-9

Dang, H.A., and A.L. Dabalen (2017). 'Is Poverty in Africa Mostly Chronic or Transient? Evidence from Synthetic Panel Data'. Policy Research Working Paper 8033. Washington, DC: World Bank. https://doi.org/10.1596/1813-9450-8033

Dang, H.-A., and P. Lanjouw. (2013). 'Measuring Poverty Dynamics with Synthetic Panels Based on Cross-Sections'. Policy Research Working Paper 6504. Washington DC: World Bank. https://doi.org/10.1596/1813-9450-6504

Dang, H.A., and P. Lanjouw. (2014). 'Welfare Dynamics Measurement: Two Definitions of a Vulnerability Line'. Policy Research Working Paper 6944. Washington, DC: World Bank. https://doi.org/10.1596/1813-9450-6944

Dang, H.A., and P.F. Lanjouw (2015). 'Poverty Dynamics in India between 2004 and 2012: Insights from Longitudinal Analysis Using Synthetic Panel Data'. Policy Research Working Paper 7270. Washington, DC: World Bank. https://doi.org/10.1596/1813-9450-7270

Dang, H.A., and P. Lanjouw (2016). 'Toward a New Definition of Shared Prosperity: A Dynamic Perspective from Three Countries'. In K. Basu and J. Stiglitz (eds), *Inequality and Growth: Patterns and Policy*. Basingstoke: Palgrave MacMillan. https://doi.org/10.1057/9781137554543_5

Dang, H.A., and P.F. Lanjouw (2018). 'Poverty Dynamics in India between 2004 and 2012: Insights from Longitudinal Analysis Using Synthetic Panel Data'. *Economic Development and Cultural Change*, 67(1): 131–170. https://doi.org/10.1086/697555

Dang, H.A., P.F. Lanjouw, J. Luoto, and D.J. McKenzie (2011). 'Using Repeated Cross-Sections to Explore Movements In and Out of Poverty'. Policy Research Working Paper 5550. Washington, DC: World Bank. https://doi.org/10.1596/1813-9450-5550

Dang, H.A., P.F. Lanjouw, and R. Swinkels (2014a). 'Who Remained in Poverty, Who Moved Up, and Who Fell Down? An Investigation of Poverty Dynamics in Senegal in the Late 2000s'. Policy Research Working Paper 7141. Washington, DC: World Bank. https://doi.org/10.1596/1813-9450-7141

Dang, H.A., P. Lanjouw, J. Luoto, and D. McKenzie. (2014b). 'Using Repeated Cross-Sections to Explore Movements In and Out of Poverty'. *Journal of Development Economics*, 107: 112–28. https://doi.org/10.1016/j.jdeveco.2013.10.008

DEEF (2016a). *Pobreza e Bem-Estar em Moçambique: Quarta Avaliação Nacional*. Maputo: Direcção de Estudos Económicos e Financeiros, Ministério de Economia e Finanças de Moçambique.

Dercon, S., and P. Krishnan (2000). 'Vulnerability, Seasonality and Poverty in Ethiopia'. *Journal of Development Studies*, 36(6): 25–53. https://doi.org/10.1080/00220380008422653

Dercon, S., and P. Krishnan (2003). 'Poverty in Rural Ethiopia 1989–95: Evidence from Household Panel Data in Selected Villages'. In P. Mosley (ed.), *The New Poverty Strategies*. London: Palgrave Macmillan. https://doi.org/10.1057/9780230371156_8

DNEAP (2010). *Poverty and Wellbeing in Mozambique: Third National Poverty Assessment*. Maputo: Ministério da Planificação e Desenvolvimento.

DNPO (1998). *Poverty and Well-Being in Mozambique: The First National Assessment*. Maputo: Ministério de Plano e Finanças.

DNPO (2004). *Poverty and Well-Being in Mozambique: The Second National Assessment*. Maputo: Ministério da Planificação e Desenvolvimento.

INE (2004). *Relatório Final do Inquérito aos Agregados Familiares sobre Orçamento Familiar, 2002/3*. Maputo: Instituto Nacional de Estatística.

INE (2010). *Relatório Final do Inquérito ao Orçamento Familiar: IOF-2008/9*. Maputo: Instituto Nacional de Estatística.

INE (2015). *Relatório Final do Inquérito ao Orçamento Familiar: IOF-2014/15*. Maputo: Instituto Nacional de Estatística.

Kumar, N., and A.R. Quisumbing (2013). 'Gendered Impacts of the 2007–2008 Food Price Crisis: Evidence Using Panel Data from Rural Ethiopia'. *Food Policy*, 38: 11–22. https://doi.org/10.1016/j.foodpol.2012.10.002

Salvucci, V., and R. Santos (2020). 'Vulnerability to Natural Shocks: Assessing the Short-Term Impact on Consumption and Poverty of the 2015 Flood in Mozambique'. *Ecological Economics*, 176: 106713. https://doi.org/10.1016/j.ecolecon.2020.106713

Appendix

Table A1: Summary statistics for the variables used in the analysis, Q1, Q2, and Q4 of the 2014/15 household survey

	2014/15 Q1					2014/15 Q2					2014/15 Q4				
Variable	Obs.	Mean	SD	Min.	Max.	Obs.	Mean	SD	Min.	Max.	Obs.	Mean	SD.	Min.	Max.
Household real consumption pc	11,505	48.654	102.66	0.578	8,211	10,372	45.155	114.67	0.047	12,062	11,308	46.090	77.329	0.491	5,607
Household head gender	11,505	0.760	0.427	0	1	10,167	0.764	0.425	0	1	10,621	0.761	0.426	0	1
Household head age	11,505	43.747	14.134	14	99	10,167	43.770	14.072	14	95	10,621	44.040	14.114	14	95
No education	11,505	0.291	0.454	0	1	10,167	0.298	0.458	0	1	10,621	0.331	0.471	0	1
Primary school (first cycle, 5y)	11,505	0.391	0.488	0	1	10,167	0.392	0.488	0	1	10,621	0.366	0.482	0	1
Primary school (second cycle, 7y)	11,505	0.152	0.359	0	1	10,167	0.150	0.357	0	1	10,621	0.150	0.357	0	1
Secondary school (first cycle, 10y)	11,505	0.100	0.299	0	1	10,167	0.091	0.288	0	1	10,621	0.086	0.280	0	1
Secondary school (second cycle, 12y)	11,505	0.045	0.206	0	1	10,167	0.046	0.209	0	1	10,621	0.045	0.207	0	1
Tertiary or higher (13+ y)	11,505	0.022	0.146	0	1	10,167	0.022	0.148	0	1	10,621	0.022	0.147	0	1
Niassa	11,498	0.064	0.244	0	1	10,353	0.064	0.246	0	1	11,301	0.064	0.245	0	1
Cabo Delgado	11,498	0.074	0.262	0	1	10,353	0.073	0.260	0	1	11,301	0.074	0.261	0	1
Nampula	11,498	0.195	0.396	0	1	10,353	0.196	0.397	0	1	11,301	0.195	0.396	0	1
Zambezia	11,498	0.187	0.390	0	1	10,353	0.187	0.390	0	1	11,301	0.190	0.392	0	1
Tete	11,498	0.097	0.296	0	1	10,353	0.097	0.296	0	1	11,301	0.099	0.298	0	1
Manica	11,498	0.076	0.264	0	1	10,353	0.075	0.263	0	1	11,301	0.074	0.262	0	1
Sofala	11,498	0.080	0.271	0	1	10,353	0.080	0.271	0	1	11,301	0.078	0.268	0	1
Inhambane	11,498	0.059	0.235	0	1	10,353	0.058	0.234	0	1	11,301	0.058	0.233	0	1
Gaza	11,498	0.055	0.229	0	1	10,353	0.056	0.229	0	1	11,301	0.055	0.227	0	1
Maputo Province	11,498	0.066	0.248	0	1	10,353	0.066	0.248	0	1	11,301	0.066	0.248	0	1
Maputo City	11,498	0.049	0.215	0	1	10,353	0.049	0.215	0	1	11,301	0.049	0.215	0	1
Rural	11,505	0.684	0.465	0	1	10,372	0.684	0.465	0	1	11,308	0.682	0.466	0	1

Note: all estimates are obtained without the restrictions on household head's age and are weighted with population weights. For education level dummies, the years in parentheses represent the corresponding education years. * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$.

Source: author's calculations based on household survey data.

Table A2: Summary statistics for the variables used in the analysis, 1996/97, 2002/03, 2008/09, 2014/15

	1996/97					2002/03				
Variable	Obs.	Mean	SD	Min.	Max.	Obs.	Mean	SD	Min.	Max.
Household real consumption pc	8,250	5.350	5.864	0.300	335.6	8,700	10.924	16.439	0.415	1,574.9
Household head gender	8,273	0.825	0.380	0	1	8,700	0.795	0.404	0	1
Household head age	8,261	43.152	13.544	14	95	8,659	43.160	14.087	15	95
No education	8,273	0.690	0.462	0	1	8,700	0.697	0.460	0	1
Primary school (first cycle, 5y)	8,273	0.203	0.402	0	1	8,700	0.177	0.382	0	1
Primary school (second cycle, 7y)	8,273	0.078	0.267	0	1	8,700	0.070	0.255	0	1
Secondary school (first cycle, 10y)	8,273	0.019	0.137	0	1	8,700	0.034	0.182	0	1
Secondary school (second cycle, 12y)	8,273	0.009	0.096	0	1	8,700	0.020	0.140	0	1
Tertiary or higher (13+ y)	8,273	0.001	0.025	0	1	8,700	0.002	0.039	0	1
Niassa	8,273	0.050	0.218	0	1	8,700	0.051	0.220	0	1
Cabo Delgado	8,273	0.077	0.267	0	1	8,700	0.084	0.278	0	1
Nampula	8,273	0.188	0.390	0	1	8,700	0.188	0.391	0	1
Zambezia	8,273	0.193	0.395	0	1	8,700	0.192	0.394	0	1
Tete	8,273	0.068	0.251	0	1	8,700	0.077	0.266	0	1
Manica	8,273	0.057	0.232	0	1	8,700	0.067	0.250	0	1
Sofala	8,273	0.101	0.302	0	1	8,700	0.084	0.277	0	1
Inhambane	8,273	0.072	0.259	0	1	8,700	0.074	0.261	0	1
Gaza	8,273	0.070	0.255	0	1	8,700	0.070	0.255	0	1
Maputo Province	8,273	0.057	0.231	0	1	8,700	0.056	0.230	0	1
Maputo City	8,273	0.067	0.250	0	1	8,700	0.057	0.233	0	1
Rural	8,273	0.789	0.408	0	1	8,700	0.679	0.467	0	1

Variable	2008/09 Obs.	Mean	SD	Min.	Max.	2014/15 Obs.	Mean	SD	Min.	Max.
Household real consumption pc	10,832	23.829	31.742	0.596	2,844.9	33,185	47.092	99.732	0.049	12,012
Household head gender	10,832	0.758	0.428	0	1	32,293	0.762	0.426	0	1
Household head age	10,813	42.592	13.973	13	105	32,293	43.854	14.107	14	99
No education	10,832	0.265	0.441	0	1	32,293	0.307	0.461	0	1
Primary school (first cycle, 5y)	10,832	0.470	0.499	0	1	32,293	0.383	0.486	0	1
Primary school (second cycle, 7y)	10,832	0.143	0.351	0	1	32,293	0.151	0.358	0	1
Secondary school (first cycle, 10y)	10,832	0.080	0.271	0	1	32,293	0.092	0.289	0	1
Secondary school (second cycle, 12y)	10,832	0.022	0.148	0	1	32,293	0.045	0.207	0	1
Tertiary or higher (13+ y)	10,832	0.020	0.139	0	1	32,293	0.022	0.147	0	1
Niassa	10,832	0.059	0.236	0	1	33,185	0.064	0.245	0	1
Cabo Delgado	10,832	0.078	0.269	0	1	33,185	0.074	0.261	0	1
Nampula	10,832	0.192	0.394	0	1	33,185	0.195	0.396	0	1
Zambezia	10,832	0.190	0.393	0	1	33,185	0.188	0.391	0	1
Tete	10,832	0.090	0.286	0	1	33,185	0.098	0.297	0	1
Manica	10,832	0.070	0.255	0	1	33,185	0.075	0.263	0	1
Sofala	10,832	0.081	0.273	0	1	33,185	0.079	0.270	0	1
Inhambane	10,832	0.061	0.240	0	1	33,185	0.058	0.234	0	1
Gaza	10,832	0.063	0.243	0	1	33,185	0.055	0.228	0	1
Maputo Province	10,832	0.063	0.243	0	1	33,185	0.066	0.248	0	1
Maputo City	10,832	0.052	0.222	0	1	33,185	0.049	0.215	0	1
Rural	10,832	0.696	0.460	0	1	33,185	0.683	0.465	0	1

Note: all estimates are obtained without the restrictions on household head's age and are weighted with population weights. For education level dummies, the years in parentheses represent the corresponding education years. * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$.

Source: author's calculations based on household survey data.

Table A3: Consumption model synthetic panel Mozambique, 1996/97, 2002/03, 2008/09, 2014/15: dependent variable is log of household consumption per capita

	1996/97	2002/03	2008/09	2014/15
Household head gender	–0.034	0.007	–0.013	–0.010
	(0.029)	(0.040)	(0.030)	(0.020)
Household head age	0.000	0.001	–0.001	0.001
	(0.001)	(0.001)	(0.002)	(0.001)
Primary school (first cycle, 5 years)	0.177	0.114	0.025	0.132
	(0.028)***	(0.033)***	(0.031)	(0.019)***
Primary school (second cycle, 7 years)	0.329	0.417	0.195	0.224
	(0.046)***	(0.055)***	(0.050)***	(0.024)***
Secondary school (first cycle, 10 years)	0.522	0.725	0.387	0.496
	(0.087)***	(0.064)***	(0.050)***	(0.030)***
Secondary school (second cycle, 12 years)	1.021	1.105	0.799	0.856
	(0.100)***	(0.079)***	(0.069)***	(0.044)***
Tertiary or higher (13+ years)	1.775	2.339	0.993	1.564
	(0.261)***	(0.296)***	(0.113)***	(0.053)***
Niassa	–0.328	0.015	–0.059	–0.885
	(0.085)***	(0.079)	(0.086)	(0.060)***
Cabo Delgado	–0.046	–0.102	–0.127	–0.584
	(0.086)	(0.100)	(0.079)	(0.059)***
Nampula	–0.273	–0.014	–0.306	–0.748
	(0.082)***	(0.084)	(0.082)***	(0.053)***
Zambezia	–0.215	–0.002	–0.460	–0.772
	(0.073)***	(0.084)	(0.077)***	(0.055)***
Tete	–0.519	–0.316	–0.270	–0.488
	(0.080)***	(0.087)***	(0.094)***	(0.059)***
Manica	–0.099	–0.032	–0.405	–0.591
	(0.103)	(0.095)	(0.082)***	(0.057)***
Sofala	–0.723	0.204	–0.450	–0.595
	(0.078)***	(0.077)***	(0.141)***	(0.067)***
Inhambane	–0.481	–0.541	–0.248	–0.594
	(0.077)***	(0.084)***	(0.083)***	(0.061)***
Gaza	–0.096	0.004	–0.435	–0.655
	(0.089)	(0.077)	(0.094)***	(0.066)***
Maputo Province	–0.197	–0.224	–0.405	–0.121
	(0.080)**	(0.071)***	(0.060)***	(0.056)**
Rural	0.024	0.011	–0.015	0.075
	(0.053)	(0.044)	(0.049)	(0.028)***
Constant	3.256	3.242	3.599	3.769
	(0.074)***	(0.088)***	(0.084)***	(0.062)***
Adjusted R^2	0.163	0.191	0.126	0.251
N	5,757	6,130	7,765	22,578

Note: for education level dummies, the years in parentheses represent the corresponding education years. * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$.

Source: authors' calculations based on household survey data.