Review of Income and Wealth
Series 53, Number 2, June 2007
doi: 10.1111/j.1475-4991.2007.00229.x

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

BY DENIS COGNEAU

IRD-DIAL Paris

AND

MICHAEL GRIMM*

University of Göttingen, ISS The Hague, DIAL Paris and DIW Berlin

We show how to account for differentials in demographic variables, in particular mortality, when performing welfare comparisons over time. The idea is to apply various ways of "correcting" estimated income distribution measures for "sample selection" due to differential mortality. We distinguish the direct effect of mortality, i.e. individuals who die leave the population and no longer contribute to monetary welfare, from the indirect effect, i.e. the impact on survivors in the deceased's household who may experience a decrease or increase in monetary welfare. In the case of Indonesia, we show that the direct and indirect effects of mortality on income distribution have opposite signs, but are roughly the same in magnitude. Moreover, the effects of other demographic changes dominate the effects of mortality, whether direct or indirect. However, in the post-crisis period these demographic changes also explain a substantial part of the overall change in the distribution of income.

1. INTRODUCTION

Demographic behavior can significantly affect the distribution of income, when it is correlated with the income measure used. If for instance poor people have greater mortality, fertility and migration rates than rich people, income distribution dynamics will be significantly impacted. When analyzing the causes of distributional change, it is useful to isolate these effects from changes in labor supply behavior and changes in returns on the labor market, which can also have a strong impact on the distribution of income, but are driven rather by structural and institutional change. Obviously, the cited transmission channels may be interdependent and therefore hard to disentangle. For instance, the death of one household member can alter the labor supply, the educational investment, and the consumption behavior of other household members. Given the lack of appropriate methods to explore the importance of the demographic channels, little is known about their empirical importance.

The purpose of our paper is twofold: first to derive instructive analytics on how to account for differentials in demographic variables, in particular mortality,

Note: We thank Gordon Anderson, two anonymous referees and the editor Bart van Ark for very useful comments and advice. We also thank participants at the ESPE conference in Bergen, the IARIW conference in Cork, the VfS conference in Bonn and the EEA meeting in Amsterdam. Of course, any remaining errors and omissions remain our own responsibility. Michael Grimm acknowledges financial support for this research by the German Research Foundation (DFG).

*Correspondence to: Michael Grimm, University of Göttingen, Department of Economics, Platz der Göttinger Sieben 3, 37073 Göttingen, Germany (mgrimm@uni-goettingen.de).

when performing welfare comparisons overtime; second, to explore the potential impact of demographic change on the distribution of welfare. The idea of the methodology we suggest is to apply various ways of "correcting" estimated welfare distributions for "sample selection" due to differential mortality. A central issue is then to derive reliable estimates for mortality rates as functions of income (or its correlates) and age. Once the conditional density of mortality is known, a reweighted welfare distribution can be calculated giving the welfare variation attributable to individual deaths. Further complications arise when the household, rather than the individual, is the unit of analysis. The key estimation problem then becomes to construct a counterfactual distribution that would have prevailed if the survivors had continued living with their former household members and had decided jointly on labor supply and consumption expenditure. The procedure we propose to address these issues is very much in the spirit of the decompositions performed by DiNardo *et al.* (1996).

We proceed as follows. In the next section, we discuss the welfare implications of differentials in demographic variables and especially differential mortality. In Section 3, we show some illustrative simulations regarding the potential impact of differential mortality on the distribution of income. In Section 4, we present our methodology. In Section 5, we apply our approach empirically using three waves of the Indonesian Family Life Survey (IFLS). In Section 6, we summarize our main results and conclude.

2. Welfare Implications of Differential Demographics

Empirical studies on income distribution dynamics typically avoid considering variations in population size. They usually provide a kind of "snapshot measure" of economic well-being. In other words, indicators are considered such as per capita GDP, the Human Development Index, the poverty headcount index and the Gini coefficient at two different points in time without taking into account whether the population size has changed over the relevant time period. The implicit ethical judgment, then, is that we are "neutral" to the population.

Such a judgment may be acceptable when variations in population size are independent from the welfare measure under consideration. But, if the demographic forces, such as fertility, mortality and migration, are correlated with the welfare measure, neutrality to these variations may lead to a biased picture of how the income distribution evolved over time. For instance, if mortality is negatively correlated with income, which indeed seems to be the case in both developing and developed countries,[1] standard poverty measures such as the FGT family headcount index (Foster *et al.*, 1984) may show an improvement over time if individuals below the poverty line die. Or, put differently, higher mortality among the poor is "good" for poverty reduction. The current AIDS epidemic in developing

[1]For empirical evidence, see Kitagawa and Hauser (1973), Deaton and Paxson (2001) and Lantz *et al.* (1998). Valkonen (2002) provides a survey of the empirical evidence of social inequalities in mortality. He finds that social inequality is found in almost all studies regardless of the fact that they consider different populations and use different indicators of socio-economic position such as social and occupational class, socio-economic status, educational attainment, income and housing characteristics.

countries, the 1918 influenza epidemic and the black plague centuries ago might have reduced poverty by increasing the capital–labor ratio, but also simply by killing the poor harder hit by the diseases.[2] Most people will agree that this kind of "repugnant conclusion" is incompatible with the principle on which poverty concepts are normally based. This point was recently also raised by Kanbur and Mukherjee (2003).

A similar problem is found if we consider fertility. Higher fertility among the poor may increase poverty due simply to differential growth rates across the income distribution. It could be concluded that minimizing fertility among the poor is a means of reducing poverty.[3] Again, this seems neither economically nor ethically reasonable or acceptable. Lastly, rural-to-urban migration may reduce rural poverty and increase urban poverty, without changing anything in the situation of those who stay in their initial place.

In the following, we suggest some general methods to account for differentials in demographic variables, especially mortality, when analyzing income distribution dynamics over time. We do not address the issue of giving a value to a lost life and hence to different population sizes, but we show how the income distribution would look if variations in population size were not selective, i.e. independent of the underlying welfare metric.

We first consider solely what we call the "direct effect" or "pure demographic effect." Then we develop in turn measures to take into account the effect a death might have on household income rather than just household income per capita, first because the deceased does not contribute to the household's income any more, and second because the death might have changed the labor supply behavior of the other household members. However, before we present our analytical method and its application to Indonesia between 1993 and 2000, it will be useful to give an approximate idea of the potential effects of differential mortality on standard income distribution indicators. For this purpose, we present some illustrative simulations.

3. The Impact of Differential Mortality on the Distribution of Income: Some Illustrative Simulations

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. We then analyze different scenarios where the selection of death events is correlated with income, but disrupted by some unobserved heterogeneity γ_i. The risk r_i of death is assumed to be given by the relationship

$$\ln r_i = \lambda \ln y_i + \gamma_i. \quad (1)$$

The death event is represented by a dichotomic variable $d_i = 0$, 1. A death occurs when r_i becomes higher than a given threshold $\tilde{r}$ (i.e. $d_i = 1 \Leftrightarrow r_i \geq \bar{r}$). The

[2]For instance, Brainerd and Siegler (2003) find empirical evidence that the 1918 influenza epidemic had a robust positive effect on per capita income growth across the U.S. in the 1920s.

[3]See, on this point, the analyses and discussions in Lam (1986) and Chu and Koo (1990).

term for unobserved heterogeneity is derived from a normal distribution $N(\mu_\gamma, \sigma_\gamma^2)$. Hence 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. We can hence write the individual probability of death, P_i, as follows:

$$(2)\qquad P_i = P(d_i = 1) = P(r_i \geq \bar{r}) = P(\lambda \ln y_i + \gamma_i \geq \ln \tilde{r}) = P\left[\frac{\gamma_i - \mu_\gamma}{\sigma_\gamma} \geq \frac{\ln \bar{r} - \mu_\gamma - \lambda \ln y_i}{\sigma_\gamma}\right]$$

and the corresponding c.d. as

$$(3)\qquad P_i = 1 - \Phi\left(\frac{\ln \bar{r} - \mu_\gamma - \lambda \ln y_i}{\sigma_\gamma}\right).$$

We examine a total of four different simulations (see Table 1), which we compare with the baseline scenario.

For computing simulations, people who die are selected by ranking the sample in descending order based on r_i and simulating a death for the d times 10,000-people for whom r_i is the highest. The incomes y_i are derived from a log-normal distribution where the mean and variance correspond to those observed in our sample used later in Section 4 to implement our approach empirically. As income distribution indicators, we consider the Gini coefficient and the poverty headcount index, i.e. the percentage of people below the poverty line. We choose two alternative poverty lines: one considers the first 10 percent and the other considers the first 50 percent at the bottom of the income distribution in the base year as poor. The simulation results are shown in Figure 1.

The first line (Simulation 1) of Figure 1a shows that, for a death rate of 3 percent and a relatively sizeable unobserved heterogeneity component, the Gini coefficient decreases by roughly one percentage point if we reduce λ from 0 to −1. A value of −1 for λ implies that a 1 percent increase in y reduces the risk of death

TABLE 1

The Impact of Differential Mortality on the Distribution of Income: 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: Simulations 1 and 2: $\varphi(r_i, y_i|\lambda = -1) = -0.333$; Simulations 3 and 4: $\varphi(r_i, y_i|\lambda = -1) = -0.441$.

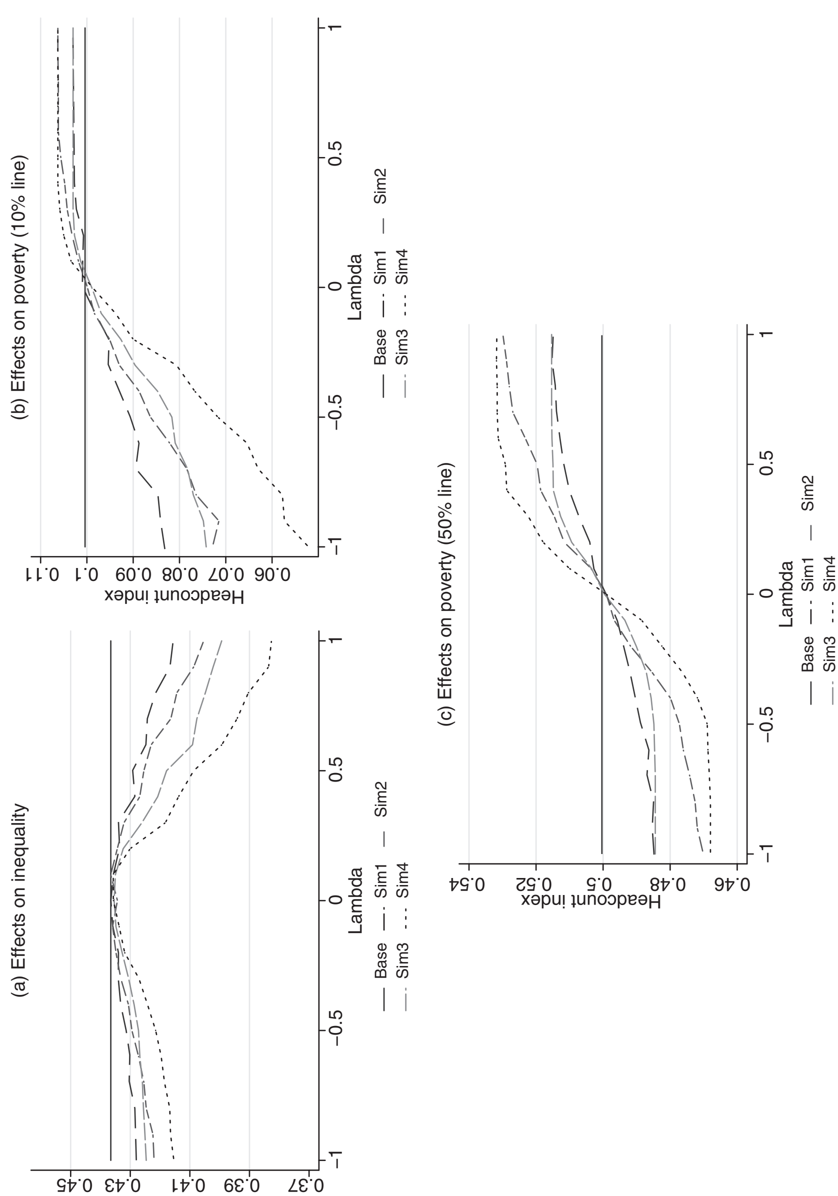

Figure 1. The Impact of Differential Mortality on the Distribution of Income: Some Illustrative Simulations

Source: Simulations by the authors.

by 1 percent. If λ is 0, i.e. there is no differential mortality, the Gini coefficient obviously corresponds to that of the baseline. If mortality is positively correlated with income, i.e. λ between 0 and 1, inequality tends to decrease. In both cases, negative and positive correlation between mortality and income, inequality decreases since, in each case, we "eliminate" individuals at the lower or upper end of the income distribution. By contrast, a scenario where middle class individuals faced higher mortality could lead to an increase in inequality. If we raise the death rate to 0.06 (Simulation 2) or reduce the error term (Simulation 3) or both (Simulation 4), we can state, as one would 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.

Figures 1b and 1c show that the poverty rate also reacts strongly to the extent of differential mortality. The assumption of a death rate of 0.03 (Simulations 1 and 3) and strong negative differential mortality reduces the poverty headcount index by roughly 2 percentage points, which corresponds to approximately 20 percent in the case of the lower poverty line. Again, the effect is greater if the death rate increases (Simulation 2), the error term is reduced (Simulation 3) or both (Simulation 4). For instance, in Simulation 4, the headcount index for the 10 percent poverty line decreases by some 50 percent. Obviously, for positive values of λ the headcount index is less affected with the lower poverty line than with the higher poverty line.

These simple simulations illustrate the potential and purely demographic effect of differential mortality on income distribution and especially its distinct effects on inequality and poverty measures. In what follows, we develop an analytical method to isolate these effects from other changes in the income distribution.

4. A General Method to Account for Differential Mortality in Welfare Comparisons Over Time

4.1. *Definitions and General Idea*

For each period t, individual welfare $w(y)$ is a non-decreasing function of income y. It is assumed that y is a continuous variable that may vary between 0 and $y_{\max}$, with a c.d.f. $F_t(y)$ and a d.f. $f_t(y) = dF_t(y)$. In the utilitarian tradition, a societal welfare index W can then be defined as

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

Likewise, a large class of monetary poverty indices corresponds to

(5) $$P(F_t) = \int_0^{z} p(y)dF_t(y)dy,$$

where z is the poverty line and p a non-increasing 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

(6) $$W\left(F^*_{t+1}\right) = \int_0^{y\max} w(y)\,dF^*_{t+1}(y)\,dy \quad \text{or}$$

(7) $$P\left(F^*_{t+1}\right) = \int_0^{z} p(y)\,dF^*_{t+1}(y)\,dy.$$

To be more precise, let us assume that we have some knowledge about the mortality process taking place between t and $t+1$. Let us further assume that individuals live in households and pool their resources. Thus individual income y stands for household income per capita or, more generally, household income per adult equivalent. Then, theoretically, the occurrence of individual deaths should have at least three kinds of effects on the distribution of income:

(1) A "direct" (arithmetical) individual effect: individuals who die leave the population and no longer contribute to societal (monetary) welfare or poverty.
(2) An "indirect" microeconomic effect on household income per surviving household member: survivors in the deceased's household may experience a decrease or increase in their income y, as the deceased's previous income contribution disappears from the household's income, the number of equivalent consumer units changes, and various labor supply and household composition adjustments occur.
(3) A "general equilibrium" or "external" macroeconomic impact on the overall income distribution caused by feedback effects from mortality and changes in the population age structure on the economy.

We will not consider the third, general equilibrium effect in the following. Hence, the construction of a counterfactual distribution of income entails looking first at the direct effect and then at the indirect effect. However, what is meant by "counterfactual" should be clarified first for both cases. Intuitively speaking, we seek to reconstruct the distribution of income as it would have been in $t+1$ if the observed deaths between t and $t+1$ had not occurred. This definition of counterfactual raises no particular problems when the mortality process can be assumed to be exogenous from the distribution of income itself. Think of a sudden epidemic originating outside the country or a natural disaster like an earthquake or a flood. Obviously, however, the exogeneity of mortality does not preclude the possibility of its correlation with income. Things become more complicated when the probability of dying is causally determined by contemporary individual income, the distribution of income within some reference group, or the overall distribution of income (see Deaton and Paxson, 2001). For instance, people whose income has fallen beneath a subsistence level (extreme poverty line) may be exposed to a probability of death that is close to one. Giving these people a "counterfactual income" beneath the subsistence level would be absurd if nobody can survive in this situation. We believe that a meaningful counterfactual distribution of income should always include the income-determined deaths or, put in another way, should only seek to discount the distribution of deaths exogenous to the final income distribution. In the rest of this paper, we 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 is worth noting. Assessing the impact of mortality between two dates is not the same thing as assessing the impact of changes in mortality.

In the first case, we need to deduce the effect of all deaths during the period, while in the second case, we need to deduce the effect of the difference between the initial and terminal pattern of deaths. We focus on the first case in the following section and then examine the second case.

4.2. *The Direct Impact of Mortality*

Let us first assume that the death of an individual has no external effects, either on other individuals such as household survivors and neighbors or on the population as a whole. We therefore seek to define a counterfactual for a purely arithmetical individual effect. Secondly, assume that mortality patterns between t and $t+1$ are described entirely by observable individual attributes x, which are either constant over time, such as gender and adult education levels, or vary with time, such as age, health and household composition. This makes the individual survival rate $s_{x,t}(x)$ independent of the distribution of attributes, i.e. the survival rate is independent of the population structure. Thirdly, assume that the income pattern specific to each attribute, i.e. the conditional density of income relating to the attributes, depends not on the distribution of attributes but on an "income schedule" that changes over time by means of redistribution policies and other changes in returns on the attributes, in keeping with the Oaxaca (1973) and DiNardo *et al.* (1996) decompositions. This means that we again assume that mortality has neither external effects nor "general equilibrium" effects. It 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 transient components of income.

The econometrician observes $f(y|t_y = t)$, that is the actual density of income for each t,

$$f\left(y, x|t_{y,x}=t+1\right)=f\left(y|x, t_y=t+1\right)dF\left(x|t_x=t+1\right). \tag{8}$$

Equation (8) shows that the bivariate distribution of income y and individual attributes x in $t+1$ can be written as the product of the income schedule conditional on x and the distribution of x in $t+1$. Alternatively, we can write equation (8) by using the distribution of x in t instead of that in $t+1$ and (i) multiplying it with survival rates $s(x)$ between t and $t+1$ conditional on x and (ii) dividing it by rates $\Psi(x)$ expressing the probability of being of characteristics x in t relative to $t+1$,

$$f\left(y, x|t_{y,x}=t+1\right)=f\left(y|x, t_y=t+1\right)\frac{s_{x,t}(x)}{\Psi_{x,t}(x)}dF\left(x|t_x=t\right). \tag{9}$$

Put differently, $\Psi_{x,t}(x)$ denotes changes in the population structure which are not due to mortality, but instead to births, migration, household composition and so on occurring over $[t;\ t+1]$.

We can then compute the counterfactual distribution of income due to deaths related to initial attributes simply by reweighting the observations with $s_{x,t}(x)$:

(10) $$f_t^*(y) = \int_{x\in\Omega_x} s_{x,t}(x) f\left(y|x, t_y = t\right) dF(x|t_x = t).$$

Semiparametric decompositions as proposed by DiNardo *et al.* (1996) take the study 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 associated with mortality) on income distribution:

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

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

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

(where $\Pr(t_x = t|x)$ can be estimated with a probit model), we can compute:[4]

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

So far, we have considered the direct impact of the *level* of individual mortality. Apart from introducing the indirect impact, the following section will also consider the impact of *changes* in mortality patterns.

4.3. *The Indirect Impact of Mortality*

When the income concept used is household income per capita or household income per adult equivalent, mortality obviously also has, as mentioned above, an indirect impact by affecting the distribution of income across the household survivors. Analysis of this indirect impact calls for the construction of a counterfactual income distribution that includes a counterfactual income pattern for survivors. We

[4] We can compute a similar counterfactual "backward," i.e. starting from the $t+1$ distribution of income: $g_{t+1}^*(y) = \int \frac{\Psi_{x,t}(x)}{s_{x,t}(x)} f(y|x, t_y = t+1) dF(x|t_x = t+1)$ and $g_{t+1}^{**}(y) = \int \Psi_{x,t}(x) f(y|x, t_y = t+1) \times dF(x|t_x = t+1)$. 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 schedule $f(y|x,\ t_y = t+1)$ and the initial distribution of attributes $dF(x|t_x = t)$. Then the double difference between $\left[g_{t+1}^{**}(y) - g_{t+1}^*(y)\right]$ and $\left[f_t^*(y) - f_t(y)\right]$ gives the mortality impact associated with the change in the income schedule from $f(y|x,\ t_y = t)$ to $f(y|x,\ t_y = t+1)$.

define $z \in \{0,1\}$ as a variable taking the value 1, if an individual under consideration has known a death in his/her household between t and $t-1$. Thus, the observed density of income in t becomes a weighted sum of conditional densities on z:

$$(14)\quad f_t(y) = \Pr(z=0|t_z=t)\, f\left(y|z=0, t_y=t\right) + \Pr(z=1|t_z=t)\, f\left(y|z=1, t_y=t\right).$$

We would like to design a counterfactual that can be written as:

$$(15)\quad f_t^{\#}(y) = \Pr(z=0)\, f\left(y|z=0, t_y=t\right) + \Pr(z=1)\, f_{z=0}\left(y|z=1, t_y=t\right).$$

This requires the estimation of the counterfactual density for survivors $f_{z=0}(y|z=1,\ t_y=t)$. This kind of estimation typically involves addressing the two main problems recurrently discussed by the recent econometric literature on handling treatment effects (see e.g. Heckman and Vytlacil, 2005). In our case, z can be seen as the treatment variable. The two problems are, first, the possible endogeneity of the treatment, and second, the heterogeneity of the responses to the treatment, i.e. the impact of the treatment on the distribution (or density function) of the outcome variable. Dealing with endogenous selection into treatment requires observation of some instrumental variable or making the assumption of conditional independence. In contrast, dealing with heterogenous responses to the treatment requires a quantile estimator. For instance, a quantile treatment instrumental variable estimator could be used (see, e.g. Abadie *et al.*, 1998). However, finding an instrument for the occurrence of a death within the household can prove difficult, even with some knowledge about the causes of death. A quantile treatment effect can be computed using matching estimators (see, e.g. Firpo, 2004), based on the assumption of conditional independence of the treatment with respect to the outcome when controlling for a set of observable variables x.[5] This latter conditional independence hypothesis is precisely the one that is used in this paper. In other words, we rely on the principle of the Oaxaca decompositions in assuming that mortality is exogenous from contemporary income (see above), conditionally on the set of attributes x. Moreover, we show that the estimation of the indirect impact of mortality does not require estimation of the counterfactual density of income. Indeed, when survivor status z is known for both periods, we can apply the DiNardo *et al.* (1996) reweighting technique to isolate the effects of *changes* in the "survival rate" between two periods. DiNardo *et al.* (1996) apply this technique in order to quantify the impact of a change in the unionization rate on the distribution of wages in the United States. Given that we have information on survivors in t and in $t+1$, that is, individuals having experienced a death within the household between $t-1$ and t and between t and $t+1$, a counterfactual for the impact of changes in mortality patterns can easily be constructed. Hence, we write:

$$(16)\quad \begin{aligned} f_t^{indir}(y) &= \iint f\left(y|x, z, t_y=t\right) dF\left(z|x, t_{z|x}=t+1\right) dF\left(x|t_x=t\right) \\ &= \iint \Psi_{z|x}(z,x)\, f\left(y|x, z, t_y=t\right) dF\left(z|x, t_x=t\right) dF\left(x|t_x=t\right), \end{aligned}$$

where

[5] $\Pr(z=1|y_{z=0}, y_{z=1}, x) = \Pr(z=1|x)$.

$$\Psi_{z|x}(z, x) = \frac{dF\left(z|x, t_{z|x} = t+1\right)}{dF\left(z|x, t_{z|x} = t\right)}$$
$$= z\frac{\Pr\left(z=1|x, t_{z|x} = t+1\right)}{\Pr\left(z=1|x, t_{z|x} = t\right)} + [1-z]\frac{\Pr\left(z=0|x, t_{z|x} = t+1\right)}{\Pr\left(z=0|x, t_{z|x} = t\right)}$$

can be estimated using a standard probit model such as:

$$\Pr\left(z=1|x, t_{z|x} = t\right) = \Phi(\beta_t' x).$$

We can then design a triple decomposition for the impact of changes in mortality patterns between $t-1$ and $t+1$.[6] First, we compute a counterfactual for the $t+1$ distribution of income isolating the direct arithmetic impact of changes in individual mortality patterns based on observable attributes:

$$(17) \qquad f_t^{**}(y) = \int_{x\in\Omega_x} \frac{s_{x,t}(x)}{s_{x,t-1}(x)} f\left(y|x, t_y = t\right) dF\left(x|t_x = t\right).$$

This latter counterfactual density is best understood in two steps. In a first step we compute $\frac{1}{s_{x,t-1}(x)} dF(x|t_x = t)$, which corresponds to the distribution of attributes x that would have prevailed in t in the absence of mortality between $t-1$ and t. Then, reweighing the distribution obtained in the first step by $s_{x,t}(x)$, i.e. computing $\frac{s_{x,t}(x)}{s_{x,t-1}(x)} dF(x|t_x = t)$ yields the distribution of attributes that would have prevailed in $t+1$ in the absence of mortality between $t-1$ and t (first step) and under the sole action of mortality between t and $t+1$ (second step). Both discounting for $[t-1; t]$ mortality patterns and accounting for $[t; t+1]$ mortality patterns results in accounting for the impact of changes in mortality on the distribution of attributes. In the absence of changes in mortality patterns between the two periods, i.e. $s_{x,t-1}(x) = s_{x,t}(x)$, one gets $f_t^{**}(y) = f_t(y)$. In other words, the difference $f_t^{**} - f_t$ would be zero, and hence would not contribute to the explanation of overall changes in the distribution of income $f_{t+1} - f_t$.

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

$$(18) \quad f_t^{\Delta}(y) = \iint \Psi_{z|x}(z, x) \frac{s_{x,t}(x)}{s_{x,t-1}(x)} \times f\left(y|x, z, t_y = t\right) dF\left(z|x, t_{z|x} = t\right) dF\left(x|t_x = t\right).$$

Here again, the weight $\Psi_{z|x}(z, x)$ is equal to one in the absence of changes in the survivors' status conditional distribution, i.e. when changes in mortality

[6]As above, decompositions are computed, in what follows, "forward," i.e. starting from the t distribution of attributes. Of course, they can alternatively be computed "backward," i.e. starting from the $t+1$ distribution. See also Footnote 3.

patterns (according to x) do not modify the distribution of attributes of survivors. The absence of changes in mortality is one special case here. Third, we compute a counterfactual for the $t + 1$ distribution of income discounting the effect of all changes in the distribution of observable attributes:

(19)

$$g_t^{\Delta}(y) = \iint \frac{\Psi_{x,t-1}(x)}{\Psi_{x,t}(x)} \Psi_{z|x}(z, x) \frac{s_{x,t}(x)}{s_{x,t-1}(x)} \times f\left(y| x, z, t_y = t\right) dF\left(z| x, t_{z|x} = t\right) dF\left(x| t_x = t\right).$$

The intuition is the same as for mortality. In a first step we discount for the overall distribution of attributes x in t. In a second step we account for the overall distribution of attributes in $t + 1$.

5. An Empirical Application to the Case of Indonesia

5.1. *Data and Economic Background*

To illustrate the methods proposed in Section 4, we use three waves of the Indonesian Family Life Survey (IFLS) conducted by RAND. The IFLS is an ongoing longitudinal socioeconomic and health survey. It is representative of 83 percent of the Indonesian population living in 13 of the nation's current 26 provinces. The first wave (IFLS1) was conducted in 1993 and covers 33,083 individuals living in 7,224 households. IFLS2 sought to re-interview the same respondents in 1997. Those who had moved were tracked to their new location and, where possible, interviewed there. A full 94.4 percent of IFLS1 households were located and re-interviewed, in that at least one person from the IFLS1 household was interviewed. This procedure added a total of 878 split-off households to the initial households. The entire IFLS2 cross-section comprises 33,945 individuals living in 7,619 households. The third wave, IFLS3, was conducted in 2000. It covered 6,800 IFLS1 households and 3,774 split-off households, totaling 43,649 individuals. In IFLS3, the re-contact rate was 95.3 percent of IFLS1 households. Hence, nearly 91 percent of IFLS1 households are complete panel households.[7] Table A1 in the Appendix presents some descriptive statistics of the full samples in 1993, 1997 and 2000. The 1997 and 2000 samples are cross-sections in that they include, in addition to the panel individuals, individuals born after 1993 or who joined a household in the initial sample or a split-off household for another reason.

We used the data to construct two longitudinal samples: 1993 to 1997 and 1997 to 2000. We included in each those individuals who were re-interviewed 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 than death and moved to provinces not covered by the survey.[8] The survey gives the exact date of the interviews and the month of death, such that a relatively detailed survival analysis can be performed. We counted 743 deaths from 1993 to 1997 and 558 deaths from 1997 to 2000 (see also Table A1).

[7]For details see Strauss *et al.* (2004).

[8]Or they migrated to provinces covered by the survey, but could not be located.

The IFLS contains detailed information on household expenditure. However, household incomes and especially individual incomes are not completely observed. We therefore use real household expenditure per capita as the welfare or income measure in the following. Expenditure is expressed in 1994 prices and adjusted by regional price deflators to the Jakarta price level.

Note that the economic crisis started to be felt in the South-East Asia region in April 1997, but that the major impact did not hit Indonesia until December 1997/January 1998, just after IFLS2 was conducted. The sustained crisis period continued in Indonesia for more than a year. Yet in 2000, when IFLS3 was conducted, the population had returned to roughly its pre-crisis standard of living, with some people even a little better off (Strauss *et al.*, 2002). When constructing the 1997 and 2000 expenditure distributions, we find precisely this dynamic, i.e. slightly lower poverty and inequality in 2000 compared to 1997. We find substantial poverty reduction in the pre-crisis period from 1993 to 1997. This is also consistent with other findings (e.g. Tjiptoherijanto and Remi, 2001) and gives a good explanation as to why Indonesian households—based on the former positive dynamic—recovered so quickly from the crisis.

However, public health expenditure fell significantly during the crisis. In addition, the 1997/98 drought, which was a consequence of *El Niño*, and some serious forest fires caused serious health problems and a sharp drop in food production in some regions. Rukumnuaykit (2003) shows that the drought and smoke pollution had significant adverse effects on infant mortality in rural areas. However, Strauss *et al.* (2002) find that adult body mass indices did not worsen and that the fraction of preschool-aged children with very low heights for their age and gender even fell over the 1997–2000 period.

5.2. *The Direct Impact of Mortality*

Using the methods outlined in Section 4.2, we construct counterfactuals showing the "direct impact" of mortality on the Indonesian distribution of log household income per capita for the years 1997 and 2000.

We start with the estimation of the $s_{x,t}(x)$ and $\Psi_{x,t}(x)$ weights for $t = 1993$ and $t = 1997$. For each gender, we estimate a probit model (weighted by cross-section sample weights) for survival from 1993 to 1997 and from 1997 to 2000 depending on a set of individual attributes and attributes related to the household to which the individual belongs, all observed in the initial year: a third degree polynomial for age, household size, dummies for the individual's and household head's level of education, the household head's gender, a third degree polynomial for the house-hold head's age, and a dummy for urban areas.[9] Table A2 (Appendix) shows the probit estimates of the $s_{x,t}(x)$ function, for both genders and both periods. We also estimate 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). Our estimates show that, over time, the sample population gets slightly older, slightly more educated, and lives more often in urban areas and in smaller households (compare also with Table A1). These

[9]We also tested duration models to estimate survival rates. However, this did not significantly change the results. We therefore retained the simple probit model.

probabilities reflect overall demographic changes including migration and educational developments occurring during both periods. They may also reflect a small sampling bias associated with the panel structure of the IFLS surveys (attrition). Combining the probability estimates for $s_{x,t}(x)$ and $\Psi_{x,t}(x)$, we can compute the counterfactual defined by equations (12) and (13).

We subsequently compute density estimates (Gaussian kernels of bandwidth 0.2) of f_{93}, f_{97} and f_{00} for the actual log income distributions. Figure 2a shows that the income distribution substantially improves from 1993 to 1997, with a large reduction in poverty and inequality. The vertical line corresponds to a constant per capita poverty line used throughout the analysis.[10]

In 2000, i.e. after the macroeconomic crisis, the income distribution merely resumed its 1997 form, as found by Strauss *et al.* (2002). Figure 2b shows the corresponding differences in the density distributions.

We then compute kernel estimates (weighted by cross-section sample weights) of f_{93}^{*} and f_{97}^{*} for the "direct mortality impact" counterfactual distributions. We also compute a $f_{93}^{(0)}$ $\left(\text{resp. } f_{97}^{(0)}\right)$ density estimated for the 1993 (resp. 1997) population from which individuals who "will die" between 1993 and 1997 (resp. 1997 and 2000) have been removed.[11] Figure 3a compares the two counterfactual impacts of individual deaths: $f_{93}^{(0)} - f_{93}$ (excluding dead individuals) and $f_{93}^{*} - f_{93}$ (1993 reweighted). Figure 3b does the same for the 1997–2000 period. The "excluding dead individuals" effects take into account individual mortality differentials associated with unobservable factors such as transient components of income. For 1993–97, the absence of a significant difference between this latter counterfactual and the first supports our choice to compute mortality impacts using reweighting techniques based on observables exogenous to income. For 1997–2000, this difference is higher in relative terms, although fairly limited in absolute terms, as can be seen by comparing the scale of Figures 3a and 3b. This contrast is in line with the already noted similarity of the 1997 and 2000 income distributions. In all cases, individual mortality directly contributes to a decrease in poverty, as argued by Kanbur and Mukherjee (2003). This finding could also be forecast from the positive correlation between initial income and survival probabilities, i.e. the extent of differential mortality with respect to income, which is presented for selected age groups in Figure 4. However, these counterfactual impacts are very small when compared to the magnitude of observed changes in the distribution from 1993 to 1997 (compare the scale of the vertical axis in Figures 2b and 3a). To see how the observable determinants of mortality are directly related to income, see also Table A5, which presents per capita income regressions using the same exogenous variables as the equations used to estimate the survival probabilities in Table A2.

Next, we compute kernel estimates of g_{93}^{*} and g_{97}^{*} for the DiNardo *et al.* (1996) counterfactual distributions with a "constant distribution of attributes." Bear in mind that these "all observable attributes" counterfactuals also include the

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

[11]Whenever we measure an impact using a difference in densities, we smooth this difference again by means of a Gaussian kernel of bandwidth 0.2.

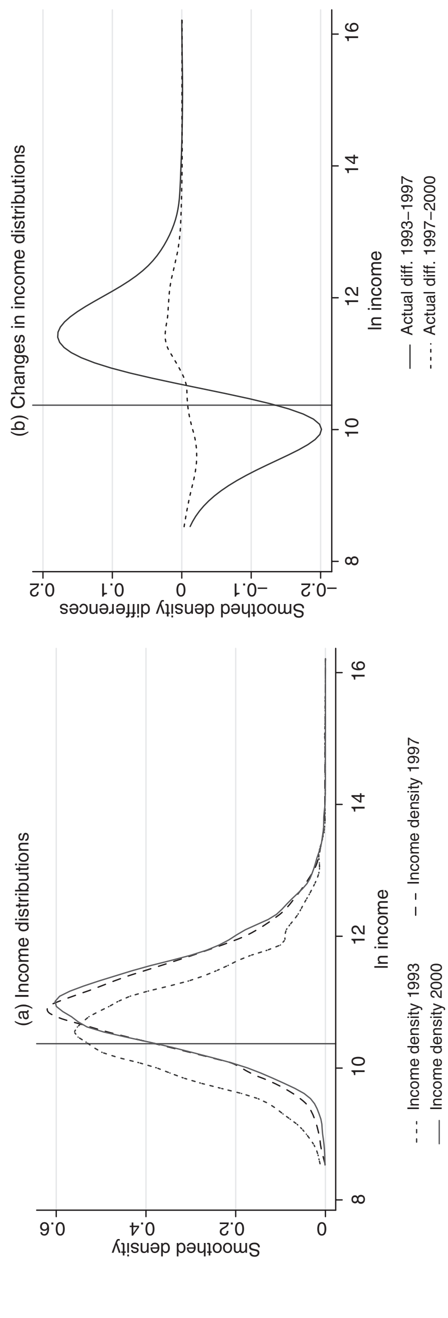

Figure 2. Income Per Capita (ln) Kernel Densities in 1993, 1997 and 2000

Source: IFIS1, IFLS2 and IFLS3; estimations by the authors.

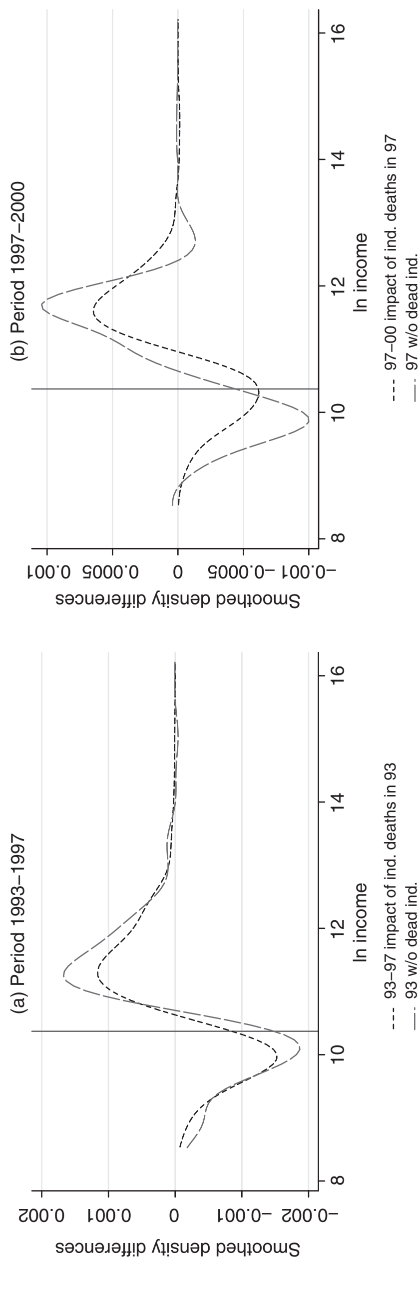

Figure 3. Smoothed Direct Impact of Mortality

Source: IFIS1, IFLS2 and IFLS3; estimations by the authors.

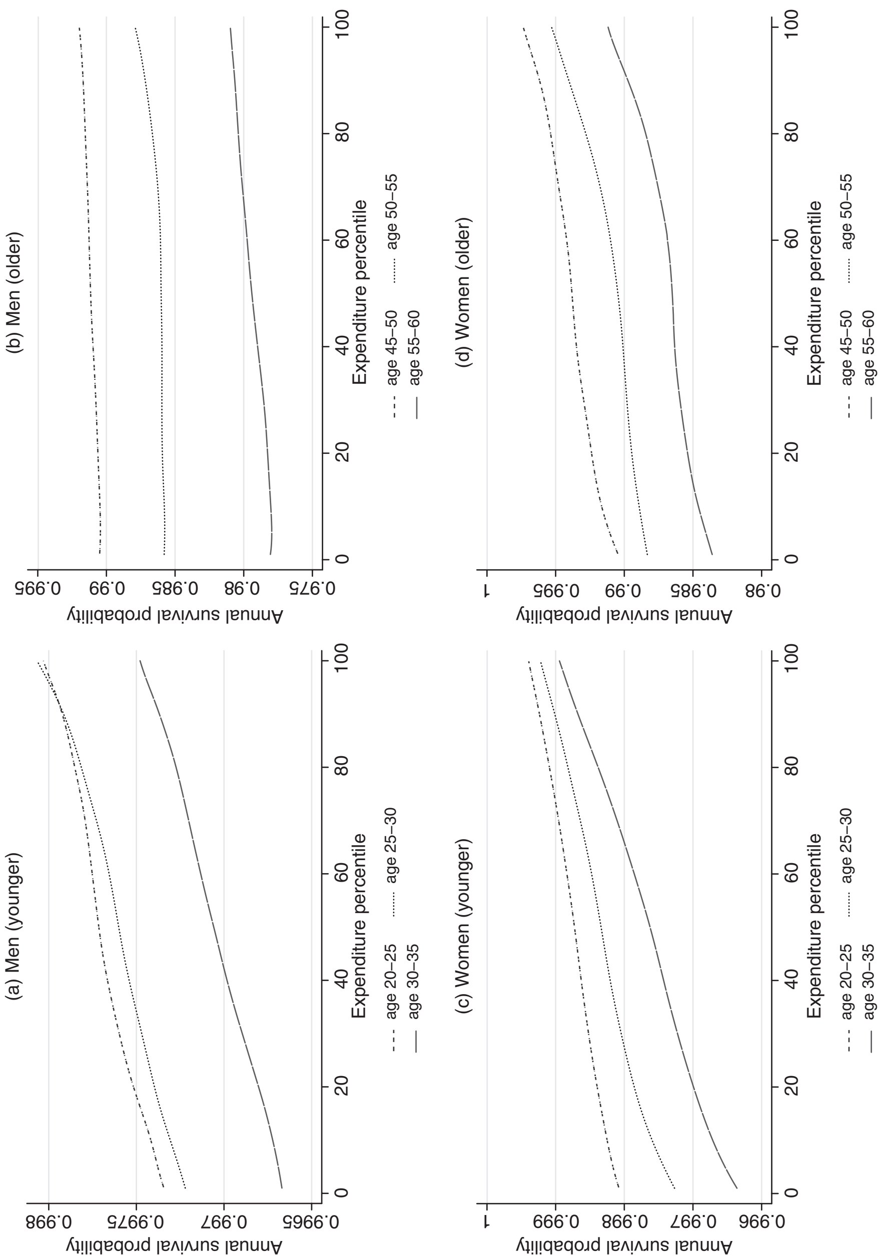

Figure 4. Smoothed Survival Probabilities by Per Capita Income Percentile for Men and Women and Selected Age Groups (means of predicted values for the 1993 sample using the model in Table A2, Column 1)

Source: IFIS1, and IFLS2; estimations by the authors.

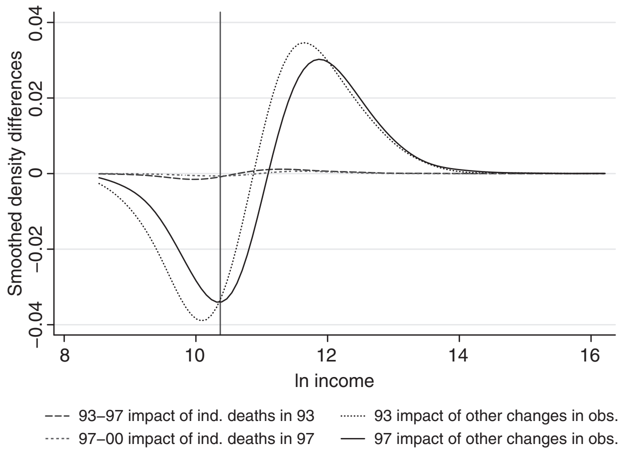

Figure 5. Smoothed Direct Impact of Mortality Compared to Impact of Changes in All Observable Attributes

Source: IFIS1, IFLS2 and IFLS3; estimations by the authors.

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 $f_{97}^{*} - f_{97}$ that we have just described. The comparison shows that individual mortality plays only a minor role in the distributional changes that can be imputed to demographic changes. The mortality impacts are ten times (in the case of 1993–97) to twenty times (1997–2000) lower than the overall demographic (including education) impacts. However, it is interesting to see that the effects of overall changes in the distribution of observable attributes correspond to the individual mortality effect, i.e. they are unambiguously poverty decreasing.

Lastly, 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 changes in the population structure do not explain very much of the change in the distribution of income per capita from 1993 to 1997. In contrast, for the 1997–2000 period, the distributional impact of demographic factors other than mortality is substantial. Reweighting indicates that demographic forces induced a shift towards the right of the income distribution, i.e. the poverty rate would have been slightly worse without the changes in the population structure. Overall demographic factors have in a way contributed to the observed recovery from the 1997–98 crisis, but do not explain to a large extent changes in inequality. Indeed, the income regressions presented in Table A5 confirm that smaller households with more educated members living in urban areas have higher real per capita expenditure. It is therefore not surprising to find that the main demographic changes we mentioned above lead to (counterfactual) poverty reduction.

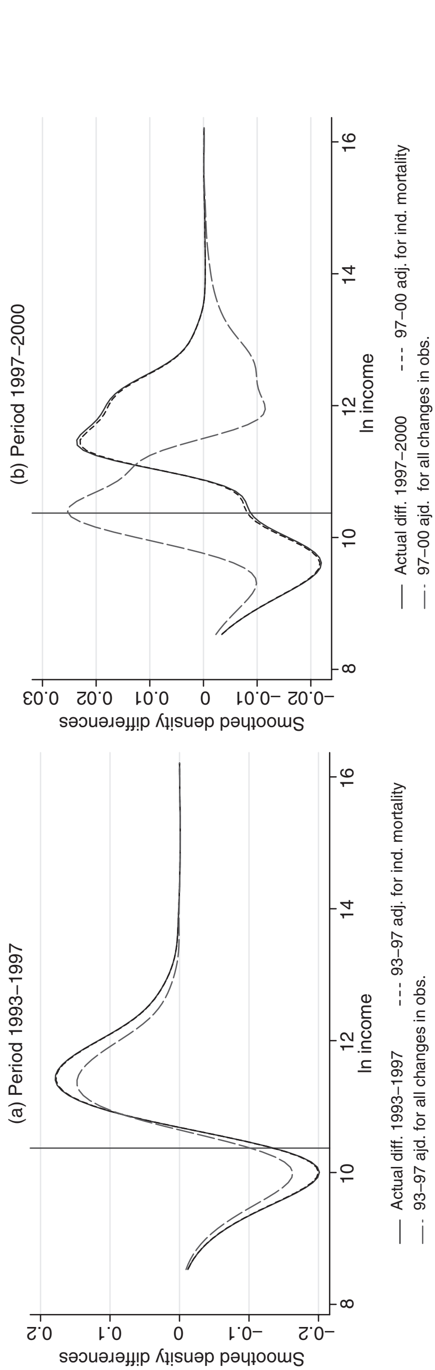

Figure 6. Smoothed Impact of Individual Mortality and Changes in All Observable Attributes Compared to Overall Change in Per Capita Income Distributions

Source: IFIS1, IFLS2 and IFLS3; estimations by the authors.

Before ending this section, it is worth noting that we checked for the path-dependency of the results. All results mentioned above are maintained, both in sign and magnitude, when we consider "backward" decompositions, i.e. counterfactuals based on the end-of-period distribution of income (1997 for 1993–97 and 2000 for 1997–2000).

5.3. *The Direct and Indirect Impact of Changes in Mortality Patterns*

We now incorporate the indirect impact of mortality on the income of household survivors using the methodology described in Section 4.3. We therefore add to our estimates of individual survival probabilities, estimates of the conditional individual probability (conditional on individual and household observables) of knowing a death in the household of origin between 1993 and 1997 and between 1997 and 2000, 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 genders and both periods. The results are presented in Table A4. All estimates show, as one can expect, that individuals in households with a female household head have more often suffered a death. As in the case of individual survival probabilities, education and household size differentials also play a role in explaining the probability of individuals to have known a death event in the household they belong to, even if measured at the end of the period, i.e. the terminal household size is positively correlated with the probability of having suffered a death (except for women between 1993 and 1997).

As the period 1993–97 has a year more than 1997–2000, the overall survival probability is higher in the latter period. When comparing income distributions, this difference in time range will generate the same effect as a decrease in mortality rates. The mortality gradients also change, as can be seen in Table A2. Likewise, as one can expect, the probability of individuals to have known a death event in the household they belong to also decreases from the first period to the second. Consequently, there are also some changes in the probability functions of being a "household survivor." That can be seen in Table A4. Using the ratio of individual survival probabilities estimated for both periods, we compute kernel estimates of the direct effect of *changes* in mortality patterns on the evolution of the income distribution $\left(f_{97}^{**} - f_{97}\right)$. In addition, based on the ratio of the "household survivor" probability functions, we compute the indirect effect of changes in mortality $\left(f_{97}^{indir} - f_{97}\right)$. The impact of these changes on mortality levels and gradients is assessed in Figure 7.

The direct effect of the change in mortality patterns $\left(f_{97}^{**} - f_{97}\right)$ is unambiguously poverty increasing. In other words, the direct effect of the downturn in mortality is to increase poverty. However, the effect is again rather small. Conversely, the indirect effect of the change in mortality patterns $\left(f_{97}^{indir} - f_{97}\right)$ is unambiguously poverty decreasing, although still very slight. It is as if individuals in households where at least one individual died were, controlling for all other observables, poorer than their "unaffected" counterparts. In other words, the indirect effect of the downturn in mortality is to reduce poverty. When the direct

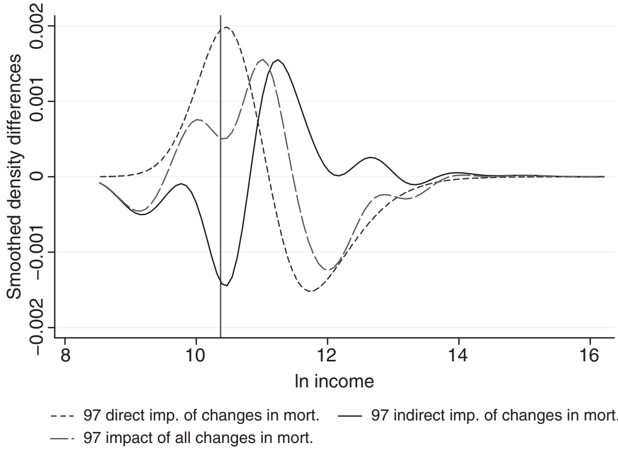

Figure 7. Smoothed Impact of Changes in Mortality Patterns Between 1993–97 and 1997–2000 on the 2000 Income Distribution

Source: IFIS1, IFLS2 and IFLS3; estimations by the authors.

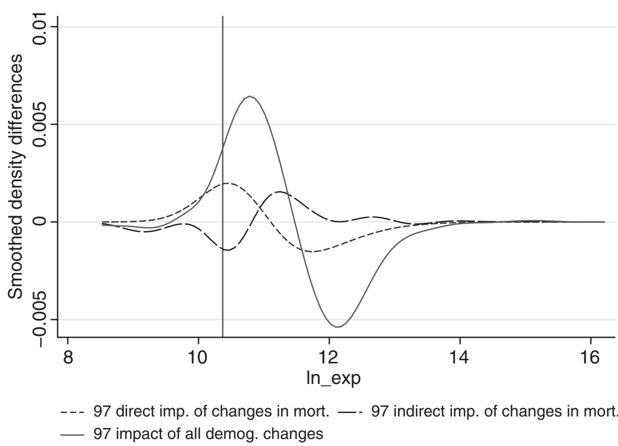

Figure 8. Smoothed Impact of Changes in Mortality Patterns Between 1993–97 and 1997–2000 on the 2000 Income Distribution Compared to the Impact of A Change in the Speed of Other Changes in the Population Structure

Source: IFIS1, IFLS2 and IFLS3; estimations by the authors.

and indirect impacts of the downturn in mortality are added together $\left(f_{97}^{\Delta} - f_{97}\right)$, the result is ambiguous: we find a slight decrease in the inequality of the income distribution rather than a change in poverty.

Lastly, we assess the impact of all changes in the population structure, including the survivor's status $\left(g_{97}^{\Delta} - g_{97}\right)$. Figure 8 shows that the effects of mortality, whether direct or indirect, are completely dominated by other demographic effects. Here again, demographic changes affect the distribution of income in the same way as the direct effect of mortality, but on a larger scale. The changes in the

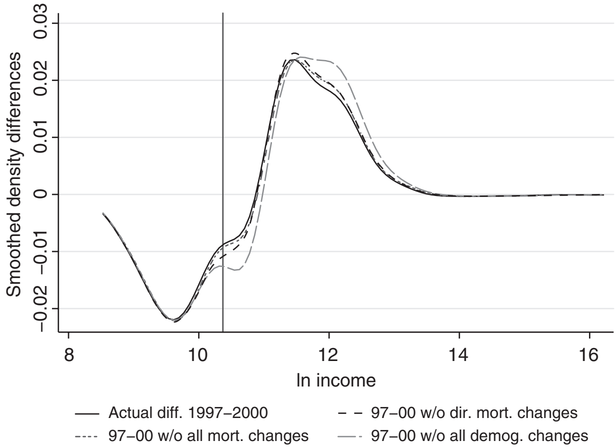

Figure 9. Smoothed Impact of Changes in Mortality Patterns and a Change in the Speed of Other Changes in the Population Structure Compared to the Overall Change in the Per Capita Income Distribution from 1997 to 2000

Source: IFIS1, IFLS2 and IFLS3; estimations by the authors.

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 change had been the same in the 1997–2000 period as in the 1993–97 period, which would in fact imply an acceleration of change given that the second period is shorter by one year, then the resulting distribution of income in 2000 would have presented slightly lower poverty and inequality.

However, Figure 9 shows that the overall impact of the change in the speed of changes in the demographic population structure is rather small, i.e. without any variation in the speed of demographic change, the observed change in the distribution of income between 1997 and 2000 would not have been very different. However, remember that the change in the population structure had, in contrast, a substantial impact, at least for the period 1997–2000 (compare Figure 6b).

Here again, we checked for the path-dependency of the results. All results hold both in sign and magnitude when "backward" decompositions are considered.

6. Conclusion

We have presented a general methodology designed to study the counterfactual effect of mortality and changes in mortality on income distribution. This methodology, inspired by the work of DiNardo *et al.* (1996), is based upon the semi-parametric reweighting of income distributions using functions of individual observable characteristics. Like Kanbur and Mukherjee (2003), we look at the direct arithmetic effect of individual deaths on poverty,[12] which is greatest when individual deaths are unevenly distributed across the income distribution. But we

[12]However, Kanbur and Mukherjee (2003) do not apply their approach empirically.

also correct for the indirect effect of an individual death on the income of survivors in the same household, which can be just as substantial. If mortality is negatively correlated with income, then, when mortality increases (resp. decreases) over time, the direct effect should be poverty decreasing (resp. increasing). Conversely, if mortality is negatively correlated with income and if a death in a household reduces household income, then, when mortality increases overtime (resp. decreases), the indirect effect should be poverty increasing (resp. decreasing). In our empirical part, we show that, in the case of Indonesia, the direct and indirect effects of a drop in mortality on the distribution of income indeed have opposite signs and are roughly the same in magnitude, such that they almost cancel out each other. We also show that the effect of other demographic changes, such as changes in the pattern of fertility, migration, and educational attainment, dominate the mortality effects regardless of whether they are direct or indirect. However, in the post-crisis period only, these changes also explain a substantial part of the overall change in the distribution of income. In the pre-crisis period other effects, e.g. institutional, seem to be more important. Moreover, *changes* in mortality patterns and *changes in the speed* of demographic change had no significant impact, but this is of course partly due to the fact that the observation horizon is rather small, and hence these changes have been rather small.

Appendix

- Descriptive statistics of the variables used: see Table A1.
- Estimated survival probabilities: see Table A2.
- Estimated "being present" probabilities: see Table A3.
- Estimated probabilities for having known a death in the household oforigin: see Table A4.
- Estimated coefficients for correlates of household income per capita: see Table A5.

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
Tracking status (shares)		*1993–97*	*1997–2000*
Survivors		0.969	0.977
Deceased		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
Tracking status (shares)		*1993–97*	*1997–2000*
Survivors		0.976	0.978
Deceased		0.024	0.022

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

TABLE A2

Estimated Probit Model for Survival Probabilities (marginal probabilities computed based on sample means)

Dependent Variable Survived (Binary)	1993–97 Coeff.	1993–97 Std. Err.	1997–2000 Coeff.	1997–2000 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	0.001
Age^2 HH-head	–5.92E-05**	2.95E-05	–9.12E-06	2.23E-05
Age^3 HH-head	4.68E-07**	2.14E-07	1.06E-07	1.41E-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	0.002
Junior high	0.008**	0.002	0.004	0.002
Senior high/coll./univ.	0.010***	0.002	0.008***	0.002
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***	0.002
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 based on sample means)

Dependent Variable Being Present (Binary)	1997 vs. 1993 Coeff.	1997 vs. 1993 Std. Err.	2000 vs. 1997 Coeff.	2000 vs. 1997 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 a Death Occurred During Past Period (marginal probabilities computed based on sample means)

Dependent Variable Survivor (Binary)	1993–97 Coeff.	1993–97 Std. Err.	1997–2000 Coeff.	1997–2000 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

Results of Household Income Per Capita Regressions (pooled sample 1993, 1997, 2000)

Dependent Variable in HH-Expend. Per Capita	*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.

References

Abadie, A., J. D. Angrist, and G. W. Imbens, "Instrumental Variables Estimation of Quantile Treatment Effects," NBER Working Paper t0229, Cambridge, 1998.

Brainerd, E. and M. V. Siegler, "The Economic Effects of the 1918 Influenza Epidemic," CEPR Discussion Paper No. 3791, CEPR, London, 2003.

Chu, C. Y. C. and H. Koo, "Intergenerational Income-Group Mobility and Differential Fertility," *American Economic Review*, 80, 1125–38, 1990.

Deaton, A. and C. Paxson, "Mortality, Education, Income and Inequality Among American cohorts," in D. Wise (ed.), *Themes in the Economics of Aging*, University of Chicago Press, Chicago, 2001.

DiNardo, J., N. M. Fortin, and T. Lemieux, "Labor Market Institutions and the Distribution of Wages, 1973–1992: A Semiparametric Approach," *Econometrica*, 64(5), 1001–44, 1996.

Firpo, S. "Efficient Semiparametric Estimation of Quantile Treatment Effects," Discussion Paper 04-01, University of British Colombia, Vancouver, 2004.

Foster J., J. Greer, and E. Thorbecke, "A Class of Decomposable Poverty Measures," *Econometrica*, 52, 761–76, 1984.

Heckman, J. J. and E. Vytlacil, "Structural Equations, Treatment Effects, and Econometric Policy Evaluation," *Econometrica*, 73, 669–738, 2005.

Kanbur, R. and D. Mukherjee, "Premature Mortality and Poverty Measurement," ISER Working Papers 2003-6, ISER, University of Essex, 2003.

Kitagawa, E. M. and P. M. Hauser, *Differential Mortality in the United States: A Study in Socioeconomic Epidemiology*, Harvard University Press, Cambridge, 1973.

Lam, D., "The Dynamics of Population Growth, Differential Fertility and Inequality," *American Economic Review*, 76, 1103–16, 1986.

Lantz, P. M., J. S. House, J. M. Lepkowski, D. R. Williams, R. P. Mero, and J. Chen, "Socioeconomic Factors, Health Behaviors, and Mortality," *Journal of the American Economic Medical Association*, 279, 1703–8, 1998.

Oaxaca R., "Male–Female Wage Differentials in Urban Labor Markets," *International Economic Review*, 14, 693–709, 1973.

Rukumnuaykit, P., "Crises and Child Health Outcomes: The Impacts of Economic and Drought/Smoke Crises on Infant Mortality and Birthweight in Indonesia," Mimeo, Michigan State University, Lansing, 2003.

Strauss, J., K. Beegle, A. Dwiyanto, Y. Herawati, D. Pattinasarany, E. Satriawan, B. Sikoki, B. Sukamdi, and F. Witoelar, "Indonesian Living Standards Three Years After the Crisis: Evidence from the Indonesia Family Life Survey, Executive Summary," RAND Corporation, 2002.

Strauss, J., K. Beegle, B. Sikoki, A. Dwiyanto, Y. Herwati, and F. Witoelar, "The Third Wave of the Indonesia Family Life Survey (IFLS3): Overview and Field Report," WR-144/1-NIA/NCHID, RAND Corporation, 2004.

Tjiptoherijanto, P. and S. S. Remi, "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, 2001.

Valkonen, T., "Social Inequalities in Mortality," in G. Caselli, J. Vallin and G. Wunsch (eds), *Demography: Analysis and Synthesis*, Paris, 51–67, 2002.