

Grimm, Michael; Cogneau, Denis

**Working Paper**

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

Diskussionsbeiträge, No. 122

**Provided in Cooperation with:**

Department of Economics, University of Goettingen

*Suggested Citation:* Grimm, Michael; Cogneau, Denis (2005) : The Measurement of Income Distribution Dynamics when Demographics are correlated with Income, Diskussionsbeiträge, No. 122, Georg-August-Universität Göttingen, Volkswirtschaftliches Seminar, Göttingen

This Version is available at:

<https://hdl.handle.net/10419/22147>

**Standard-Nutzungsbedingungen:**

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

**Terms of use:**

*Documents in EconStor may be saved and copied for your personal and scholarly purposes.*

*You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.*

*If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.*

# **Diskussionsbeiträge / Discussion Paper Series**

**Volkswirtschaftliches Seminar  
Universität Göttingen**

**/ Department of Economics  
/ University of Goettingen, Germany**

Beitrag Nr. 122 / Issued No. 122

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

**Michael Grimm**  
University of Göttingen  
**Denis Cogneau**  
DIAL-IRD, Paris  
**January 2005**

**ISSN 1611-7522**

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

Denis Cogneau\* and Michael Grimm†

First version: June 1, 2004 / This version: January 25, 2005

## Abstract

The purpose of our paper is to derive instructive analytics on 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 implement our approach empirically using three waves of the Indonesian Family Life Surveys (IFLS). 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. We also show that the effects of other demographic changes—such as 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 is substantial enough to explain a significant part of the change in income distribution, regardless of whether the pre-crisis period (1993-1997) or the post-crisis period (1997-2000) is considered.

**JEL Classification:** D10, D63, J17.

**Key words:** Differential Mortality, Income Distribution Dynamics, Welfare Comparisons, Decomposition.

---

\*DIAL-IRD, 4, rue d’Enghien, 75010 Paris, France, email: cogneau@dial.prd.fr.

† University of Göttingen, Department of Economics, Platz der Göttinger Sieben 3, 37073 Göttingen, Germany, DIW Berlin and DIAL Paris, email: mgrimm@uni-goettingen.de.

We sincerely thank the participants at the 2004 ESPE (Bergen) and IARIW (Cork) conferences. We are particularly grateful to Gordon Anderson for his comments and suggestions. Lastly, we would like to thank BPS Statistics Indonesia for giving us access to price data.

# 1 Introduction

Demographic behavior can significantly affect the distribution of income, when it is correlated with the income measure used. For instance, assume poor people are more likely to die than rich people, have more children than rich people and are more likely to migrate than rich people. All of these factors can have significant effects on income distribution dynamics. 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, 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 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 a function 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 semiparametric procedure we propose 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 following section, we discuss the welfare implications of differentials in demographic variables and especially differential mortality and present a brief overview 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

Variations in population size pose a well-known problem in welfare comparisons over space and time. Dasgupta, Sen and Starrett (1973) raised this problem in their note on Atkinson’s seminal paper on the measurement of inequality (Atkinson, 1970). It also appears in literature on the general form of social welfare functions (see, for example, Blackorby and Donaldson, 1984;

Blackorby, Bossert and Donaldson, 1995). Two aspects are important here. First, which aspects of personal well-being we include in the individual welfare function, i.e. should the length of life matter. Second, whether a social welfare function should take into account the number of members in the society at a given point in time.

The standard welfaristic approach usually disregards non-material sources of personal well-being and has a strong focus on annual income flows, at least in the empirical literature. Under these assumptions, interpersonal utility comparisons are not affected by the fact that two individuals have different life expectancies. In other words, two people 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 their life expectancy differences. Anderson (2004) recognizes these shortcomings in the standard welfaristic approach and suggests a framework that incorporates life expectancy into the calculation of economic welfare comparisons. He implements his approach at country level and compares how GDP performs over time with and without accounting for changes in life expectancy. In the case of Africa, he finds a substantial downturn in welfare over time when life expectancy is included, which is not the case with usual GDP comparisons.

In contrast to the welfaristic approach, Sen's capability approach has a much wider focus and is much more flexible from this point of view (Sen, 1985). This approach can easily be defined in such a way that factors or 'functionings' allowing for a more or less long life are explicitly taken into account, by assuming that health or a certain length of life can be produced in a complementary manner by means of commodities  $q$ , and personal characteristics and societal and environmental circumstances  $z$ . Therefore, if  $q$  and  $z$  are conducive to health, they will give rise to longer life and hence extend the range of 'capabilities'.

Turning now to the second point, the classical utilitarian (or *Benthamite*) social welfare function is regarded as 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 derived by individual  $i$  from  $x_i$ . So, clearly, the number of individuals in society,  $N$ , can be seen here as a source of social welfare. Yet in most cases, we think of a constant population when invoking such a utility function or we simply use it in per capita terms ( $W/N$ ) and sidestep the issue. The implicit ethical judgment, then, is that we are 'neutral' to the population. At the same time, the focus on per capita welfare means that we are indifferent to the unborn and are even biased towards keeping population growth down if it affects per capita welfare adversely.

Empirical studies on the dynamics of inequality and poverty generally avoid properly addressing this issue by implicitly assuming a constant population. They usually provide a kind of 'snapshot measure' of economic well-being. In other words, we consider indicators such as per capita GDP, the Human Development Index, the poverty headcount 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.

When considering a single country, three demographic forces drive variations in population size over time: fertility, mortality and migration. If these

forces are correlated with the welfare measure used, welfare comparisons may become complex and sometimes ambiguous. For instance, if mortality is negatively correlated with income, which indeed seems to be the case in both developing and developed countries,<sup>1</sup> standard poverty measures such as the FGT family headcount index (Foster, Greer and Thorbecke, 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 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.<sup>2</sup> 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 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.<sup>3</sup> 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.

Kanbur and Mukherjee (2003) propose calculating FGT-poverty measures based on an individual’s lifetime income curve. They define a normative lifetime to account for premature mortality among the poor, which has a positive effect on the poverty measure. However, there are two crucial issues in their procedure. First, the choice of the normative lifetime, which can influence 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 assuming 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 he or she is replaced by exactly one individual. The critical assumption concerning the hypothetical income to be imputed raises the general question of which ‘value’ we might want to place on a lost life. Even if we exclude issues of personal pain and loss, the purely material loss can only be calculated arbitrarily.

In the following, we suggest some general methods to account for differentials in demographic patterns, especially mortality, when making poverty comparisons over time. We do not address the issue of giving a value to a lost life. We first consider solely what we call the ‘direct effect’ or ‘pure demographic effect’. Then we develop in turn measures to take into account

---

<sup>1</sup>For empirical evidence, see Kitagawa and Hauser (1973), Deaton and Paxson (2001) and Lantz, House and Lepkowski *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.

<sup>2</sup>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 US states in the 1920s.

<sup>3</sup>See, on this point, the analyses and discussions in Lam (1986) and Chu and Koo (1990).

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 anymore and, second, because the death might have changed the 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  $t$ , welfare indicator  $y$  is defined for a population of individuals. It is assumed that  $y$  is a continuous variable that 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. Likewise, a large class of monetary poverty indices 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

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

To be more precise, 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 effect: people who die leave the population and no longer contribute to monetary welfare or poverty;
2. An indirect microeconomic effect on household income: survivors in the deceased's household may experience a decrease or increase in  $y$ , 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.

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.

Last but not least, it is worth noting another important aspect regarding the construction of counterfactuals. 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.

### 3.1 The direct arithmetical impact of individual mortality

Let us first assume that individual deaths have 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 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  $t$

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

where  $\Psi_{x,t}(x)$  denotes population structure changes 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  $s_{x,t}(x)$ :

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

Semiparametric decompositions as proposed by DiNardo *et al.* (1996) 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:

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

Using then DiNardo *et al.* (1996) and the reweighting 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 can compute:<sup>4</sup>

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

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) calls solely for an additional preliminary reweighting of the  $t$  income distribution with 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 will come back to this latter decomposition when taking into consideration the indirect impact of changes in the distribution of household survivors.

---

<sup>4</sup>We can 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 associated with the change in income schedule from  $f(y|x, t_y = t)$  to  $f(y|x, t_y = t + 1)$ .

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

When the income concept used is household income per capita or per adult equivalent unit, mortality obviously has, as mentioned above, an indirect impact on 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. Let  $z \in \{0, 1\}$  be a variable indicating whether somebody has suffered 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 that can be written as

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

This requires the estimation of the counterfactual density for survivors  $f_{z=0}(y|z = 1, t_y = t + 1)$ . It is very hard to calculate such a counterfactual. Quantile treatment IV estimators can be used (Abadie, Angrist and Imbens, 1998) if an instrument is available for the occurrence of a death within the household (some knowledge about the causes of death could prove useful in this respect). If we assume conditional independence on a set of attributes  $x$ ,<sup>5</sup> quantile treatment effects can 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 may prove easier to construct. 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’. Hence, we write

$$f_{t+1}^{indir}(y) = \int \int f(y|x, z, t_y = t + 1)dF(z|x, t_{z|x} = t)dF(x|t_x = t + 1) = \int \int \Psi_{z|x}(z, x)f(y|x, z, t_y = t + 1)dF(z|x, t_x = t + 1)dF(x|t_x = t + 1), \quad (12)$$

where

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

can be estimated using a standard probit model such as

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

---

<sup>5</sup> $\Pr(z = 1|y_{z=0}, y_{z=1}, x) = \Pr(z = 1|x)$ .

We can 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:

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

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

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

## 4 An empirical application to the case of Indonesia

### 4.1 Data and economic background

To illustrate the methods proposed in Section 3, we use three waves of the Indonesian Family Life Survey (IFLS) conducted by RAND, UCLA and the University of Indonesia's Demographic Institute. The IFLS is an ongoing longitudinal socioeconomic and health survey. It is representative of 83% 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% 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% of IFLS1 households. Hence, nearly 91% of IFLS1 households are complete panel households.<sup>6</sup> 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.

---

<sup>6</sup>For details see Strauss, Beegle, Sikoki *et al.* (2004).

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 and moved to provinces not covered by the survey.<sup>7</sup> 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).

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. Intertemporal price variations are taken into account by household-specific price deflators, i.e. using disaggregated price deflators and each household’s observed budget shares as weights.

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 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, Beegle, Dwijanto *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.

## 4.2 Some illustrative simulations

Before we empirically apply the approach suggested in Section 3, it appears useful to give an approximate idea 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.

---

<sup>7</sup>Or they migrated to provinces covered by the survey, but could not be located.

We then analyze different scenarios where the selection of death events is correlated with income, but disrupted by some unobserved heterogeneity  $\gamma_i$ . 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 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  $\lambda$ ,  $\mu_\gamma$  and  $\sigma_\gamma^2$ . 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. We can hence 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

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

We examine a total of 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 derived from a log-normal distribution where the mean and variance correspond to those observed in our sample drawn from IFLS1. 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% and the other considers the first 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 1a shows that, for a death rate of 3% and a relatively sizeable unobserved heterogeneity component, the Gini coefficient decreases by roughly one percentage point if we reduce  $\lambda$  from 0 to -1. A value of -1 for  $\lambda$  implies that a 1% increase in  $y$  reduces the risk of death by 1%. If  $\lambda$  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.  $\lambda$  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  $\lambda$ . 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 (Simulation 1 and 3) and strong negative differential mortality reduces the poverty headcount index by roughly 2 percentage points, which corresponds to approximately 20% 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% poverty line decreases by some 50%. Obviously, for positive values of  $\lambda$  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 the next section, we endeavor to isolate this effect from the overall changes in inequality observed in Indonesia over the 1993-1997 and 1997-2000 periods. The empirical application obviously poses a lot of additional problems, such as the fact that individuals are grouped in households taking joint decisions on labor supply, household production and expenditure.

### 4.3 Results for Indonesia for the period 1993-2000

#### 4.3.1 Estimates of the direct arithmetical impact of mortality

In the following, we construct ‘direct mortality impact’ counterfactuals of the Indonesian distribution of log income per capita for 1997 and 2000 using the methods outlined in Section 3.1.

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 observed in the initial year: a third degree polynomial for age, household size, dummies for the individual and household head’s level of education, the household head’s gender, a third degree polynomial for the household head’s age, and a dummy for urban areas.<sup>8</sup> 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 (see also Table A1). These probabilities reflect overall demographic

<sup>8</sup>We also tested duration models to estimate survival rates. However, this did not significantly change the results. We therefore retained the simple probit model.

changes including migration and educational developments occurring during both periods. They may also reflect a sampling bias associated with the panel structure of the IFLS surveys (attrition).

We subsequently compute density estimates (Gaussian kernels with 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 poverty line used throughout the analysis.<sup>9</sup> 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.

[please insert Figure 2 about here]

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  $f_{93}^{(0)}$  (resp.  $f_{97}^{(0)}$ ) density estimated for the 1993 (resp. 1997) population from which (future) dead individuals between 1993 and 1997 (resp. 1997 and 2000) have been removed.<sup>10</sup> 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 both periods, the absence of significant differences between these latter counterfactuals and the first two supports our choice to compute mortality impacts using reweighting techniques based on observables exogenous to income. 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 2a 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.

[please insert Figures 3 and 4 about here]

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 impact of individual mortality on the distribution of observable attributes

<sup>9</sup>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.

<sup>10</sup>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.

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 in the distributional changes that can be imputed to demographic changes. The mortality impacts are ten times (in the case of 1993-1997) 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.

[please insert Figure 5 about here]

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 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 changes other than mortality is roughly the same as that for the observed distributional changes. Reweighting indicates that demographic changes induced a shift towards the right of the income distribution, i.e. the poverty rate would have been slightly worse without such changes than that observed in 2000. Overall demographic changes have in a way contributed to the observed recovery from the 1997/98 crisis, but do not explain many of the 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 the main demographic changes we mention above leading 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 incorporate the indirect impact of mortality on the income of household survivors using the methodology described in Section 3.2. We therefore add to our estimates of individual survival probabilities, estimates of the conditional individual probability (conditional on individual and household observables) of suffering a death in the household of origin from 1993 to 1997 or from 1997 to 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 genders and both periods. The results are presented in Table A4. All estimates show that households headed by a woman have more often suffered a death, which can easily be understood. As in the case of individual survival probabilities, education and household size differentials also play a role in explaining the probability of death events, 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.

As the period 1993-1997 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 expected, the occurrence of death events in households also decreases from one period to the next. Consequently, 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 probability estimates measured for both periods, we compute kernel estimates of the direct effect of changes in mortality patterns on the evolution of the income distribution ( $f_{97}^{**} - f_{97}$ ). Based on the ratio of the household survivor probability functions, we compute the indirect effect of changes in mortality ( $f_{97}^{indir} - f_{97}$ ). The impact of these 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 is again very slight. Or the direct effect of the downturn in mortality is to increase monetary poverty.

Conversely, the indirect effect of the change in mortality patterns ( $f_{97}^{indir} - f_{97}$ ) is unambiguously poverty decreasing, although still very slight. It is as if households of ‘survivors’, controlling for all other observables, were poorer than their ‘unaffected’ counterparts. Or the indirect effect of the downturn in mortality is to reduce monetary poverty.

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

Lastly, we assess the impact of all changes in the population structure, including the survivor’s status ( $g_{97}^{\Delta} - g_{97}$ ). 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 arithmetic effect of mortality, but on a larger scale. The changes in 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 been the same in the 1997-2000 period as in the 1993-1997 period, which would imply an acceleration of changes given the difference in time range, then the resulting distribution of income in 2000 would have presented slightly 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 minor, i.e. without any variation in the population

structure, the observed change from 1997 to 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 effect of mortality and changes in mortality on income distribution. This methodology is inspired by the work of DiNardo *et al.* (1996). It is based on the non-parametric reweighting of income distributions using functions of individual observable attributes. Like Kanbur and Mukherjee (2003), we look at the direct arithmetic effect of individual deaths on poverty changes,<sup>11</sup> which is greatest when individual deaths are unevenly distributed across the income distribution. But we 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 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 reduces 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 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, we find that none of these demographic changes is substantial enough to explain a significant part of the changes in income distribution, whether in the pre-crisis period (1993-1997) or the post-crisis period (1997-2000).

## Appendix

### Descriptive statistics for 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]

---

<sup>11</sup>However, Kanbur and Mukherjee (2003) do not apply their approach empirically.

## **Estimated equations for living in a household in which no death has occurred in the 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. (2004), Life Expectancy and Economic Welfare: The Example of Africa in the 1990's. *Review of Income and Wealth*, forthcoming.
- 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. Siegler (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 (2001), Mortality, education, income and inequality among American cohorts. In D. Wise (ed.), *Themes in the Economics of Aging*, Chicago: University of Chicago Press.
- 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, Discussion Paper 04-01, University of British Columbia, Vancouver.
- Foster J., J. Greer and E. Thorbecke (1984), A Class of Decomposable Poverty Measures. *Econometrica*, 52: 761-776.
- 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.
- 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, 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 for the variables used

	1993	1997	2000
<b>BOYS/MEN</b>			
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>			
Survivors		0.969	0.977
Deaths		0.031	0.023
<b>GIRLS/WOMEN</b>			
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>			
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 based on sample means)

<i>Dependent variable</i>	1993—1997		1997—2000	
<i>Survived (binary)</i>	Coeff.	Std. Err.	Coeff.	Std. Err.
BOYS/MEN				
Age	4.35E-04	0.001	0.001	4.40E-04
Age <sup>2</sup>	-1.68E-05	1.98E-05	-2.30E-05	** 1.14E-05
Age <sup>3</sup>	-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	*	1.49E-04	0.001
Age <sup>2</sup> HH-head	-5.92E-05	** 2.95E-05	-9.12E-06	2.23E-05
Age <sup>3</sup> 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 <sup>2</sup>	0.192		0.196	
GIRLS/WOMEN				
Age	8.02E-05	4.03E-04	0.001	*** 2.82E-04
Age <sup>2</sup>	-9.20E-06	1.09E-05	-2.54E-05	*** 7.00E-06
Age <sup>3</sup>	-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 <sup>2</sup> HH-head	-1.64E-05	1.79E-05	-2.78E-06	2.04E-05
Age <sup>3</sup> 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 <sup>2</sup>	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)

<i>Dependent variable being present (binary)</i>	1993/1997		1997/2000	
	Coeff.	Std. Err.	Coeff.	Std. Err.
<b>BOYS/MEN</b>				
Age	-1.08E-04	0.002	-0.011 ***	0.001
Age <sup>2</sup>	2.11E-05	4.69E-05	3.01E-04 ***	4.09E-05
Age <sup>3</sup>	-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 <sup>2</sup> HH-head	-9.01E-05	9.93E-05	1.30E-05	7.43E-05
Age <sup>3</sup> 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 <sup>2</sup>	0.011		0.009	
<b>GIRLS/WOMEN</b>				
Age	-0.004 ***	0.001	-0.014 ***	0.001
Age <sup>2</sup>	1.03E-04 ***	3.90E-05	4.02E-04 ***	3.91E-05
Age <sup>3</sup>	-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 <sup>2</sup> HH-head	-1.16E-04	9.66E-05	5.52E-05	7.30E-05
Age <sup>3</sup> 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 <sup>2</sup>	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 based on 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 <sup>2</sup>	-1.07E-05	3.63E-05	-2.85E-05	3.30E-05
Age <sup>3</sup>	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 <sup>2</sup> HH-head	1.68E-05	5.85E-05	2.24E-05	5.36E-05
Age <sup>3</sup> 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 <sup>2</sup>	0.029		0.021	
GIRLS/WOMEN				
Age	-0.001	0.001	0.001	0.001
Age <sup>2</sup>	2.16E-05	2.70E-05	-2.01E-05	2.21E-05
Age <sup>3</sup>	-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 <sup>2</sup> HH-head	1.94E-05	4.46E-05	-1.20E-04 ***	4.43E-05
Age <sup>3</sup> 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 <sup>2</sup>	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)

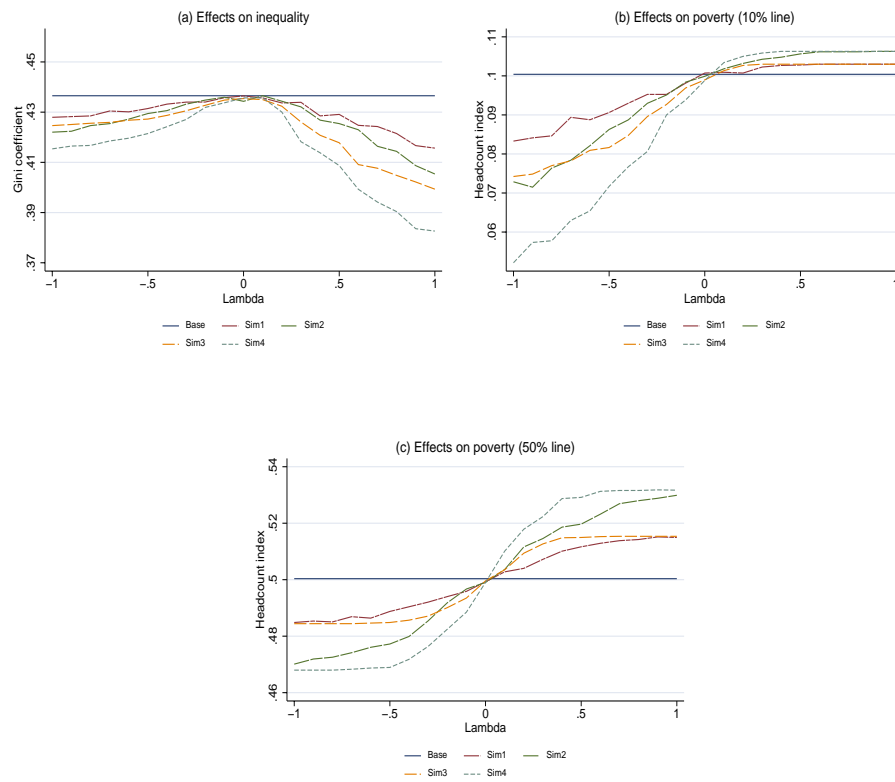
<i>Dependent variable</i> <i>ln HH-expend. per capita</i>	Boys/Men			Girls/Women		
	Coeff.		Std. Err.	Coeff.		Std. Err.
Age	-0.006	***	0.001	-0.011	***	0.001
Age <sup>2</sup>	1.45E-04	***	4.15E-05	3.22E-04	***	3.89E-05
Age <sup>3</sup>	-9.58E-07	***	3.35E-07	-2.40E-06	***	3.05E-07
Education						
No 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 <sup>2</sup> HH-head	-4.04E-04	**	1.90E-04	-3.31E-04	*	1.82E-04
Age <sup>3</sup> 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 <sup>2</sup>	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.

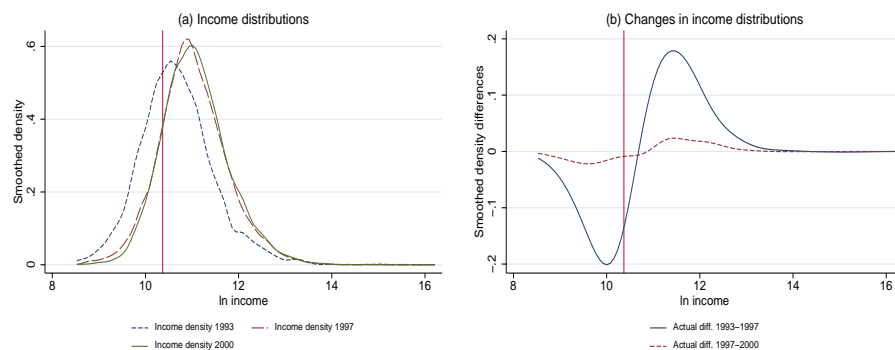
## Figures

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



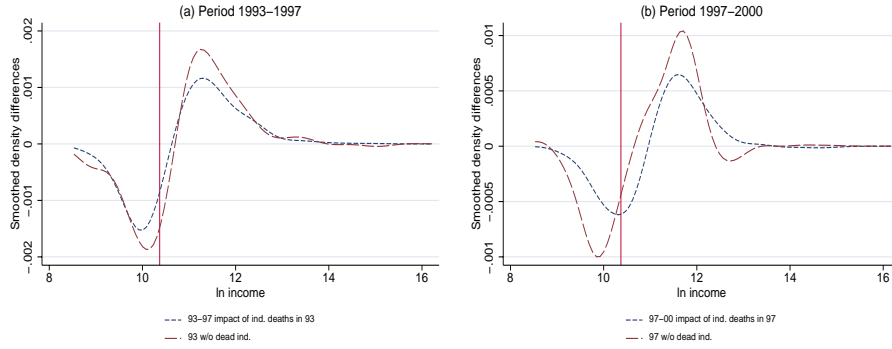
Source : Simulations by the authors.

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



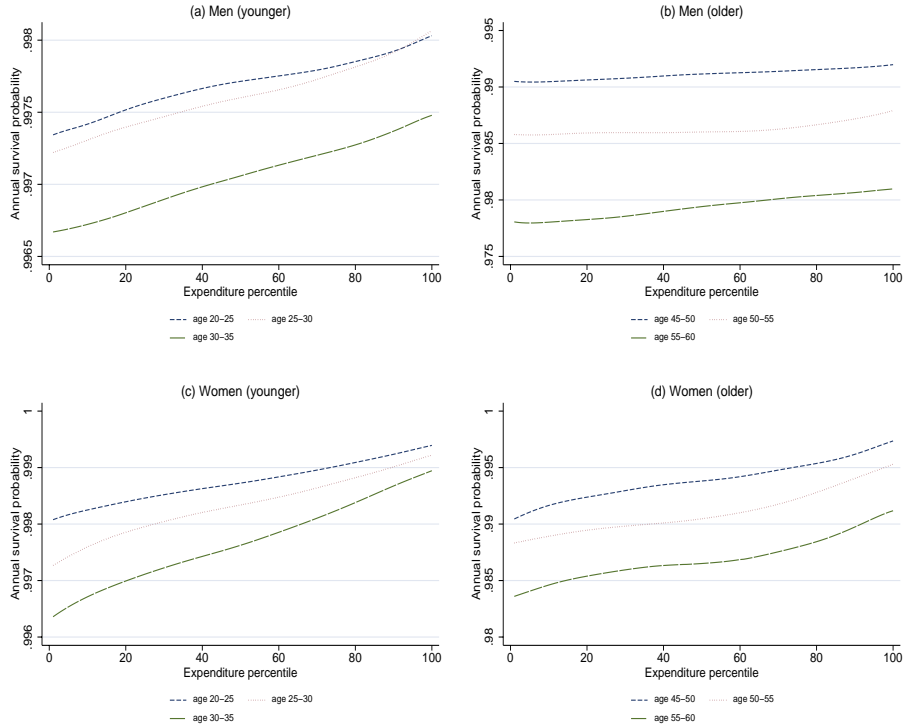
Source : IFLS1, IFLS2 and IFLS3; estimations by the authors.

Figure 3  
Smoothed impact of individual deaths



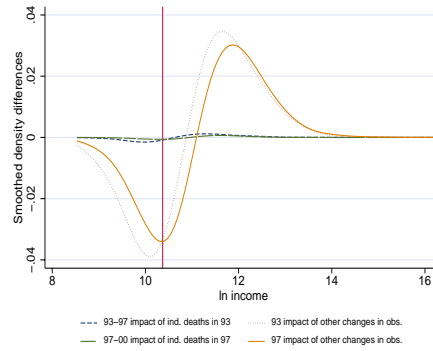
Source : IFLS1, IFLS2 and IFLS3; estimations by the authors.

Figure 4  
Smoothed survival probabilities by 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, col. 1)



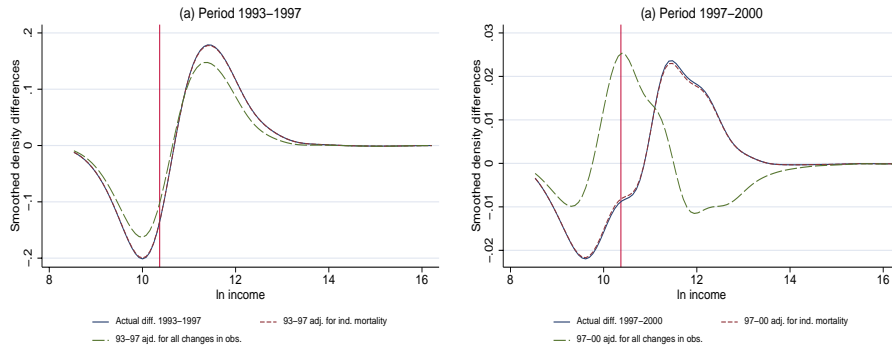
Source : IFLS1 and IFLS2; estimations by the authors.

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



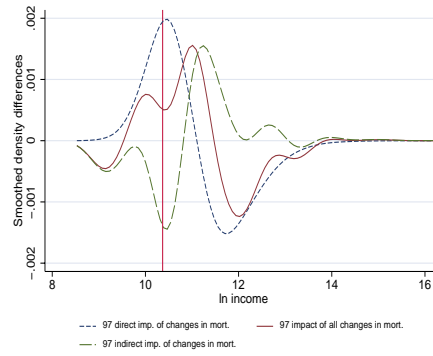
Source : IFLS1, IFLS2 and IFLS3; estimations by the authors.

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



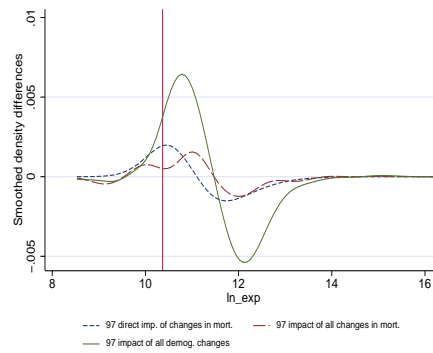
Source : IFLS1, IFLS2 and IFLS3; estimations by the authors.

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



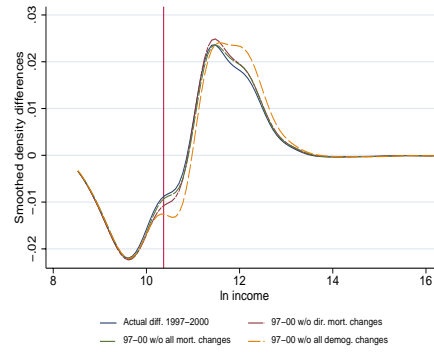
Source : IFLS1, IFLS2 and IFLS3; estimations by the authors.

Figure 8  
Smoothed impact of changes in mortality patterns  
between 1993/1997 and 1997/2000 on the 2000 income distribution  
compared to the impact of 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  
compared to the overall change in the per capita income distribution from 1997 to 2000



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

## Recent Discussion Papers

98. Gerloff, Axel: Stabilization during the Early Years of Transition - Some Stylized Facts. March 1999
99. Ahrens, Joachim: Governance, Conditionality and the Transformation of Post-Socialist Countries. April 1999
100. Mohsen, Fadi: Technischer Fortschritt und Humankapitalbildung in der Neuen Wachstumstheorie. September 1999
101. Sieg, Gernot: A political business cycle with boundedly rational agents. March 2000
102. Jarchow, Hans-Joachim: Geldpolitik bei extrapolativen, semirationalen und rationalen Inflationserwartungen. June 2000
103. Georgopoulos, Antonios / Salavrakos, Ioannis-Dionysios: Griechische Joint-Ventures in Ost-Europa - Ein erfolversprechendes Kooperationsmodell? November 2000
104. Haufler, Andreas / Schjelderup, Guttorm / Stähler, Frank: Commodity Taxation and International Trade in Imperfect Markets. January 2001
105. Barbaro, Salvatore: Gibt es eine Umverteilung von den *Armen* zu den *Reichen* durch die öffentliche Hochschulfinanzierung? Tragen Akademiker die Kosten ihres Studiums? Einige methodologische Anmerkungen zum Forschungsstand über die Verteilungswirkungen der öffentlichen Hochschulfinanzierung in der Bundesrepublik Deutschland. May 2001
106. Krieger, Tim: Intergenerational Redistribution and Labor Mobility: A Survey. May 2001
107. Kalbitzer, Ute: Das Schweigen der Ökonomik. Wissenschaftliche Politikberatung als wirtschaftspolitischer Diskurs. May 2001
108. Haufler, Andreas / Pflüger, Michael: International Commodity Taxation under Monopolistic Competition, June 2001
109. Südekum, Jens: Home Goods and Regional Price Indices: A Perspective from New Economic Geography, July 2001
110. Onozaki, Tamotsu / Sieg, Gernot / Yokoo, Masanori: Stability, Chaos and Multiple Attractors: A Single Agent Makes a Difference, November 2001
111. Rühmann, Peter / Südekum, Jens: Severance Payments and Firm-Specific Human Capital, November 2001
112. Krieger, Tim: Immigration, Public Pensions, and Heterogenous Voters, April 2002
113. Lambsdorff, Johann Graf / Sitki Utku Teksoz: Corrupt Relational Contracting, May 2002



114. Barbaro, Salvatore: The Distributional Impact of Subsidies to Higher Education - Empirical Evidence from Germany, September 2002
115. Sauer, Christoph / Schratzenstaller, Margit: Strategies of international fiscal competition for foreign direct investment in a model with impure public goods, December 2002
116. Barbaro, Salvatore: The Combined Effect of Taxation and Subsidization on Human Capital Investment, January 2003
117. Südekum, Jens: Increasing Returns and Spatial Unemployment Disparities, April 2003
118. Krieger, Tim / Sauer, Christoph: Will Eastern European Migrants Happily Enter the German Pension System after the EU Eastern Enlargement? May 2003
119. Thies Büttner / Robert Schwager: Regionale Verteilungseffekte der Hochschulfinanzierung und ihre Konsequenzen. October 2003
120. Barbaro, Salvatore / Südekum, Jens: Reforming a complicated income-tax system: The political economics perspective. January 2004
121. Barbaro, Salvatore: Tax Distortion, Countervailing Subsidies and Income Redistribution. April 2004
122. Cogneau, Denis / Grimm, Michael: The Measurement of Income Distribution Dynamics when Demographics are correlated with Income. January 2005.