

ATTRITION AND FOLLOW-UP RULES IN PANEL SURVEYS: INSIGHTS FROM A TRACKING EXPERIENCE IN MADAGASCAR

BY JULIA VAILLANT*

DIAL, Université Paris-Dauphine, and IRD

Most longitudinal surveys recontact households only if they are still living in the same dwelling, producing very high attrition rates, especially in developing countries where rural–urban migration is prevalent. In this paper, we discuss the implications of the various follow-up rules used in longitudinal surveys in the light of an original tracking survey from Madagascar. This survey attempted in 2005 to search and interview all individuals who were living in the village of Bepako in 1995, the baseline year of a yearly survey, the Rural Observatories. The tracking survey yielded an individual recontact rate of 78.8 percent, more than halving attrition compared to a standard dwelling-based follow-up rule. The tracking reveals a very high rate of out-migration (38.8 percent) and household break-ups, as three-quarters of recontacted households had divided between 1995 and 2005. The average income growth of the sample over the period increases by 28 percentage points when follow-up is extended to those who moved out of their household or village, suggesting that dwelling-based panels give a partial view of the welfare dynamics of the baseline sample. A higher baseline income per capita is associated with a higher probability of staying in Bepako and of being found in the tracking if one moved out. The hardest people to find are the poorest and most isolated. Special attention should be paid to collecting data that enable the identification and follow-up of individuals, without which attrition is likely to remain a source of bias even after a tracking procedure is carried out.

JEL Codes: C81, I32, O12, O15

Keywords: attrition, mobility, panel data, tracking surveys

1. INTRODUCTION

This paper is a methodological contribution to the literature on economic mobility and poverty dynamics in developing countries. The essential input to carry out quantitative and rigorous studies of economic mobility or poverty dynamics is longitudinal data. Repeated interviews of the same units, households for example, using the same questionnaire, at more or less distant time intervals, give the opportunity to look at the evolution of individual well-being indicators, study the persistence of poverty, and the reasons of entrance into and exit from poverty. Although panel data are practically indispensable to study these phenomena, they carry a certain number of disadvantages and limitations such as measurement and sampling issues. Household longitudinal surveys require a rigorous

Notes: The tracking survey used in this study was carried out as part of the research project “Dynamics of Rural Poverty,” which has received financial support from the French Ministry of Research under the ACI initiative, “Sociétés et cultures dans le développement durable.” I would like to extend thanks to Flore Gubert and Anne-Sophie Robilliard for making the data available and for their very helpful comments during the preparation of this paper. I also thank two anonymous referees for their helpful comments and suggestions.

*Correspondence to: Julia Vaillant, UMR DIAL, 4 rue d’Enghien, 75010 Paris, France (vaillant@dial.prd.fr).

definition of who is the successor of a given household, to ensure that the same unit is effectively re-interviewed, which is not straightforward. In addition, given that a particular definition of the longitudinal household is chosen, attrition and panel aging make panel data less representative than cross-sections over time (Glewwe and Jacoby, 2000).

Considering these limitations, Ashenfelter wondered, in 1986, “whether collecting panel data in developing countries . . . made sense” (Ashenfelter *et al.*, 1986). The increasing use since then of panel data econometrics has shown that researchers in development microeconomics consider them to be, at the very least, reliable enough to make useful inferences. However, they have often overlooked these limitations, and in particular, not enough attention has been paid to the issue of attrition (Dercon and Shapiro, 2007). This problem is further aggravated in developing countries where rural–urban migration is prevalent (Alderman *et al.*, 2001). High spatial mobility causes attrition, which in turn causes neglect of these particular economic trajectories in panel data.

In this paper, we discuss attrition and its effect on the measurement of poverty dynamics and economic mobility in the light of an original survey from Madagascar. The dataset is the combination of the Rural Observatories (ROR¹) and a tracking survey. The ROR is a unique survey that has been conducted every year since 1995 in several selected Malagasy villages. A simple follow-up rule was used to create the panel, which was based on the head of the household. If the household head had left the village between two rounds but his spouse had not, she was re-interviewed in his place. This widespread follow-up rule and other factors created large attrition, and, in particular, made the study of migration impossible. One of the eight villages of the ROR, Bepako, was chosen in 2005 as the field for an innovative survey that attempted to find all individuals that belonged to the baseline sample in 1995, even if they had left the village. By cutting attrition by half the tracking survey should improve the study of the dynamics of poverty by reducing the attrition bias and including those who moved out of the village.

Because the data comes from one village only and is not statistically representative of a larger area, we believe it has little external validity. What it can tell us, however, is how much we can learn from some specific methodological choices made in the data collection process. Taking the data as two cross-sections from 1995 and 2005 yields an increase in the average village per capita income of 6.4 percent, while using the panel dimension (without tracking) results in an average per capita income growth of individuals of 37 percent. As the first methodology does not take into account individual dynamics, it is clearly not adapted to the context which requires longitudinal data. Besides, the addition of migrants found in the course of the tracking to the standard panel of individuals who stayed in the same household and in the same village yields an average per capita income growth of 65.7 percent, which compared to the panel without tracking, increases the average income growth over the period by 28 percentage points. Therefore, the

¹The Rural Observatories were subsequently renamed the Rural Observatories Network (Réseau des Observatoires Ruraux), hence the ROR acronym used in this work.

choice to collect longitudinal data and the extent to which sample units are followed are critical as they can determine to a large extent the results found in the studies of income mobility and poverty dynamics.

In this paper, we attempt to bring some insights on the benefits and limitations of a tracking survey in terms of analysis and measurement of poverty dynamics. We try to answer the following questions: Can the attrition bias be reduced by tracking movers? What are the characteristics of those not found in the course of the tracking? What are the implications of using various follow-up rules in longitudinal surveys? In the next section, we review the literature on attrition, follow-up rules in panel surveys, and tracking surveys in developing countries. Section 3 presents the dataset and the tracking experiment carried out in Madagascar. Results of the tracking are discussed in Section 4. In Section 5, we explore the extent of the attrition bias in the ROR, with and without the tracking. Tests of attrition on observables are run, considering first the sample without tracking, then comparing it to the sample with tracking, that is, including individuals who changed households or moved out of the village. In particular, we run probit models of attrition with income at baseline as an explanatory variable. The importance of post-tracking attrition is discussed. We estimate a multinomial model of post-tracking attrition with the various types of attrition as outcomes. We discuss the desirability of such tracking devices, their benefits and limitations, and offer some suggestions as to how these could be improved in the conclusion.

2. CONCEPTUAL FRAMEWORK AND LITERATURE

2.1. *Panel Follow-Up Rules and Attrition*

Attrition has been called the “Achilles heel of longitudinal household surveys” (Thomas *et al.*, 2001). It is the process by which a panel loses its sampling units over time. This can occur if the surveyed unit dies, refuses to participate in the survey, or moves to another dwelling (if the identification of the surveyed household is dwelling-based), village or region. The break-up of a household can create attrition as well, depending on how the follow-up rule is defined. If attrition is not random with respect to the outcome studied, that is, if there is selection of attriters, there will be an attrition bias in the estimates because the balanced sample will not be representative of the population.

A related but slightly different issue is “panel aging,” which also reduces representativeness as the individuals of the sample become older than the population every year. In a basic panel design, sampling is done once, at the onset of the survey, which has the advantage of reducing costs, compared to repeated cross-sections. But as panel members grow older, as individuals leave their households to join or create new ones, the sample will no longer be a random draw of the population. This is not because the sample itself is selected, as in attrition, but because the population out of which it was drawn has changed while the sample has not. Following individual income dynamics, as it requires longitudinal data, implies therefore a loss of representativeness (Ashenfelter *et al.*, 1986; Ravallion, 1992; Deaton, 1997).

A particular case of attrition occurs when respondents are lost because they move out of their initial location, where the survey is run. If migrants are not followed and interviewed, an attrition bias can be potentially generated because it is likely that migrants are not a random sample of the population. Besides, it makes the study of the welfare dynamics of migrants impossible. This creates an important limitation in the study of poverty dynamics and economic mobility using panel data because spatial mobility has been shown to be one of the strategies that individuals implement to improve their welfare and cope with risk (Ashenfelter *et al.*, 1986; Dercon and Shapiro, 2007).

While part of attrition is completely independent of the surveyor's will (deaths, for example), the survey methodology and in particular the choice of a follow-up rule can crucially determine the extent of attrition in panel data. The World Bank's Living Standard Measurement Surveys program recommends a dwelling-based follow-up rule, which means that the sampling unit is a dwelling that is randomly drawn in the population of enumerated dwellings, and whoever lives there will be the interviewed household. If the household leaves the house between survey rounds, then it will be lost to follow-up, which will result in attrition in the data (Glewwe and Jacoby, 2000). The sample is updated at each wave by adding a random sample of newly-built dwellings. A major inconvenience of these widely used residence-based panels is that they cannot take into account geographical mobility: migrants are automatically excluded from the panel after they move.

Several papers have brought arguments against this type of follow-up. Rosenzweig (2003) or Thomas *et al.* (2001), for example, find that it creates significant biases in the study of economic mobility, and recommend following movers if they leave the baseline location. This is the purpose of tracking surveys, which can be local, national, or even international. These surveys try to reduce attrition due to spatial mobility by searching and interviewing those who have moved away from their initial location and would be lost to follow-up if no search was attempted. A trade-off has to be made between a rigorous and comprehensive follow-up of respondents, through tracking surveys, and the cost of the survey. This will depend on how much mobility there is in the sample. If an extremely small fraction of the sample has moved far away, it might not be beneficial to try to find them, given the poor representativeness of these units compared to the cost of follow-up. A very mobile sample would gain from a local or national tracking as migration is an important characteristic of the population studied (Glewwe and Jacoby, 2000).

In theory, all attrition due to spatial mobility could vanish with the use of extensive tracking surveys (regardless of their cost). But the problem is more complex than that as the choice of a sampling unit (between individuals and households, for example) is not straightforward and results in different attrition patterns.² Most living standard surveys are household panels, and although they usually collect information on members of the households, they do not include exactly the same individuals each year as they move in and out of the household. Panel of individuals, on the other hand, attempt to interview the same individuals

²In this paper we consider only household surveys. Firms, for example, can be sampling units in other types of surveys.

in each round even if the household they belong to changes or breaks-up between waves. The choice between individual- and household-based panels will have different consequences in terms of attrition and panel aging. If one attempts to interview all the individuals of a household, attrition is likely to be higher than if the household is followed as a global entity. However, a household-based panel can create more panel aging as it does not account for household break-ups and formations.

The central issue here is the definition of the household and how to take into account its dynamic nature, especially in developing countries. Taking the apparently simple example of a nuclear family in which the couple divorces in the subsequent round and the children live with the mother in another location, we see how the problem is far from obvious. Which of the two new households is the successor of the initial one? Should it be the one where the initial household head lives or where the majority of the initial members reside? It could also be none of the above, according to a definition based on household type (single individual, couple with children, single parent family). Previous works by demographers on developing countries have proposed “longitudinal household” definitions, which set clear rules to decide which households are the same across time and which ones are different (McMillen and Herriot, 1985). This methodology is subject to criticism because studies using longitudinal households as the unit of analysis ignore individuals who do not belong to longitudinal households, that is, those who live in split-offs that have been classified as “different” from the antecedent household. This will create a selection bias as those excluded from the panel have disproportionately experienced events such as migration or divorce (Duncan and Hill, 1985). Besides, if household division is non-random and depends on individual endowments, a sampling rule that drops split-off households will produce substantial biases in estimates of economic mobility (Rosenzweig, 2003). Duncan and Hill (1985) recommend the alternative approach of defining the individual as the unit of analysis rather than the household and attributing to the individual the characteristics of the household in which he lives. More recently, Dercon and Shapiro (2007) make a convincing case for basing sampling and tracking strategies on individuals rather than households in developing countries, based on findings in works on poverty mobility reviewed in their paper.

Surveys based on individuals lead to an increasing number of households in the sample, as, in theory, every household in which a member of a baseline household lives should be interviewed. Unlike classic household- or residence-based panels, they give the opportunity to analyze household formations and break-ups over time. Until fairly recently, this issue was largely ignored in the development microeconomics literature, as the household structure was taken as exogenous (Foster and Rosenzweig, 2002). Changes in household size and structure, through marriage and divorce, births and deaths, and child fostering, are being increasingly viewed by researchers, not only as factors of welfare changes, but also as strategies to move out of poverty and cope with risk. This is particularly true in rural economies, where the household is both the production and consumption locus (Fafchamps and Quisumbing, 2007; DeVreyer *et al.*, 2008).

2.2. Testing the Attrition Bias

Attrition can potentially introduce biased estimates in the econometric analysis of welfare dynamics if the probability of being lost to follow-up is not random with respect to the outcome studied. The extent to which it is more than a theoretical threat to consistent estimates is not entirely clear. It will largely depend on the setting and the survey itself. Several procedures exist to test the existence of attrition.

Attrition is a particular case of sample selection in the context of panel data, whereby in the first period, the sample is random, and in the subsequent rounds, the panel loses units. Let y be the outcome of interest, and x a vector of covariates:

$$(1) \quad y = \beta_0 + \beta_1 x + \varepsilon.$$

We define the dummy variable A which is equal to 1 if a household attrited between baseline and the period under study. Thus, y is observed only if $A = 0$. Let A^* be a latent index, where $A = 0$ if $A^* < 0$ and $A = 1$ if $A^* \geq 0$. The selection process can be modeled as:

$$(2) \quad A = \delta_0 + \delta_1 x + \delta_2 z + v,$$

where z is an auxiliary variable observed for all units but not included in x . Lagged values of y can play the role of z .

To test the existence of attrition on observables,³ one can estimate equation (2) using a probit model, and test whether $\delta_2 = 0$, that is, whether baseline y really affects attrition. In addition, the BGLW test (based on the procedure developed by Becketti *et al.*, 1988), closely related to the first one, estimates equation (2) at baseline, adding the dummy A , indicating future attrition and a set of interactions of x and A . This test indicates whether the slope coefficients and the intercept are significantly different between attriters and non-attriters, and how representative the former is of the full sample in terms of initial behavioral relationships.

These tests were run by Becketti *et al.* (1988) and Fitzgerald *et al.* (1998) on the Panel Study of Income Dynamics, an American panel started in 1968. They do not find a significant attrition bias in the data, despite large attrition. Similar results are found by Alderman *et al.* (2001) in developing countries (Bolivia, Kenya, and South Africa), and they suggest that attrition does not necessarily create inconsistent estimates, which gives support to the collection of longitudinal data, despite large attrition in the samples. These results are at odds with evidence from tracking surveys carried out in several developing countries, as described in the next section.

³We consider in this paper tests of attrition on observables, which occurs when observable determinants of attrition are correlated with the outcome studied, conditional on the covariates used in the main equation estimated. Attrition on unobservables requires either instruments or census data, both of which are difficult to find. See Fitzgerald *et al.* (1998) and Alderman *et al.* (2001) for in-depth discussions of the different types of attrition.

2.3. Tracking Surveys in Developing Countries

Ashenfelter *et al.* (1986) stated that “if panel surveys in developing countries are intended to generate accurate information on income dynamics, they must attempt to follow movers or else devise methods to correct for the bias.” Even though this idea is not recent and the potential selection bias created by not following movers demonstrated, surveys tracking movers were almost never carried out until recently. Tracking surveys aim to find and interview households or individuals in a panel, even if they have moved away from their initial location, which ultimately reduces attrition and, hopefully, the associated statistical bias. We have summarized in Table 1 the main surveys in which the follow-up implied some sort of tracking of households or individuals who moved out of their village.⁴ The table covers socioeconomic, demographic, and health surveys run in developing countries in which the household is the main survey unit and for which the information on recontact rates using both a dwelling-based follow-up rule and a tracking of migrants is available. Tracking surveys were carried out 4 to 19 years after baseline.

We should note here that we did not include cohort studies in the review. Cohort surveys are longitudinal and follow a group of individuals who share a common characteristic, such as the birth year. They are widely used, especially in medical research or to follow the health and socioeconomic status of children, adolescents, or women for example, such as Young Lives, Birth to Twenty, or the Cebu Longitudinal Health and Nutrition survey (Norris *et al.*, 2007; Outes-Leon and Dercon, 2008). Although similar to panel surveys, they are somewhat different in the initial design: individual follow-up is planned at the outset, and often there is some kind of tracking of migrants, as in the Young Lives survey (Outes-Leon and Dercon, 2008). Unlike household panels, as they are based on individuals, the problem of defining the successor of the sampling unit is irrelevant. In addition, cohort studies do not face the problem of an ever-growing sample size as is the case when all households including an individual from the baseline sample are interviewed at follow-up.

Using the LSMS follow-up rule as a benchmark, we define in Table 1 the recontact rate without tracking as the proportion of the sample still living in the same dwelling at the time of follow-up. Individual tracking yields lower recontact rates than household tracking because the former attempts to find all individuals targeted, whereas the latter considers a household recontacted if at least one member was found. This is not true for Bangladesh, where individuals were not actually physically searched and interviewed, but data were collected from key informants still residing in the initial village, which explains the 100 percent recontact rate. Tracking protocols managed to increase the recontact rate by 4–48 percentage points depending on the survey (Table 1). For comparative purposes, a line summarizing the results of the survey described in this paper is added at the bottom of the table.

Studies cited in Table 1 analyze the characteristics of movers that were found thanks to the tracking, and would have been attriters otherwise. They generally

⁴This table updates a previous, similar table in Dercon and Shapiro (2007).

TABLE 1
SUMMARY OF TRACKING SURVEYS IN DEVELOPING COUNTRIES

Survey	Baseline	Baseline Sample	Follow-Up	Target	Recontact Rate		Source
					Without Tracking	With Tracking	
Bangladesh Nutrition Survey	1981–82	308 households	2000	Males aged <30 at baseline	57%	100%	Rosenzweig (2003)
Indonesia Family Life Survey	1993	7000 households	1997	Households	84%	94%	Thomas <i>et al.</i> (2001)
Malaysian Family Life Survey	1976–77	1262 families including an ever-married woman	1988–89	Ever-married women	46%	72%	Haaga <i>et al.</i> (1994)
KwaZulu-Natal Income Dynamics Study (South Africa)	1993	1354 households	1998	Households	80%	84%	Alderman <i>et al.</i> (2001)
Kagera Health and Development Survey (Tanzania)	1991–94	912 households; 6204 individuals	2004	Individuals	52%	82%	Maluccio (2000)
Rural Household Survey (Philippines)	1996	940 households	2003	Households	89%	96%	Beegle <i>et al.</i> (2011)
Rural Observatories, Bepako (Madagascar)	1995	307 households; 1490 individuals	2005	Individuals	49%	79%	Fuwa (2011) Author's calculation

Notes: “Target” is the sample sought to be recontacted at follow-up. “Individuals” means all members of baseline households. “Households” means at least one member of baseline household. Recontact rates without tracking represent the percentage of target living in the same dwelling as in baseline and recontacted. In Bangladesh, individuals were not actually tracked; information on movers was obtained from interviews of key informants in the initial village, which explains the 100% recontact rate.

show that the baseline characteristics of movers are significantly different from those of the stayers. But more importantly, these surveys reveal quite different economic dynamics between movers and stayers, suggesting the presence of attrition on unobservables in the data. As noted by Dercon and Shapiro (2007), the analysis of baseline correlates of attrition does not give sufficient proof of the presence or absence of an attrition bias in the data. Positive and negative shocks that occur between baseline and subsequent rounds can trigger attrition, by encouraging migration, for example. However, these shocks are by definition not observed for attriters. This creates selection on unobservables, which will generate an attrition bias if the outcome studied is correlated with the shock.

Several studies using tracking survey data offer information on the extent of non-random attrition and the benefits of tracking migrants on the reduction of the attrition bias with respect to economic variables. The Indonesia Family Life Survey is a panel survey which tracked migrants in the follow-up round four years after baseline (Thomas *et al.*, 2001). The authors find that attrition is positively related to the household baseline per capita expenditures, but this effect disappears when community characteristics are controlled for. Within a community, the poorest households are most likely to move and least likely to be found through the tracking. Results also suggest that households found in the extensive tracking have similar characteristics to those who were not found, while local movers are more similar to non-movers. Using South African data from the Kwa-Zulu Natal Income Dynamics Study, Maluccio (2000) finds that it is the wealthier that are most likely to move and be found. Attrition appears to bias coefficients of the determinants of household-level expenditures. This study illustrates how the attrition bias is specific to the outcome studied, as, using the same dataset but a different outcome (child nutritional status), Alderman *et al.* (2001) do not find an attrition bias. In the Kagera Health and Development Survey, migrants were tracked in 2004, 10–13 years after baseline. The average change in consumption or in the poverty headcount over the period is significantly different between those who stayed in the village or the area and those who moved further away. Migrants improved their consumption by a large percentage more than stayers, even though baseline characteristics between the two groups were similar (Beegle *et al.*, 2011). The Rural Household Survey in the Philippines also tried to recontact migrants in 2003, with a baseline in 1996. Results show that migration in the sample is selective. Regressions of per capita consumption at baseline and follow-up on a set of covariates interacted with a migrant indicator (the BGLW test) suggest that coefficients are not affected by whether or not attriters are included in the analysis (Fuwa, 2011). These studies suggest that the extent of the attrition bias and the benefits of tracking migrants is an empirical issue, which depends on both the setting and the outcome studied.

3. PRESENTATION OF THE SURVEY AND THE TRACKING EXPERIMENT

3.1. *Rural Observatories: An Original Longitudinal Survey Design*

The Rural Observatories of Madagascar were set up in 1995. A socioeconomic observatory is a statistical tool whose aim is to follow and monitor the

population of a particular area in order to identify the dynamics of improvement or worsening of the living conditions of that population.⁵ The Rural Observatories are intended to illustrate particular key issues of Malagasy agriculture and were selected according to several criteria: the agro-climatic zone, the dominant production system, demographic characteristics, remoteness, etc. (Droy *et al.*, 2000).⁶ In 1995, four observatories were set up. Two villages were chosen in each observatory, across which 500 households were randomly drawn and interviewed. These households have been interviewed every year since then. The village of Bepako was chosen in 2005 as the field for a tracking survey, therefore we will focus the presentation and discussion on the ROR survey in this village.

Bepako is located about 80 km from the third most important city of the country, Mahajanga, and 5 km away from Marovoay, the nearest town. It is situated at the heart of one of the irrigated perimeters of Madagascar, a large, flat, paddy growing area, where irrigation infrastructures such as canals and pumps were developed during the colonial era. This area grows a very significant amount of the total rice produced in Madagascar, and most of the population in Bepako is involved in the paddy growing sector, as farmers, landowners, or day laborers. The indigenous ethnic group, the Sakalava, is traditionally a tribe of nomadic cattle raisers rather than farmers. Most of the inhabitants of the region are migrants from the East and the Central Highlands of the country, who came to the Mahajanga region as farm workers when the irrigated perimeter was set up. When the infrastructure and land were not longer state owned and run, they had to buy the land and became smallholders. The irrigation scheme faced a serious crisis in the 1980s and the farmers no longer had an easy access to inputs and agricultural tools. The Observatory attempts to collect data to analyze the strategies implemented by the farmers to deal with these issues.

The questionnaire is around 10–15 pages long and contains modules on the demographic and social characteristics of the household, its dwelling and level of comfort, expenditures on food, non-food and durables, and comprehensive and detailed modules on the farm inputs and outputs of the households. There are specific sections for paddy production, consumption and sales, and for other crops, livestock, and livestock products.

3.2. *Follow-up Methodology and Attrition in the ROR*

The main advantage of the Rural Observatories survey is that it is a yearly household panel. A household is defined in the interviewer's guide as "a group of persons related or not, who live under the same roof or on the same compound,

⁵The Rural Observatories are similar to village level studies such as the ICRISAT panel in India, or the IFPRI survey in Ethiopia (Dercon and Krishnan, 1998; Badiani *et al.*, 2007). See also Lanjouw (2000) for an example of how village studies can be used to study long term rural growth and development in specific areas.

⁶The Rural Observatories were created in 1995 as part of the MADIO project. In 1994, the Malagasy National Statistical Institute (INSTAT) and the French Research Institute for Development (IRD), and in particular DIAL, a research center which belongs to the IRD, put together a four-year research program in Madagascar, called MADIO. MADIO had several goals: to promote economic research in Madagascar; to produce analyses of the transition that was taking place in the 1990s; and to renovate the statistical system by creating several different specific surveys, among which were the Rural Observatories (Droy *et al.*, 2000).

take their meals together or in small groups, share part or all of their income for the good functioning of the group, and whose expenditures depend on a common authority called the ‘household head’” (ROR, 1995; author’s translation). Households are interviewed every year. At the start of each round, the survey team takes a census of the village to count the population, see who moved out, and whether there are new households. The follow-up methodology is based on the dwelling. Enumerators look for panel households in the house in which they lived the previous year. However, the survey zone is very small, so in practice, enumerators would search the household within the village if it had changed houses. However, in the context of a small Malagasy village, for an entire household to change dwelling while remaining in the village is uncommon, therefore, for simplicity, we will call the rule “dwelling-based,” which will refer to a follow-up of households within the village. If the head of the household dies or moves out, the remaining spouse is surveyed instead. Households are lost to follow-up if both the head and its spouse leave the village: all migrants are thus excluded from the panel. If a household surveyed the previous year is not found, refuses to answer, or has died, it is replaced by another one randomly drawn from the village census.

The follow-up methodology used in the ROR creates attrition if only by excluding migrants. Besides, the number of villages surveyed in each observatory was progressively extended, while keeping sample sizes constant (500 households per observatory). This created additional attrition. As households were excluded randomly from the panel to reduce the sample size in each village, this should not create an attrition bias. However, it does substantially reduce the size of the samples.

Figure 1 shows yearly recontact rates of baseline households in Bepako. The recontact rate in a given year is calculated as the proportion of households included in the panel every year since baseline. By this definition, re-entries are not

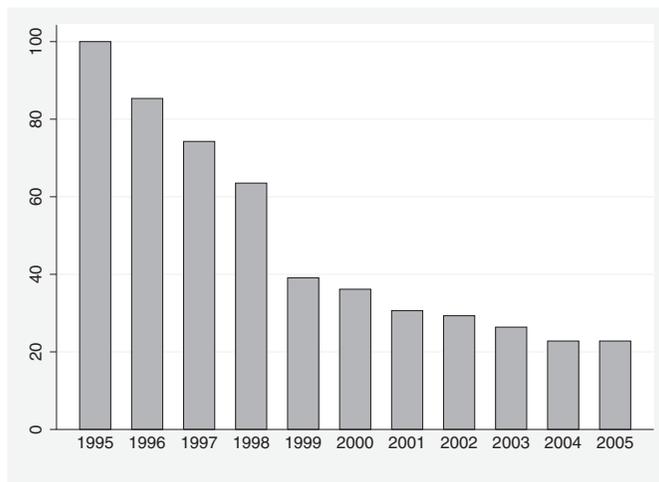


Figure 1. Recontact Rate of Baseline Households by Year in Bepako

Source: ROR data from 1995 to 2005.

allowed, so any household who is not interviewed (because he refuses or is absent for example) in a round is excluded from the panel permanently. The recontact rate is therefore necessarily decreasing. As shown by Figure 1, attrition so defined is very high. A particularly visible drop in recontact rates occurred in 1999 because of the exogenous reduction in sample size mentioned above. The recontact rate at the end of the 10-year period is 24.6 percent: a quarter only of baseline households were interviewed every single year since 1995.

The ROR actually does not exclude re-entries in the panel, but it attempts to follow-up those interviewed the previous year only. Households who re-enter the panel are replacement households, drawn randomly from the village census. They are not identified by the survey team as previous sampled households, and are attributed a new identifier. Thus a merging of rounds based on the household identifier only will yield these very low recontact rates. Only a careful examination of the households in the sample using a dedicated survey, such as the one we will detail below, can enable the construction of a more complete panel, in which households that left and re-entered the panel are identified as the same units. In itself, this operation will lead to an important reduction in the 1995–2005 household attrition rate in the village of Bepako.

Although some data are collected on individuals, such as their age, sex, education, or activity, the survey was not designed to create an individual panel: a person who leaves the surveyed household is not followed. In addition, individual identifying codes are not intended for matching across rounds; they are not reported on pre-filled questionnaires from one year to the next. Matching of individuals across waves is thus a laborious, and sometimes impossible task. In the Malagasy context, this is particularly complex, because “last names” are not “family” names: there is no way of identifying members of a family based on the name. Furthermore, nicknames are often used as substitutes to the official given name and dates of births are declared approximately. As a consequence, creating a panel of individuals *ex post*, by manually matching across names and dates of birth, is extremely difficult in this survey. Any kind of assessment of individual attrition will thus be imprecise.

The design of the ROR survey, which follows households only locally, generates the loss of two categories of individuals, those who move out of their household and those who move out of the village. First, as it attempts to re-interview households as a global entity, defined by its household head, it excludes all individuals who change households between survey rounds. Second, as the households are interviewed only if they still live in the village where they were initially surveyed, it loses all those who move to another area (village, region, country).

3.3. Tracking Survey Design

While the data contained in the ROR panel are very rich, they cannot be used to study migrant welfare dynamics, or to analyze household formations and break-ups in time. To fill in this gap, a tracking survey was implemented in 2005, aiming to find all baseline individuals from Bepako. The choice of this village was motivated by the size of the 1995 sample (307 households) and because it was in fact a

census, as all households in the village were surveyed. This rather experimental enterprise reduced attrition by half. The other motivations for running this survey were to include in the data individual trajectories that are naturally excluded from panels without follow-up, such as migration or household changes, and to account for household structure instability as death, marriages, or divorce occur.⁷ The fieldwork in Bepako was carried out in four steps, described in detail in the following sections.

3.3.1. Census of Bepako

The tracking survey started with a census of the population of Bepako, in July 2005. The list of households living in Bepako obtained through the census was then compared to the theoretical household roster, that is, all households ever included in the ROR since 1995 (for one or more years). Households were classified according to whether they were still living in the village, had moved away, were new in the village, or were deceased, if all the members of the household had died since the last interview.

3.3.2. Individual Trajectories Survey

The next step was to administer an Individual Trajectory questionnaire to households still residing in the village. The questionnaire used was designed to collect information on the composition of the household every year since 1995. The form was pre-filled using the information available in the ROR panel, with the years of presence and absence of each member of the household. The table was completed during the interview with the household head. He was asked to confirm or correct the information, and complete with new members, their entry date, age, sex, link to head, etc. Reasons and years of exit, entrance, and absence of each member were recorded. There were also questions on transfers from the members who had left the households, a detailed module on child fostering, and a module on the shocks undergone in the previous years by the household, such as bad crops or deaths in the family.

If the household head could not be found, another member of the household answered the questionnaire. If the entire household was gone, it was filled in by a neighbor or the village chief. In this case, a Household Tracking Form was used to record the date the household had left the village and their new address, as well as the maximum information that could be given to help the survey team find these households. If the household was found in Bepako, but some (baseline) individual members had left the village, then an Individual Tracking Form was used to collect the same type of information. Even without tracking individuals who moved out of the village, this phase brings valuable information to the data as it identifies the households that had left and re-entered the panel and thus were recorded as distinct households. Thanks to this, the household attrition rate between 1995 and 2005 could be greatly reduced, just by matching households across waves.

⁷The tracking survey was carried out by Flore Gubert and Anne-Sophie Robilliard (UMR 225 DIAL/IRD), in the context of the project “Dynamique de la pauvreté rurale en Afrique sur longue période: le cas de Madagascar.”

3.3.3. Tracking and Interview of Migrants

The tracking forms were the basis of the third step of the project, the search and interview of movers. The aim was to find all individuals who were living in Bepako in 1995 (and were thus included in the baseline sample) and who had left the village since then. Once they were found, they were interviewed using a standard ROR type questionnaire if they had stayed in a rural area or if their main activity was farming or cattle-raising. If they had left the agricultural sector and lived in a town or city, they answered a different, specific urban questionnaire.

A particular choice that was made in this tracking was to limit the search to individuals who had stayed in the region: short- and medium-distance movers, but not long-distance movers. A major reason for restricting the tracking to the region is to be found in the specificity of migrants in Madagascar, where it is a very widespread practice to be buried in one's region of origin, in the same vault as the rest of the family. Many inhabitants of Bepako are in fact migrants or children of migrants, and they originate from far away regions, often the East of the country or the Central Highlands. When they reach old age, if they can afford it, they return to their native region to die and be buried there. The chances of finding them still alive were thus very slim for a high cost of research, which motivated the decision not to try to find them. We will see whether this particular choice was justified *ex post* in the next sections.

3.3.4. ROR Survey in Bepako

Finally, all households living in Bepako (and enumerated in the census mentioned above) were surveyed during the regular ROR round, in the following months of December and January. This means that in both 1995 and 2005, all households of Bepako were interviewed. This is beneficial from the point of view of attrition as it includes all those who were lost from the panel but still living in the village, either because they were excluded in the exogenous sample size reduction in 1999, or because of identifiers mix-ups, or because they had left the panel once and re-entered but were not identified as the same households. This also adds to the wealth of data because it creates two pictures of a village, ten years apart, which can be used to study the demographic evolution of the village, identify newcomer characteristics, etc.

4. TRACKING RESULTS

Results of the tracking survey are considered at the individual level in terms of recontact rates, as the tracking was undertaken on an individual basis, trying to locate not only all baseline households, but also all baseline individuals wherever their location in 2005.⁸ There were 1490 individuals in the 1995 sample that can be classified in three broad categories in the tracking survey: deceased, stayers, and movers. We call "stayer" an individual who, in 2005, would be included in the panel if a dwelling-based follow-up rule were used. This includes individuals who are both still residing in Bepako in 2005 and belong to the same household as in

⁸See Vaillant (2010) for recontact rates at the household level.

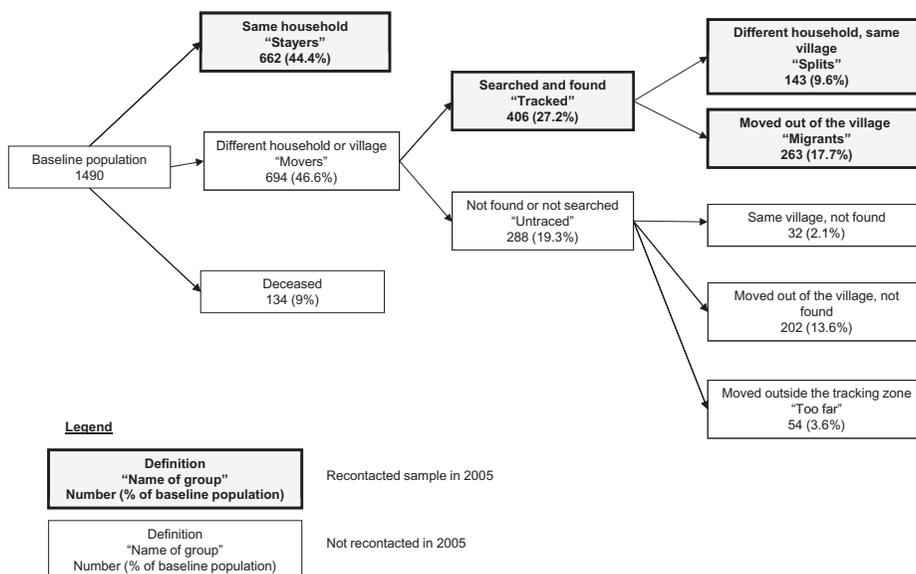


Figure 2. Individual Tracking Results

1995. Once again the difficulty lies in the definition of the “same household.” What is called a stayer, therefore, is someone still living in the household that would be followed by enumerators on the field, who use as a definition the same head or spouse if the head died.⁹ Although there may be objections to this definition, our goal is not to define the successor of a household, which as mentioned previously is a somewhat vain endeavor, but rather to examine the impact of specific follow-up rules on the quality of the data obtained. A “mover” is a person who moved out of his household and either stayed in Bepako or migrated. For the time being, we do not differentiate between local movers (inside the village) and migrants, because with regard to attrition, they would all be lost to follow-up without the tracking. Even though the survey was not designed to create an individual panel, this does not create attrition at the household level but it does at the individual level. Besides, such a person might be living in another household included in the panel but not identified as such. As a consequence, stayers actually represent the panel without tracking, that is, the 1995–2005 panel that would have been obtained without any attempt to track movers. Figure 2 recapitulates frequencies and percentages of the population belonging to each category, as well as the definition of each group.

The tracking survey enabled us to find and interview a total of 1068 “recontacted” individuals, among which 662 were stayers, and 406 movers, as shown in Table 2. Excluding those 134 (9 percent) who had died in the ten-year period, this represents a recontact rate of 78.8 percent. This figure is comparable to the track-

⁹As mentioned earlier, in the infrequent case when an entire household changes dwelling but stays in the village, enumerators interview the household in its new house. So the follow-up rule is a “soft” dwelling-based rule.

TABLE 2
RECONTACT RATE AT THE INDIVIDUAL LEVEL

	Number	%
Recontacted	1068	71.6
Stayers (same household)	662	44.4
Tracked movers (different household or village)	406	27.2
Untraced	288	19.3
Deceased	134	9.0
Total	1490	100.0

Source: ROR data from 1995, tracking survey data from 2005; author's calculation.

ing rate of 82.2 percent in the Tanzania KHDS survey (Beegle *et al.*, 2011; see also Table 1). The remaining 288 individuals were not recontacted, either because they could not be found in the second phase of the survey (“untraced”), or because they had moved outside of the tracking zone (“too far”). The tracking survey thus reduced the attrition rate from 55.6 percent without the tracking procedure, to a final rate of 28.3 percent. If one excludes the deceased, for which the attrition rate could obviously not be reduced, we can see that the tracking reduced by more than half the initial attrition of the panel.

As implied above, there are two types of movers, with regard to attrition: those who moved out of their household and those who moved out of the village. Those of the first type become attritors because the panel is meant to be at the household level and does not attempt to follow individuals who change households. Those of the second type, migrants, are attritors because the follow-up is restricted to the village. Among the movers that could not be found in the tracking, a small number are supposedly still living in Bepako (32 individuals). While this might seem incoherent, it actually means that the person who filled out the Individual Tracking Form declared them to be residing there, but they could not be found or matched in the tracking; 3.6 percent of individuals from the baseline sample were not searched because they lived outside the tracking zone.

Figure 2 shows in particular how numerous migrants are in the sample. Combining tracked, not found, and long-distance movers, those 519 individuals represent an outmigration rate (of survivors) of 38.3 percent. We believe this makes a strong case in favor of such tracking protocols. Migrants are not marginal in the village and excluding them from studies on poverty dynamics and economic mobility is likely to produce biases in the estimates.

An additional argument in favor of this type of survey can be seen in Table 3, which shows the number of households generated by the initial 307 in the baseline sample. There are 69 intact households in 2005, meaning that no member left the household since 1995 (except individuals who died). In most cases, intact households did grow in size, due for example to births or entrances by marriage, but they did not generate any new household over the period studied. The number of splits refers to the number of households in which there is at least one member from the initial one. These figures are actually a lower bound, as for individuals who were

TABLE 3
HOUSEHOLD SPLITS IN 2005

	Number	%
All members deceased	7	2.3
Not found	42	13.7
Recontacted households	258	84.1
Intact household	69	22.5
Two splits	98	31.9
Three splits	41	13.4
Four or more splits	50	16.3
Total	307	100.0

Source: ROR data from 1995 and 2005, tracking survey data from 2005; author's calculation.

not found, we do not know how many households they joined, so they are counted as one split only. For example, suppose household A had 10 members in 1995, 3 of which were still in A in 2005, 2 had created B and 2 had created C. The remaining 3 who were not found are considered to have joined a single household, although it could be 2 or 3. Household A thus generated a minimum of 4 and a maximum of 7 households in 2005. Table 3 demonstrates how much households change, break-up, and recombine in a period of ten years, and how, notwithstanding the definition difficulties of a longitudinal household, following only the initial household considerably restricts the analysis to a particular type of sample.

5. ANALYSIS OF ATTRITION WITH AND WITHOUT TRACKING

The particular structure of the sample enables us to assess the benefits of the tracking in terms of reducing the attrition bias. In this paper, we specifically look at potential biases in the estimates of poverty dynamics and income mobility, although other outcomes are possible and have been studied such as education or health (Falaris, 2003). The selection bias is specific to a particular outcome studied. A biased sample with respect to schooling attainment could be unbiased with respect to labor force participation. In this section, we assess the selection of attritors with respect to income per capita only. Using data collected by the ROR and the tracking survey, this section attempts to find evidence that supports or does not support the use of tracking surveys in developing countries. We inquire into how much can be learned from the supplement of information gained through the tracking survey, and whether the global picture of economic mobility in Bepako is thereby modified.

The analysis is carried out in three steps. First, we compare baseline characteristics of attritors and non-attritors according to a dwelling-based follow-up rule, a village-based and an extensive tracking. The sign and significance level of the per capita income coefficient will indicate whether there is an attrition bias in the dataset when a tracking is not carried out and whether the tracking managed to reduce that bias. Ideally, the tracking should reduce or make the bias disappear, if such a bias exists initially. Next, we compare characteristics of the different groups of migrants, those re-interviewed, those not found in the tracking, and those who

moved outside the recontact zone. We try to find out whether migrants who could not be found are similar to those who were tracked. Finally, we include information retrieved through the tracking in 2005 by comparing the change in income over the period of nested groups, from stayers to movers within the village and outside the village. Beyond the analysis of baseline characteristics, this should inform us on biases in the analysis of income dynamics over the period implied by the lack of tracking.

5.1. *Does the Tracking Bring More Representativity to the ROR Sample?*

5.1.1. Descriptive Statistics

Keeping in mind that the baseline sample is a census of the village, and therefore it is the population, we start by comparing the distribution of observable characteristics of the population to stayers (deceased excluded). Stayers are individuals who stayed in the same household between 1995 and 2005.¹⁰ We add first the splits (those who stayed in Bepako but changed household), then the tracked migrants to the stayers and compare these respective nested groups to the population. The goal is to see whether the sample without tracking was initially representative of the population, and, if not, to what extent and in what way do various follow-up rules increased its representativity. As shown in Table 4, there are numerous differences between stayers and the baseline population: stayers are older, and accordingly, more likely to be married and the head of a household (or his spouse), than single and the child in the household. The place of birth is also a significant difference: those born in the district are overrepresented in the stayers group, which is intuitive, as they are less likely to move back to their region of origin.

Adding individuals who left their initial household but stayed in the village to the group of stayers considerably reduces the differences in demographic characteristics between samples. In a dwelling-based panel, the mere fact of changing households, even while staying in the same village, excludes individuals who leave their parents' home to marry, start up new households, etc. This explains the younger age of splits, as at baseline they are much more often single persons and children of household heads. It is therefore clear that there is a certain type of trajectory that is not accounted for in the original panel (stayers): beginning an adult life, marriage, having young children, etc. The obvious lifecycle effect in these statistics could have a significant impact on the analysis of the dynamics of income. Aging is also very unlikely to be well represented in the data, as individuals who die of old age between 1995 and 2005 are not in the panel. These descriptive statistics suggest that following household heads only and not individuals will lead to biased samples, as shown for instance by Rosenzweig (2003). However, for cost reasons, most surveys attempt to recontact only one person from the original household, while others set rules for the definition of a successor household. Individual follow-up is necessary to ensure demographic representativity of the sample, especially over a rather long period as is the case here. Otherwise, it is

¹⁰See also Figure 2 for definitions of the different groups.

TABLE 4
 BASELINE CHARACTERISTICS OF INDIVIDUALS USING DIFFERENT FOLLOW-UP RULES

Follow-Up Rule	Dwelling-Based	Village-Based	Tracking	
Sample	Baseline Population	Same Household	Stayed in Bepako	All Recontacted
Female	0.50	0.49	0.48	0.49
Age	20.6	22.3**	21.2	20.7
Education	3.04	2.93	3.02	2.99
Married	0.32	0.39***	0.34	0.32
Never married	0.62	0.55***	0.60	0.62
Widowed	0.021	0.024	0.021	0.018
Divorced	0.037	0.036	0.036	0.036
Household head	0.19	0.25***	0.21	0.20
Wife of household head	0.16	0.20**	0.17	0.16
Child of household head	0.51	0.47*	0.51	0.52
Not related	0.13	0.080***	0.11	0.12
Betsileo	0.40	0.42	0.44*	0.40
Merina	0.22	0.25	0.23	0.22
Sakalava	0.13	0.12	0.14	0.13
Antesaka	0.21	0.17**	0.16***	0.21
Other ethnic group	0.047	0.038	0.034	0.037
Born in Marovoay district	0.74	0.79**	0.81***	0.78**
Born in Mahajanga region	0.083	0.069	0.065	0.076
Born in another region	0.18	0.14**	0.13***	0.15**
Head is female	0.035	0.048	0.040	0.037
Age of household head	43.3	42.6	43.8	43.6
Education of head (years)	4.28	4.30	4.16	4.12
Education of wife of head (years)	3.02	3.18	3.03	2.96
Household size	5.96	5.93	6.10	6.14*
Dependency ratio	1.02	1.02	1.01	1.06
Income (per capita)	565.8	602.7	607.2	583.4
Household grows paddy	0.90	0.93**	0.94***	0.92*
Household owns farmland	0.63	0.61	0.64	0.63
Household owns cattle	0.29	0.33*	0.33**	0.31
Dwelling all good quality material	0.17	0.16	0.17	0.16
Dwelling all low quality material	0.31	0.27*	0.27**	0.29
Student	0.21	0.18	0.20	0.21
Main activity is farming	0.50	0.51	0.52	0.50
Main activity not farming	0.032	0.024	0.021	0.027
Unemployed or inactive retired	0.049	0.036	0.037	0.046
Observations	1356	662	805	1068

Notes: Sample means of each group. Asterisks indicate the significance level of mean comparison test of "Population" against sample of "Same household," "Stayed in Bepako," and "All recontacted" respectively. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. "Population" indicates the entire sample of individuals living in Bepako at baseline excluding the deceased. "Same household": individuals who stayed in the same household (stayers). "Stayed in Bepako": "Stayers" + re-interviewed individuals who changed households and stayed in Bepako (splits). "All recontacted": "Stayed in Bepako" + recontacted individuals who moved out of Bepako (migrants).

Source: ROR data from 1995, tracking survey data from 2005; author's calculation.

essential to take lifecycle effects into account when designing panel surveys to avoid these biases.

The tracking of migrants, on the other hand, seems to be important in reducing biases relative to economic characteristics. The proportion of paddy growers is higher among stayers than the whole population, and is still higher after migrants were recontacted. Paddy growing is the main income-generating activity

in Bepako, and not producing rice can be the result of a household not owning or finding land to rent. In this case, its members may be less attached to the land and thus more mobile, or they might be pushed away to find land elsewhere. Another explanation could be that the individual simply is not involved in agriculture, and could be more likely to find a job in an urban area. Similar factors can explain the higher proportion of cattle-owners among stayers. In addition, they have less frequently bad quality dwellings, which could be an indicator of higher wealth and being more permanently settled in the village. In these two dimensions, tracking migrants helped gain some representativity in the sample. The size of the household is significantly higher among the recontacted than the baseline population, which directly stems from the fact that migrants with a large household have a higher chance of having a member remaining in the Bepako, able to inform enumerators on the location of the migrant.¹¹

5.1.2. Attrition Probability Model

As variables in mean comparisons tend to be highly correlated, we now run probit regressions of the determinants at baseline of the probability of being in the sample in 2005, without and with the tracking. We start by defining the dependent variable as indicating whether the individual is a stayer, that is, we look at the determinants of being in the sample using a dwelling-based follow-up rule. Following Thomas *et al.* (2001), we first run the regression with one explanatory variable, the log per capita income,¹² which is our outcome of interest (Table 5, column 1). Higher income per capita increases the probability of staying in Bepako. Adding household size and other controls to the regression actually increases the effect of income (columns 2–3). Therefore, without any kind of correction of attrition, analysis carried out on the non-attriting sample will be biased toward the initially better off. There is a quadratic relationship between age and attrition: the effect of age on the probability of staying in the same household decreases until 34.5 years old, then increases. This is intuitive as the sample includes both household heads, that are unlikely to split off, and children of heads, who move out and get married before 35 years old. Women are less likely to stay in the same household, which we assume is linked to feminine exogamous marriages. Finally, individuals who are more educated than average in their age group are less likely to leave. This is in contradiction with findings in the literature, where education is a factor of mobility (Thomas *et al.*, 2001).

As in the descriptive statistics, we now change the definition of the dependent variable. It is equal to 1 in columns 4–6 if the individual was recontacted in Bepako (includes splits), and in columns 7–9 if the individual was recontacted, whether in or outside of Bepako. Individuals belonging to larger households in 1995 are more likely to be recontacted in 2005, and this effect is particularly strong for migrants. This can be explained by the fact that, as large households are less likely to have

¹¹The results presented in this paper also hold for analyses carried out at the household level (see Vaillant, 2010).

¹²Use of income rather than consumption as a baseline welfare measure is motivated by the absence of sufficient expenditure data in the 1995 questionnaire, while farm income data were collected in detail. For more detail on the construction of the income aggregate, see Vaillant (2010).

TABLE 5
 PROBIT MODEL OF REMAINING IN THE SAMPLE BEFORE AND AFTER TRACKING, INDIVIDUAL LEVEL

Dependent Variable = 1 if	Stayed in the Same Household			Stayed in Bepako			Recontacted		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Log income (per capita)	0.177** (0.078)	0.186** (0.078)	0.259*** (0.086)	0.233*** (0.083)	0.284*** (0.082)	0.340*** (0.086)	0.162* (0.094)	0.233** (0.093)	0.317*** (0.095)
Log size of household		0.050 (0.117)	0.359** (0.151)		0.286*** (0.125)	0.512*** (0.154)		0.470*** (0.111)	0.739*** (0.138)
Female			-0.116* (0.068)			-0.138** (0.068)			-0.084 (0.068)
Age			-0.061*** (0.009)			-0.028*** (0.009)			-0.036*** (0.011)
Age squared			0.001*** (0.000)			0.000*** (0.000)			0.001*** (0.000)
Education (deviation to age mean)			0.045** (0.021)			0.039* (0.021)			0.057** (0.026)
Married			0.893*** (0.162)			0.394** (0.159)			0.361* (0.197)
Head is female			1.398*** (0.273)			0.938*** (0.270)			1.065*** (0.328)
Age of household head			-0.011 (0.027)			-0.046 (0.028)			-0.036 (0.030)
Age of household head squared			0.000 (0.000)			0.000 (0.000)			0.000 (0.000)
Education of head (years)			-0.034 (0.025)			-0.041 (0.025)			-0.083*** (0.026)
Constant	-1.096** (0.489)	-1.231** (0.539)	-1.269* (0.718)	-1.162** (0.518)	-1.954*** (0.551)	-1.223* (0.719)	-0.161 (0.572)	-1.371** (0.645)	-0.684 (0.810)
N	1342	1342	1342	1342	1342	1342	1342	1342	1342
Pseudo-R ²	0.008	0.008	0.072	0.014	0.021	0.046	0.007	0.028	0.070

Notes: * p < 0.1, ** p < 0.05, *** p < 0.01. Robust standard errors adjusted for clustering at the household level in parentheses. Deceased individuals excluded from the regression. "Stayed in the same household" means that the individual stayed in the same household and in Bepako between 1995 and 2005. "Stayed in Bepako" means that the individual stayed in Bepako, either in the same household or in another household. "Recontacted" means the individual is either stayed in Bepako or moved away and was recontacted through the tracking.

Source: ROR data from 1995, tracking survey data from 2005; author's calculation.

entirely moved away from the village, informants on the new location of movers are easier to find. The gender effect disappears when migrants are tracked (column 9), while being a woman significantly reduces the likelihood of being recontacted in the village (column 6). Tracking outside the village thus appears necessary to retrieve a more gender balanced sample, as women are more likely to marry outside the village.

The coefficient for per capita income is higher when splits and migrants are tracked than without tracking. The positive effect of income on the probability of being recontacted suggests that individuals who are found through the tracking are closer to non-movers in terms of economic conditions than those not found. These were situated at the lower end of the income distribution at baseline. This unexpected result suggests that the with-tracking sample is less representative of the population than the without-tracking sample. In addition, the stronger effect of the income variable in the village-based follow-up model (column 6) than in the migrant tracking (column 9) indicates that only following individuals who split off but do not migrate may yield larger biases than not following them at all! Attempting to track migrants in a panel survey therefore does not guarantee that the attrition bias will be substantially reduced. Long-distance and untraced movers may potentially have different characteristics, which would be correlated both with baseline welfare and the likelihood of being found. We investigate this matter in the next section.

5.2. *How Serious is Post-Tracking Attrition?*

5.2.1. Descriptive Statistics

The previous tests of attrition due to selection on observables failed to reject the absence of a bias even after the tracking. A finer analysis of the characteristics of the groups who were not interviewed in 2005 is needed to understand the reasons and implications of this finding. “Post-tracking” attritors belong to three categories: those who died in the period, those who moved in the tracking zone but could not be found, and those who moved outside the tracking zone.¹³ While the first two categories are independent of the will of the survey team, the last category is a methodological choice, motivated by the assumption that individuals moving far away are older and wish to spend the end of their life in their region of origin, and thus the chance of finding them alive is slim compared to the cost of the search. Table 6 shows the mean characteristics of each of these groups, and the significance level of the test of equality of means of each group compared to all recontacted individuals.

Starting with the group of deceased, we see from Table 6 that they are, as expected, much older than average. They are less educated, which is likely to be correlated to their age, and are more likely to be married or ever married than the tracked. The group of deceased is heterogeneous because it also includes the very young, among which death rates are higher than in the rest of the population. Besides the fact that a particular age group, the elderly, is not well represented in

¹³There is also a small group of individuals who supposedly stayed in Bepako but could not be found in the tracking, which we leave out of the analysis.

TABLE 6
BASELINE CHARACTERISTICS OF POST-TRACKING ATTRITORS, INDIVIDUAL LEVEL

	Recontacted	Deceased	Untraced	Too Far
Female	0.49	0.46	0.55	0.39
Age	20.7	37.8***	19.7	22.7
Education	2.99	2.40**	3.08	3.80*
Married	0.32	0.43**	0.31	0.35
Never married	0.62	0.41***	0.62	0.59
Widowed	0.018	0.097***	0.038**	0
Divorced	0.037	0.067*	0.038	0.056
Household head	0.20	0.35***	0.16	0.19
Wife of household head	0.16	0.14	0.17	0.13
Child of household head	0.53	0.32***	0.48	0.48
Not related	0.12	0.19**	0.20***	0.20*
Betsileo	0.40	0.50**	0.38	0.31
Merina	0.22	0.14**	0.24	0.093**
Sakalava	0.13	0.16	0.13	0.056
Antesaka	0.21	0.11***	0.19	0.35**
Other ethnic group	0.037	0.090***	0.060	0.19***
Born in Marovoay district	0.78	0.67***	0.65***	0.31***
Born in Mahajanga region	0.076	0.090	0.10	0.13
Born in another region	0.15	0.24***	0.24***	0.56***
Head is female	0.037	0.060	0.030	0
Age of household head	43.6	52.0***	42.7	41.1
Education of head (years)	4.12	3.20***	4.97***	4.57
Education of wife of head (years)	2.95	2.04***	2.73	5.61***
Household size	6.14	5.84	5.31***	5.31**
Dependency ratio	1.06	0.86***	0.92**	0.64***
Income (per capita)	584.1	588.0	470.4***	630.5
Household grows paddy	0.92	0.89	0.85***	0.69***
Household owns farmland	0.63	0.71*	0.64	0.50**
Household owns cattle	0.31	0.29	0.21***	0.26
Dwelling all good quality material	0.16	0.17	0.18	0.24
Dwelling all low quality material	0.29	0.30	0.39***	0.30
Student	0.21	0.082***	0.21	0.15
Main activity is farming	0.49	0.57	0.52	0.44
Main activity not farming	0.027	0.052	0.030	0.13***
Unemployed or inactive retired	0.046	0.13***	0.051	0.11**
Observations	1068	134	234	54

Notes: Sample means of each group. Stars indicate the significance level of mean comparison test of “Recontacted” against “Deceased,” “Untraced,” “Too far” respectively. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. “Recontacted” indicates sample of all recontacted individuals in 2005, whether in or outside of Bepako (stayers + tracked movers).

Source: ROR data from 1995, tracking survey data from 2005; author’s calculation.

the sample because of the mortality, the group of deceased does not seem to have very different economic characteristics than the tracked.

Turning to the characteristics of individuals who moved too far away to be tracked, we see from Table 6 that they are more often men, with more education. As males are more educated in general, this needs to be verified in the multivariate analysis. However, this suggests that, as those with the most human capital move the furthest, it could be linked to their ability to adapt to a foreign environment. Age does not seem to play a role in this type of move, indicating that the assumption that those who move far away are the elderly is not true after all. This group is however more likely to be born outside the region, which means that they are

potentially returning migrants. With a smaller household size and dependency ratio and a higher income per capita, individuals who move far away seem to be initially better off, giving them the means to migrate. The proportion of paddy growers is the lowest in this group, as well as of landowners, consistent with the fact that they arrived somewhat recently, and are less involved in the traditional economic activities of the village.

The third and last group comprises those individuals who moved inside the tracking zone and could not be found or for whom no information at all on their new location was collected. The issue here is whether this group is very different from the group of tracked movers. If, conditional on having moved out, being found or not is a random event, then the fact that about half of the movers are still attritors after the tracking will not create a bias in the estimates. Comparing observables at baseline of this group and the entire tracked sample, we see that they are more likely to be widowed and not related to the head of the household in which they live, suggesting fewer ties to this household and potentially a weaker social network. The informants from Bepako would then have less precise, if any, information on the whereabouts of that person, making the search difficult or impossible. Consistent with this is the proportion of this group that is born in another region. They migrated to Bepako in the course of their lifetime, and would thus have a smaller social network than those who have been there for longer. They might also be long-distance movers returning to their region of origin. In this case, they would not have been searched anyway. They could also have moved away from the location indicated by the informants in Bepako who are less well informed of their movements. Smaller household sizes are consistent with smaller chances of being found, with less people from the baseline household remaining in the village to give information on movers. Turning to their living conditions in 1995, we see in Table 6 that their income per capita is lower than in the tracked sample. Even though they have fewer members, their income per capita is lower, and their economic conditions seem inferior to average. This could be a push factor for migration. Weak social ties are a factor for lower well-being if the individual is a wage laborer, for example (as landowners tend to favor their kin and friends as day laborers). Clearly these individuals are not well represented in the sample after the tracking. This is one limitation of such survey devices: it is not always possible to recontact individuals, even with time, means, and good will. After a long period of time, some sampling units are bound to be lost from the panel, with little one can do about it.

5.2.2. Multinomial Logit Model of Post-Tracking Attrition

Next, we run a multinomial logit regression, which attempts to find the determinants of the different types of post-tracking attrition, as they may be correlated in the bivariate analysis. The model is run at the individual level, on the sample of movers only, for which we define a dependent variable with three outcomes: moved and tracked, moved but not found, and moved outside the tracking zone. Covariates included in the regression are age, education, sex, and marital status of the individual, the same characteristics of the head of the household, as well as the size and the income per capita in the household. Table 7 shows

TABLE 7
MULTINOMIAL LOGIT REGRESSION

	Tracked Migrant	Untraced Migrant	Too Far
Log income (per capita)	0.0522 (0.0416)	-0.0528 (0.0442)	0.000641 (0.0366)
Female (d)	0.00233 (0.0447)	0.0316 (0.0421)	-0.0339** (0.0167)
Age	-0.0157** (0.00784)	0.0154** (0.00723)	0.000357 (0.00435)
Age squared	0.0229** (0.0109)	-0.0240** (0.0111)	0.00112 (0.00544)
Education	0.0133 (0.0120)	-0.0192 (0.0127)	0.00589 (0.00452)
Married (d)	0.0420 (0.107)	-0.0206 (0.101)	-0.0214 (0.0436)
Head is female (d)	0.243 (0.148)	-0.139 (0.148)	-0.104*** (0.0220)
Age of household head	-0.00942 (0.0180)	-0.00373 (0.0179)	0.0131 (0.00853)
Age of household head squared	0.00597 (0.0175)	0.00950 (0.0174)	-0.0155* (0.00868)
Education of head (years)	-0.0339** (0.0145)	0.0363** (0.0143)	-0.00240 (0.00716)
Log size of household	0.282*** (0.0965)	-0.243*** (0.0868)	-0.0392 (0.0395)

Notes: Marginal effects. Robust standard errors adjusted for clustering at the household level in parentheses. (d) for discrete change of dummy variable from 0 to 1. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Age squared and age of household head squared are /100 for readability

Source: ROR data from 1995, tracking survey data from 2005; author's calculation.

the marginal effects of the regression with the reference category being those who moved and were found.

There is a significant quadratic relationship between the age and the probability of not being found, which is positive until 32 years old, then negative. Consistent with findings from the IFLS and KIDS datasets, the education of the head and the size of the household are both significant (Maluccio, 2000; Thomas *et al.*, 2001). The effect of the household size has already been discussed. The education of the household head may have a positive impact on the mobility of an individual and his propensity to move to an urban area. Obviously the dataset cannot tell us whether urban migrants are less likely to be found than rural migrants, as none of these were in fact found, but it is plausible that tracking in urban areas is rendered difficult by a higher degree of anonymity and weaker social ties in cities. Although a decline in per capita income increased the probability of being lost to the panel in the descriptive analysis, and the sign of the coefficient is negative, it is no longer significant in this model, when compared to movers that were found. This suggests that this variable was highly correlated to other observable characteristics in the descriptive statistics. The non-significance of the coefficient may also be due to relatively small sample sizes. As in Indonesia and South Africa, attrition is somewhat higher among lower income individuals (Maluccio, 2000; Thomas *et al.*, 2001).

Lastly, women are less likely to be found in the group of movers outside the tracking zone, but education is no longer significant, indicating that, in the descriptive analysis, it was mainly an effect of gender. Age is not a significant determinant of being a long-distance mover and neither is per capita income. The assumption on which was based the choice not to follow long-distance movers was thus probably incorrect, but in terms of economic conditions, this group is not significantly different from movers who were found.

5.3. *Economic Mobility with Alternative Follow-Up Rules*

So far, we have tried to assess the existence of an attrition bias based on observable characteristics of individuals at baseline. Evidence of the reduction of the attrition bias thanks to the tracking is mixed, based on the above tests. However, as noted by Dercon and Shapiro (2007), evidence on attrition that relies on baseline characteristics only is potentially flawed, if negative (or even positive) shocks force people to leave their village. These events would be unobserved for attritors and could generate a sample that is non-representative with respect to income mobility and poverty dynamics. This is a motivation for running tracking surveys, to include in the sample economic trajectories that are excluded from traditional, residence-based panels.

Table 8 summarizes growth of per capita income, as well as baseline and final levels, of individuals for nested samples, constructed using alternative follow-up

TABLE 8
INCOME GROWTH USING ALTERNATIVE FOLLOW-UP RULES AT THE INDIVIDUAL LEVEL

Follow-Up Rule	Dwelling-Based	Village-Based	Tracking	
	Bepako Census	Same Household	Stayed in Bepako	All Recontacted
Income 1995 (per cap.)	568.3 (15.7) [537.6; 599.0]	603.2 (24.9) [554.5; 652.0]	607.7 (22.9) [562.8; 652.7]	584.1 (19.5) [545.9; 622.4]
Income 2005 (per cap.)	604.9 (17.8) [569.9; 639.9]	655.9 (37.5) [582.5; 729.4]	659.6 (31.7) [597.4; 721.7]	679.4 (25.9) [628.7; 730.2]
Change 1995–2005 (absolute)	36.6	52.7 (28.0) [–2.20; 107.6]	51.9 (26.1) [0.63; 103.1]	95.3 (24.0) [48.3; 142.3]
Change 1995–2005 (%)	6.44	37.0 (4.54) [28.1; 45.9]	42.0 (4.63) [33.0; 51.1]	65.7 (5.67) [54.6; 76.8]
Observations	1490–1783	662	805	1068

Notes: Income is per capita income of household to which the individual belongs in the indicated year. All figures are in 1000 FMg, except relative change in %. Standard errors in parentheses, 95% confidence interval in brackets. In the first column, the income change (both absolute and in %) is the overall difference between the 2005 and the 1995 income. In the last 3 columns, the change in income is the individual change averaged over all households. The number of observations in the first column are the number of inhabitants of Bepako in 1995 and 2005 respectively. “Same household”: individuals who stayed in the same household (stayers). “Stayed in Bepako”: “Stayers” + re-interviewed individuals who changed households and stayed in Bepako (splits). “All recontacted”: “Stayed in Bepako” + recontacted individuals who moved out of Bepako (migrants).

Source: ROR data from 1995 and 2005, tracking survey data from 2005; author’s calculation.

rules. The first column shows those statistics for the two cross-sections, ignoring the panel dimension of the data. They summarize per capita income in 1995 and 2005 in the entire village (as both years were censuses of the population of Bepako). Income growth in this column is thus the difference in the means of the two years. These figures give a more “macroeconomic” view of the economic conditions in the village. They can be considered a benchmark against which to compare panel data. The second column shows summary statistics for the group of individuals obtained using a residence-based follow-up rule (stayers). It is included in the sample of the next column, which is built according to an individual follow-up inside the village. This group comprises all individuals who stayed in Bepako and were recontacted (stayers and splits, see Figure 2). The last column adds recontacted movers, yielding the entire sample obtained through the tracking process. In these last three columns, the change (absolute and in percentage) is the mean of the change experienced by each household.

The sample obtained using the dwelling-based follow-up rule shows an average improvement of the income per capita of 37 percent, while including those who changed households but stayed in Bepako yields a mean income growth of 42 percent. The overlapping 95% confidence intervals suggest that there is only a slight difference between the two groups. When income growth is averaged over the entire sample of recontacted individuals, it is equal to 65.7 percent. In this case, the confidence intervals obtained using the dwelling- or village-based follow-up rule do not overlap with the intervals obtained for the entire sample. It is clear from Table 8 that adding 406 observations (those who moved out of Bepako and were recontacted) changes the global picture of the evolution of income between 1995 and 2005. This simple result is an indication of the usefulness of such tracking devices, as also shown by Beegle *et al.* (2011) in Tanzania, using the KHDS data.

It is clear that longitudinal data give a very different picture of the evolution of economic conditions in Bepako than two cross-sections. The average per capita income in Bepako has grown by 6.4 percent only using the data from the two censuses. Assuming that they were appropriately deflated, this would mean that living conditions were stable or had slightly increased between those two years in Bepako. However, the panel shows that for the specific individuals followed, income grew much faster, all the more so if they moved away.¹⁴

6. CONCLUSION

The increasing use and availability of panel data in developing countries have been essential in fostering studies on economic mobility that are free of certain types of biases present in cross-sectional data. However, panel datasets also have their own drawbacks and limitations, not least of which is attrition. Traditional living standard surveys have often generated attrition by not attempting to re-interview migrants, who may well have different observable and unobservable characteristics, thus creating an attrition bias in the analysis. More recently, track-

¹⁴This also suggests that newcomers to Bepako in that period (who are, by construction, not included in the panel) have a lower income than those established for a longer period. This is related to the issue of panel aging, briefly discussed earlier. As the population of Bepako changes and welcomes new migrants, the panel becomes increasingly less representative of the village population.

ing migrants has been attempted in several surveys and recommended in the literature. The extent to which these devices are necessary to reduce a potential attrition bias is an empirical matter, depending on the setting and the outcome variable studied. This paper is an attempt to bring more evidence on the usefulness of tracking movers using data from a tracking survey carried out in Madagascar in 2005.

Despite the complexity of locating individuals a decade after the baseline interview, the tracking experience carried out in Madagascar resulted in a high recontact rate. It highlighted the limitations of dwelling- and household-based follow-up rules in developing countries. However, while tracking movers largely reduced *attrition*, tests of baseline characteristics show only a slight mitigation of the *attrition bias*. It appears that the hardest people to find are the poorest and most isolated. In countries with such underdeveloped communication infrastructures as Madagascar, it is quite possible that those bearing the worst hardships, in economic and human terms, are never included in any panel dataset. The implications obviously go beyond the pure statistician's interest for unbiased estimations. Ultimately, results obtained using longitudinal data enable one to design targeted poverty reduction policies, but one cannot design programs for the worse-off if they do not emerge from the data. Those being more mobile and less easily followed would also hinder the implementation of continuing poverty reduction programs.

However, the results on income growth of tracked individuals show that tracking movers does change the global picture of income mobility of those living in Bepako in 1995 over the following decade. This is consistent with other results in the literature. The longitudinal aspect of the data complemented by a tracking of migrants shows the extent of migratory movements and households split-ups. Thanks to the tracking, it is possible to use data on both income at destination and at origin, and to compare its growth with those who did not migrate. This pleads in favor of following individuals rather than households, and this follow-up must be extended outside the initial village. As the sample used is very small, and comes from one village only, with its idiosyncrasies, generalizing the results to all settings and panel surveys is beyond our scope. However, we contend that it illustrates well issues pertaining to regions with high geographical mobility, such as the one studied. The region has indeed been a traditional migratory destination in Madagascar, but it is also common for these migrants to go back home at the end of their lives.

Figures of income growth obtained using the Bepako 1995 and 2005 censuses as two cross-sections also plead for repeated annual surveys which will take into account newcomers to a village and the evolution of their living conditions. The censuses bring a wealth of information on the evolution of a village over a decade, which enable the analysis of the evolution of poverty and inequality at a more global level. It should be noted that following a specific group of people, in our case, the population of Bepako in 1995, and monitoring a specific geographical area, namely, the village of Bepako, over time, are two different approaches that complement rather than oppose each other.

Even though the ROR and most other surveys in developing countries are household surveys, special attention should be paid to collecting data that identify individuals. We believe that with little cost and effort, individual panels can be

easily constructed from household panels, provided the data allow consistent identification of individuals across survey rounds. A permanent identifying code should be attributed to each individual in the household and each member should be matched from the previous round roster, while new members should be attributed a new code. In the ROR case, this would entail making sure that the interviewer records full names as well as nicknames, and that the link to the household head is collected precisely, to avoid confusion between biological children, foster children, and grandchildren for example. Identifying who remained and who left between rounds would be strongly helped by accuracy in the collection of these individual variables. Such a procedure would make the use of household panels for individual studies possible and more reliable.

Questions on the measurement of welfare at the individual level are the natural corollary to the construction of individual panels. The ROR, like the great majority of panel surveys in developing countries, collects income and consumption data at the household level. These aggregates can then be adjusted for household size and composition, using per capita measures, or adult equivalence scales, to infer individual welfare from household measures. This however does not take individual bargaining power inside the household into account. Measuring individual consumption is technically possible, but, assuming the results are reliable, the main drawbacks of such survey devices are their cost and the length of the interview, which do not seem applicable to an annual, agricultural, panel survey such as the ROR. Further exploration of this issue is necessary to design cost- and time-effective questionnaires capable of apprehending resource distribution within households.

Finally, this discussion should recognize the progress that has been made in the collection of panel data in developing countries in recent years. The use of electronic devices such as cell phones or GPS to localize respondents has greatly improved the success of migrant tracking. The ROR was started in 1995, when issues of attrition were less well acknowledged. The duration of this survey was unknown at the start, nor was the intensity of the household formation and dissolution phenomena. It is fair to say that studies of income dynamics undertaken nowadays would design the data collection process differently, to avoid some of the limitations of the ROR, such as matching on names, not allowing re-entries, or having randomly reduced the sample size at some point. However, in the context of collecting data to monitor living conditions in rural areas, and particularly in Africa, the issues raised in this paper remain relevant. As explained above, the appropriate unit of study of income dynamics is still the household, yet to avoid the biases discussed in this paper, surveys also need to follow individuals. In this background, this study aimed to illustrate that monitoring living conditions in rural areas still exhibits specific problems due to the complexity of dealing with these two different sampling units simultaneously.

REFERENCES

- Alderman, H., J. R. Behrman, H. P. Kohler, J. A. Maluccio, and S. C. Watkins, "Attrition in Longitudinal Household Survey Data," *Demographic Research*, 5, 79–124, 2001.
- Ashenfelter, O., A. Deaton, and G. Solon, "Collecting Panel Data in Developing Countries: Does it Make Sense?" LSMS Working Paper No. 23, World Bank, 1986.

- Badiani, R., S. Dercon, P. Krishnan, and K. P. C. Rao, "Changes in Living Standards in Villages in India 1975–2004: Revisiting the ICRISAT Village Level Studies," CPRC Working Paper No. 85, 2007.
- Beckett, S., W. Gould, L. Lillard, and F. Welsch, "The Panel Study of Income Dynamics After Fourteen Years: An Evaluation," *Journal of Labor Economics*, 6, 472–92, 1988.
- Beegle, K., J. De Weerd, and S. Dercon, "Migration and Economic Mobility in Tanzania: Evidence from a Tracking Survey," *Review of Economics and Statistics*, 93, 1010–33, 2011.
- Deaton, A., *The Analysis of Household Surveys: A Microeconomic Approach to Development Policy*, World Bank Publications, 1997.
- Dercon, S. and P. Krishnan, "Changes in Poverty in Rural Ethiopia 1989–1995: Measurement, Robustness Tests and Decomposition," Centre for the Study of African Economies, Working Paper Series No. 98-7, 1998.
- Dercon, S. and J. S. Shapiro, "Moving On, Staying Behind, Getting Lost: Lessons on Poverty Mobility from Longitudinal Data," in Deepa Narayan and Patti Petesch (eds), *Moving Out of Poverty. Cross-Disciplinary Perspectives on Mobility*, World Bank, Washington DC, 77–126, 2007.
- DeVreyer, P., S. Lambert, A. Safir, and M. Sylla, "Pauvreté et structure familiale: pourquoi une nouvelle enquête?" *Stateco*, 108, 55–70, 2008.
- Droy, I., R. Ratovoariny, and F. Roubaud, "Les observatoires ruraux à Madagascar 1995–1998: une méthodologie originale pour le suivi des campagnes," *Stateco*, 95, 123–40, 2000.
- Duncan, G. J. and M. S. Hill, "Conceptions of Longitudinal Households: Fertile or Futile," *Journal of Economic and Social Measurement*, 13, 361–75, 1985.
- Fafchamps, M. and A. R. Quisumbing, "Household Formation and Marriage Markets in Rural Areas," in T. Paul Schultz and John A. Strauss (eds), *Handbook of Development Economics*, Elsevier, 3187–247, 2007.
- Falaris, E. M., "The Effect of Survey Attrition in Longitudinal Surveys: Evidence from Peru, Côte d'Ivoire and Vietnam," *Journal of Development Economics*, 70(1), 133–57, 2003.
- Fitzgerald, J., P. Gottschalk, and R. Moffitt, "An Analysis of Sample Attrition in Panel Data: The Michigan Panel Study of Income Dynamics," *Journal of Human Resources*, 33, 251–99, 1998.
- Foster, A. D. and M. R. Rosenzweig, "Household Division and Rural Economic Growth," *Review of Economic Studies*, 69, 839–69, 2002.
- Fuwa, N., "Should We Track Migrant Households When Collecting Household Panel Data? Household Relocation, Economic Mobility and Attrition Biases in the Rural Philippines," *American Journal of Agricultural Economics*, 93(1), 56–82, 2011.
- Glewwe, P. and H. Jacoby, "Recommendations for Collecting Panel Data," in Margaret Grosh and Paul Glewwe (eds), *Designing Household Survey Questionnaires for Developing Countries: Lessons from Fifteen Years of LSMS Experience. Volume 2*, World Bank, Washington DC, 275–314, 2000.
- Haaga, J., J. DaVanzo, C. Peterson, and N. Tey, "Twelve-Year Follow-Up of Respondents in a Sample Survey in Peninsular Malaysia," *Asia-Pacific Population Journal*, 9(2), 61–72, 1994.
- Lanjouw, P., "Economic Development in Palanpur 1957–1993: A Sort of Growth," William Davidson Institute Working Papers Series No. 330, 2000.
- Maluccio, J. A., "Attrition in the KwaZulu Natal Income Dynamics Study, 1993–1998," FCND Discussion Papers, IFPRI, 2000.
- McMillen, D. B. and R. Herriot, "Toward a Longitudinal Definition of Households," *Journal of Economic and Social Measurement*, 13, 349–60, 1985.
- Norris, S. A., L. M. Richter, and S. A. Fleetwood, "Panel Studies in Developing Countries: Case Analysis of Sample Attrition Over the Past 16 Years Within the Birth to Twenty Cohort in Johannesburg, South Africa," *Journal of International Development*, 19, 1143–50, 2007.
- Outes-Leon, I. and S. Dercon, "Survey Attrition and Attrition Bias in Young Lives," Young Lives Technical Note No. 5, 2008.
- Ravallion, M., "Poverty Comparisons: A Guide to Concepts and Methods," LSMS Working Paper No. 88, World Bank, 1992.
- ROR, *Manuel de l'enquêteur*, Projet MADIO, 1995.
- Rosenzweig, M. R., "Payoffs from Panels in Low-Income Countries: Economic Development and Economic Mobility," *American Economic Review*, 93, 112–7, 2003.
- Thomas, D., E. Frankenberger, and J. P. Smith, "Lost But Not Forgotten: Attrition and Follow-Up in the Indonesia Family Life Survey," *Journal of Human Resources*, 36, 556–92, 2001.
- Vaillant, J., "The Methodological Challenge of Monitoring Living Conditions: Insights from a Tracking Experience in Madagascar," DIAL Working Paper No. 2010-13, 2010.