

The Measurement of Income Distribution Dynamics when Demographics are correlated with Income

Denis COGNEAU
Michael GRIMM

THE MEASUREMENT OF INCOME DISTRIBUTION DYNAMICS WHEN DEMOGRAPHICS ARE CORRELATED WITH INCOME

Denis Cogneau
DIAL - UR CIPRÉ de l'IRD
cogneau@dial.prd.fr

Michael Grimm
University of Göttingen, Department of Economics, DIW (Berlin)
and DIAL
mgrimm@uni-goettingen.de

Document de travail DIAL / Unité de Recherche CIPRÉ

Novembre 2004

ABSTRACT

The purpose of our paper is to derive instructive analytics on how to account for differentials in demographic variables, and in particular mortality, when performing welfare comparisons over time. The idea is to “correct” in various ways estimated income distribution measures for “sample selection” due to differential mortality. We implement our approach empirically using three waves (1993, 1997 and 2000) of the Indonesian Family Life Surveys (IFLS). We distinguish the direct effect of mortality, i.e. individuals who die are withdrawn from the population and no longer contribute to monetary welfare, from the indirect effect, i.e. the impact on survivors pertaining to the same household of dead individuals, who may experience a decrease or an increase in monetary welfare. For the case of Indonesia, we show that the direct and indirect effects of mortality on the income distribution have opposite signs, but have the same order of magnitude. We also show that the effect of other demographic changes, like changes in the structure of fertility, migration, and educational attainment, dominate the effects of mortality, whether direct or indirect. However, we find that none of these demographic developments are large enough to explain a significant part of the change in income distribution, whether the pre-crisis period (1993-1997) or the post-crisis period (1997-2000) are considered.

Keywords: Differential Mortality, Income Distribution Dynamics, Welfare Comparisons, Decomposition.

RÉSUMÉ

L'objectif de ce papier est de proposer des instruments analytiques permettant de prendre en compte les différentiels relatifs aux variables démographiques, en particulier la mortalité, lorsqu'on effectue des comparaisons de pauvreté au cours du temps. L'idée de base consiste à « corriger » les estimations de la distribution du revenu de la sélection liée à la mortalité différentielle. Nous mettons en œuvre notre approche sur les trois vagues (1993, 1997 et 2000) de l'Indonesian Family Life Survey (IFLS). Nous distinguons l'effet direct de la mortalité, à savoir la disparition des individus décédés de la population de calcul du bien-être monétaire, de l'effet indirect, à savoir l'impact sur les survivants appartenant au même ménage qu'un individu décédé. Dans le cas de l'Indonésie, nous montrons que les effets directs et indirects de la mortalité sur la distribution du revenu ont des signes opposés mais environ le même ordre de grandeur. Nous montrons aussi que l'effet des autres changements démographiques (comme ceux de la structure de la fécondité, de la migration ou de l'éducation), dominant les effets de la mortalité qu'ils soient directs ou indirects. Cependant, nous trouvons enfin qu'aucun de ces changements démographiques n'est assez important pour expliquer une part significative du mouvement de la distribution du revenu, que l'on regarde la période précédant la crise économique (1993-97) ou la période suivante (1997-2000).

Mots-clefs : Mortalité différentielle, Dynamique de la distribution du revenu, comparaisons de bien-être, décomposition.

JEL codes : D10, D63, J17

Contents

1. INTRODUCTION	4
2. WELFARE IMPLICATIONS OF DIFFERENTIAL DEMOGRAPHICS	5
3. SOME GENERAL METHODS TO ACCOUNT FOR DIFFERENTIAL MORTALITY IN POVERTY COMPARISONS OVER TIME	8
3.1 The direct arithmetical impact of individual mortality	10
3.2 The indirect micro-impact of mortality on survivors income distribution	12
4. AN EMPIRICAL IMPLEMENTATION FOR THE CASE OF INDONESIA.....	14
4.1 Data and economic context	14
4.2. Some illustrative simulations.....	16
4.3. Results for Indonesia for the period 1993 to 2000	19
4.3.1. Estimates of the direct arithmetical impact of mortality	19
4.3.2. Estimates of the direct and indirect impact of changes in mortality patterns.....	22
5. CONCLUSION	24
APPENDIX	25
REFERENCES	25

List of tables

<i>Table 1 : The effects of differential mortality on standard income distribution indicators</i>	<i>28</i>
<i>Table A1 : Descriptive statistics of the variables used</i>	<i>29</i>
<i>Table A2 : Estimated probit model for survival probabilities.....</i>	<i>30</i>
<i>Table A3 : Estimated probit model for “being present” probabilities</i>	<i>31</i>
<i>Table A4 : Estimated probit model for being a survivor in a household in which no death occurred during past period.....</i>	<i>32</i>
<i>Table A5 : Estimated household income per capita regressions</i>	<i>33</i>

List of figures

Figure 1 : The effects of differential mortality on standard income distribution indicators	34
Figure 2 : Income per capita (ln) Kernel densities in 1993, 1997 and 2000.....	34
Figure 3 : Smoothed impact of individual deaths	35
Figure 4 : Smoothed survival probabilities by income per capita percentile for men and women and selected age groups.....	35
Figure 5 : Smoothed impact of individual mortality compared to impact of changes	36
Figure 6 : Smoothed impact of individual mortality and of changes in all observable attributes on the total change of the income per capita distributions	36
Figure 7 : Smoothed impact of changes in mortality patterns between 1993 / 1997 and 1997 / 2000 on the income distribution in 2000	37
Figure 8 : Smoothed impact of changes in mortality patterns on the income distribution in 2000: Changes in mortality patterns between 1993 / 1997 and 1997 / 2000 compared to other changes in the development of the population structure.....	37
Figure 9 : Smoothed impact of changes in mortality patterns and other observables on the total change of the income per capita distribution between 1997 and 2000.....	38

1 Introduction¹

Demographic behavior may significantly affect the income distribution, whenever it is correlated with the used income measure. Poor people who are more likely to die than rich people, poor people who have more children than rich people or poor people who are more likely to migrate than rich people, are all channels which can have significant and even substantial effects on income distribution dynamics. When analyzing the causes of distributional change, it seems worthwhile to isolate these effects from changes in labor supply behavior or changes in the returns on the labor market which in turn can also have a strong impact on the income distribution, but which are rather driven by structural and institutional change. Of course, the cited transmission channels can be interdependent and therefore hard to disentangle. For instance, the death of one household member can alter labor supply, educational investment, and consumption behavior of other household members. Given the lack of appropriate methods to explore the importance of the demographic channels, not much is known concerning their empirical importance.

The purpose of our paper is twofold: first to derive instructive analytics on how to account for differentials in demographic variables, and in particular mortality, when performing welfare comparisons over time; second, to explore the potential impact of demographic change on the distribution of welfare. The idea of the methodology we suggest, is to ‘correct’ in various ways estimated welfare distributions for ‘sample selection’ due to differential mortality. A central issue is then to derive reliable estimates for mortality rates as a function of income or its correlates and age. Once the conditional density of mortality is known, one can compute a reweighted welfare distribution giving the welfare variation attributable to individual deaths. Further complications arise when the household and not the individual is the unit of analysis. The key estimation problem becomes then to construct a counterfactual distribution that would have prevailed if the survivors would still live with their former

¹Many thanks are due to participants at the ESPE (Bergen) and IARIW (Cork) conferences 2004. We are particularly grateful for the comments and suggestions received by Gordon Anderson. Finally, we thank BPS Statistics Indonesia for having given us access to price data.

household members and would decide jointly on labor supply and consumption expenditure. The semiparametric procedure we suggest to address these issues is very much in the spirit of the decompositions performed by DiNardo, Fortin and Lemieux (1996).

We proceed as follows. In the next section, we discuss the welfare implications of differentials in demographic variables and in particular differential mortality and provide a quick review of the related literature. In Section 3, we present our methodology. In Section 4 we implement our approach empirically using three waves of the Indonesian Family Life Surveys (IFLS). In Section 5, we summarize our main results and conclude.

2 Welfare implications of differential demographics

A well known problem of welfare comparisons over space and time are variations in population size. This problem was raised, for instance, by Dasgupta, Sen and Starrett (1973) in their note on Atkinson's seminal paper on the measurement of inequality (Atkinson, 1970). It appears also in the literature on the general form of social welfare functions (see e.g. Blackorby and Donaldson (1984) and Blackorby, Bossert and Donaldson (1995)). Two aspects are of importance here. First, which dimensions of personal well-being we allow to enter the individual welfare function, i.e. should the length of life matter. Second, should a social welfare function take into account the number of members in the society at a given point in time.

The standard welfaristic approach usually neglects non-materialistic sources of personal well-being and has—at least in the empirical literature—a strong focus on annual income flows. Under these assumptions interpersonal utility comparisons are not affected by the fact that two individuals have to expect a different length of life. In other words, two persons receiving the same income over a given period of time and having the same individual utility function are regarded as enjoying the same utility, irrespective of differences in their expected length of life. Recognizing these drawbacks of the standard welfaristic approach, Anderson (2003) suggested recently a framework which includes life expectancy into the calculus of economic welfare comparisons.

He implements his approach at the level of countries and compares how GDP performs over time with and without accounting for changes in life expectancy. In the case of Africa, when life expectancy is included he finds a substantial decline in welfare over time, which is not the case if usual GDP comparisons are undertaken.

In contrast to the welfaristic approach, Sen's capability approach has in this sense a much wider focus and is much more flexible (Sen 1985). This approach can easily be defined in a way, that factors or *functionings* which allow a less or more long length of life are explicitly taken into account, by assuming that health or a certain length of life can be produced in a complementary way through commodities q , personal characteristics and societal and environmental circumstances z . Therefore if q and z are favorable for health they will map into longer life and by this channel enlarge the so-called 'capability' set.

Turning now to the second point; the classical utilitarian (or *Benthamite*) social welfare function is given by the sum of individual utilities $W = \sum_{i=1}^N u_i(x_i)$, where N is the total number of individuals, x_i are commodities and u_i is the utility drawn by individual i from x_i . So, clearly, here the number of individuals in the society N can be seen as a source of social welfare. But, in most cases we think of a constant population when invoking such an utility function or simply use it in per capita terms (W/N) and sidestep this issue. The implicit ethical judgement, then, is that we are 'neutral' toward population. At the same time, the focus on per capita welfare means that we are indifferent to the unborn and are even biased toward keeping population growth down if it affects per capita welfare adversely.

Empirical studies on the dynamics of inequality and poverty generally do not really address this issue by supposing implicitly a constant population. They usually provide a kind of 'snapshot-measure' of economic well-being. In other words, we consider indicators such as GDP per capita, the Human Development Index, the poverty headcount or the Gini coefficient at two different points in time without asking if the population size has changed during the relevant time period.

When considering a single country, variations in population size over time are driven by three demographic forces: fertility, mortality and migrations. If these forces are correlated with the used welfare measure, welfare comparisons may become complex and sometimes ambiguous. For instance, if mortality is negatively correlated with income which seems indeed to be the case in developing as well as in developed countries,² standard poverty measures as the headcount-index of the FGT-family (Foster, Greer, and Thorbecke, 1984), for instance, may show an improvement over time if individuals under the poverty line die. Or, put differently, higher mortality among the poor is ‘good’ for poverty reduction. The current AIDS epidemic in developing countries, the 1918 influenza epidemic or the black plague centuries ago might thus have reduced poverty, not only by increasing the capital-labor ratio, but also simply by killing the poor if they are more than others affected by these diseases.³ Most people will agree that this kind of ‘repugnant conclusion’ is not compatible with the axiomatic on which poverty concepts are normally based. This point was recently raised by Kanbur and Mukherjee (2003).

The problem is alike if we consider fertility. Higher fertility among the poor may increase poverty simply due to differential growth rates over the income distribution. One might conclude that minimizing fertility among the poor is a mean to reach poverty reduction.⁴ Again, this seems neither economically nor ethically reasonable or acceptable. Finally, migration from rural to urban migration might reduce rural and increase urban poverty, without having changed anything in the situation of those who stayed at their initial place.

Kanbur and Mukherjee (2003) proposed to compute FGT-poverty measures based on the lifetime income profile of an individual. They define a

²For empirical evidence see e.g., Kitagawa and Hauser (1973), Deaton and Paxson (1999) or Lantz *et al.* (1998). Valkonen (2002) provides a survey of the empirical evidence concerning social inequalities in mortality. He finds that social inequality was observed in almost all studies using different populations and using different indicators of socio-economic position, such as social or occupational class, socio-economic status, educational attainment, income and housing characteristics.

³For instance Brainerd and Siegler (2003) found empirical evidence that the 1918 influenza epidemic had a robust positive effect on per capita income growth across US states during the 1920s.

⁴See on this issue the analytics and discussions in Lam (1986), Chu and Koo (1990) and Lam (1997).

normative measure for the length of life to account for premature mortality among the poor which affect the poverty measure positively. There are however two crucial issues in their procedure. First, the choice of the normative length of life, which can influence on the poverty ranking of different populations. Second, the hypothetical income which has to be imputed for the years between the actual age of death and the normative age of death. This issue is handled by supposing constant income levels over time, no mobility across income levels, and that each individual at income level Y_i lives for l_i periods, after which time her or she is replaced by exactly one individual. The critical assumption concerning the hypothetical income that should be imputed, raises the general question of which ‘value’ we might want to attribute to a foregone life. Even if we exclude here issues of personal pain and loss, the pure materialistic loss can only arbitrarily be computed.

In what follows we suggest some general methods to account for differentials in demographic ‘behavior’, and especially mortality, when performing poverty comparisons over time and circumvent the problem of giving a value to lost life. We first consider only what we call the ‘direct effect’ or ‘pure demographic effect’ and then develop successively measures which take into account the effect that a deaths might have changed household income (not only household income per capita), first because the person who died does not anymore contribute to the household income and, second, because the death might have changed labor supply behavior of the other household members.

3 Some general methods to account for differential mortality in poverty comparisons over time

For each period t , the welfare indicator y is defined over a population of individuals. It is assumed that y is a continuous variable which may vary between 0 and $\max(y)$, with a c.d.f. $F_t(y)$ and a d.f. $f_t(y) = dF_t(y)$. In the utilitarian tradition, a monetary welfare index is then defined as

$$W(F_t) = \int_0^{y^{\max}} w(y) dF_t(y) dy \quad (1)$$

w being a non-decreasing function of income. In the same sense, a large class of monetary poverty indexes corresponds to

$$P(F_t) = \int_0^z p(y) dF_t(y) dy \quad (2)$$

where z is the poverty line and p a non-decreasing non-negative function of income defined over $[0; z]$.

Expressed in its most general form, our problem is to design counterfactual distributions of y , $F_{t+1}^*(y)$ under alternative mortality processes taking place between t and $t + 1$, and then to compute

$$W(F_{t+1}^*) = \int_0^{y^{\max}} w(y) dF_{t+1}^*(y) dy, \quad (3)$$

or

$$P(F_{t+1}^*) = \int_0^z p(y) dF_{t+1}^*(y) dy \quad (4)$$

More precisely, let us assume that we have some knowledge about the mortality process taking place between t and $t + 1$. The occurrence of individual deaths should theoretically have at least three kinds of effects on the distribution of income:

1. a direct ‘arithmetical’ individual impact: people who die are withdrawn from the population and do not longer contribute to monetary welfare or to poverty;
2. an indirect micro-economic impact on household income: survivors pertaining to the same household than those who died may experience a decrease or an increase in y , with the previous income contribution of the dead being withdrawn from household income, with the number of equivalent consumption units being modified, and with various labor supply or household composition adjustments occurring;
3. a ‘general equilibrium’ or ‘external’ macroeconomic impact on the overall income distribution.

In the following, we shall not consider the third, general equilibrium, effect. Hence, the construction of a counterfactual distribution of income requires to

deal with, first, the direct impact, and, second, the indirect impact. However, what is meant by ‘counterfactual’ should first be clarified for both cases. Intuitively speaking, we seek to reconstruct what would be the distribution of income in $t + 1$ if the observed deaths between t and $t + 1$ had not occurred. This definition of a counterfactual raises no particular difficulty when the mortality process can be assumed exogenous from the distribution of income itself. Think at a sudden epidemics coming from outside the country or at a natural catastrophe like an earthquake or a flood. Of course the exogeneity of mortality does not preclude that it can be correlated with income. Things are more intricate when the probability of dying is causally determined by the contemporary individual income, by the distribution of income within some reference group, or by the overall distribution of income (see Deaton and Paxson, 1999). For instance, people whose income has fallen under a subsistence level (extreme poverty line) may have been exposed to death with a probability close to one. Giving these people a ‘counterfactual income’ under the subsistence level would have no meaning at all if nobody can survive in this situation. It seems to us that a meaningful counterfactual distribution of income should always include those deaths that are income determined, or, put in another way, should only try to discount the distribution of deaths that is exogenous to the final income distribution. In the remainder, we shall always make the assumption that mortality is exogenous to transient components of contemporary income, but may vary with permanent income determinants.

Finally another important aspect regarding the construction of counterfactuals has to be emphasized. Assessing the impact of mortality between two dates does not mean the same than assessing the impact of changes in mortality. In the first case we need to subtract the impact of all deaths during the period, while in the second case we need to subtract the impact of the difference between the ex-post and ex-ante pattern of occurred deaths. We focus on the first case in the following section and examine the second case afterwards.

3.1 The direct arithmetical impact of individual mortality

Let us first assume that individual deaths have no external effects, neither on other individuals like household survivors or neighbors nor on the whole population. We therefore seek to design a counterfactual for a pure arithmetical individual effect. Assume second that mortality patterns between t and $t + 1$ are totally described by observable individual attributes x which are either constant over time like sex and education of adults, or varying with time like age, health status and household composition. This makes the survival rate $s_{x,t}(x)$ independent from the distribution of attributes, i.e. the survival rate is independent from the population structure. Assume third that the income pattern belonging to each attribute, i.e. the conditional density of income with respect to attributes, does not depend on the distribution of attributes but instead only on some ‘income schedule’ that changes over time through redistribution policies and other changes in the returns to attributes (in the spirit of the Oaxaca (1973) or DiNardo *et al.* (1996) decompositions). This means that we assume again that mortality has neither external effects nor ‘general equilibrium’ effects. This also means that we exclude the possibility of non-random selection of deaths by contemporary unobservable determinants of income (y_{t+1}), i.e. that mortality is caused by the transient components of income.

The econometrician observes $f(y|t_y = t)$ that is the actual density of income for each t . He may write the following conditionnal density:

$$f(y, x|t_{y,x} = t + 1) = f(y|x, t_y = t + 1)dF(x|t_x = t + 1) = f(y|x, t_y = t + 1) \frac{s_{x,t}(x)}{\Psi_{x,t}(x)} dF(x|t_x = t), \quad (5)$$

where $\Psi_{x,t}(x)$ characterizes the changes in the population structure which are not due to mortality, but instead due to births, migrations, household composition etc. having occurred over $[t ; t + 1]$. We may then compute the counterfactual distribution of income due to deaths related to initial attributes, by simply reweighing observations with $s_{x,t}(x)$.

$$f_t^*(y) = \int_{x \in \Omega_x} s_{x,t}(x) f(y|x, t_y = t) dF(x|t_x = t) \quad (6)$$

Semiparametric decompositions as proposed by DiNardo *et al.* (1996) allow to go a bit further by isolating the impact of changes in the distribution of all attributes. Hence, we can compute the following counterfactual which gives the overall impact of all changes in attributes (including changes linked to mortality) on income distribution:

$$g_t^*(y) = \int_{x \in \Omega_x} f(y|x, t_y = t) dF(x|t_x = t + 1) \quad (7)$$

using DiNardo *et al.* (1996) and using the reweighing technique based on Bayes' rule:

$$\frac{s_{x,t}(x)}{\Psi_{x,t}(x)} = \frac{dF(x|t_x = t + 1)}{dF(x|t_x = t)} = \frac{\Pr(t_x = t + 1|x)}{\Pr(t_x = t|x)} \cdot \frac{\Pr(t_x = t)}{\Pr(t_x = t + 1)},$$

where $\Pr(t_x = t|x)$ can be estimated with a probit model. We then obtain, for instance:⁵

$$g_t^*(y) = \int_{x \in \Omega_x} \frac{s_{x,t}(x)}{\Psi_{x,t}(x)} f(y|x, t_y = t) dF(x|t_x = t). \quad (8)$$

So far, we have considered the counterfactual impact of the level of individual mortality. Computing the impact of *changes* in mortality patterns (based on individual observables) just calls for an additional preliminary reweighing of the t income distribution by past survival rates:

$$f_t^{**}(y) = \int_{x \in \Omega_x} \frac{s_{x,t}(x)}{s_{x,t-1}(x)} f(y|x, t_y = t) dF(x|t_x = t) \quad (9)$$

We shall come back to this latter decomposition when taking into consideration the indirect impact of the changes in the distribution of household survivors.

3.2 The indirect micro-impact of mortality on survivors income distribution

When the income concept that is used is household income per capita or per adult equivalent unit, mortality has, as mentioned above, obviously an indirect

⁵We may also compute: $g_{t+1}^*(y) = \int_{x \in \Omega_x} f(y|x, t_y = t + 1) dF(x|t_x = t)$ and: $g_{t+1}^{**}(y) = \int_{x \in \Omega_x} f(y|x, t_y = t + 1) s_{x,t}(x) dF(x|t_x = t)$.

The difference between $g_{t+1}^*(y)$ and $g_{t+1}^{**}(y)$ should indicate the impact of mortality on a distribution of income characterized by the final income schedules $f(y|x, t_y = t + 1)$ and the initial distributions of attributes $dF(x|t_x = t)$. Then the double difference between $[g_{t+1}^{**}(y) - g_{t+1}^*(y)]$ and $[f_t^*(y) - f_t(y)]$ gives the mortality impact linked to the change in income schedule from $f(y|x, t_y = t)$ to $f(y|x, t_y = t + 1)$.

impact on the distribution of income over household survivors. In order to analyze this indirect impact we have to construct a counterfactual income distribution which includes a counterfactual income pattern for survivors. Let $z \in \{0, 1\}$ be a variable indicating whether somebody has experienced a death in his/her household between t and $t + 1$. The observed density of income in $t + 1$ is a weighted sum of conditional densities on z

$$f_{t+1}(y) = \Pr(z = 0|t_z = t + 1)f(y|z = 0, t_y = t + 1) + \Pr(z = 1|t_z = t + 1)f(y|z = 1, t_y = t + 1). \quad (10)$$

We would like to design a counterfactual which can be written as

$$f_{t+1}^\#(y) = \Pr(z = 0)f(y|z = 0, t_y = t + 1) + \Pr(z = 1)f_{z=0}(y|z = 1, t_y = t + 1). \quad (11)$$

It requires the estimation of the counterfactual density for survivors $f_{z=0}(y|z = 1, t_y = t + 1)$. Computing such a counterfactual is very difficult. Quantile treatment IV estimators could be used (Abadie, Angrist, Imbens, 1998) if some instrument was available for the occurrence of a death within the household (some knowledge about death causes could prove useful in this respect). Under the assumption of conditional independence on a set of attributes x ,⁶ quantile treatment effects may also be computed using matching estimators (Firpo, 2004).

However, given that we have information on survivors at period t , that is people having experienced a death within the household between $t - 1$ and t , a counterfactual for the impact of *changes* in mortality patterns can prove easier to construct. Indeed, when survivor status z is known for both periods, we may apply the DiNardo *et al.* (1996) reweighing technique to isolate the effects of changes in the ‘survival rate’. Hence, we write

$$f_{t+1}^{indir}(y) = \int \int f(y|x, z, t_y = t + 1)dF(z|x, t_z|x = t)dF(x|t_x = t + 1) = \frac{\int \int \Psi_{z|x}(z, x)f(y|x, z, t_y = t + 1)dF(z|x, t_x = t + 1)dF(x|t_x = t + 1)}{\Pr(z = 1|y_{z=0}, y_{z=1}, x)} = \Pr(z = 1|x). \quad (12)$$

where

$$\begin{aligned} \Psi_{z|x}(z, x) &= dF(z|x, t_{z|x} = t)/dF(z|x, t_{z|x} = t + 1) = \\ &= z \frac{\Pr(z = 1|x, t_{z|x} = t)}{\Pr(z = 1|x, t_{z|x} = t + 1)} + [1 - z] \frac{\Pr(z = 0|x, t_{z|x} = t)}{\Pr(z = 0|x, t_{z|x} = t + 1)} \end{aligned}$$

can be estimated using a standard probit model such as

$$\Pr(z = 1|x, t_{z|x} = t) = 1 - \Phi(-\beta'_t H(x)).$$

We may then design a triple decomposition for the impact of changes in mortality patterns between t and $t + 1$. First, we compute a counterfactual for the $t + 1$ distribution of income discounting the direct arithmetic impact of *changes* in individual mortality patterns based on observable attributes

$$f_{t+1}^{**}(y) = \int_{x \in \Omega_x} \frac{s_{x,t+1}(x)}{s_{x,t}(x)} f(y|x, t_y = t + 1) dF(x|t_x = t + 1). \quad (13)$$

Second, we compute a counterfactual for the $t + 1$ distribution of income discounting both the direct and the indirect impact of *changes* in mortality patterns based on observable attributes:

$$\begin{aligned} f_{t+1}^{\Delta}(y) &= \int \int \Psi_{z|x}(z, x) \frac{s_{x,t+1}(x)}{s_{x,t}(x)} \times \\ &= f(y|x, z, t_y = t + 1) dF(z|x, t_{z|x} = t + 1) dF(x|t_x = t + 1). \end{aligned} \quad (14)$$

Third, we compute a counterfactual for the $t + 1$ distribution of income discounting the effect of all changes in the distribution of observable attributes:

$$\begin{aligned} g_{t+1}^{\Delta}(y) &= \int \int \frac{\Psi_{x,t}(x)}{\Psi_{x,t+1}(x)} \Psi_{z|x}(z, x) \frac{s_{x,t+1}(x)}{s_{x,t}(x)} \times \\ &= f(y|x, z, t_y = t + 1) dF(z|x, t_{z|x} = t + 1) dF(x|t_x = t + 1). \end{aligned} \quad (15)$$

4 An empirical implementation for the case of Indonesia

4.1 Data and economic context

To illustrate the methods proposed in Section 3, we use the three waves of the Indonesian Family Life Survey (IFLS) conducted by RAND, UCLA and the Demographic Institute of the University of Indonesia. The IFLS is a continuing longitudinal socioeconomic and health survey. It is representative for

83% of the Indonesian population living in 13 of the nation's 26 provinces. The first wave (IFLS1) was carried out in 1993 and covers 33,083 individuals living in 7,224 households. IFLS2 sought to reinterview the same respondents in 1997. Movers were tracked to their new location and if possible interviewed there. Finally a full 94.4% of IFLS1 households were relocated and reinterviewed, in the sense that at least one person from the IFLS1 household was interviewed. This procedure added a total of 878 *split-off* households to the *origin* households. The whole cross-section of IFLS2 includes 33,945 individuals living in 7,619 households. The third wave, IFLS3, was conducted in 2000. It covered 6,800 IFLS1 households plus 3,774 split-off households, in total comprising 43,649 individuals. In IFLS3 the re-contact rate was 95.3% of IFLS1 households. Hence, nearly 91% of IFLS1 households are complete panel households.⁷ Table A1 in the appendix presents some descriptive statistics of the complete samples in 1993, 1997 and 2003. The samples of 1997 and 2003 are cross-sections in the sense that they include besides the panel-individuals also the individuals who were born after 1993 or joined a household of the initial sample or a *split-off* household by another reason.

Using the data we constructed two longitudinal samples: 1993 to 1997 and 1997 to 2000. We included in each those individuals who were reinterviewed at the end of the respective period or for whom a death or another reason for an 'out-migration' was declared. Out-migration means here that these individuals left their households for other reasons and moved to provinces not covered by the survey.⁸ The survey informs about the exact date of the interviews and the month of death, such that a relatively detailed survival analysis can be performed. Between 1993 and 1997 we counted 743 deaths and between 1997 and 2000 558 deaths (see also Table A1).

The IFLS contains detailed information on household expenditure. In contrast, household incomes and especially individual incomes are not completely observed, therefore we use in what follows real household expenditure per capita as welfare or income measure. Expenditures are expressed in prices

⁷For details see Frankenberg and Karoly (1995), Frankenberg and Thomas (1997) and Strauss, Beegle, Sikoki *et al.* 2004.

⁸Or they migrated to provinces covered by the survey but have not been relocated.

of 1994 and adjusted by regional price deflators to the price level of Jakarta. Inter-temporal price variations are taken into account by household-specific price deflators, i.e. using disaggregated price deflators and the observed budget shares of each household as weights.

It is worth to emphasize that the economic crisis began to be felt in the southeast Asia region in April 1997, but that the major impact did not hit Indonesia until December 1997/January 1998, that is just after the IFLS2 was carried out. The sustained crisis period lasted in Indonesia more than one year. But, in the year 2000, year of the IFLS3, the population had again achieved roughly the pre-crisis living standard, with some people even a little better off (Strauss, Beegle, Dwijanto *et al.*, 2002). When constructing the expenditure distributions for 1997 and 2000 we find exactly this dynamic, i.e. slightly lower poverty and inequality in 2000 compared to 1997. For the pre-crisis period 1993 to 1997 we state a substantial poverty reduction. This is also consistent with other findings (see e.g. Tjiptoherijanto and Remi, 2001) and explains very well why Indonesian households—using the former positive dynamic—recovered so quickly from the crisis.

However, public health expenditures fell significantly during the crisis. In addition the 1997/98 drought, which was a consequence of *El Nino* and serious forest fires led in some regions to serious health problems and a sharp drop in food production. Rukumnuaykit (2003) shows that the drought and smoke pollution had significant adverse effects on infant mortality in rural areas. However, Strauss, Beegle, Dwijanto *et al.* (2002) found that adult Body-Mass-Indices did not worsen and that the fraction of preschool-aged children who have very low heights for their age and sex even dropped over the whole period between 1997 and 2000.

4.2 Some illustrative simulations

Before we implement empirically the approach suggested in Section 3, it seems useful to give an order of magnitude of the potential effects of differential mortality on standard income distribution indicators. For this purpose we use a fictitious sample of 10,000 individuals i where the only observed heterogeneity

stems from income y_i . To this sample we apply a crude death rate of d . In the baseline scenario, deaths are drawn randomly, i.e. independent of income. Then we analyze various scenarios where the selection of death events is correlated with income, but perturbed by some unobserved heterogeneity γ_i . The relative risk r_i of death is assumed to be given by the relationship

$$\ln r_i = \lambda \ln y_i + \gamma_i. \quad (16)$$

The term for unobserved heterogeneity is drawn from a normal distribution $N(\mu_\gamma, \sigma_\gamma^2)$. Therefore the correlation coefficient between r_i and income y_i , $\varphi(r_i, y_i)$, depends, for a given distribution of y_i , on λ , μ_γ and σ_γ^2 . Persons who die are selected by ranking in a descending order the sample according to r_i and simulating a death for the d times 10,000-persons for whom r_i is the highest. Therefore, we can write the individual probability of death, P_i , as follows

$$P_i = P(d_i = 1) = P(r_i \geq \tilde{r}) = P(\lambda \ln y_i + \gamma_i \geq \ln \tilde{r}) = P \left[\frac{\gamma_i - \mu_\gamma}{\sigma_\gamma} \geq \frac{\ln \tilde{r} - \mu_\gamma - \lambda \ln y_i}{\sigma_\gamma} \right] \quad (17)$$

and the corresponding c.d. as

$$P_i = 1 - \Phi \left(\frac{\ln \tilde{r} - \mu_\gamma - \lambda \ln y_i}{\sigma_\gamma} \right). \quad (18)$$

In total we examine four different simulations, which we compare with the baseline scenario. The various sets of parameters assumed are noted in Table 1.

[please insert Table 1 about here]

The incomes y_i are drawn from a log-normal distribution where the mean and the variance correspond to those observed in our sample drawn from the IFLS1. As income distribution indicators we consider the Gini-coefficient and the poverty headcount-index, i.e. the percentage of persons below the poverty line. We chose two alternative poverty lines. A first which considers the 10% and a second which considers the 50% at the bottom of the income distribution in the base year as poor. The corresponding simulation results are shown in Figure 1.

[please insert Figure 1 about here]

The first line (Simulation 1) of Figure 1 shows that for a death rate of 3% and a relatively important unobserved heterogeneity component the Gini-coefficient decreases by roughly one percentage point if we decrease λ from 0 to -1. A value for -1 for λ implies that an increase of y by one percent decreases the risk factor of death by also one percent. If λ is 0, i.e. there is no differential mortality, the Gini-coefficient corresponds of course to that of the baseline. If mortality was positively correlated with income, i.e. λ between zero and one, inequality tends also to decrease. In both cases, negative and positive correlation between mortality and income, inequality decreases, because in each case we ‘eliminate’ persons at the lower or upper end of the income distribution. In contrast, a scenario, where in particular individuals of the middle class faced higher mortality could lead to an increase of inequality. If we increase the death rate to 0.06 (Simulation 2) or decrease the importance of the error term (Simulation 3) or both (Simulation 4), we can state, as one can expect, that variations in inequality become correspondingly stronger. The effects on inequality are not symmetric for negative and positive values of λ . This is due to the fact that the initial distribution is skewed to the left, i.e. is normal in $\ln(y)$, not in y .

The second and third row in Figure 1 show, that the poverty rate reacts also strongly to the degree of differential mortality. Assuming a death rate of 0.03 (Simulation 1 and 3) and strong negative differential mortality, results in a decrease of the poverty headcount-index by roughly 2 percentage points, which corresponds in the case of the lower poverty line to roughly 20%. Again, the effect is stronger if the death rate increases (Simulation 2), the error term becomes less important (Simulation 3), or both (Simulation 4). For instance in Simulation 4, the headcount-index for the 10% poverty line decreases by roughly 50%. For positive values of λ and the lower poverty line the headcount-index is of course less affected than with the higher poverty line.

These simple simulations illustrate the potential and pure demographic effect of differential mortality on income distribution, and in particular the distinct effects on inequality and poverty measures. In the next section, we try

to isolate this effect from the total changes in inequality observed in Indonesia during the periods 1993 to 1997 and 1997 to 2000. The empirical application poses of course a lot of additional problems, such as the fact that individuals are grouped in households taking joint decisions on labor supply, household production and expenditures.

4.3 Results for Indonesia for the period 1993 to 2000

4.3.1 Estimates of the direct arithmetical impact of mortality

In what follows we construct ‘without individual deaths’ counterfactuals of the Indonesian distribution of log income per capita for 1997 and 2000 using the methods outlined in Section 3.

We begin with the estimation of the $s_{x,t}(x)$ and $\Psi_{x,t}(x)$ weights, for $t = 1993$ and $t = 1997$. For each sex, we first estimate a (by cross-section sample weights weighted) probit model of surviving between 1993 and 1997 and between 1997 and 2000,⁹ depending on a set of individual attributes: a third degree polynomial of age in the initial year, the size of the household, dummies for the education level of the individual and of the household head, sex of the household head, a third degree polynomial for the age of the household head, and a dummy for urban areas. Table A2 (Appendix) shows the probit estimates of the $s_{x,t}(x)$ function, for both sexes and for both periods. In both periods for women, differentials linked to individual education and to household size are significant. We also estimated probit models for ‘being present in 1997 rather than in 1993’ and for ‘being present in 2000 rather than in 1997’, in order to compute the $\Psi_{x,t}(x)$ weights (see Table A3). Estimations show that, over time, the sample population gets a bid older, a bid more educated, lives more often in urban areas and in households of smaller sizes (see also Table A1). These probabilities reflect overall demographic changes including migration and the educational expansion having occurred during both periods. They may also reflect some sampling bias linked to the panel structure of the IFLS surveys (attrition).

We then first compute density estimates (gaussian kernels with band-

⁹We also tested duration models to estimate survival rates. However, this did not change significantly the following results. Therefore we kept the simple probit model.

width=0.2) of f_{93} , f_{97} and f_{00} , for the actual distributions of log income (Figure 2). Figure 2a reveals that the income distribution has substantially improved between 1993 and 1997, with a large reduction in poverty and inequality. The vertical line corresponds to a constant poverty line used throughout the analysis.¹⁰ In 2000, that is after the macro-economic crisis, the income distribution came only back to its shape of 1997, in line with the results already found by Strauss *et al.* (2002). Figure 2b shows the corresponding differences of the density distributions.

[please insert Figure 2 about here]

We next compute (weighted by cross-section sample weights) kernel estimates of f_{93}^* and f_{97}^* for the ‘direct mortality impact’ counterfactual distributions. We also compute a $f_{93}^{(0)}$ (resp. $f_{97}^{(0)}$) density estimated on the 1993 (resp. 1997) population from which (future) dead individuals between 1993 and 1997 (resp. 1997 and 2000) have been removed (Figures 3a and 3b).¹¹ Figure 3a compares the two counterfactual impacts of individual deaths: $f_{93}^{(0)} - f_{93}$ (without dead individuals) and $f_{93}^* - f_{93}$ (1993 reweighted). Figure 3b does the same for the 1997-2000 period. The ‘without dead individuals’ impacts take into account individual mortality differentials linked to unobservable factors like for instance transient components of income. For both periods, the absence of significant differences between these latter counterfactuals and the two others gives confidence in our choice to compute mortality impacts through reweighing techniques based on observables that are exogenous to income. In all cases, individual mortality directly contributes to a decrease in poverty, as argued by Kanbur and Mukherjee (2003). This result could also be forecast from the positive correlation between initial income and survival probabilities, i.e. the extend of differential mortality with respect to income, that is shown for selected age groups in Figure 4. However, the magnitude of all these counterfactual impacts is very small when compared to the magnitude of observed

¹⁰This poverty line was determined such that we matched exactly the headcount-index computed by Strauss *et al.* (2002) using the IFLS data of 1997, i.e. 32,041 Rupiahs per month in 1994 Jakarta prices.

¹¹Whenever we measure an impact by a difference in densities we smooth this difference again by a gaussian kernel of bandwidth 0.2.

changes in the distribution between 1993 and 1997 (compare the scale of the vertical axis in Figures 2a and 3a). To see how the observable determinants of mortality are related directly to income one can also refer to Table A5, which presents income per capita regressions using the same exogenous variables as the equations used to estimate the survival probabilities in Table A2.

[please insert Figures 3 and 4 about here]

Next, we compute kernel estimates of g_{93}^* and g_{97}^* for the ‘constant distribution of attributes’ DiNardo *et al.* (1996) counterfactual distributions. Remember that these ‘all observable attributes’ counterfactuals also include the impact of individual mortality on the distribution of observable attributes in the population. In Figure 5, we then present the corresponding differences $g_{93}^* - f_{93}$ and $g_{97}^* - f_{97}$ and compare them to the direct mortality impacts $f_{93}^* - f_{93}$, and to the difference $f_{97}^* - f_{97}$ that we have just described. The comparison shows that individual mortality plays only a minor role for the distributional changes that can be imputed to demographic changes. The mortality impacts are ten (in the case of 1993-97) to twenty times (1997-2000) lower in magnitude than overall demographic (including education) impacts. However, it is interesting to see that overall changes in the distribution of observable attributes go in the same direction than the effects of individual mortality, i.e. they are unambiguously poverty decreasing.

[please insert Figure 5 about here]

Finally, Figures 6a and 6b summarize the results by sequentially discounting from the $f_{97} - f_{93}$ (resp. $f_{00} - f_{97}$) density difference, first the impact of mortality, and second the impact of all changes in the distribution of attributes (including mortality). Obviously, changes in mortality and in the population structure do not contribute to the explanation of the change in the distribution of income per capita between 1993 and 1997. In contrast for the period 1997-2000, the distributional impact of demographic changes other than mortality has the same order of magnitude than the observed distributional changes. Reweighing indicates that demographic changes induce a shift

towards the right of the distribution of income, i.e. without such changes the poverty rate would have been a bit worse than those observed in 2000. In a way, overall demographic changes have contributed to the observed recovery from the 1997/98 crisis, but do not explain much of the changes in inequality. Indeed, the income regressions presented in Table A5 confirm that households with more educated members, with smaller sizes and living in urban areas have higher real expenditures per capita. It is therefore not surprising to find that the main demographic changes that we underlined above led to some (counterfactual) poverty reduction.

[please insert Figure 6 about here]

4.3.2 Estimates of the direct and indirect impact of changes in mortality patterns

We now integrate in our analysis the indirect impact of mortality on the income of household survivors using the methodology described in Section 3. We therefore add to our estimates of individual survival probabilities estimates of the conditional individual probability (conditional on household observables) of having experienced a death in the household of origin between 1993 and 1997 or between 1997 and 2000 respectively, i.e. estimates of $\Pr(z = 1|x, t_{z|x} = 1997)$ and $\Pr(z = 1|x, t_{z|x} = 2000)$. This estimation is performed using a probit model (weighted by cross-section sampling weights) for both sexes and both periods. The results are given in Table A4. All estimates show that households headed by a woman have more often experienced a death event, which can easily be understood. Like in the case of individual survival probabilities, education and household size differentials also play a role in explaining the probability of death events.¹²

Between 1993/97 and 1997/2000, overall individual mortality rates decrease. The mortality gradients also change, as can be seen in Table A2. Likewise, the occurrence of death events in households decreases between the two periods; there are also some changes in the probability functions of being a household survivor, as can be seen in Table A4. Using the ratio of survival

¹²Even if measured at the end of the period, the household size is positively correlated with the probability of having experienced a death.

probability estimates measured for both periods, we compute kernel estimates of the direct effect of changes in mortality patterns on the evolution of the distribution of income ($f_{97}^{**} - f_{97}$). With the ratio of the household survivors probability functions, we compute the indirect effect of changes in mortality ($f_{97}^{indir} - f_{97}$). The impact of such such changes in mortality levels and gradients are assessed in Figure 7.

[please insert Figure 7 about here]

The direct effect of the change in mortality patterns ($f_{97}^{**} - f_{97}$) is unambiguously poverty increasing, but has again a very small order of magnitude. All happens as if the mortality decrease between 1993/97 and 1997/2000 had not modified the correlation between the distribution of income and the distribution of death events (income inequalities in front of death). Put differently, poor individuals have not benefited more than others from mortality improvements. Or, the direct effect of the mortality decline increases monetary poverty!

Conversely, the indirect effect of the change in mortality patterns ($f_{97}^{indir} - f_{97}$) is unambiguously poverty decreasing, although still with a very small order of magnitude. In contrast to the direct impact, the indirect impact of mortality changes ($f_{97}^{indir} - f_{97}$) indeed contributes to a decrease in the weight of the poorest. All happens as if households of ‘survivors’, controlling for all other observables, were poorer than their ‘unaffected’ counterparts. In other words, a decrease in mortality rather makes the poorest households richer and improves the distribution of income. Or, the indirect effect of the mortality decline decreases monetary poverty!

Hence, when the direct and indirect impact of changes in mortality patterns are summed together ($f_{97}^{\Delta} - f_{97}$), the result is more ambiguous. The overall changes in mortality patterns seem to have contributed to a slight decrease in the inequality of the income distribution rather than to the change in poverty. This implies, that the overall effect of the mortality decline on monetary poverty is ambiguous.

Finally, we assess the impact of all changes in the population structure, in-

cluding survivor's status ($g_{97}^{\Delta} - g_{97}$). Figure 8 shows that the mortality effects, whether direct or indirect, are completely dominated by other demographic effects. Here again, demographic changes influence the distribution of income in the same direction than the direct arithmetic effect of mortality, but at a higher level of magnitude. The changes in the evolution of the population structure in terms of age, education and place of residence (urban/rural) have again a slight poverty increasing effect. If the speed of demographic changes had remained the same in the period 1997/2000 than in the period 1993/97, then the resulting distribution of income in the year 2000 would have shown a bid lower poverty and inequality. Instead, some deceleration of positive demographic changes occurred, which worsened the impact of the 1997/98 economic crisis.

[please insert Figure 8 about here]

However, Figure 9 shows that the overall impact of these changes in the demographic structure is rather small, i.e. without any variation in the population structure, the observed evolution between 1997 and 2000 would not have been very different.

[please insert Figure 9 about here]

5 Conclusion

We have presented a general methodology designed to study the counterfactual impact of mortality and changes in mortality on income distribution. This methodology is inspired by the work of DiNardo *et al.* (1996). It relies on the non-parametric reweighing of income distributions using functions of individual observable attributes. Like Kanbur and Mukherjee (2003), we are able to deal with the direct arithmetic effect of individual deaths on poverty changes,¹³ which is most important when individual deaths are unevenly distributed across the income distribution. But we are also able to correct for the indirect effect of an individual death on the income of survivors pertaining

¹³However, Kanbur and Mukherjee (2003) did not implement their approach empirically.)

to the same household, which might be as much important. If the mortality risk is negatively correlated with income, then, when mortality increases (resp. decreases) over time, the direct effect is usually poverty decreasing (resp. increasing). Conversely, if mortality is negatively correlated with income and if a death in a household decreases household income, then, when mortality increases over time (resp. decreases), the indirect effect should be poverty increasing (resp. decreasing). In our empirical part we show for the case of Indonesia, that the direct and indirect effects of a mortality decline on the distribution of income have indeed opposite signs and show roughly the same order of magnitude, so that they almost cancel out each other. We also show that the effect of other demographic changes, like changes in the pattern of fertility, migration, and educational attainment, dominate the effects of mortality, whether direct or indirect. We however find that none of these demographic developments are large enough to explain a significant part of the changes in income distribution, whether the pre-crisis period (1993-1997) or the post-crisis period (1997-2000) is considered.

Appendix

Descriptive statistics of the variables used

[please insert Table A1 about here]

Estimated equations for survival probabilities ($s_{x,t}(x)$)

[please insert Table A2 about here]

Estimated equations for ‘being present’ probabilities

[please insert Table A3 about here]

Estimated equations for living in a household in which no death occurred during past period

[please insert Table A3 about here]

Estimated coefficients for correlates of household income per capita

[please insert Table A3 about here]

References

- Abadie A., J. D. Angrist, G. W. Imbens (1998), *Instrumental Variables Estimation of Quantile Treatment Effects*, NBER Working Paper t0229, Cambridge.
- Anderson G. (2003), Life Expectancy and Economic Welfare: The Example of Africa in the 1990's. Mimeo, University of Toronto, Department of Economics, Toronto.
- Atkinson A.B. (1970), On the Measurement of Inequality. *Journal of Economic Theory*, 2: 244-263.
- Blackorby C., W. Bossert and D. Donaldson (1995), Intertemporal population ethics: critical-level utilitarian principles. *Econometrica*, 63: 1303-1320.
- Blackorby C. and D. Donaldson (1984), Social criteria for evaluating population change. *Journal of Public Economics*, 25: 13-33.
- Brainerd E. and M.V. Sieglar (2003), The economic effects of the 1918 influenza epidemic. CEPR Discussion Paper No. 3791, CEPR, London.
- Chu C.Y.C. and H. Koo (1990), Intergenerational income-group mobility and differential fertility. *American Economic Review*, 80: 1125-1138.
- Dasgupta P., A. Sen and D. Starrett (1973), Notes on the Measurement of Inequality. *Journal of Economic Theory*, 6: 180-187.
- Deaton A. and C. Paxson (1999), Mortality, education, income and inequality among American cohorts. Mimeo, Princeton, Princeton University.
- DiNardo J., N.M. Fortin and T. Lemieux (1996), Labor market institutions and the distribution of wages, 1973-1992: A semiparametric approach. *Econometrica*, 64 (5): 1001-1044.
- Firpo S. (2004), *Efficient Semiparametric Estimation of Quantile Treatment Effects*, University of British Columbia, Vancouver, Discussion Paper 04-01.
- Foster J., J. Greer and E. Thorbecke (1984), A Class of Decomposable Poverty Measures. *Econometrica*, 52: 761-776.
- Frankenberg E. and L. Karoly (1995), The 1993 Indonesian Family Life Survey: Overview and Field Report, RAND, DRU-1195/1-NICHD/AID.
- Frankenberg E. and D. Thomas (2000), The Indonesian Family Life Survey: Study Design and Results from Waves 1 and 2, RAND, DRU-2238/1-NIA/NICHD.
- Kanbur R. and D. Mukherjee (2003), Premature Mortality and Poverty Measurement. ISER Working Papers 2003-6, ISER, University of Essex.

- Kitagawa E.M. and P.M. Hauser (1973), *Differential Mortality in the United States: A Study in Socioeconomic Epidemiology*. Cambridge: Harvard University Press.
- Lam D. (1986), The dynamics of population growth, differential fertility and inequality. *American Economic Review*, 76: 1103-1116.
- Lam D. (1997), Demographic variables and income inequality. In M.R. Rosenzweig and O. Stark (eds.), *Handbook of Population and Family Economics* (pp. 1015-1059), Amsterdam: North-Holland.
- Lantz P.M., J.S. House, J.M. Lepkowski, D.R. Williams, R.P. Mero and J. Chen (1998), Socioeconomic factors, health behaviors, and mortality. *Journal of the American Economic Medical Association*, 279: 1703-1708.
- Oaxaca R. (1973), Male-Female Wage Differentials in Urban Labor Markets. *International Economic Review*, 14: 693-709.
- Rukumnuaykit P. (2003), Crises and Child Health Outcomes: The Impacts of Economic and Drought/Smoke Crises on Infant Mortality and Birth-weight in Indonesia. Mimeo, Michigan State University, Lansing.
- Sen A.K. (1985), *Commodities and Capabilities*. Amsterdam: North-Holland.
- Strauss J., K. Beegle, A. Dwiyanto, Y. Herawati, D. Pattinasarany, E. Satriawan, B. Sikoki, B. Sukamdi and F. Witoelar (2002), Indonesian living standards three years after the crisis: evidence from the Indonesia Family Life Survey, Executive Summary, unpublished, Michigan State University, World Bank, University of Jakarta, and RAND Corporation.
- Strauss J., K. Beegle, B. Sikoki, A. Dwiyanto, Y. Herwati and F. Witoelar (2004), The Third Wave of the Indonesia Family Life Survey (IFLS3): Overview and Field Report. WR-144/1-NIA/NCHID, RAND Corporation.
- Tjiptoherijanto P. and S.S. Remi (2001), Poverty and Inequality in Indonesia: Trends and Programs. Paper presented at the International Conference on the Chinese Economy 'Achieving Growth with Equity', Beijing, July 4-6.
- Valkonen T. (2002), Social inequalities in mortality. In G. Caselli, J. Vallin and G. Wunsch (eds.), *Demography: Analysis and Synthesis*, (pp. 51-67), INED, Paris.

Tables

Table 1
The effects of differential mortality
on standard income distribution indicators
(some illustrative simulations)

	$d = 0.03$	$d = 0.06$
$\mu_\gamma = \overline{\ln(y_i)}$ $\sigma_\gamma = \sigma_{\ln(y_i)}$ $-1 \leq \lambda \leq 1$	Sim. 1	Sim. 2
$\mu_\gamma = 0.5\overline{\ln(y_i)}$ $\sigma_\gamma = 0.5\sigma_{\ln(y_i)}$ $-1 \leq \lambda \leq 1$	Sim. 3	Sim. 4

Notes: For $\lambda = -1$, the constellation of the noted parameters yields the following correlation coefficients, $\varphi(r_i, y_i)$, between the risk factor r_i and income y_i : Simulation 1 and 2: $\varphi(r_i, y_i|\lambda = -1) = -0.333$; Simulation 3 and 4: $\varphi(r_i, y_i|\lambda = -1) = -0.441$.

Table A1
Descriptive statistics of the variables used

	1993	1997	2000
BOYS/MEN			
Age	25.9	27.3	27.5
Education			
No education	0.216	0.188	0.182
Elementary educ.	0.500	0.460	0.435
Junior High.	0.130	0.155	0.155
Senior High./Coll./Univ	0.154	0.197	0.228
HH-head male	0.921	0.908	0.915
Age HH-head	45.4	46.9	46.0
Education of HH-head			
No education	0.179	0.144	0.110
Elementary educ.	0.560	0.535	0.504
Junior High.	0.106	0.120	0.132
Senior High./Coll./Univ	0.155	0.201	0.254
HH-size	5.549	5.407	5.189
Urban	0.352	0.401	0.440
Death in HH in 1993-97. 1997-2000		0.102	0.072
No. of observations	16,058	16,325	20,966
<i>Tracking status (shares)</i>			
		<i>1993-1997</i>	<i>1997-2000</i>
Survivors		0.969	0.977
Deaths		0.031	0.023
GIRLS/WOMEN			
Age	26.6	27.8	28.6
Education			
No education	0.302	0.254	0.241
Elementary educ.	0.486	0.455	0.443
Junior High.	0.104	0.136	0.139
Senior High./Coll./Univ	0.108	0.154	0.178
HH-head male	0.852	0.840	0.834
Age HH-head	45.5	47.0	46.3
Education of HH-head			
No education	0.202	0.161	0.126
Elementary educ.	0.539	0.522	0.495
Junior High.	0.101	0.115	0.128
Senior High./Coll./Univ	0.158	0.202	0.251
HH-size	5.415	5.309	5.095
Urban	0.357	0.403	0.444
Death in HH in 1993-97. 1997-2000		0.105	0.073
No. of observations	16,970	17,487	21,985
<i>Tracking status (shares)</i>			
		<i>1993-1997</i>	<i>1997-2000</i>
Survivors		0.976	0.978
Deaths		0.024	0.022

Source: IFLS1, IFLS2 and IFLS3; computations by the authors.

Table A2
 Estimated probit model for survival probabilities ($s_{x,t}(x)$)
 (marginal probabilities computed at sample means)

<i>Dependent variable</i> <i>Survived (binary)</i>	1993—1997		1997—2000	
	Coeff.	Std. Err.	Coeff.	Std. Err.
BOYS/MEN				
Age	4.35E-04	0.001	0.001	4.40E-04
Age ²	-1.68E-05	1.98E-05	-2.30E-05	** 1.14E-05
Age ³	-4.20E-08	1.59E-07	9.64E-08	8.25E-08
Education				
No education	Ref.		Ref.	
Elementary educ.	0.003	0.005	0.003	2.94E-03
Junior High.	-0.003	0.008	0.005	0.003
Senior High./Coll./Univ	0.005	0.007	0.003	0.004
HH-head male	-0.007	0.005	0.004	0.004
Age HH-head	0.002	*	0.001	1.49E-04
Age ² HH-head	-5.92E-05	**	2.95E-05	-9.12E-06
Age ³ HH-head	4.68E-07	**	2.14E-07	1.06E-07
Education of HH-head				
No education	Ref.		Ref.	
Elementary educ.	0.003	0.005	-0.002	3.05E-03
Junior High.	0.008	0.005	-0.002	0.005
Senior High./Coll./Univ	0.006	0.006	-0.008	0.006
ln HH-size	-0.003	0.003	0.002	0.002
Urban	-0.004	0.003	-0.001	0.002
No. of observations	13,548		14,490	
Pseudo R ²	0.192		0.196	
GIRLS/WOMEN				
Age	8.02E-05	4.03E-04	0.001	*** 2.82E-04
Age ²	-9.20E-06	1.09E-05	-2.54E-05	*** 7.00E-06
Age ³	-9.15E-09	8.12E-08	1.08E-07	** 4.72E-08
Education				
No education	Ref.		Ref.	
Elementary educ.	0.009	***	0.003	0.001
Junior High.	0.008	**	0.002	0.004
Senior High./Coll./Univ	0.010	***	0.002	0.008
HH-head male	9.73E-05	0.003	-0.003	* 0.002
Age HH-head	0.001	0.001	-1.38E-04	0.001
Age ² HH-head	-1.64E-05	1.79E-05	-2.78E-06	2.04E-05
Age ³ HH-head	9.28E-08	1.15E-07	4.25E-08	1.25E-07
Education of HH-head				
No education	Ref.		Ref.	
Elementary educ.	-0.002	0.003	-4.20E-04	0.002
Junior High.	0.002	0.004	0.002	0.003
Senior High./Coll./Univ	0.004	0.003	0.001	0.003
ln HH-size	-0.008	***	0.002	-0.005
Urban	1.67E-04	0.002	0.002	0.001
No. of observations	14,429		15,583	
Pseudo R ²	0.204		0.246	

Notes: *** coefficient significant at the 1% level, ** 5% level, * 10% level.

Source: IFLS1, IFLS2 and IFLS3; estimations by the authors.

Table A3
 Estimated probit model for 'being present' probabilities
 (marginal probabilities computed at sample means)

<i>Dependent variable being present (binary)</i>	1993/1997		1997/2000			
	Coeff.	Std. Err.	Coeff.	Std. Err.		
BOYS/MEN						
Age	-1.08E-04	0.002	-0.011	***	0.001	
Age ²	2.11E-05	4.69E-05	3.01E-04	***	4.09E-05	
Age ³	-2.36E-07	3.84E-07	-2.15E-06	***	3.28E-07	
Education						
No education	Ref.		Ref.			
Elementary educ.	-0.010	0.012	0.038	***	0.011	
Junior High.	0.038	**	0.015	0.055	***	0.014
Senior High./Coll./Univ	0.016	0.017	0.071	***	0.015	
HH-head male	-0.050	***	0.012	0.001	0.010	
Age HH-head	0.007	0.005	-0.004		0.004	
Age ² HH-head	-9.01E-05	9.93E-05	1.30E-05		7.43E-05	
Age ³ HH-head	5.63E-07	6.60E-07	2.47E-07		4.95E-07	
Education of HH-head						
No education	Ref.		Ref.			
Elementary educ.	0.074	***	0.011	0.035	***	0.010
Junior High.	0.098	***	0.014	0.058	***	0.013
Senior High./Coll./Univ	0.141	***	0.014	0.076	***	0.013
ln HH-size	-0.046	***	0.008	-0.052	***	0.007
Urban	0.021	***	0.007	0.019	***	0.006
No. of observations	33,383		37,291			
Pseudo R ²	0.011		0.009			
GIRLS/WOMEN						
Age	-0.004	***	0.001	-0.014	***	0.001
Age ²	1.03E-04	***	3.90E-05	4.02E-04	***	3.91E-05
Age ³	-5.05E-07	*	3.04E-07	-2.79E-06	***	3.11E-07
Education						
No education	Ref.		Ref.			
Elementary educ.	0.043	***	0.009	0.062	***	0.009
Junior High.	0.116	***	0.013	0.083	***	0.012
Senior High./Coll./Univ	0.123	***	0.014	0.100	***	0.012
HH-head male	-0.020	**	0.009	-0.001		0.008
Age HH-head	0.009	**	0.005	-0.007	*	0.004
Age ² HH-head	-1.16E-04		9.66E-05	5.52E-05		7.30E-05
Age ³ HH-head	5.68E-07		6.26E-07	1.73E-08		4.75E-07
Education of HH-head						
No education	Ref.		Ref.	***		
Elementary educ.	0.065	***	0.010	0.039	***	0.009
Junior High.	0.089	***	0.013	0.064	***	0.012
Senior High./Coll./Univ	0.110	***	0.012	0.080	***	0.011
ln HH-size	-0.034	***	0.008	-0.042	***	0.007
Urban	0.005		0.007	0.018	***	0.006
No. of observations	34,457		39,472			
Pseudo R ²	0.013		0.011			

Notes: *** coefficient significant at the 1% level, ** 5% level, * 10% level.

Source: IFLS1, IFLS2 and IFLS3; estimations by the authors.

Table A4
 Estimated probit model for living in a household
 in which no death occurred during past period
 (marginal probabilities computed at sample means)

<i>Dependent variable survivor (binary)</i>	1993—1997		1997—2000	
	Coeff.	Std. Err.	Coeff.	Std. Err.
BOYS/MEN				
Age	4.26E-04	0.001	0.002	0.001
Age ²	-1.07E-05	3.63E-05	-2.85E-05	3.30E-05
Age ³	5.11E-08	2.86E-07	1.10E-07	2.56E-07
Education				
No education	Ref.		Ref.	
Elementary educ.	-0.004	0.010	-0.012	0.008
Junior High.	-0.020 *	0.011	-0.014	0.010
Senior High./Coll./Univ	-0.014	0.013	-0.024 *	0.010
HH-head male	-0.156 ***	0.013	-0.144 ***	0.010
Age HH-head	-0.005	0.003	-0.005	0.003
Age ² HH-head	1.68E-05	5.85E-05	2.24E-05	5.36E-05
Age ³ HH-head	3.44E-07	3.91E-07	2.86E-07	3.53E-07
Education of HH-head				
No education	Ref.		Ref.	
Elementary educ.	-8.85E-04	0.009	0.004	0.008
Junior High.	0.004	0.012	0.009 **	0.011
Senior High./Coll./Univ	0.015	0.013	0.015	0.011
ln HH-size	0.015 **	0.007	0.038	0.006
Urban	-0.002	0.006	0.001 **	0.005
No. of observations	16,325		20,966	
Pseudo R ²	0.029		0.021	
GIRLS/WOMEN				
Age	-0.001	0.001	0.001	0.001
Age ²	2.16E-05	2.70E-05	-2.01E-05	2.21E-05
Age ³	-5.38E-08	2.11E-07	7.23E-08	1.64E-07
Education				
No education	Ref.		Ref.	
Elementary educ.	0.008	0.008	-0.008	0.006
Junior High.	0.011	0.011	-0.006	0.008
Senior High./Coll./Univ	0.019 **	0.012	-0.010	0.008
HH-head male	-0.104 ***	0.010	-0.093 ***	0.007
Age HH-head	-3.65E-04 **	0.002	0.007 ***	0.002
Age ² HH-head	1.94E-05	4.46E-05	-1.20E-04 ***	4.43E-05
Age ³ HH-head	-1.87E-07	2.93E-07	6.84E-07 **	2.69E-07
Education of HH-head				
No education	Ref.		Ref.	
Elementary educ.	0.011	0.007	0.018 ***	0.007
Junior High.	0.022	0.011	0.032 ***	0.010
Senior High./Coll./Univ	0.002	0.009	0.008	0.009
ln HH-size	-0.002 ***	0.005	0.010 **	0.004
Urban	-0.008	0.004	-0.006	0.004
No. of observations	17,487		21,985	
Pseudo R ²	0.035		0.029	

Notes: *** coefficient significant at the 1% level, ** 5% level, * 10% level.

Source: IFLS1, IFLS2 and IFLS3; estimations by the authors.

Table A5
 Estimated household income per capita regressions
 (Pooled sample 1993, 1997, 2000)

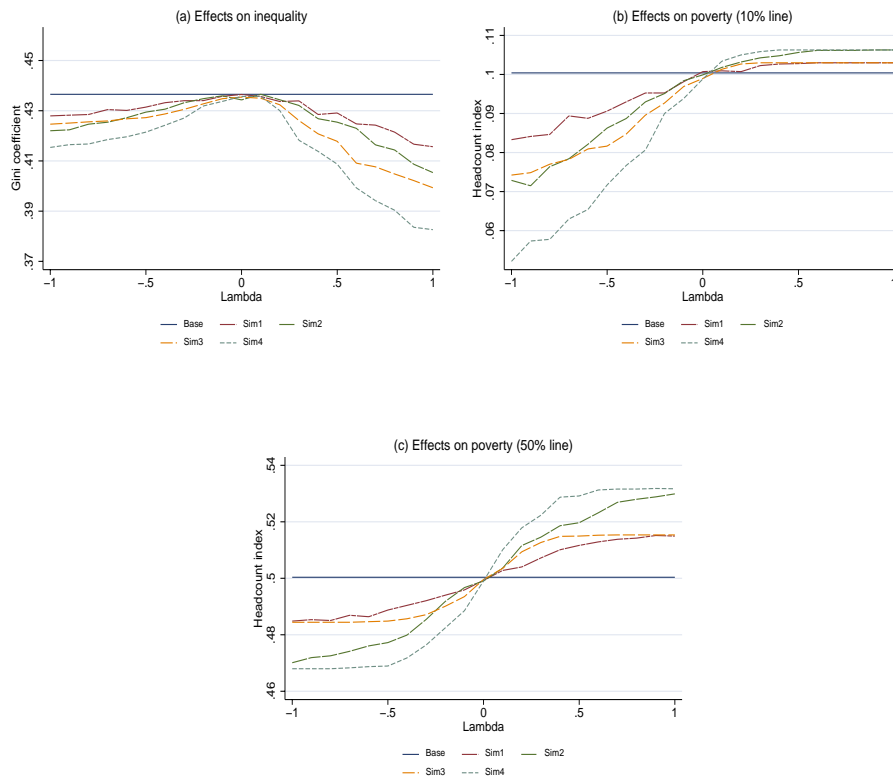
<i>Dependent variable</i> <i>ln HH-expend. per capita</i>	Boys/Men			Girls/Women		
	Coeff.		Std. Err.	Coeff.		Std. Err.
Age	-0.006	***	0.001	-0.011	***	0.001
Age ²	1.45E-04	***	4.15E-05	3.22E-04	***	3.89E-05
Age ³	-9.58E-07	***	3.35E-07	-2.40E-06	***	3.05E-07
Education						
No education	Ref.			Ref.		
Elementary educ.	0.076	***	0.011	0.123	***	0.010
Junior High.	0.164	***	0.015	0.260	***	0.014
Senior High./Coll./Univ	0.293	***	0.017	0.382	***	0.016
HH-head male	0.063	***	0.018	0.064	***	0.016
Age HH-head	0.037	***	0.010	0.032	***	0.009
Age ² HH-head	-4.04E-04	**	1.90E-04	-3.31E-04	*	1.82E-04
Age ³ HH-head	7.23E-07		1.21E-06	5.38E-07		1.15E-06
Education of HH-head						
No education	Ref.			Ref.		
Elementary educ.	0.150	***	0.020	0.189	***	0.017
Junior High.	0.386	***	0.026	0.429	***	0.023
Senior High./Coll./Univ	0.667	***	0.026	0.744	***	0.023
ln HH-size	-0.467	***	0.016	-0.435	***	0.014
Urban	0.220	***	0.013	0.195	***	0.013
IFLS 1993 dummy	Ref.			Ref.		
IFLS 1997 dummy	0.246	***	0.011	0.243	***	0.010
IFLS 2000 dummy	0.220	***	0.011	0.213	***	0.010
Intercept	10.200	***	0.148	10.179	***	0.142
No. of observations			53,349			56,442
R ²			0.299			0.301

Notes: *** coefficient significant at the 1% level, ** 5% level, * 10% level. Huber/White/sandwich estimators used for standard errors to account for dependent observations within households.

Source: IFLS1, IFLS2 and IFLS3; estimations by the authors.

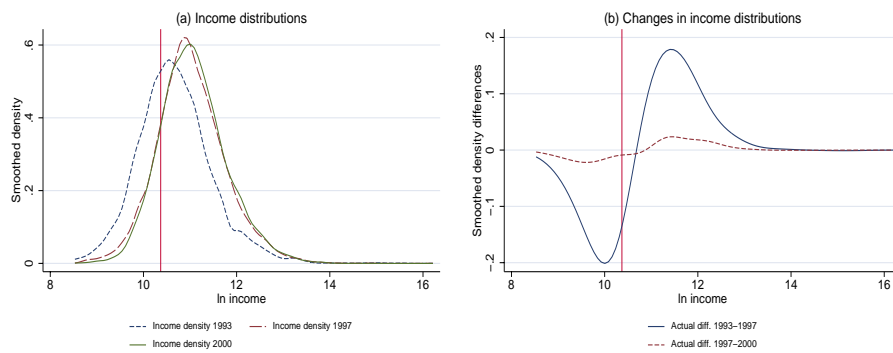
Figures

Figure 1
The effects of differential mortality on standard income distribution indicators
(some illustrative simulations)



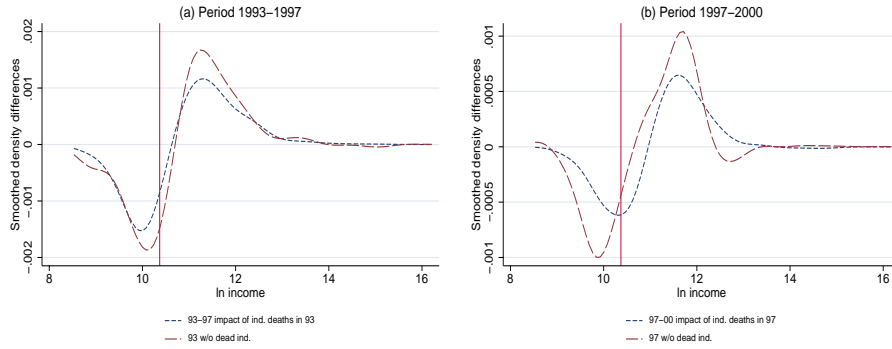
Source : Simulations by the authors.

Figure 2
Income per capita (ln) kernel densities in 1993, 1997 and 2000



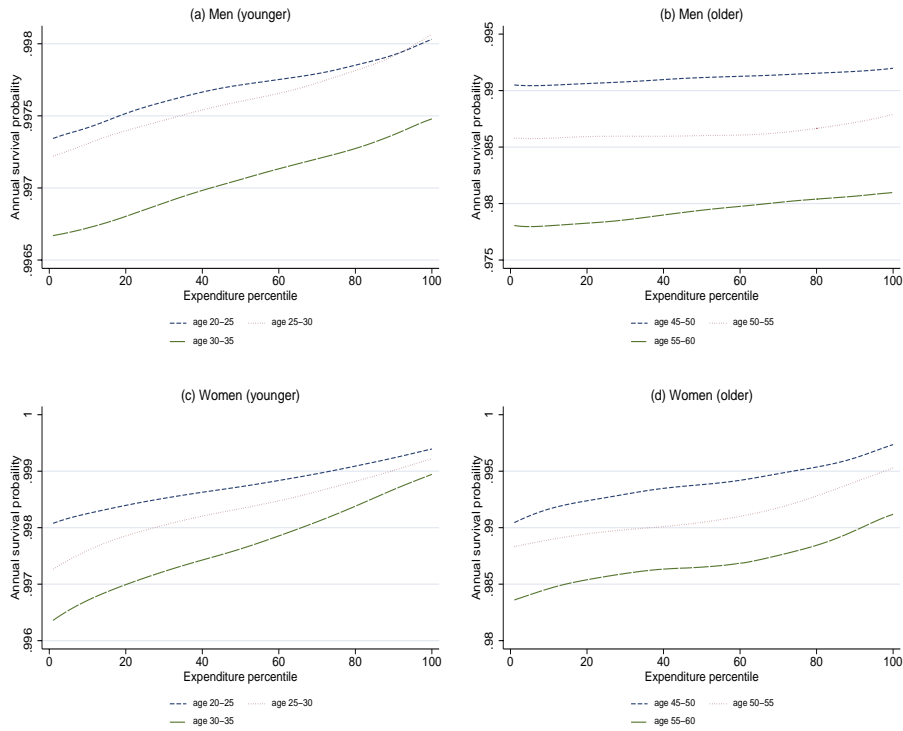
Source : IFLS1, IFLS2 and IFLS3; estimations by the authors.

Figure 3
Smoothed impact of individual deaths



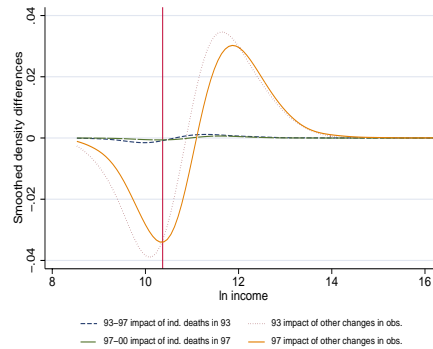
Source : IFLS1, IFLS2 and IFLS3; estimations by the authors.

Figure 4
Smoothed survival probabilities by income per capita percentile
for men and women and selected age groups
(means of predicted values for the 1993 sample using the model of Table A2, col. 1)



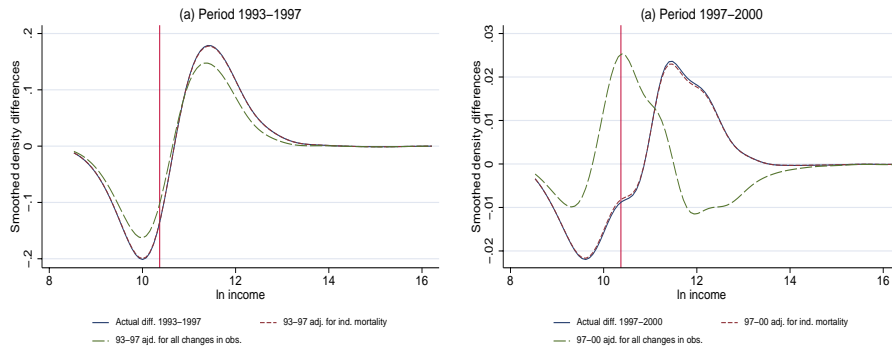
Source : IFLS1 and IFLS2; estimations by the authors.

Figure 5
Smoothed impact of individual mortality
compared to impact of changes in all observable attributes



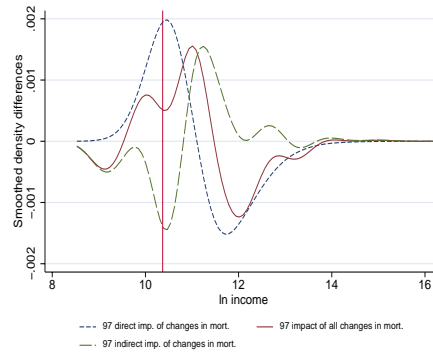
Source : IFLS1, IFLS2 and IFLS3; estimations by the authors.

Figure 6
Smoothed impact of individual mortality and of changes in all observable attributes
on the overall change of the income per capita distributions



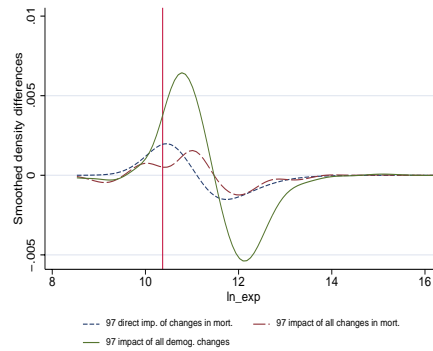
Source : IFLS1, IFLS2 and IFLS3; estimations by the authors.

Figure 7
 Smoothed impact of changes in mortality patterns
 between 1993/1997 and 1997/2000 on the 2000 income distribution



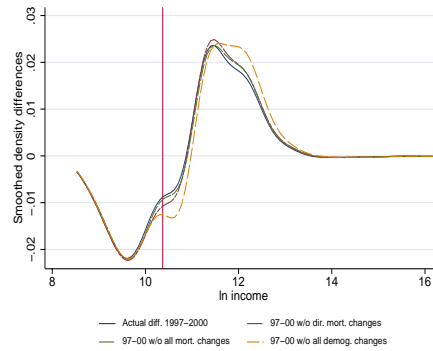
Source : IFLS1, IFLS2 and IFLS3; estimations by the authors.

Figure 8
 Smoothed impact of changes in mortality patterns
 on the income distribution in 2000:
 Changes in mortality patterns between 1993/1997 and 1997/2000
 compared to other changes in the evolution of the population structure



Source : IFLS1, IFLS2 and IFLS3; estimations by the authors.

Figure 9
Smoothed impact of changes in mortality patterns and other observables
on the overall change of the income per capita distribution between 1997 and 2000



Source : IFLS1, IFLS2 and IFLS3; estimations by the authors.