UNIVERSITÀ DEGLI STUDI DI ROMA "TOR VERGATA"

FACOLTA' DI ECONOMIA

DOTTORATO DI RICERCA IN
ECONOMIA INTERNAZIONALE

XXII. CICLO DEL CORSO DI DOTTORATO

Essays on Economics and Demography

Tilman Tacke

A.A. 2008/2009

Docente Guida/Tutor: Prof. Robert J. Waldmann

Coordinatore: Prof. Giancarlo Marini

Essays on Economics and Demography

Updated version of a doctoral thesis presented by

Tilman Tacke

to the Department of Economics

Università degli Studi di Roma "Tor Vergata"

Supervisor: Prof. Robert J. Waldmann

June 2009

Contents

Abstract

Differences in demographics between countries and over time are caused by many political, economic and cultural circumstances. This research analyzes two of these circumstances and their possible influence on infant mortality.

The infant mortality rate varies internationally between 3 and more than 150 deaths per thousand live births. Differences in per capita income explain these variations to a large extent. Chapter one and two discuss how the distribution of per capita income and the source of health care expenditure, both independent from their absolute levels, might help to explain differences in mortality between countries with the same level of per capita GDP.

Chapter 1 indicates the relevancy of the share of national income concentrated among a country's richest 20% for health outcomes of the lowest three quintiles of the income distribution independent of their personal absolute income. A comparison of 93 countries suggests that the income distance between an individual and the people the individual directly observes at work or in the neighborhood might not be the only form of relative deprivation, but that the distance to the wealthy might matter as well. The results are not caused by the distribution of absolute income: previous research has shown that the level of GDP per capita is the single most important determinant of health outcomes (see Prichett and Summers, 1996), which our results confirm. The chapter discusses possible explanations for the link between income distribution and health outcomes. Suggestions for two explanations are found: public disinvestment in human capital in countries where income is unequally distributed and relative deprivation, i.e. social comparison resulting in stress, risk taking behavior, or unwise consumption expenditure.

Chapter 2 investigates the relative efficiency of public and private health care spending in reducing infant and child mortality using cross-national data for 163 countries. The analysis shows that an increase in public funds is both, significantly correlated with lower mortality rates and significantly more efficient in reducing mortality than private health care expenditure. The results indicate that an increase in private health care expenditure might even be associated with higher, not lower, mortality. Although some of the estimated difference in the efficiency of public and private health care expenditure can be explained by geographies and socioeconomic factors, chapter 1 concludes that the indications for such difference are robust.

Given the comments from the PhD committee, both chapters will be subject to changes. The author will be pleased to send on request (email to tacke@economia.uniroma2.it) the most recent version of the chapters.

Acknowledgements

I am deeply grateful for the support and encouragement provided by my supervisor, Professor Robert J. Waldmann. His guidance and constructive suggestions have shaped my thesis for the better in many ways. It is probably difficult to find another supervisor as instructive and at the same time entertaining as him, still less one who is available for gmail-chats about added-variable plots on Sunday afternoon.

I would also like to thank my thesis committee Prof. Giovanni Andrea Cornia, Prof. Andrea Brandolini, and Prof. Giovanni Vecchi for their excellent comments during and after my finale examination. I am sure that these comments will help to substantially improve the quality of my work.

During the doctorate I was fortunate to find myself within an extremely encouraging environment. I want to thank the coordinator of the Ph.D. Programme, Professor Giancarlo Marini, for his support and advice. I am particularly grateful to him and Professor Pasquale Scaramozzino for providing me the opportunity to gain teaching experience from early on, and to Dott.ssa Michela Carnevali for giving me a great deal of help in my everyday life at Tor Vergata. I would also like to acknowledge Dott. Piergiuseppe Cossu's help throughout the three years of the Ph.D. Programme, but especially at the final stage of my dissertation.

I would like to show my gratitude to all the teaching staff of the Department of Economics and Institutions as well as to Aldo Christian, Daniele, Oege, Gianluca, Melody and my other colleagues at the University of Rome. I would also like to thank my friends Frank, Maxim, and Khaled for providing me feedback to some of my papers and life in general.

Finally, it is my pleasure to thank those who I enjoy the most spending my time with, my parents Uta and Maurus, my brothers Moritz, Leo, and Fritz, my sister Clara, and my girlfriend Laura.

Chapter 1

Infant Mortality and Relative Income

Do health outcomes depend on relative income as well as on an individual's absolute level of income? We use infant mortality as a health status indicator and find a significant and positive link between infant mortality and income inequality using cross-national data for 93 countries. Holding constant the income of each of the three poorest quintiles of a country's population, we find that an increase in the income of the upper 20% of the income distribution is associated with higher, not lower infant mortality. Our results are robust and not just caused by the concave relationship between income and health. The estimates imply a decrease in infant mortality by 1.5 percent for a one percentage point decrease in the income share of the richest quintile. The overall results are sensitive to public policy: public health care expenditure, educational outcomes, and access to basic sanitation and safe water can explain the inequality-health relationship. Our findings support the hypothesis of public disinvestment in human capital in countries with high income inequality. However, we are not able to determine whether public policy is a confounder or mediator of the relationship between income distribution and health. Relative deprivation caused by the income distance between an individual and the individual's reference group is another possible explanation for a direct effect from income inequality to health. Our results suggest two sorts of reference groups: an individual's social peer group, visible at work or in the neighborhood, and the wealthy.

1.1 Introduction

Is income inequality associated with bad health outcomes? If so, what causes the relationship between distribution of income and health? This paper presents some new empirical evidence related to both questions. Even given the income of each of the three lowest income quintiles, we find a statistically significant association of higher income inequality and worse health outcomes. Our findings support the relative income hypothesis, which asserts that the association between inequality and poor health outcome is not just due to absolute poverty. The results suggest that a one percentage point decrease in the income share of the highest quintile of the income distribution ("the rich") would be associated with a decrease in infant mortality by 1.5 percent. Given an estimated annual infant mortality of 5.86 million in the 93 countries covered (2000), the one percentage point reduction in the income share of the rich would prevent nearly 90,000 infant deaths each year. This estimated effect does not require an increase in the absolute income of the poor as a result of income redistribution. We find evidence for public disinvestment in human capital in countries with high income inequality. The estimated link between inequality and health outcomes is sensitive to public health care expenditure, education outcomes and access to safe water and basic sanitation, but we have no ability to determine which are independent causes and which transmission channels. A further possible explanation for the income inequality-health relationship is relative deprivation: the distance in income between an individual and its reference group might affect personal health outcomes. Most existing literature defines the neighborhood ("keeping up with the Joneses") as the reference group: low relative income compared to the local peer group may cause social stress and increase the probability of taking health risks such as smoking and alcohol consumption. We find suggestive evidence of an effect of relative deprivation between similar social classes. However, our analysis indicates that the neighborhood might not be the only reference group that matters for health outcomes. We find a relative deprivation effect also if we define the top end of the income distribution as the reference group: controlling for the income of the poor, infant mortality is higher where the income distance between the poor and the rich is greater. Relative deprivation due to income distance to the rich may matter for health not because of stress or health risk taking, but if the non-rich try to emulate the consumption behavior of the rich by allocating a larger share of consumption from non visible goods (such as health care) to status goods.

A direct effect of income inequality on health outcomes is due to individuals' absolute level of income (absolute income hypothesis). Suppose two countries with identical average income per capita but differences in income distribution. For a concave relationship between health and income - additional income has a smaller effect on the health of rich people than on the health of poor people - we expect the unequal country to have worse average health outcomes. In addition to the absolute income hypothesis, the relative income hypothesis has been formulated to describe a further effect of income distribution on health outcomes: the health of an individual might not only depend on the individual's absolute income, but also on income compared to others in the same country. There are various theoretical rationales for the relative income hypothesis (for a discussion see Kawachi and Kennedy, 1999): income inequality may affect health outcomes through (a) public disinvestment in human capital in countries with high income inequality, (b) reduced social cohesion, or (c) negative effects of social comparison ("relative deprivation") such as increased stress, risk taking behavior, or unwise consumption. Also the link between inequality and health might not be due to a casual effect of inequality on infant mortality, but due to omitted variable bias.

The idea that relative poverty can cause absolute mortality is not new. It was noted by Adam Smith in The Wealth of Nations (1776). Smith argues that "necessaries" include "whatever the custom of the country renders it indecent for credible people to be without" and that such custom affects the level of absolute poverty at which the population is stable due to insufficient "ability of the inferior ranks of people to bring up families".

A large body of literature analyzes the relationship between income distribution and health outcomes with controversial results (for a comprehensive review of existing literature see Wilkinson and Pickett (2006)). Wilkinson (1992) shows an association between high infant mortality and high income inequality for industrialized countries both in levels and over time. Waldmann (1992) uses data for 57 countries in the 1960s and 1970s and derives similar results. Marmot and Bobak (2000) show higher mortality rates as income inequality increases in Eastern Europe and the former Soviet Union. De Volgi et al. (2005) find support for a link between higher income inequality and lower life expectancy in Italy (1995-2000) and in 21 industrialized countries (2001). Some research finds statistically significant and adverse associations of income inequality and infant mortality that become statistically insignificant when controls are added: Judge et al.'s (1998) association does not remain statistically significant when they control for female labor participation, Lynch et al.'s (2001) findings are largely driven by the United States. Other studies fail to find a statistically significant relationship between income distribution and health outcomes: Gerdtham and

Johannesson (2004) cannot find a statistically negative effect from income inequality on infant mortality within Swedish counties or municipalities; Mellor and Milyo (2001, 2002) find no support for Wilkinson's (1992) results; Wildman et al. (2003) replicate Waldmann's (1992) model with more recent data, but are not able to achieve similar results.

Two general patterns can be observed when comparing results from existing research: first, the aggregation level of the reference group matters, positive findings are usually found when income inequality is measured in larger areas (e.g. Mellor and Milyo, 2002; Soobader et al., 1999; Wilkinson and Pickett, 2006); second, the relationship between health and income distribution in international comparisons was temporarily lost during the decade after the mid-1980s (e.g. Mellor and Milyo, 2001; Wildman et al., 2003; Wilkinson and Pickett, 2006). Time-series analysis seems therefore difficult: if the relationship between health and income distribution disappeared in the mid-1980s, we cannot describe this relationship with a constant coefficient. Furthermore, a within estimator of the relationship between health and income inequality within countries over time would require assumptions concerning a time-lag between changes in the distribution of income and their possible effects on health outcomes. To anticipate problems related to time-series analysis and low levels of aggregation we investigate the link between income distribution and health outcomes using an international cross-section analysis.

Existing research questions the robustness and reliability of findings concerning the relative income hypothesis from international cross-section analysis mainly for two reasons: data quality and the aggregation problems. The comparability of international data describing income distributions is limited due to varying definitions of inequality and measurement problems (see Atkinson and Brandolini (2001) for a discussion of pitfalls in the use of income distribution data). Gravelle (1998) and Wildman et al. (2003) discuss the aggregation problem in aggregated level studies of the relative income hypothesis: when aggregated data is used, it is difficult to distinguish between the absolute and the relative income hypotheses. The concave relationship between health and income could cause such aggregation problem. If the benefit of one additional Dollar spent on health care is higher for the poor (e.g. spent for vaccination) than for the rich (e.g. spent for whitening of teeth), we would expect average health of a society to be lower in a more unequal country not because of adverse effects of a lower relative income, but because of a lower absolute income of the poor. Many aggregate cross-section studies are subject to such aggregation problems (Wildman et al., 2003). Some authors suggest not using aggregated data for studies of income distribution and population health, but favor micro-data instead. While this approach definitely solves the aggregation problem, it inevitably restricts the reference groups, which can be considered. If one uses micro data on health outcomes to study relative deprivation, it is necessary that reference group income is different for different families in the same micro data set. Wilkinson (1997) argues that health outcomes in deprived neighborhoods are poorer not because of the inequality within their neighborhoods, but compared to other neighborhoods. Such effects might not be visible in studies of micro-data. Waldmann (1992) and Wildman et al. (2003) address the aggregation problem by using the income level of the poorest quintile and the income of the population between the twentieth and the ninety-fifth percentile to allow for the effect of absolute income on health; they then test the relative income hypothesis by using the income share of the rich.

Relative deprivation has been discussed as a possible explanation for the link between income inequality and health outcomes. The relative deprivation hypothesis assumes that the income distance between an individual and his or her reference group affects the individual's health. Reference groups are usually defined by individual characteristics and geographic coverage: Eibner and Evans (2005) for example define the reference group as a combination of the state of residence, age, race, and education. Soobadera and LeClere (1999) investigate different levels of geographic aggregation of the reference group and find independent effects of inequality on health at higher levels of aggregation. Deaton (2001) shows that mortality depends on income inequality within and between states. Runciman (1966) proposes two types of reference groups: individuals in the same neighborhood/social class and individuals serving as role models from higher social classes. Current literature focuses on relative deprivation between individuals from similar social classes resulting in health affecting stress or risk taking behavior. Little attention has so far been given to the rich as a reference group for the non-rich and relative deprivation leading to unwise consumption. A further explanation for the health-inequality relationship is public investment in human capital. The idea is that countries with higher income inequality might spend less on public education or public health care. A reason for lower public investment might be that high income inequality leads to a divergence of the interest of the rich and the non-rich, with the rich trying to lobby for lower taxes and public services (Kawachi and Kennedy, 1999). Wilkinson and Pickett (2006) point to the difficulty to determine whether public investments are pathways or confounders of the relationship between income inequality and health outcomes.

Given the contradictory results of earlier research and concerns regarding their design and data quality it seems worthwhile to reinvestigate the relationship between income distribution and infant mortality. The paper is divided into 5 sections. The section following the introduction discusses data, specification, and estimation technique. The third section presents the results of the estimation and their robustness concerning subsample stability and choice of specification. The fourth section discusses alternative justifications of the relationship between income distribution and health outcomes. The paper closes with a concluding section.

1.2 Income Distribution and Health

Income is the single most important determinant of a country's health status. People live longer and infant mortality is lower in rich countries than in poor countries (Prichett and Summers, 1996). The health-wealth relationship does not only hold among countries, but also within countries. While the income of a country increases over time, the health status improves accordingly, and at a moment in time wealthier people are healthier than their poorer fellow citizens (Deaton, 2003). Income distribution within a country might influence average health outcomes in two ways: (a) the relationship between health and income is concave: an increase in income has a larger effect on health among poor people than among rich people. The literature has referred to effects of absolute income on health outcomes as the absolute income hypothesis; (b) holding constant an individual's income, a more unequal distribution of income might reduce health because of, among many possible explanations, unwise expenditure or social stress. Effects of the relative income position on health outcomes have been described by the relative income hypothesis. In this chapter we discuss both hypotheses, starting with the effect of income distribution via absolute levels of income.

Table 1.1 shows Gini coefficients as a proxy for income distribution, infant mortality as a health status indicator, and GDP per capita in purchasing power parity (PPP) for 125 countries in 1999-2003. Averages for the 25 countries with the highest Gini coefficient are displayed at the left hand side, for the lowest Gini coefficients at the right hand side. The table illustrates the association between higher infant mortality and both lower per capita income and higher income inequality. The most unequal countries are at the same time on average the poorest (average GDP per capita of $4,486) and have the highest infant mortality with 61 deaths per 1,000 live births. Infant mortality falls and average income increases monotonically with more equal income distribution to 21 deaths per 1,000 live births and an average GDP per capita of $14,522 in countries with the lowest income inequality.

Table 1.1: Infant Mortality and Per Capita Income, by Income Distribution Quintiles (1999-2003)

	Most unequal 25 countries	Second 25	Third 25	Fourth 25	Most equal 25 countries
Gini	56	46	39	34	29
Infant Mortality	61	59	55	29	21
Per Capita Income	4,486	4,490	5,751	12,005	14,522

Notes: Infant mortality, per capita income ($, in PPP), and gini coefficients are from World Bank's World Development Indicators 2005.

The left graph in Figure 1.1 shows the concave relationship between income and health in an international comparison of 162 countries (mortality and income data for 1999-2003): infant mortality in countries with income per capita levels of $2,000 is on average 75 deaths per 1,000 live births; infant mortality decreases to (on average) approximately 12 deaths for income levels around $10,000; countries with an average income per capita around $25,000 have an infant mortality of on average 4 deaths per 1,000 live births. I.e., while infant mortality decreases on average by 8 deaths for a $1,000 increase in income in countries with a GDP per capita between $2,000 and $10,000, it decreases by 0.5 deaths for the same income increase in countries with a GDP per capita between $10,000 and $25,000. The right graph shows that the log of infant mortality of a country appears to be a linear function of the log of per capita GDP (in PPP) for the same country. We therefore use a log-log specification in regressions of real income on infant mortality.

In addition to an association of richer countries with better health outcomes, the absolute income hypothesis (Wildman et al., 2003) postulates, controlling for average income, a lower infant mortality in countries with low income inequality. To test the relationship between income distribution and infant mortality, we use different measures of average income as determinants of infant mortality. Using income distribution by quintiles we find that the explanatory power of average GDP per capita of the whole population is significantly lower than the explanatory power of the average GDP per capita of the bottom 80% of the income distribution or of the lowest 60% of the income distribution. The highest explanatory power is associated with the third quintile: the median income seems to be more relevant for health outcomes then the average income. Table 1.2 shows these results based on income distribution and infant mortality data for 93 countries.[1] $\ln(\text{GDP}_{q2-q4})$ denotes the log of the average GDP of the second, plus the third, plus the forth quintile of the income distribution.

These results support the use of median income (in Table 1.2 represented by the average income of the third quintile $\ln(\text{GDP}_{q3})$) as a proxy for social welfare. The regression of $\ln(\text{GDP}_{q3})$ yields the highest significance and R^2 in explaining international differences in infant mortality.

1.2.1 Specification

This paper tries to overcome the aggregation problem and disentangle the relative from the absolute income hypothesis by using income data by quintiles. The advantage of using income per quintiles as opposed to average income of the whole population is that regression estimates for income per quintile capture the relationship between absolute income and infant mortality: infant mortality is expected to be higher in more unequal countries because the absolute income of the poor is lower. To isolate relative income effects from absolute income effects we introduce a measure of income distance between lower and upper quintiles. Holding constant the income of each of the three lower quintiles, the absolute income hypothesis predicts lower infant mortality for a larger distance between the income of lower and upper quintiles (as the larger income distance represents generalized Lorenz dominance). On the contrary, the relative income hypothesis would predict *higher* infant mortality associated with a larger income distance (as the larger income distance represents higher income inequality).

Our empirical strategy to test the relative income hypothesis is very simple: we specify multivariate regressions that try to explain the level of infant mortality in 93 countries using the log of per capita GDP for each of the lower three quintiles and an explanatory variable ΔInc representing the relative difference between the income of quintile i and quintile j.

$$\begin{aligned} \ln(infant\ mortality) = {} & \beta_1 \ln(per\ capita\ GDP_{q1}) + \beta_2 \ln(per\ capita\ GDP_{q2}) \\ & + \beta_3 \ln(per\ capita\ GDP_{q3}) + \beta_4 \Delta Inc_{qi,qj} \end{aligned} \tag{1.1}$$

Infant mortality is the number of deaths in the first year of life per thousand live births. We calculate ΔInc as the difference in income between two quintiles relative to the income of the higher income quintile. E.g., for $i = 1$ and $j = 4$,

[1] The reduced sample size is due to the lack of data on income distribution by quintiles for some countries.

Figure 1.1: Infant Mortality and GDP per Capita (left), Log Infant Mortality and Log GDP per Capita (right)

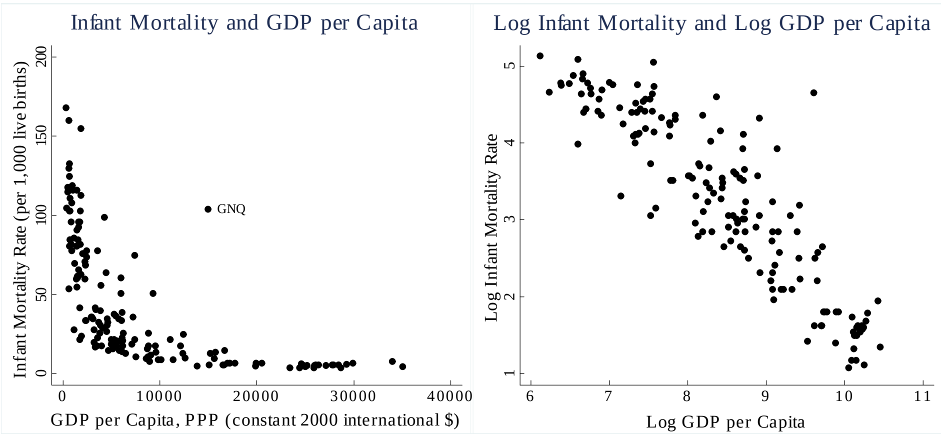


Table 1.2: Infant Mortality and Per Capita Income, OLS Estimates (1999-2005)

	(I).	(II)	(III)	(IV)	(V)
$\ln(\text{GDP}_{q1-q5})$	-0.87				
	(-22.37)				
$\ln(\text{GDP}_{q1-q4})$		-0.85			
		(-24.98)			
$\ln(\text{GDP}_{q1-q3})$			-0.83		
			(-24.96)		
$\ln(\text{GDP}_{q2-q4})$				-0.85	
				(-25.06)	
$\ln(\text{GDP}_{q3})$					-0.85
					(-25.10)
Constant	10.58	13.73	13.13	13.70	12.67
	(32.05)	(32.61)	(33.04)	(32.74)	(33.61)
R^2	0.84	0.87	0.87	0.87	0.87
N	93	93	93	93	93

Dependent variable: log infant mortality; t-statistics in parentheses.

$\Delta Inc_{q1,q4}$ denotes the difference in income between the lowest quintile ($q1$) and the fourth quintile ($q4$)

$$\Delta Inc_{q1,q4} = \frac{q4 - q1}{q4} \tag{1.2}$$

where $q1$ is the income share held by the bottom 20% of the income distribution. A value of $\Delta Inc_{q1,q4} = \frac{2}{3}$, for example, indicates that the difference between the average income of a person in the lowest quintile and an average person in the fourth quintile amounts to $\frac{2}{3}$ of the average income of the wealthier people. β_4 is not only an indicator for the validity of the relative income hypothesis, but also for the absolute income hypothesis in upper quintiles. A positive β_4 is evidence for the relative income hypothesis: the higher the distance between the average income levels of two quintiles, the higher the infant mortality. The other coefficients should be positive. However, due to high multicollinearity between the first three variables, the estimates of their standard errors are large, so β_4 is the only coefficient of interest. Our specification allows to investigate the reference group that might matter for relative deprivation. Relative deprivation between individuals in similar social classes would be indicated by a positive and statistically significant β_4 when we compare the income share of a quintile with the share of the next higher income quintile (e.g. $q3$ with $q4$). Relative deprivation resulting from a comparison with the rich can be analyzed by using distant quintiles (e.g. $q2$ and $q5$) of the income distribution.

We use further specifications to evaluate the relative income hypothesis. Similar to Waldmann (1992), we evaluate the relationship between infant mortality and the real income of the highest quintile or the income share of the highest quintile. The use of the real income of the highest quintile is problematic due to collinearity between the real incomes across quintiles.

1.2.2 Data

To estimate equation (1.1) we collect data on 93 countries for the years 1999 to 2005.[2] The aggregate population of these 93 countries is 87% of the world

[2] The countries included in the data set are: Albania, Armenia, Austria, Azerbaijan, Bangladesh, Belarus, Belgium, Benin, Bolivia, Bosnia and Herzegovina, Brazil, Bulgaria, Burkina Faso, Cambodia, Cameroon, Canada, Chile, China, Colombia, Cote d'Ivoire, Croatia, Ecuador, Egypt, El Salvador, Estonia, Ethiopia, Finland, Georgia, Germany, Ghana, Guatemala, Guinea, Guyana, Haiti, Honduras, Hungary, India, Indonesia, Iran, Ireland, Israel, Italy, Jamaica, Jordan, Kazakhstan, Kyrgyz Republic, Lao PDR, Latvia, Lithuania, Luxembourg, Macedonia, Madagascar, Malawi, Mali, Mauritania, Mexico, Mongolia, Morocco, Mozambique, Nepal, Netherlands, Nicaragua, Nigeria, Norway, Pakistan, Panama, Paraguay, Philippines, Poland, Romania, Russian Federation, Senegal, Slovak Republic, Slovenia, South

population. In the following we describe the construction of our data-set.

The starting point is the UNU/WIDER World Income Inequality Database (WIID), version WIID2c (downloaded the 31st of July 2008). Compared to other secondary data-sets on income distribution the WIID offers the advantage of including information about the coverage (area, population, ages), the definitions of income (consumption, expenditure, gross income, net income etc.), the income sharing unit (households, family, person etc.), the unit of analysis (indicating weighting of data with person or household weight), the equivalence scale used for weighting, and an assessment of data quality. Atkinson and Brandolini (2001) discuss problems of quality and consistency in income distribution data within and across countries. They demonstrate the dependency of regression estimates on the choice of income distribution data.

To limit consistency and quality problems we select income distribution data by quintiles (in some cases calculated using data by deciles) only in case the whole population is covered across all age classes and including rural and urban areas (the WIID2c labels these specifications as "All" for the criteria area coverage (AreaCovr), population coverage (PopCovr), and age coverage (AgeCovrse)). We restrict the data to surveys that assign per capita household income to each member of the household (i.e. "Household" as the income sharing units (IncSharU), "Person" as the unit of analysis (UofAnala), and "Household per capita" as the equivalence scale (Equivsc)). To obtain a sufficiently large sample of countries we cannot restrict ourselves to one definition of income only. Observations include the distribution of consumption/expenditure, as well as of disposable income, disposable monetary income, and income without specification of the treatment of taxes (from the Socio-Economic Database for Latin America and the Caribbean). In case there is more than one observation per country that satisfies our criteria, we prefer observations with a high quality ranking (WIID ranks the data quality from 1 to 3). Within a quality class we prefer consumption/expenditure data over disposable income data, and disposable income data over income data without specification of tax treatment. For multiple observations of a country within a quality and definition class we try to collect data for the year 2000, in case there are no estimates for 2000, we use the estimate for the closest year to 2000 in the range 1999-2005. Out of the 93 observations, 56 use expenditure/consumption as the definition of income, 22 disposable income, 9 disposable monetary income, and 6 income without speci-

Africa, Spain, Sri Lanka, Swaziland, Sweden, Tajikistan, Tanzania, Thailand, Tunisia, Turkey, Uganda, Ukraine, United Kingdom, United States, Uzbekistan, Venezuela, Vietnam, Yemen, and Zambia.

fication of the treatment of taxes. There are 39 observations of the year 2000, 12 for 1999, 20 for 2001, 9 for 2002, 8 for 2003, 3 for 2004, and 2 for 2005 in the data-set.

To account for differences in the income definition we introduce dummy variables. One dummy variable is constructed for each income definition. A shortcoming of the introduction of simple dummy variables for income definitions (e.g. the consumption dummy has the value of 1 for income distribution data based on consumption, otherwise a value 0) is that there is a selection bias: e.g. if all African countries use the same definition of income in their income distribution statistics, the dummy variable captures at the same time fixed effects of differences in the income definition and a fixed continent effect for Africa. To overcome the selection bias, Deininger and Squire (1996) construct dummy variables comparing Gini coefficients with different income definitions for the same country. Unfortunately only 20 out of the 93 countries in our data set provided income distribution data by quintiles (or deciles) with different definitions of income for the same year. We thus have to use simple dummy variables to account for definitional differences: "Exp." for income distributions based on consumption/expenditure, "M. Inc." for disposable monetary income, and "NS Inc." income without specification. Controlling for income definitions leads to substantial adjustments of our regression estimates, the dummy variable for data based on distribution of consumption/expenditure is always statistically significant and positive. All regression estimates shown in this paper are therefore obtained using income definition dummy variables. Including a dummy variable for the year of the observation has almost no effect on the estimates (changes of the point estimate are $<2\%$, of its statistical significance $<1\%$). We therefore do not include a dummy for the year of the observation in our regressions.

We use infant mortality as an indicator of health status as it is available for a large number of countries, and as it avoids potential problems of reverse causation associated with the relationship between adult health and income distribution (bad health outcomes cause income inequality). Unfortunately, definitions of infant mortality vary internationally; the comparability of infant mortality data is therefore limited. We therefore use child mortality (number of infants dying before reaching five years of age per 1,000 live births in a given year) as an additional variable. Our data for infant/child mortality and GDP per capita (in PPP, constant 2000 international $) is primarily provided by the World Bank's World Development Indicators 2005 (CD version). Figures for infant mortality reported by the World Bank are often based on interpolations, extrapolations, or comparisons with other countries. However, the estimations are found to be reasonably accurate for cross-national comparisons at a point in time (Prichett and Summers, 1996). One observation of infant mortality (Canada) is taken from the Globalization–Health Nexus Database (Version 1).

We use other variables as controls: public and private health care expenditure as a share of GDP, access to basic sanitation and safe water, and the ratio of total female enrollment in primary school to the female population of the age group are taken from the World Bank's World Development Indicators 2005; obesity data (BMI>=30 kg/m^2, age group: 15-100 years) is provided by the World Health Organization for 2005; female literacy is taken from the UNESCO Institute for Statistics' Literacy and Adult Education Statistics Programme, if not available - which is the case for some OECD countries - from the CIA World Factbook 2007; female smoking prevalence (prevalence of current tobacco use among female adults (>=15 years) in 2005) is provided by the World Health Organization; HIV prevalence is taken from the World Bank's World Development Indicators 2005, if not available from the UNAIDS' "2008 Report on the Global Aids Epidemic" (2001 estimates).

1.3 Results

This section reports first the results of ordinary least squares (OLS) regressions corresponding to equation (1.1). We then discuss the robustness of the results concerning subsample stability and model specification.

1.3.1 OLS Regressions

Table 1.3 shows the results of estimating equation (1.1). The regression estimates support the validity of the relative income hypothesis as all β_4 coefficients (i.e. the coefficients on $\Delta Inc_{qi,qj}$) are positive and most of them statistically significant: holding constant the income of each of the lower three quintiles, a higher income of the upper two quintiles is associated with higher infant mortality. Note that the estimates for the income quintiles in Table 1.3 are imprecise due to multicollinearity as GDP per capita is a common factor used in the calculation of ln(GDP$_{q1}$), ln(GDP$_{q2}$), and ln(GDP$_{q3}$). Multicollinearity does not bias coefficient estimates, but causes large standard errors.[3]

[3] We calculate the variance inflation factor (VIF) to analyze potential problems arising from multicollinearity. A large VIF indicates that it is difficult to distinguish between the effects of collinear variables. For a VIF of 10 or above, the estimated significance of a variable is not trustworthy (Neter et al., 1996). In all regressions multicollinearity is a serious problem

The average value of $\Delta Inc_{q2,q5}$ in our sample is 75.5%: the average income in the second poorest quintile of the population is therefore less than 25% of the average income in the riches income quintile. The results in (VIII) suggest that infant mortality increases with income inequality described by $\Delta Inc_{q2,q5}$ with semi-elasticity of 2.33 (t-statistic of 2.61). Holding the income share of the second poorest quintile constant at 10.8%, a one percentage point drop in the income share of the richest quintile from 46.5% to 45.5% would imply a decreased infant mortality of 1.5% in a country at the sample mean.[4] If the same relationship were true for all countries in our sample, a one percentage point decline in the income share of the richest quintile in all 93 countries would avert annually nearly 90,000 deaths.[5] This decrease would be due to effects as postulated by the relative income hypothesis, not because of an increase in the absolute income of the lower quintiles. The other columns of Table 1.3 show that we find similar relationships between infant mortality and income distribution if we use the distance between either the highest or second highest quintile and any of the lower three quintiles. In the following we will use the distance between the second and the highest quintile of the income distribution as it achieves the statistically most stable results.

1.3.2 Robustness of Regression Estimates

In this section we investigate the robustness of the results reported in Table 1.3 concerning the stability of our estimates with respect to influential observations and model specification. Controlling for outlying observations or leverage points and using different model specifications does not change our initial finding: given the income of each of the three lower quintiles of the income distribution, an increase in the income of the higher quintiles is statistically significantly associated with higher infant mortality.

Regression estimates are sensitive to influential observations, which can be divided into outliers and leverage points. Outliers are observations with large residuals: the value of the dependent variable is unusual given the values of its predictor variables; leverage points are observation with an extreme value on a

among $\ln(\text{GDP}_{q1})$, $\ln(\text{GDP}_{q2})$, and $\ln(\text{GDP}_{q3})$ with VIFs always above 100. However, the collinearity between our measures of income distance and the other variables is low. All but one VIF (the VIF of $\Delta Inc_{q1,q4}$ in regression (VII) equals 13.9, this might be the reason why the estimate for β_4 in (VII) is statistically not significant) of our income distance measures are below the critical threshold of 10. This implies that the significance of the variables testing the relative income hypothesis is not overly sensitive to the inclusion of the other variables.

[4] Note that the one percentage point decrease in the income share of the wealthiest quintile results in a decrease in $\Delta Inc_{q2,q5}$ from 75.5% to 74.9%.

[5] The 93 countries reported 5.86 million child deaths in 2000.

Table 1.3: Infant Mortality, Per Capita Income, and Income Distance, OLS Estimates (1999-2005)

	(VI)	(VII)	(VIII)	(IX)	(X)	(XI)
ln(GDP$_{q1}$)	0.13	0.14	-0.05	0.12	0.17	-0.08
	(0.31)	(0.23)	(-0.14)	(0.28)	(0.38)	(-0.20)
ln(GDP$_{q2}$)	-0.16	-0.45	0.28	0.36	-0.19	0.26
	(-0.18)	(-0.51)	(0.31)	(0.37)	(-0.22)	(0.28)
ln(GDP$_{q3}$)	-0.67	-0.41	-0.91	-1.19	-0.67	-0.89
	(-0.94)	(-0.50)	(-1.30)	(-1.38)	(-0.96)	(-1.19)
$\Delta Inc_{q1,q5}$	2.71					
	(1.95)					
$\Delta Inc_{q1,q4}$		1.50				
		(0.96)				
$\Delta Inc_{q2,q5}$			2.33			
			(2.61)			
$\Delta Inc_{q2,q4}$				2.54		
				(2.08)		
$\Delta Inc_{q3,q5}$					2.83	
					(2.06)	
$\Delta Inc_{q3,q4}$						2.83
						(2.16)
Exp.*	0.40	0.41	0.40	0.36	0.40	0.38
	(2.98)	(2.97)	(3.04)	(2.59)	(2.96)	(2.79)
M.Inc.*	0.06	0.04	0.08	0.03	0.06	0.06
	(0.40)	(0.25)	(0.56)	(0.24)	(0.40)	(0.40)
NS Inc.*	-0.05	-0.06	-0.11	-0.12	-0.06	-0.14
	(-0.29)	(-0.33)	(-0.61)	(-0.63)	(-0.30)	(-0.75)
Constant	8.51	9.91	8.87	9.89	8.41	10.04
	(6.90)	(10.61)	(9.62)	(13.20)	(6.85)	(13.98)
R^2	0.89	0.89	0.89	0.89	0.89	0.89
N	93	93	93	93	93	93

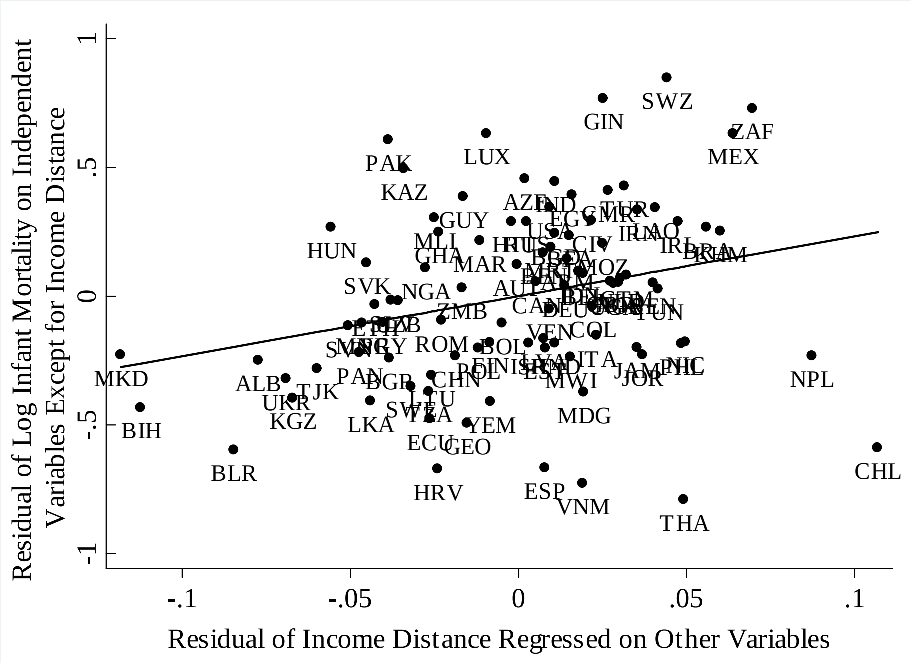
Dependent variable: log infant mortality; t-statistics in parentheses.
*Dummy variables for income distribution definitions.

predictor variable, i.e. leverage is a measure of how far an independent variable deviates from its mean. Figure 1.2 shows a partial scatter of infant mortality against the income distance between the second lowest and the highest income quintile. The residuals from a regression of the log of infant mortality on all explanatory variables except the income distance are plotted on the vertical axis, the residuals from a regression of the log of the income distance on all other variables are plotted on the horizontal axis. The slope of the line corresponds to the estimates of $\beta_4 = 2.33$ in regression (VIII) in Table 1.3. We define an outlier as an observation with a studentized residual, a type of standardized residual, exceeding +2 or -2.[6] Two outlying observations are evident: Thailand (THA) with a studentized residual of -2.8 and Chile (CHL) with a studentized residual of -2.7. Further outlying observations include Swaziland (SWZ; studentized residual 2.3), Vietnam (VNM; -2.3), Spain (ESP; -2.1), Guinea (GIN; 2.1), Pakistan (PAK; 2.1), and Luxembourg (LUX; 2.0). The coefficient on the income distance in regressions not including all eight outlying observations is 3.37 (compared to 2.61 in regressions including all observations) with a *t*-statistic of 4.56 (all observations: 2.61). By calculation of leverage values we can identify influential observations. The following observations have leverage values exceeding $(2k+2)/n$, with k the number of independent variables: Brazil (BRA; leverage value of 0.20), Colombia (COL; 0.17), Haiti (HTI; 0.19), Honduras (HND; 019), Macedonia (MKD; 0.27), Netherlands (NLD; 0.18), Panama (PAN; 0.21), and Paraguay (PRY, 0.19). Excluding these observations we obtain a coefficient on the income distance of 2.26 with a *t*-statistic of 2.39. Cook's D can be used to detect influential observations in general as it is affected by large residuals and/or large leverage. Values exceeding $4/n$ indicate such influential observations: Chile (CHL; 0.13), Thailand (THA; 0.09), Spain (ESP; 0.06), South Africa (ZAF 0.06), and Luxembourg (LUX; 0.05). Excluding these observations, the estimated coefficient on the income distance becomes 3.18 with a *t*-statistic of 3.77.

We use dummy variables for the countries identified as influential observations (using Cook's D). If we include dummy variables for Chile, Thailand, Spain, South Africa, and Luxembourg into regression (VIII) all but South Africa and Luxembourg are statistically significant: Thailand (*t*-statistic of -3.07), Chile (-2.86), Spain (-2.26), Luxembourg (1.96), and South Africa (1.37). For the rest of our analysis we therefore include dummy variables for Chile, Spain,

[6] A studentized residual is calculated by division of a residual by its estimated standard deviation. A studentized residual of +2 indicates that the observed value of an observation is 2 standard deviations above prediction.

Figure 1.2: Partial Scatter of Infant Mortality against Income Distance between Second Lowest and Highest Income Quintile



and Thailand, the exclusion leads to a larger and statistically more significant relationship between infant mortality and income distribution: the estimated coefficient on the income distance increases from 2.33 to 3.40, the t-statistic from 2.61 to 4.05 (see regression (XII) in Table 1.4).[7]

We investigate the general model specification of equation (1). The set of explanatory variables in regression (XII) in Table 1.4 accounts for 92% of the variability in infant mortality between 93 countries. A computed F(10,81)=98.54 provides evidence that our basic set of predictor variables accounts for a statistically significant amount of the variability in the model (p-value for $>$ F $= 0.000$). The Shapiro-Wilk W test is designed to investigate whether errors are normally distributed. The computed W is 0.995, the p-value of 0.980 indicates that the assumption of normally distributed residuals cannot be rejected. In addition, a Cook-Weisberg test for heteroscedasticity (Cook and Weisberg, 1983) did not reject the null hypothesis of constant variance at any conventional significance

[7] Another way to limit the influence of outlying observations and leverage points on regression estimates is to use robust regressions. Robust regressions confirm our findings: the estimated coefficient on the income distance is 2.93 with a t-statistic of 3.24.

level (chi-square 2.04 with p-value of 0.154).

The association between higher infant mortality and higher income inequality is not due to the chosen specification (1.1). Table 1.4 shows the positive coefficients on the income share of the highest quintile in regression (XIII) and on the log of real income of the highest quintile in regression (XIV). The results are statistically significant with t-statistics of 3.90 and 3.92. Regressions (XV) and (XVI) include the log of real income of the fourth quintile, the association of higher infant mortality with higher income inequality remains statistically significant. Regression (XVIII) has both a similar setup and similar findings to Waldmann (1992). Wildman et al. (2003) re-examine Waldmann's investigations. One of Wildman et al.'s regressions is identical to regression (XVIII) but they find no statistically significant partial correlation between the income share of the highest quintile and infant mortality. Note that the log of the aggregate real per capita income of the second, third, and fourth quintile ($\ln(\mathrm{GDP}_{q2-q4})$) in regressions (XVII) and (XVIII) captures the association of higher real absolute income and lower infant mortality, i.e support the absolute income hypothesis for these quintiles. The coefficients on log real income of the poorer quintiles in all other regressions are statistically not significant due to multicollinearity between $\ln(\mathrm{GDP}_{q1})$, $\ln(\mathrm{GDP}_{q2})$, and $\ln(\mathrm{GDP}_{q3})$.

By including the log of real per capita income of each of the three lower quintiles of the income distribution we control for the effect of differences in absolute income on infant mortality. However, using income by quintiles allows only to control for differences between the income quintiles, not within quintiles. We include quadratic terms of the log of the real income of each of the lower three quintiles in the equation to correct for the curvilinear nature of the absolute income – infant mortality relationship. We would expect that the effect of inequality due to curvature of the risk of infant mortality as a function of absolute income has implications for the estimated coefficients on the real income of the lower quintiles, but not for the income distance. Regression (XIX) in Table 1.5 shows that our prediction is correct: the coefficient on the income distance between the highest and the second lowest quintile remains unchanged. It also shows that the new variables are not statistically significant and add very little to R^2. The reason is that we already use three or four parameters to correct for the curvilinear nature of the relationship between absolute income and infant mortality. Due to multicollinearity already the sign of the coefficients on the real per capita income of the lower quintiles does not represent the relationship between higher absolute income and lower infant mortality. Adding more parameters based on the same raw numbers is thus very unlikely to improve the specification. The quadratic terms are not included in the rest of our analysis.[8]

The definition of infant mortality varies internationally. We repeat our basic regression replacing infant mortality by child mortality (mortality before reaching five years of age) and "net child mortality" (mortality of children aged between 1 and 5 years of age).[9] The income distance between the second and the highest quintile of the income distribution is in both cases statistically significantly associated with higher child mortality.

Our results support both, the relative and the absolute income hypothesis. The relationship between income and health of an individual is concave: higher individual income leads to better health but with a decreasing marginal effect. High infant mortality in a country with high income inequality is not only due to low income of the poor: holding constant the absolute income of the poor, a higher distance to the income of upper quintiles is always associated with higher infant mortality. We have shown that the effect can neither be explained with outlying observations, nor with our specification or the definition of infant mortality.

1.4 Possible Explanations

It is impossible to prove the existence of relative deprivation in data aggregated on country levels. Income distribution and infant mortality depend on many geographic, demographic, political, economic and cultural circumstances and we are not able to rule out that the association between income inequality and infant mortality is due to omitted variable bias. But even if we find controls that explain the relationship between health and income distribution, we do not know whether these factors are genuine confounders or pathways (see Wilkinson and Pickett, 2006). This section discusses indications for relative deprivation and the relevance of public policy for the link between income inequality and health.

Although there is no obvious link between HIV and either the relative deprivation hypothesis or public policy, we add the prevalence of HIV as a control variable because it significantly affects our regression estimates. Column (B) in

[8] The RESET test (Regression Specification Error Test) for right hand side variables can be used to test the specification of equation (1). The test tries to significantly improve the model by including powers of the predictions of the model. The F-values of the RESET test for right hand side variables is 0.99 with a p-value of 0.47, well above the conventional significance level of 0.05. The RESET test supports our log-log specification.

[9] The sample for regression (XXII) in Table 1.5 ("net child mortality") is reduced to 92 countries as Honduras reports identical numbers for infant and child mortality.

Table 1.4: Different Specifications for Infant Mortality, Per Capita Income, and Income Distribution, OLS Estimates (1999-2005)

	(XII)	(XIII)	(XIV)	(XV)	(XVI)	(XVII)	(XVIII)
ln(GDP$_{q1}$)	0.15	0.06	0.03	0.10	-0.03	0.29	0.17
	(0.44)	(0.19)	(0.08)	(0.27)	(-0.08)	(1.19)	(0.77)
ln(GDP$_{q2}$)	0.23	0.08	-0.17	0.35	-0.10		
	(0.28)	(0.09)	(-0.21)	(0.41)	(-0.13)		
ln(GDP$_{q3}$)	-1.01	-0.77	-1.28	-0.63	-0.99		
	(-1.56)	(-1.24)	(-1.84)	(-0.66)	(-1.04)		
ln(GDP$_{q4}$)				-0.43	-0.37		
				(-0.52)	(-0.44)		
ln(GDP$_{q2-q4}$)						-0.91	-0.80
						(-3.69)	(-3.43)
$\Delta Inc_{q2,q5}$	3.40			3.78		3.66	
	(4.05)			(3.39)		(4.58)	
q5		0.04					0.04
		(3.90)					(4.47)
ln(GDP$_{q5}$)			0.79		0.87		
			(3.92)		(3.24)		
Exp.*	0.43	0.49	0.49	0.44	0.51	0.41	0.48
	(3.54)	(4.05)	(4.07)	(3.56)	(4.04)	(3.56)	(4.17)
M. Inc.*	0.03	0.07	0.06	0.03	0.07	0.01	0.06
	(0.20)	(0.52)	(0.49)	(0.24)	(0.54)	(0.06)	(0.48)
NS Inc.*	-0.16	-0.24	-0.26	-0.14	-0.25	-0.11	-0.21
	(-0.95)	(-1.45)	(-1.51)	(-0.84)	(-1.49)	(-0.70)	(-1.29)
Chile**	-0.98	-1.11	-1.11	-0.98	-1.12	-0.94	-1.11
	(-2.92)	(-3.17)	(-3.19)	(-2.90)	(-3.19)	(-2.92)	(-3.31)
Spain**	-0.82	-0.81	-0.81	-0.82	-0.82	-0.85	-0.83
	(-2.49)	(-2.44)	(-2.46)	(-2.49)	(-2.46)	(-2.64)	(-2.57)
Thailand**	-1.06	-1.02	-1.02	-1.06	-1.01	-1.13	-1.07
	(-3.26)	(-3.11)	(-3.12)	(-3.22)	(-3.07)	(-3.56)	(-3.38)
Constant	7.59	8.41	9.02	7.41	9.00	8.50	9.24
	(8.75)	(11.15)	(13.24)	(7.91)	(13.12)	(10.03)	(11.60)
R^2	0.91	0.91	0.91	0.91	0.91	0.92	0.92
N	93	93	93	93	93	93	93

Dependent variable: log infant mortality; t-statistics in parentheses.
*Dummy variables for income distribution definitions.
**Dummy variables for influential observations.

Table 1.5: Further Specifications for Infant/Child Mortality, Per Capita Income, and Income Distribution, OLS Estimates (1999-2005)

	(XIX)	(XX)	(XXI)	(XXII)
$\ln(\text{GDP}_{q1})$	0.15	-0.03	0.25	0.53
	(0.44)	(-0.01)	(0.64)	(0.84)
$\ln(\text{GDP}_{q2})$	0.23	1.09	0.64	1.82
	(0.28)	(0.14)	(0.69)	(1.19)
$\ln(\text{GDP}_{q3})$	-1.01	-1.76	-1.59	-3.28
	(-1.56)	(-0.31)	(-2.20)	(-2.77)
$(\ln(\text{GDP}_{q1}))^2$		0.01		
		(0.06)		
$(\ln(\text{GDP}_{q2}))^2$		-0.04		
		(-0.11)		
$(\ln(\text{GDP}_{q3}))^2$		0.03		
		(0.13)		
$\Delta Inc_{q2,q5}$	3.40	3.40	4.10	6.54
	(4.05)	(3.88)	(4.35)	(4.18)
Exp.*	0.43	0.44	0.42	0.38
	(3.54)	(3.16)	(3.04)	(1.68)
M. Inc.*	0.03	0.02	0.11	0.36
	(0.20)	(0.17)	(0.74)	(1.51)
NS Inc.*	-0.16	-0.16	-0.21	-0.35
	(-0.95)	(-0.91)	(-1.15)	(-1.15)
Chile**	-0.98	-0.98	-1.20	-2.11
	(-2.92)	(-2.83)	(-3.20)	(-3.43)
Spain**	-0.82	-0.83	-0.74	-0.43
	(-2.49)	(-2.39)	(-2.02)	(-0.71)
Thailand**	-1.06	-1.06	-1.18	-1.86
	(-3.26)	(-3.17)	(-3.21)	(-3.09)
Constant	7.59	8.08	8.40	8.18
	(8.75)	(2.14)	(8.65)	(5.03)
R^2	0.91	0.91	0.91	0.85
N	93	93	93	92

Dependent variable: log infant mortality (XIX), (XX);
log child mortality (XXI); log net child mortality (XXII).
t-statistics in parentheses.
*Dummy variables for income distribution definitions.
**Dummy variables for influential observations.

Table 1.6 shows estimates for regressions including the real income of each of the lowest three quintiles, the income distance between the highest and the second lowest quintile, and the prevalence of HIV. Higher HIV prevalence is statistically significantly partially correlated with higher infant mortality. The association between higher infant mortality and higher income distance decreases by 15% from 3.31 to 2.84, but remains statistically significant with a *t*-statistic of 3.24.[10]

1.4.1 Relative Deprivation

We use the common definition of relative deprivation as the difference between an individual's income and the income of his/her reference group (Jones and Wildman, 2008) and follow Runciman's (1966) duality of reference groups: individuals in the same neighborhood/social class and individuals serving as role models from higher social classes. There is a large body of literature investigating relative deprivation within a social class and between similar social classes. Negative effects of relative deprivation between and within social classes on heath are well documented. The Whitehall study (Marmot et al., 1984), e.g., finds mortality differences of British civil servants relative to their rank. Eibner and Evans (2005) use interview surveys to investigate relative deprivation and show, among other effects, higher probabilities of death, and higher body mass index of relatively deprived individuals. The adverse effects on individual health are usually explained by increased stress and risk taking behavior.

However, relative deprivation in such sense does not fully explain our or Waldmann's (1992) results. It seems unlikely that the income distance between individuals with very limited personal contact causes health affecting stress or risk taking behavior. This might explain why existing research on relative deprivation usually neglects the top end of the income distribution as a possible reference group. However, our findings indicate that the income distance between individuals from distant social classes does matter for health outcomes. Our hypothesis is that, in addition to stress or health risk taking behavior of relatively deprived individuals, their consumption patterns might change depending on the concentration of income. Individuals might want to imitate the consumption patterns of high-income earners to increase their self-esteem. A high concentration of income at the top end of the income distribution leads to a visible divergence of consumption patterns between the rich and the non-rich. The non-rich might therefore increase their consumption of positional goods as

[10] Note that we have to compare results of regression (B) to regression (A). The specification of regression (A) is identical to regression (XIX) in Table 1.5, but the estimates differ slightly because of a reduced sample size (89 countries instead of 93).

they wish to appear to be rich. Frank (2003) finds a connection between higher expenditure for positional goods and higher income inequality. He argues that concerns about relative position lead to too much expenditure on positional goods and too little on nonpositional goods (Frank, 2005). A higher expenditure for positional goods implies a lower budget for nonpositional goods. Frank and Sunstein (2001) lists health care among these nonpositional goods.

We are not able to directly test either of the hypotheses concerning the reference group underlying relative deprivation. However, our regression estimates support Runciman's (1966) duality of reference groups: (a) we find indication for similar social classes as the reference group: the estimates for the income distance to the fourth quintile become statistically more significant the closer the quintile of comparison. The distance between the lowest and the fourth quintile of the income distribution in Table 1.3 is not significant (t-statistic of 0.96): the poor consider individuals in the fourth income quintile not to be their reference group. On the contrary the income distance to the fourth quintile becomes statistically significant if compared with the second (t-statistic of 2.08) or third quintile (t-statistic of 2.16) of the income distribution; (b) in addition there are also indication for the rich as the reference group: the estimates for a link between higher infant mortality and higher income inequality are more significant if we use comparisons of lower quintiles with the highest quintile, not with the fourth quintile. Table 1.3 shows that both, coefficient and significance of the income distance of each of the two lowest quintiles compared with the wealthiest quintile are higher than the same comparisons with the second wealthiest quintile. Our results suggest that both reference groups, similar social classes and the rich, matter for relative deprivation.

Eibner and Evans (2005) show that relative deprivation affects health via risky behavior such as increased smoking and poor eating or exercising habits. We use the prevalence of smoking among female adults, and female obesity as possible transmission channels of relative deprivation. Both variable are at the same time directly related to the risk of infant mortality. Regression (D) and (F) show estimations for the relationship between obesity/smoking and infant mortality in our data set. The results show no statistically significant partial correlation between infant mortality and either obesity or smoking, nor do obesity or smoking explain the health-inequality relationship.[11] We therefore

[11] Our results are surprising as the relationship between obesity/smoking and infant mortality is well documented in existing literature (e.g. Kleinman et al. (1988) relate 10% of infant deaths to maternal cigarette smoking; Chen et al. (2009) find an increased probability of neonatal death and overall infant death for obese women compared to normal weight woman of between 40% and 180%).

fail to identify smoking or obesity as mediators of relative deprivation.

Table 1.6: Infant Mortality, Per Capita Income, Income Distribution, HIV, and Risk Factors, OLS Estimates (1999-2005)

	(A)	(B)	(C)	(D)	(E)	(F)	(G)
$\ln(\text{GDP}_{q1})$	0.10	-0.03	0.15	0.20	0.18	0.17	0.11
	(0.28)	(-0.07)	(0.44)	(0.55)	(0.53)	(0.50)	(0.31)
$\ln(\text{GDP}_{q2})$	0.33	0.52	0.23	0.19	0.05	0.01	0.07
	(0.37)	(0.60)	(0.28)	(0.22)	(0.06)	(0.01)	(0.08)
$\ln(\text{GDP}_{q3})$	-1.06	-1.12	-1.01	-1.02	-0.87	-0.79	-0.81
	(-1.57)	(-1.71)	(-1.56)	(-1.57)	(-1.37)	(-1.24)	(-1.25)
$\Delta Inc_{q2,q5}$	3.31	2.84	3.40	3.40	3.23	3.00	2.36
	(3.76)	(3.24)	(4.05)	(4.02)	(3.86)	(3.56)	(2.66)
HIV		2.13					2.24
		(2.38)					(2.51)
Obesity (f)				0.15			0.29
				(0.42)			(0.81)
Smoking (f)						-0.74	-0.72
						(-1.60)	(-1.51)
Exp.*	0.44	0.42	0.43	0.42	0.41	0.39	0.35
	(3.40)	(3.38)	(3.54)	(3.29)	(3.44)	(3.23)	(2.68)
M. Inc.*	0.04	0.02	0.03	0.03	0.03	0.07	0.07
	(0.28)	(0.13)	(0.20)	(0.25)	(0.25)	(0.58)	(0.55)
NS Inc.*	-0.16	-0.11	-0.16	-0.15	-0.24	-0.25	-0.19
	(-0.93)	(-0.70)	(-0.95)	(-0.92)	(-1.42)	(-1.51)	(-1.11)
Chile**	-0.97	-0.91	-0.98	-0.99	-0.96	-0.80	-0.76
	(-2.85)	(-2.74)	(-2.92)	(-2.93)	(-2.94)	(-2.37)	(-2.22)
Spain**	-0.82	-0.80	-0.82	-0.80	-0.80	-0.71	-0.66
	(-2.44)	(-2.46)	(-2.49)	(-2.41)	(-2.46)	(-2.18)	(-2.04)
Thailand**	-1.06	-1.00	-1.06	-1.05	-1.06	-1.11	-1.02
	(-3.18)	(-3.10)	(-3.26)	(-3.18)	(-3.31)	(-3.50)	(-3.22)
Constant	7.72	7.93	7.59	7.72	7.89	7.80	8.39
	(8.65)	(9.10)	(8.75)	(8.31)	(8.64)	(8.61)	(8.56)
R^2	0.91	0.92	0.91	0.91	0.92	0.92	0.92
N	89	89	93	93	86	86	83

Dependent variable: log infant mortality; t-statistics in parentheses.
*Dummy variables for income distribution definitions.
**Dummy variables for influential observations.

1.4.2 Public Policy

Health outcome and income inequality will be affected by public policy, which we allow for by including public health care expenditure, education outcomes, and access to basic sanitation and safe water in our regressions.

Countries with a more egalitarian income distribution may at the same time have better public health care systems. The idea is that, rather than income inequality itself, differences in the extent of and access to public health care services cause the divergence of infant mortality. In countries with large differences between the income share of the lowest and the highest quintile of the income distribution, private expenditure on health care tends to be high and public expenditure tends to be low. Health care expenditure seems thus to be part to the link between income inequality and infant mortality: when we include the share of GDP that is privately and publicly spent for health care in our regression the association becomes statistically insignificant (regression (I) in Table 1.7). The coefficient on the income distance between the richest and the poorest quintile decreases from 2.84 to 2.19, its t-statistic from 3.24 to 2.53. The estimate for the association of public health care expenditure and infant mortality is statistically significant, negative, and large: -7.86 with a t-statistic of -3.29. The estimation implies an association of an increase in public health care expenditure of one percentage point of GDP with a decrease in infant mortality of nearly eight percent. On the contrary, private health care expenditure does not have a statistically significant impact on infant mortality. These results indicate that the correlation between infant mortality and income inequality arises partly because income inequality is associated with low public expenditure and presumably limited access to health care for the poor. Public health care expenditure or access to health care services in general have already been used in attempts to explain the relationship between bad health outcomes and high income inequality. However, previous research has failed to explain such relation with cross-national differences in private and public health care funding. Le Grand (1987) cannot find a statistically significant relationship between public or private expenditure on medical care that could account for the relationship between inequality and health. Adda et al. (2003) provide indications that medical insurance coverage and access to medical services might not be the reason for a link between low socioeconomic status and bad health outcomes. Our findings suggest that differences in health care expenditure can partially explain the association of income inequality and infant mortality.

Education has often been found to have a strong effect on infant mortality. Regression (K) shows the results of including female literacy and the ratio of total female enrollment in primary school to the female population of the age group into a regression of the log of average income by quintiles and income distance on the log of infant mortality. The inclusion of female literacy has the expected effect: higher rates of female literacy are associated with lower infant mortality, the results are statistically significant (t of -2.25). Given the estimates, we expect infant mortality to be 0.6 percent lower when literacy increases by one percentage point. Female primary school enrollment has a statistically insignificant negative partial correlation with infant mortality (the association becomes significant when school enrollment is added without literacy). Including female literacy and primary school enrollment in the regression has an impact on the estimated coefficient on the income distance (3.29 versus 2.61), but income inequality remains statistically significant (t-statistic of 2.77).

The estimates in regression (M) show a statistically significant and negative relationship between infant mortality and access to basic sanitation in rural areas. The same holds for basic sanitation in urban areas when sanitation in rural areas is not included in the regression. Access to safe water in rural and urban areas is statistically not significant in regression (M), but are statistically significant when added separately into a regression including average income per quintiles and income inequality. The results suggest that an increase in the access to basic sanitation in rural areas by one percentage point corresponds to an infant mortality that is 0.5 percent lower. The insertion of access to basic sanitation and safe water into our base regression leads to a decline of the coefficient on the income distance from 3.36 to 1.86, the t-statistic decreases from 3.92 to 2.03. Basic sanitation and safe water explain almost one half of the coefficient on income inequality. However, the association of higher income inequality and higher infant mortality remains statistically significant.

In regression (N) we include all additional explanatory variables related to public policy. The partial correlation between higher income inequality and higher infant mortality becomes small and statistically insignificant. Around 20% of the decline in the significance of the relationship is due to the loss of 8 out of 93 countries in the data set, when we reduce the sample size from 93 to 85 countries, the coefficient on the income distance declines from 3.40 to 3.10, the t-statistic decreases from 4.05 to 3.25. The additional decrease in the coefficient on the income distance from 3.10 to 0.53 is related to the inclusion of the additional variables. The coefficient of the partial correlation does not only become smaller, but is also statistically not significant (t-statistic to 0.52). However, the fact that there are many correlates of income distribution does not clarify whether the controls we added are confounders or mediators in the relationship between income distribution and health outcomes (Wilkinson and Pickett (2006) discuss the usage of controls in regressions including income distribution and health outcomes). E.g. our hypothesis concerning an increased consumption of positional goods in societies with an unequal distribution of income would imply a lower consumption of nonpositional goods such as education or health care. The controls used would thus rather be mediators than confounders.

Table 1.7: Infant Mortality, Per Capita Income, Income Distribution, and Public Policy, OLS Estimates (1999-2005)

	(H)	(I)	(J)	(K)	(L)	(M)	(N)
ln(GDP$_{q1}$)	-0.03	-0.09	0.18	-0.15	0.11	-0.20	-0.54
	(-0.07)	(-0.26)	(0.52)	(-0.45)	(0.33)	(-0.58)	(-1.60)
ln(GDP$_{q2}$)	0.52	0.50	0.05	0.00	0.31	0.10	0.30
	(0.60)	(0.61)	(0.06)	(0.00)	(0.36)	(0.12)	(0.39)
ln(GDP$_{q3}$)	-1.12	-0.97	-0.85	-0.36	-1.05	-0.39	-0.20
	(-1.71)	(-1.56)	(-1.27)	(-0.57)	(-1.61)	(-0.60)	(-0.33)
$\Delta Inc_{q2,q5}$	2.84	2.19	3.29	2.61	3.36	1.86	0.53
	(3.24)	(2.53)	(3.72)	(2.77)	(3.92)	(2.03)	(0.52)
HIV	2.13	3.33					3.47
	(2.38)	(3.61)					(3.84)
Public Health		-7.86					-5.95
		(-3.29)					(-2.58)
Private Health		-0.38					-0.21
		(-0.15)					(-0.09)
Literacy (f)				-0.55			-0.61
				(-2.25)			(-2.23)
School (f)				-0.32			-0.01
				(-1.13)			(-0.03)
Sanitation (u)						0.07	0.48
						(0.22)	(1.49)
Sanitation (r)						-0.51	-0.44
						(-2.03)	(-1.51)
Water (u)						-0.17	-0.16
						(-0.31)	(-0.29)
Water (r)						-0.36	-0.21
						(-1.18)	(-0.67)
Exp.*	0.42	0.37	0.45	0.50	0.42	0.41	0.40
	(3.38)	(3.05)	(3.63)	(4.38)	(3.38)	(3.45)	(3.49)
M. Inc.*	0.02	0.10	0.01	-0.04	0.02	-0.08	-0.03
	(0.13)	(0.75)	(0.11)	(-0.35)	(0.12)	(-0.62)	(-0.22)
NS Inc.*	-0.11	-0.01	-0.22	-0.15	-0.17	-0.30	-0.17
	(-0.70)	(-0.06)	(-1.26)	(-0.93)	(-1.01)	(-1.87)	(-1.00)
Chile**	-0.91	-0.89	-0.96	-0.86	-0.99	-0.94	-0.79
	(-2.74)	(-2.83)	(-2.83)	(-2.72)	(-2.92)	(-2.88)	(-2.60)
Spain**	-0.80	-0.74	-0.85	-0.88	-0.82	-0.76	-0.74
	(-2.46)	(-2.40)	(-2.56)	(-2.91)	(-2.46)	(-2.41)	(-2.61)
Thailand**	-1.00	-1.04	-1.08	-0.97	-1.05	-0.77	-0.77
	(-3.10)	(-3.37)	(-3.31)	(-3.21)	(-3.20)	(-2.38)	(-2.65)
Constant	7.93	7.96	7.53	7.28	7.63	7.66	8.09
	(9.10)	(9.59)	(8.55)	(8.91)	(8.61)	(8.55)	(8.74)
R^2	0.92	0.93	0.91	0.93	0.91	0.92	0.94
N	89	89	91	91	91	91	85

Dependent variable: log infant mortality; t-statistics in parentheses.
*Dummy variables for income distribution definitions.
**Dummy variables for influential observations.

1.5 Conclusion

This paper draws three main conclusions. First, income inequality has a statistically significant and positive association with infant mortality, even if we control for the absolute income of each of the three lower quintiles of the income distribution. The correlation between income distribution and health cannot be explained solely by the non-linear relationship between mortality and income distribution. Second, we find indications that relative deprivation might occur as individuals make comparisons with two types of reference groups: similar social classes (e.g. neighborhood) and the rich. Third, the correlation between higher income inequality and higher infant mortality seems to arise, in part, because income inequality is at the same time associated with low public expenditure on health care, education, and water and sanitation infrastructure. These factors might either be mediators between income inequality and health or confounders of the relationship.

The limited size of our data set did not allow a separate analysis of countries by development status. This might be possible for developed countries if regional data is available. Although using regional data decreases the area in which income inequality is measured, the discussion of possible explanations might gain from a separate analysis: not all possible explanations might apply to all countries across development status (e.g., unwise consumption of status goods could be relevant for developed countries, but most likely not for developing countries). Using developed countries only could also offer the possibility to directly analyze differences in the allocation of consumption depending on income inequality. We would be interested in learning about them as high consumption of positional goods in unequal countries might support the hypothesis of relative deprivation using the rich as the reference group.

Bibliography

[1] Adda, Jérôme, Tarani Chandolab, and Michael Marmot (2003). Socio-economic status and health: causality and pathways, *Journal of Econometrics*, 112, pp. 57–63.

[2] Atkinson, Anthony. B. and Andrea Brandolini (2001). Promise and Pitfalls in the Use of "Secondary" Data-Sets: Income Inequality in OECD Countries as a Case Study, *Journal of Economic Literature*, Vol. XXXIX, No. 3, pp. 771–799.

[3] Chen, Aimin, Shingairai A. Feresu, ; Cristina Fernandez, ; Walter J. Rogan (2009). Maternal Obesity and the Risk of Infant Death in the United States, *Epidemiology*, Vol. 20, No. 1, pp. 74-81.

[4] Cook, R. Dennis, and Sanford Weisberg (1983). Diagnostics for heteroscedasticity in regression, *Biometrika*, 70(1), pp. 1–10.

[5] Deaton, Angus (2001). Relative deprivation, inequality and mortality. NBER Working Paper 8099.

[6] Deaton, Angus (2003). Health, Inequality, and Economic Development, *Journal of Economic Literature,* Vol. XLI, pp. 113–158.

[7] Deininger, Klaus, and Lyn Squire (1996). A New Data Set Measuring Income Inequality, *The World Bank Economic Review*, Vol. 10, No. 3, pp. 565–591

[8] De Vogli, Roberto, Ritesh Mistry, Roberto Gnesotto, and Giovanni Andrea Cornia (2005). Has the relation between income inequality and life expectancy disappeared? Evidence from Italy and top industrialised countries, *Journal of Epidemiol and Community Health,* No. 59, pp. 158–162.

[9] Eibner, Christine E., and William N. Evans (2005). Relative Deprivation, Poor Health Habits and Mortality, *Journal of Human Resources*, Vol. 40, pp. 591–620.

[10] Frank, Robert H., and Cass R. Sunstein (2001). Cost–Benefit Analysis and Relative Position, *University of Chicago Law Review*, 68(2), pp. 323–375.

[11] Frank, Robert H. (2003). Are Positional Externalities Different From Other Externalities?, Draft for presentation at The Brookings Institution.

[12] Frank, Robert H. (2005). Positional externalities cause large and preventable welfare losses, *American Economic Review Papers and Proceedings*, Vol. 95., No. 2, pp. 137–151.

[13] Gerdtham, Ulf-G., and Magnus Johannesson (2004). Absolute Income, Relative Income, Income Inequality, and Mortality, *Journal of Human Resources*, Vol. 39, No. 1, pp. 228–247.

[14] Gravelle, Hugh (1998). How much of the relation between population mortality and unequal distribution of income is a statistical artefact?, *British Medical Journal*, 314, pp. 382–385.

[15] Jones Andrew M., and John Wildman (2008). Health, income and relative deprivation: evidence from the BHPS. *Journal of Health Economics*, 27, pp. 308–324.

[16] Judge, Ken, Jo-Ann Mulligan, and Michaela Benzeval (1998). Income Inequality and Population Health, *Social Science and Medicine*, 46, pp. 567–679.

[17] Kawachi, Ichiro, and Bruce P Kennedy (1999). Income Inequality and Health: Pathways and Mechanisms, *Health Services Research*, No. 34:1, pp. 215-227.

[18] Kleinman, Joel C., Mitchelle B. Pierre Jr.; Jennifer H. Madans, Garland H. Land, and Wayne F. Schramm (1988). The effects of maternal smoking on fetal and infant mortality, *American Journal of Epidemiology*, 127(2), pp. 274–282.

[19] Le Grand, Julian (1987). Inequalities in health: some international comparisons, *European Economic Review*, Vol. 31, pp. 182–191.

[20] Lynch, John, George Davey Smith, Marianne Hillemeier, Mary Shaw, Trivellore Raghunathan, and George Kaplan (2001). Income Inequality, the Psychosocial Environment, and Health: Comparisons of Wealthy Nations, *Lancet,* 358, pp. 194–200.

[21] Marmot, Michael, and Martin Bobak (2000). International Comparators and Poverty and Health in Europe, *British Medical Journal,* 321, pp. 1124–1128.

[22] Marmot, Michael, Martin Shipley, and Geoffrey Rose (1984). Inequalities in Death-Specific Explanations of a General Pattern, *Lancet,* 1:8384, pp. 1003–1006.

[23] Mellor, Jennifer M., and Jeffrey Milyo (2001). Reexamining the Evidence of an Ecological Association between Income Inequality and Health, *Journal of Health Politics, Policy and Law*, 26, pp. 487–522.

[24] Mellor, Jennifer M., and Jeffrey Milyo (2002). Income Inequality and Health Status in the United States: Evidence from the Current Population Survey, *The Journal of Human Resources*, Vol. 37, No. 3, pp. 510-539

[25] Neter, John, Michael H. Kuter, Christopher J. Nachtscheim, and William Wasserman (1996). Applied Linear Regression Models. Irwin, Chicago, IL.

[26] Pritchett, Lant, and Lawrence H. Summers (1996). Wealthier is healthier, *Journal of Human Resources*, Vol. 31, pp. 841–868.

[27] Runciman, Walter G. (1966). Relative Deprivation and Social Justice, Routledge and Kegan Paul, London.

[28] Smith, Adam (1776). The Wealth of Nations, Book 5, Chapter 2, Article 4, Section 2, "Taxes on consumable commodities".

[29] Soobader, Mah-Jabeen, and Felicia B. LeClere (1999). Aggregation and the measurement of income inequality: effects on morbidity, *Social Science and Medicine*, Vol. 48, pp. 733–744.

[30] Waldmann, Robert J. (1992). Income Distribution and Infant Mortality, *Quarterly Journal of Economics*, Vol. 107, No. 4., pp. 1283–1302.

[31] Wildman, John, Hugh Gravelle, and Matthew Sutton (2003). Health and income inequality: attempting to avoid the aggregation problem, *Applied Economics*, 35:9, pp. 999–1004.

[32] Wilkinson, Richard G. (1992). Income Distribution and Life-Expectancy, *British Medical Journal*, 304, pp. 165–168.

[33] Wilkinson, Richard G. (1997). Income, inequality and social cohesion. *American Journal of Public Health*, Vol. 87, pp. 104–106.

[34] Wilkinson, Richard G., and Kate E. Pickett (2006). Income inequality and health: A review and explanation of the evidence. *Social Science & Medicine*, Vol. 62, pp. 1768–1784.

Chapter 2

The Relative Efficiency of Public and Private Health Care

A health care system is efficient when an increase in spending results in significant improvements in the health of a population. We test the relative efficiency of public and private health care spending in reducing infant and child mortality using cross-national data for 163 countries. There are two remarkable findings: First, an increase in public funds is both, significantly correlated with a lower mortality and significantly more efficient in reducing mortality than private health care expenditure. Second, private health care expenditure is in all estimations associated with higher, not lower, mortality, although this association is often not statistically significant. The results suggest, holding total health care expenditure constant, a potential decrease in total infant mortality in the 163 countries from 6.9 million deaths (2002) to 4.2-5.3 million deaths for completely publicly financed health care systems, but an increase to 9.0-10.0 million deaths for completely privately financed health care. We can explain some of the estimated difference in the efficiency of public and private health care expenditure by geographies and socioeconomic factors such as HIV prevalence, sanitation standards, corruption, and income distribution. However, the efficiency difference remains large and statistically significant in all regressions.

2.1 Introduction

One might think that, holding everything else constant, for an additional dollar spent on health care, the health of a population improves regardless of the source of the funds. Our data appears to contradict this idea. In a comparison of the health care expenditure of 163 countries, we find a statistically significant correlation between public spending and lower infant and child mortality, but an statistically insignificant association of private health care expenditure with *higher*, not lower, infant and child mortality. The results imply a potential reduction of one million child deaths for a reallocation of one percentage point of GDP from private to public health care expenditure. The difference in efficiency of public and private funds is not only due to international divergence in per capita GDP or income distribution. Geography, demographics, HIV prevalence, sanitation standards, and corruption have some additional explanatory power for the efficiency difference. However, the estimated gap between the efficiency of public and private health care expenditure remains large and statistically significant in all regressions, t-statistics vary between -2.59 and -4.22. Moreover, the apparent adverse effect of private health care expenditure does not disappear when controls are added. The effect is statistically significant in some regressions, in some not, but the correlation between private health care spending and infant mortality is always positive.

Countries with higher GDP per capita have generally better health outcomes (see for example Pritchett and Summers, 1996) and the wealth of a nation is an almost perfect predictor for a country's health care expenditure (see Newhouse, 1977). Consequently, existing literature struggles to account health improvements to health care expenditure when controlling for income. Some studies find such relation: Hitiris and Posnett (1992) relate mortality in OECD countries negatively to total per capita health spending; Crémieux et al. (1999) confirm these results for Canada by identifying a statistically significant link between lower health care spending and increased infant mortality and decreased life expectancy. Rajkumar and Swaroop (2008) find that public health spending lowers child and infant mortalities only in countries with good governance. On the contrary, Leu (1986), controlling for income, finds no significant effect of health care expenditure on health outcomes for OECD countries. Moreover, Berger and Messer (2002) find an association of the share of public health care expenditure with *increased* infant mortality and a link between higher income inequality and lower infant mortality for OECD countries, they recommend "[...] that as countries increase the level of their health expenditures, they may want to avoid increasing the proportion of their expenditures that are publicly financed [...]". Self and Grabowski (2003) as well as Grosskopf et al. (2006) derive similar results: public funding of health expenditure does not always lead to better health outcomes. In both analyses, the authors find a dependence of the efficiency of public health care expenditure on a country's level of development: only middle-income countries benefit from public health care funding. They identify a negative link between public funds and health improvements for less developed and highly developed countries. Filmer and Pritchett (1999) cannot find a statistically significant effect of public expenditure on child mortality in a broader study of 104 countries. Others derive similar findings of a lack of significance of public health care expenditure: Kim and Moody (1992), and Musgrove (1996).

An investigation of the efficiency of different sources of health care expenditure might provide interesting insights for a discussion of health care reforms. Given the contradictory results of existing studies about the efficiency of public spending, the focus of most studies on OECD countries, as well as a lack of an explicit investigation of relative efficiency of public and private health care expenditure, we hope to add to such discussion.

2.2 Methods and Empirical Results

We begin by briefly describing our empirical strategy. We then outline our basic results and test for stability. The section closes with an interpretation of the results.

2.2.1 Estimating health care efficiency

The question we consider is whether public or private health care expenditure is more efficient in reducing infant mortality. Our empirical strategy is fairly simple: we specify multivariate regressions that try to explain the level of infant mortality and child mortality in 163 countries with per capita GDP, public and private health care expenditure and additional explanatory variables X_j. Our basic equation relates the natural log of infant mortality in a country i, to the natural log of per capita real GDP (corrected for purchasing power parity) as well as to public and private health care expenditure (as a percent of GDP) in the same country.

$$\begin{aligned}\ln(infant\ mortality_i) = {} & \beta_1 \ln(per\ capita\ GDP_i) + \beta_2\,(Public\ Hex_i/GDP_i) \\ & + \beta_3\,(Total\ Hex_i/GDP_i) + \sum \beta_{j+3} X_{ij} + \epsilon_i\end{aligned} \tag{2.1}$$

β_1 represents the effect of income per capita on infant mortality. β_2 and β_3 capture the efficiency of public and private health care expenditure in reducing infant mortality. Note that the efficiency of private health care (β_3) is estimated using total health care expenditure as a variable, not private health care expenditure. Given that total health care expenditure is the sum of public and private expenditure and given that the public expenditure is included in the regression, β_3 is invariant whether we use total or private health care expenditure. But this choice does matter for the coefficient on public health care expenditure β_2. If we include total health care expenditure in the regression, the coefficient and t-statistic associated with public health care expenditure express the *relative* efficiency of public health care expenditure compared to private health care expenditure: a coefficient of -0.10 would indicate that a reallocation of one percentage point of GDP from private to public health care expenditure would reduce infant mortality by 10%. In case we use private health care expenditure (instead of total health care expenditure), the coefficient on public health care expenditure measures the *absolute* efficiency of public expenditure: a coefficient of -0.10 would in this case indicate a decrease in infant mortality of 10% for an increase in public health care expenditure by one percentage point of GDP, holding private health care expenditure constant.

Public health care expenditure comprises public outlays from central/federal, state/provincial/regional, and local/municipal authorities, plus the health care expenditure of social security schemes operated by extrabudgetary agencies, and external resources such as grants and credits with high grant components and governmental involvement. The sources of private expenditure are mainly out of pocket, private insurance schemes, and commercial and non-profit insurance schemes; furthermore private health care expenditure includes corporate health care plans and the expenditure of non-profit institutions providing health goods and services to households.[1]

[1] Definitions are taken from the WHO World Health Report 2004. Further details can be found in its explanatory notes of the statistical annex (pp. 97-101).

2.2.2 Basic results

We estimate (2.1) using health care expenditure data for 163 countries.[2] The results indicate a higher efficiency of public health care expenditure in reducing infant mortality as compared to private health care expenditure.

The difference in the efficiency of public and private expenditure, measured in (2.1) by β_2, varies in our regressions between -0.129 and -0.149. The difference is always statistically significant with t-statistics between -3.50 and -4.10. The results imply a decrease in infant or child mortality by roughly 13% to 15% for a reallocation of one percentage point of GDP from private to public health care expenditure. Private health care expenditure is correlated with higher, not lower, infant mortality. The coefficient of private health care expenditure (β_3) varies between +0.034 and +0.052. The estimations of β_3 are not robust, t-statistics vary between +1.31 and +2.07. The results estimate an association of an increase in private health care expenditure of one percentage point of GDP with an increase in infant or child mortality between 3% and 5%. The implied efficiency of public health care expenditure suggests a decrease in infant or child mortality between 8% and 11% for an increased public expenditure of one percentage point of GDP. GDP per capita is statistically significant in all regressions and negatively correlated with infant and child mortality.

We present our basic results in Table 2.1. The dependent variable in model (I)-(III) is log infant mortality ($lnIM$), in (IV) log child mortality up to five years of age ($lnCM$). The independent variables are log per capita GDP corrected for purchasing power parity ($lnGDP$), the year, public health care expenditure as a share of GDP ($PubHex$) and either private ($PrivHex$, in (I)) or total health care spending ($TotalHex$, in (II)-(IV)), again as a share of GDP. We use log transformations of GDP per capita and infant mortality to account for their non-linear relationship - this is the standard approach in existing literature.[3] Scatter plots of real per capita GDP and infant mortality support

[2] To test our basic equation we primarily use data provided by the World Bank's World Development Indicators 2005. The 163 countries contained in our data set include all countries for which both, infant mortality and health care funding data is provided between 1999 and 2003. All other variables used, but economic growth, the share of the population above 65 years, life expectancy, population growth (all taken from the United Nations Development Programme's Human Development Index), female obesity (WHO), female literacy (World Bank, UNESCO, and CIA World Factbook), and corruption (Transparency International) are taken from the World Bank, too. An overview of all variables used - we by and large use the standard data and sources - including definition, source and the usual quantitative description can be found in the appendix.

[3] According to Filmer and Pritchett (1999) "[...] every study that has examined the issue has shown a non-linear relationship between mortality and income [...]".

the log-linear relationship for our data set. The first two regressions confirm that the results for private health care expenditure ($PrivHex$) and total health care expenditure ($TotalHex$) are identical when public health care expenditure is included in the regression. The coefficient on $PubHex$ in (I) equals the coefficient on $TotalHex$ in (II). The estimated coefficients of $PubHex$ differ in (I) and (II): the coefficient in (I) measures the *absolute effect* of public health care expenditure on infant mortality, while the coefficient on public spending in (II)-(IV) measures the *difference* between $PubHex$ and $PrivHex$ (subtract the coefficient of $PrivHex$ from $PubHex$ in (I) to obtain the coefficient of $PubHex$ in (II)). The coefficient on $PubHex$ in (I) is negative (-0.116) and statistically clearly significant (t-statistic of -4.10). The results imply that for an increase of one percentage point in public health care expenditure (i.e. instead of 3% of the GDP, the government now spends 4% for health care) we expect infant mortality to drop by roughly 11%. On the contrary the statistically insignificant coefficient on private health care expenditure (+0.034) in (I) indicates the association of an increase in the private spending of one percentage point of GDP with an increase in infant mortality of 3%.

Filmer and Pritchett (1999) find similar (but not statistically significant) results for public health care expenditure: an increase in public health care expenditure is associated with a decrease in child mortality with an elasticity of -0.135. They translate this elasticity into costs per averting a child death: public spending would need to increase in a range between $50,000 and $100,000 for each child death to be prevented. Compared to the estimated medical costs of averting a child death (between $10 and $4000), they conclude that public spending is not efficient. This conclusion is surprising, it postulates a decrease in infant mortality as the only effect of an increase in public health care expenditure. In reality an increase in public spending would not only (if at all) have an impact on infant mortality, but also on the prevalence and successful treatment of many other ill health conditions. We therefore interpret our result somewhat differently. We are concerned with the relative efficiency of public versus private health care expenditure: what would happen if the funds privately spent for health care were directed via public institutions or vice versa? The answer to this question is given by the coefficients on $PubHex$ in the regressions (II)-(IV). The coefficient of -0.142 in regression (III) for example implies a decrease in infant mortality of 14% for an increase in public health care by one percentage point of GDP and a corresponding decrease in private health care expenditure by the same amount. Our data set includes 163 countries with an aggregated population of 5,993 million people (2002). Given the underlying health care spending structure, 6.9 million newborns died before reaching the age of one year in the countries covered. Using our regression results we estimate the number of infant deaths if total health care expenditure would be publicly funded and no expenditure privately, and vice versa. Although these are purely theoretic scenarios, the results indicate the importance of the source of health care expenditure for infant mortality. Keeping total health care expenditure constant, the regression estimates in (I) imply a total infant mortality of 4.2 million per year for a 100% publicly financed health care system, and an infant mortality of 9.0 million for a 100% privately financed health care system. Roughly 40% or 2.7 million infant deaths would thus be averted in an entirely publicly funded system. A pure private system would increase infant mortality by 30%, i.e. induce 2.1 million additional infant deaths per year.[4] These extreme values are based on out-of-sample extrapolations, but they indicate the sensitivity of infant mortality to the source of health care spending. As a rule of thumb, our estimates imply a decrease in infant mortality of one million deaths for a reallocation of one percentage point of GDP from private to public health care expenditure.

Heteroskedasticity, normality, and data quality Are our findings robust? Heteroskedasticity is often a problem with cross-sectional data, as the estimated standard errors might be inconsistent. In (III) we use robust standard errors (Huber-White standard errors), a way to obtain heteroskedasticity consistent t-statistics. The standard errors of the coefficients on public and total health care expenditure remain almost unchanged. An additional indication for homoscedasticity is the result from a Cook-Weisberg test (Cook and Weisberg, 1983). The test has the null hypothesis of a homogeneous residual variance. A p-value of 0.38 leads to no rejection of the hypothesis. We analyze the normality of the residuals. A Shapiro-Wilk W test for normality with a p-value of 0.0002 rejects normality at any conventional significance level. We therefore use robust regressions with M estimators in (IV) and the following regressions. Robust regressions (we use Stata's robust regressions command rreg) use iteratively reweighted least squares with Huber and biweight functions. Weights are assigned to residuals, with a weight equal to one for small residuals and

[4] The estimation of averted and additional infant mortality obviously depends on the estimated coefficients used in the calculation. Using the lowest absolute coefficient on public health care expenditure found in our regressions (-0.068), the global infant mortality would drop to 5.3 million infant deaths (-29%) in case of 100% public expenditure. Using the highest coefficient on private expenditure (+0.054), infant mortality would increase to 10.0 million infant deaths (+35%) in an entirely privately financed system.

decreasing weights for increasing residuals. Estimates from regressions with robust standard errors (III) and robust regression (IV and all following regressions) confirm our initial results for the efficiency advantages of public health care expenditure. Furthermore the positive correlation between infant mortality and private health care expenditure becomes statistically significant at a 5% level when we use robust regressions (IV).

As the definition of infant mortality varies internationally, the comparability of infant mortality statistics has been questioned. Regression (V) in Table 2.1 repeats (IV), but with child mortality as the dependent variable. Under-five mortality includes infant mortality, hence differences in the definition of a liveborn are included in both, but less important for child mortality. The results are almost identical, we thus feel confident to use infant mortality in the following - which is also the standard measure for health outcomes in existing international comparisons of health care expenditure.[5]

The estimates in Table 2.1 show that income per capita and health care expenditures together have an explanatory power between 82% and 85%. The first scatter plot in Figure 2.1 depicts the residuals of log infant mortality on log per capita GDP, the year and total health care spending on residuals of public health care spending on those variables. Figure 2.2 shows the equivalent to Figure 2.1, but for private health care expenditure. The presence of outlying observations is evident (as already indicated by the Shapiro-Wilk W test).

Subsample stability In this section we investigate the robustness of the results reported in Table 2.1 along two dimensions: the first issue concerns the stability of our estimates with respect to leverage points and outlying observations shown in the scatter plots; the second issue involves a potential correlation of the development status and the effect of health care expenditure on infant mortality in addition to the effect captured by income per capita. Controlling for both issues confirms our initial results: public health care expenditure

[5] We replicate regression (IV) substituting child mortality by "net child mortality", which we obtain by subtracting infant mortality from child mortality (i.e. "net child mortality" is the mortality of children aged between 1 and 5 years). The results are similar to the once obtained in (I)-(IV): $PubHex$ is significant (only at a 10% level) and has a negative coefficient, $TotalHex$ is insignificant but with a positive coefficient. We furthermore find a strong and positive link between public health care expenditure and the share of the population aged 65 and over, while there is no correlation of the same share with private expenditure. However, the causation of the share of the population aged 65 and over and public health care expenditure might well run from aging to expenditure and not the other way around. Life expectancy at birth is neither correlated with public nor with private expenditure.

Figure 2.1: Infant Mortality and Public Health Care Expenditure

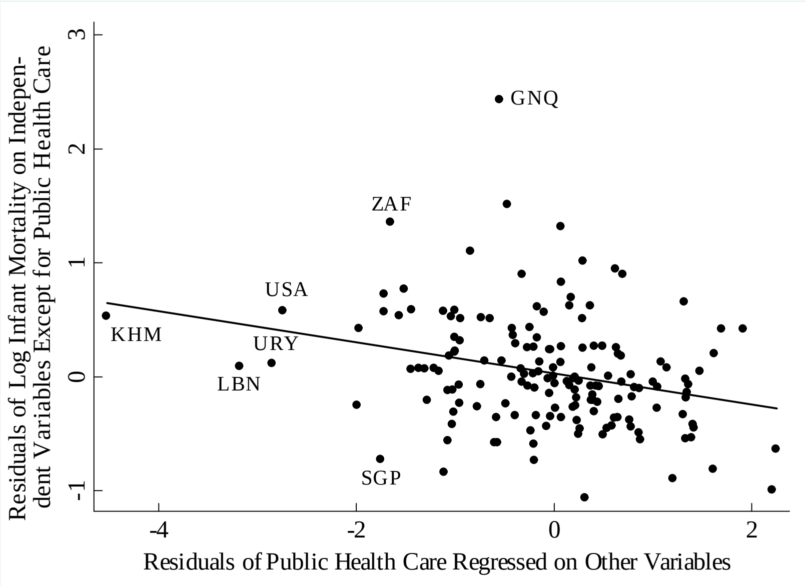


Figure 2.2: Infant Mortality and Private Health Care Expenditure

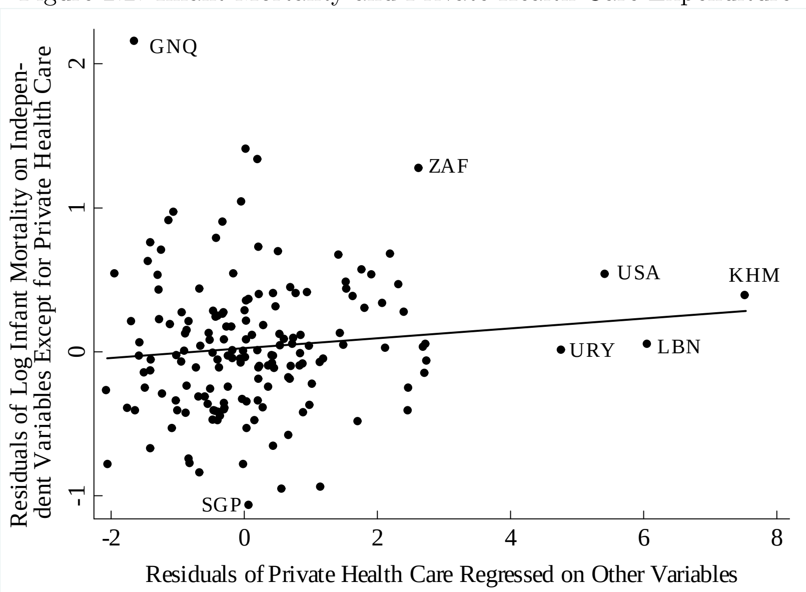


Table 2.1: Health Care Expenditure on Infant Mortality (I-III) and Child Mortality

	(I) IM	(II) IM	(III) IM[a]	(IV) IM[b]	(V) CM[b]
year	0.020	0.020	0.020	0.035	0.127
	(0.08)	(0.08)	(0.71)	(0.16)	(0.53)
$lnGDP$	-0.757***	-0.757***	-0.757***	-0.793***	-0.903***
	(-17.49)	(-17.49)	(-13.10)	(-20.10)	(-20.89)
$PubHex$	-0.116***	-0.149***	-0.149***	-0.142***	-0.129***
	(-4.10)	(-4.04)	(-3.92)	(-4.22)	(-3.50)
$PrivHex$	0.034				
	(1.31)				
$TotalHex$		0.034	0.034	0.049*	0.052*
		(1.31)	(1.46)	(2.07)	(2.02)
_cons	-29.369	-29.377	-29.377	-60.236	-243.381
	(-0.06)	(-0.06)	(-0.53)	(-0.14)	(-0.51)
R-squared	0.821	0.821	0.825	0.851	0.851
N	163	163	163	163	163

Dependent variables: ln infant mortality (I)-(III), ln child mortality (IV)
* p<0.05, ** p<0.01, *** p<0.001; t-statistics in parentheses
[a]t-statistics using Huber-White standard errors
[b]Estimates from robust regression

remains statistically significantly more efficient than private health care expenditure; private health care expenditure appears to be correlated with higher infant mortality.

A common fear in cross-national comparisons is the influence of influential observations on coefficient estimates. We use leverage values to identify leverage points and standardized residuals to identify outlying observations. Four leverage points are identified with leverage values exceeding $(4k+2)/n$ (k is the number of independent variables). Cambodia (KHM, leverage value of 0.18) is an obvious leverage point: it reports a total health care expenditure of 11.8% of GDP, while countries with similar levels of per capita income (Haiti, Sudan, Cameroon, and Georgia) spend between 3.4% and 6.8%. Further leverage points are Lebanon (LBN, leverage value of 0.11), the United States (USA, 0.11), and Uruguay (URY, 0.07) because of their high private health care expenditure (between 7.0% and 8.8%, while the average country spends 2.5%). We define outlier as observations with a standardized residual in excess of 3.5 and -3.5. Using this definition, Equatorial Guinea is the only outlier with a standardized residual of 4.6. Equatorial Guinea (GNQ) is an extreme outlier because of its abnormal high infant mortality rate: given GDP per capita and health care spending the regression estimates would predict an infant mortality of 13 deaths per 1,000 liveborns, the reported number is 103. Four of the five influential observations fit well into the regression results; it is therefore not surprising that an inclusion of dummy variables for these countries causes a drop in the significance of the estimates. However, the magnitude of the difference between $PubHex$ and $PrivHex$ remains almost unchanged and the estimates are still statistically significant (t-statistic of -3.78). Furthermore, $PrivHex$ does not lose its positive sign (Table 2.2, "Outliers"). Using Cook's D we indentify two more influential observations, South Africa (ZAF) and Singapore (SGP). Their influence on the regression estimates goes in opposite directions; an exclusion of South Africa and Singapore would have minor effects on the estimated coefficient on the efficiency gap between public and private health care expenditure.

The average infant mortality in Sub-Saharan Africa (S-S Africa) is as high as 94 infant deaths per 1,000 liveborns and compares to approximately 26 in the rest of the world. The share of total health care spending that is private amounts to 53% in Sub-Saharan Africa as compared to 40% in the rest of the world. We might therefore assume a fixed continent effect to cause the apparent difference in efficiency and the positive sign on private health care expenditure. Indeed, controlling not only for outliers, but also for continents leads to a decrease in the coefficient on $PubHex$ (i.e. the difference in the effect of public and private expenditure on infant mortality). Compared to the "Outliers" regression the efficiency gap decreases from -0.148 to -0.120, that is by 0.028 or approximately 19%. However, the estimation of the efficiency difference between public and private spending remains solid with a t-statistic of -3.25. The reduced coefficient of $TotalHex$, which decreases by 0.015 from its initial value of +0.054, accounts for more than half of the total drop in the efficiency gap. I.e., fixed continent effects explain roughly one fourth of the positive association of private health care expenditure and infant mortality. The "Continents" regression implies a statistically significantly higher infant mortality in Sub-Saharan Africa, and a lower mortality in Europe/Central Asia and East Asia/Pacific than expected given GDP per capita and health care expenditure in these regions. Including dummy variables for outlying observations and continents increases the explanatory power of the regression to 88%.

To furthermore verify the robustness of the results we use least absolute deviations (LAD) estimators that are less sensitive to outliers than OLS. The parameters are derived by minimizing the sum of the absolute deviations as opposed to the sum of the square deviations for OLS regressions. The results from these median regressions are consistent with OLS regression estimates: while the coefficient on $PubHex$ remains constant, the t-statistic decreases only slightly from -3.25 ("Continents") to -3.15. The positive coefficient on $TotalHex$

increases from +0.039 to +0.045, but stays statistically insignificant (Table 2.2, "Median Reg").

Table 2.2: Health Care Expenditure on Infant Mortality - Outliers and Continents

	Outliers[a]	Continents[a]	Median Reg
year	0.036	-0.151	-0.179***
	(0.16)	(-0.75)	(-3.57)
lnGDP	-0.794***	-0.658***	-0.697***
	(-19.43)	(-15.12)	(-15.44)
PubHex	-0.148***	-0.120**	-0.120**
	(-3.78)	(-3.25)	(-3.15)
TotalHex	0.054	0.039	0.045
	(1.72)	(1.33)	(1.47)
GNQ	2.255***	0.000	1.899***
	(5.01)	.	(12.77)
KHM	-0.014	0.454	0.330
	(-0.03)	(0.97)	(1.24)
LBN	0.000	-0.208	-0.213
	.	(-0.47)	(-1.02)
URY	-0.237	-0.353	-0.463**
	(-0.51)	(-0.85)	(-3.08)
USA	0.260	-0.064	-0.136
	(0.55)	(-0.15)	(-0.77)
East Asia&P		-0.295*	-0.349**
		(-2.50)	(-2.89)
Europe&CA		-0.428***	-0.504***
		(-4.44)	(-5.12)
Middle East		-0.056	-0.197
		(-0.45)	(-1.55)
South Asia		-0.097	-0.305
		(-0.50)	(-1.63)
S-S Africa		0.258*	0.056
		(2.42)	(0.50)
_cons	-62.625	310.298	367.225***
	(-0.14)	(0.77)	(3.67)
R-squared	0.848	0.876	
N	162	162	163

Dependent variable: ln infant mortality
* p<0.05, ** p<0.01, *** p<0.001; t-stat. in parentheses
[a]Estimates from robust regression

Previous studies find a negative relationship between public health care expenditure and health outcomes especially for highly developed countries (Grosskopf et al., 2006; Berger and Messer, 2002). Our results remain valid if we exclude low income and lower middle income countries (World Bank classifications) or if we look at high income countries or OECD countries only: public health care expenditure is statistically significantly more efficient, higher private health care expenditure is associated in all cases with higher infant mortality, although the association is often not statistically significant (Table 2.3).[6]

Overall, the results are stable: controlling for potentially inconsistent standard error estimates by heteroskedasticity consistent t-statistic estimates, robust and median regressions, adding dummy variables for outlying observations and continents as well as separating countries by development level does not fundamentally alter the results.

Table 2.3: Health Care Expenditure on Infant Mortality - Income

	High/Upper Middle	High	High (OECD)
year	0.000	0.032	0.000
	.	(0.21)	.
lnGDP	-0.684***	-0.214	0.084
	(-6.77)	(-1.03)	(0.47)
PubHex	-0.172***	-0.122*	-0.116*
	(-3.61)	(-2.25)	(-2.32)
TotalHex	0.047	0.045	0.069*
	(1.41)	(1.08)	(2.07)
_cons	9.153***	-59.870	0.802
	(10.10)	(-0.19)	(0.45)
R-squared	0.676	0.182	0.122
N	62	32	23

Dependent variable: ln infant mortality
* p<0.05, ** p<0.01, *** p<0.001
Estimates from robust regression (t-statistics in parentheses)

2.2.3 Interpretation

Why should public expenditure be more efficient than private funds spent for health care? Why should private expenditure be weakly associated with an increased infant mortality? As many factors other than health care expenditure or GDP per capita influence infant mortality we need to worry about omitted variables. We aim at finding explanatory variables to either give reasons for the higher efficiency of public expenditure or to explain the differences with omitted variable bias. We test for income distribution, reverse causation, primary health care factors such as education and sanitation standards, demographic factors, corruption, and the actual size of the health care sector (indicated by the number of hospital beds and physicians). Although some of these variables explain parts of the efficiency gap between public and private health care expenditure, the

[6] Note that as we decrease the standard deviation of GDP per capita by dropping low and middle income countries from the regression, *lnGDP* becomes insignificant and only health care expenditure remains solid.

general result remains unchanged. The magnitude of the efficiency difference between public and private expenditure, measured in (1) by β_2, remains large and varies in the regressions including further explanatory variables between -0.103 and -0.172. The efficiency difference is always statistically significant with t-statistics between -2.25 and -4.22. These results imply a decrease in infant mortality by ca. 10% to 17% if one percentage point of GDP is reallocated from private to public health care expenditure.

In the following we discuss the main variables that are either usually used to predict infant mortality or could have an impact on the efficiency of public and/or private health care spending. Most results are consistent with findings in existing literature. The following variables result as statistically significantly correlated with infant mortality (sign of correlation with infant mortality in parentheses) when separately added to regression (III): Gini index (+)[7], HIV prevalence (+), female literacy (-), improved sanitation facilities in rural areas (-), share of population aged 65 and over (-), life expectancy (-), population growth (+), corruption (+), physicians per 1,000 people (-), hospital beds per 1,000 people (-), and births attended by skilled staff (-). Variables, when added to regression (III), that increase (+) or decrease (-) the difference in efficiency between private and public health care spending by at least 10% are: income distribution (-), HIV (+), improved sanitation facilities in rural areas (-), share of population aged 65 and over (-), life expectancy (+), population growth (-), corruption (-), and the number of hospital beds per 1,000 people (-).

Table 2.4 gives an overview of the effect of additional variables on infant mortality and the efficiency of health care expenditure when added separately to regression (III). Some of the additional variables loose their significance when we control for outliers and continents, and/or when we add more than one variable at a time. Variables with very stable effects as described above are: income distribution, corruption, HIV prevalence, and improved sanitation facilities in rural areas.[8] The significance of the Gini index does not disappear when we

[7] The Gini Index is a measure of income inequality: the "Gini index [...] measures the extent to which the distribution of income (or, in some cases, consumption expenditure) among individuals or households within an economy deviates from a perfectly equal distribution. A Lorenz curve plots the cumulative percentages of total income received against the cumulative number of recipients, starting with the poorest individual or household. The Gini index measures the area between the Lorenz curve and a hypothetical line of absolute equality, expressed as a percentage of the maximum area under the line. Thus a Gini index of 0 represents perfect equality, while an index of 100 implies perfect inequality." (World Bank definition).

[8] Furthermore, life expectancy and the share of population aged 65 and over remain statistically significant. However, we have to assume endogeneity bias as the infant mortality rate

control for geographies and add HIV or sanitation standards separately, but only if we add them altogether. We present results for the stability of the explanatory variables in more detail in Table 2.7 of the Appendix.

Table 2.4: Health Care Expenditure on Infant Mortality - Additional variables
The table reports the results from regression (III) when the variables outlined in the first column are added. The second column reports the coefficient of the variable and its significance. The third and forth column present the change in the coefficient of private and public health care expenditure when the variable is added. Column 5 reports the changes in the efficiency gap and column 6 the stability of the estimated coefficient when we include further explanatory variables or dummies for outlying observations.

	Coeff.[a]	$\Delta PrivHex^a$	$\Delta PubHex^a$	ΔGap^a	Stab.[b]
Gini	0.020***	-0.025	-0.011	-0.014	.-**
Eco growth	-0.030	+0.001	-0.005	+0.006	.
HIV	0.037***	+0.004	-0.038	+0.042	***
Female obesity	0.003	+0.001	+0.003	-0.002	.
Female smoking	-0.006	+0.003	-0.008	-0.005	.
Female literacy	-0.008***	-0.002	+0.004	-0.006	.
Sanitation rural	-0.010***	-0.011	+0.025	-0.036	***
Undernourishm.	0.002	+0.001	+0.003	+0.002	.
Population 65+	-0.064***	+0.000	+0.059	-0.059	***
Life expectancy	-0.043***	-0.008	-0.050	+0.042	***
Pop growth	0.185***	-0.009	+0.014	-0.023	.
Corruption	0.086**	-0.005	+0.025	-0.030	**-***
Physicians	-0.108***	+0.010	+0.013	-0.003	.
Hospital beds	-0.029*	-0.010	+0.012	-0.021	.
Attended birth	-0.006**	-0.008	-0.002	-0.006	.

Dependent variable: ln infant mortality
* $p<0.05$, ** $p<0.01$, *** $p<0.001$
[a]Estimates when variable is added seperately to (III)
[b]Significance when further explanatory variables are added

In the following we divide explanatory variables into omitted variables (i.e. variables that replace the explanatory power of health care spending) and transmission variables (i.e. variables that explain the higher efficiency of public expenditure).

Omitted variables

Is public expenditure more efficient? The associations between health care funding and infant mortality might be due to omitted variables bias, rather than to

is used to calculate the life expectancy at birth. Furthermore ceteris paribus changes in infant mortality lead to a modification of the share of the population aged 65 and above, too.

a causal effect from expenditure to health outcome. To mitigate a possibility of misspecification as a result of excluding relevant variables, we test the robustness of our results when including further explanatory variables in regression (III). The additional explanatory variables are interesting in two ways: first, the extent and significance of their direct effect on infant mortality measured by the coefficient and t-statistic; second, the changes in the estimated effects of public and private health care expenditure on infant mortality they cause. In the case of omitted variable bias our regression estimates for β_2 and β_3 would be inconsistent. Including the omitted variable we expect the explanatory power of health care expenditure to either drop or increase. I.e. in case female education would drive infant mortality and public health care spending would be correlated with it, we would expect a drop in the explanatory power of public expenditure when female education is included in the regression. In the following we discuss the effects of variables representing income distribution, income growth, reverse causation, primary health care, demographics, and corruption.

Income distribution and income growth In case income is relevant for infant mortality, income distribution should be relevant, as well. Income is generally assumed to have a diminishing marginal effect on health; increased income inequality should therefore lead to a lower health of the population as the reduction in health of the poor is higher than the increase in health of the rich. Consequently, a negative relationship has been found between income inequality and health outcomes (e.g. Le Grand, 1987; Waldmann, 1992). Berger and Messer (2002) question these results. They suggest an association of higher inequality with *better* health outcomes. Our results confirm what seems logical: higher income inequality is correlated with higher infant mortality. The result is especially strong for countries with lower GDP per capita (in a regression of $lnGDP$, $PubHex$, $TotalHex$ and the Gini index on the log of infant mortality for countries with a GDP per capita below $10,000 the t-statistic of the Gini index is +4.85), but holds also for richer countries (t-statistic of +2.04), i.e. countries with an income per capita above $10,000. The inclusion of the Gini index as a measure of income inequality leads to an increase of the estimated efficiency of both, public and private health care expenditure. The improvement of the coefficient on private expenditure (from +0.077 to +0.052, i.e. income distribution explains one third of the positive coefficient of private expenditure) exceeds the improvement of the coefficient of public expenditure (from –0.054 to -0.065), such that the efficiency difference decreases by roughly 10%. The association of higher income inequality and higher infant mortality remains stable when we control for outliers and fixed continent effects, but becomes statistically insignificant when we include sanitation standards in the regression.

The coefficient on economic growth (represented by the average economic growth rate from 1975 to 2005) is not statistically significant, but as expected negatively correlated with infant mortality. It has nearly no effect on the magnitude of the efficiency gap between public and private health care expenditure.

Reverse causation High private health care spending might be an indicator for poor health of a society and an inefficient public health care system. Private health care expenditure should thus be correlated with poor health indicators such as infant mortality, obesity, smoking, and prevalence of HIV. However, our estimates do not support a reversed causation from bad health to higher private health care expenditure.

HIV prevalence, obesity, and smoking influence infant mortality and health care costs directly. Our results confirm an association of HIV with higher infant mortality. In contrast to existing literature we fail to find a statistically significant association of obesity or smoking with infant mortality. Crochane et al. (1978) report such association of mortality rates in 18 developed countries and calorie intake/protein consumption. Crémieux et al. (1999) find such link only for some mortality rates in Canada. The lack of a statistically significant correlation between obesity prevalence and infant mortality maybe due to the higher diversity of countries in our data set.

We control for changes in the effects of private health care expenditure on infant mortality when we add HIV prevalence, obesity, or smoking to the regression. If reverse causation - private health care expenditure is high because the population is in a bad shape - is the reason for the positive (but statistically insignificant) correlation between private health care expenditure and infant mortality, we would expect a decrease in the coefficient when we add further health indicators. The results do not confirm the reverse causation hypothesis: the coefficient and significance of private health care expenditure *increases* slightly when we add HIV prevalence, obesity, or smoking - which is not what we expected. The difference in efficiency of public and private health care expenditure remains statistically significant on a 0.1% level and of similar magnitude as in our initial regressions.

Primary health care According to the declaration of Alma-Ata in 1978, primary health care includes, along maternal and child health care, education, proper nutrition, safe water and basic sanitation. Grossman (1975) finds a positive correlation between education and health outcomes. He concludes that better education causes better health because of an increased efficiency of the health care process. Many studies confirm this link (e.g. Grossman and Kaestner, 1997). We represent education by female literacy and find the expected statistically significant correlation of better education and lower infant mortality (t-statistic of -4.38). Education hardly explains the difference in the efficiency of public and private health care expenditure. An inclusion of the female literacy reduces the difference between the coefficients by only 5%.

Increases in income lead to higher public and private expenditure on health promoting goods such as safe water and improved sanitation (Pritchett and Summers, 1996). As a consequence, some studies fail to prove an independent effect of these factors on health outcomes in regressions controlling for per capita income (Filmer and Pritchett, 1999). Our data shows such an effect. Added separately, any of the variables urban water quality, rural water quality, urban sanitation standards, and rural sanitation standards is statistically significantly correlated with lower infant mortality. Adding them altogether shows a higher importance of access to improved sanitation facilities, especially in rural areas (the access to safe water becomes statistically insignificant). The correlation, first of all of rural sanitation, is statistically very significant (t-statistic of -6.28 when added to (III), but even if controlled for outliers, continents, and other explanatory variables, the t-statistics remain around -4). Access to improved water sources and sanitation facilities explain 20% of the efficiency gap. An improvement in the access to clean water and improved sanitation facilities would not only improve population health, but also decrease income inequality (Deaton, 2003). We find the link between health, income inequality, and sanitation in our data: rural sanitation, when added to a regression on infant mortality that also includes income inequality as an explanatory variable, reduces coefficient and t-statistic of the Gini index significantly.

Including the prevalence of undernourishment in regression (III) surprisingly neither increases the explanatory power of the regression nor reduces the efficiency difference between public and private expenditure.

Demographics Demographic factors are important determinants of population health. Hitiris and Posnett's (1992) most important determinant for mortality rates is the share of the population over 65 years. We control the influence of this share as well as of life expectancy at birth on our regression results. High life expectancy indicates a healthy population; high shares of the population above 65 indicate an aging population. Both variables have a statistically very significant association with lower infant mortality (coefficients of -0.064 for the share of the population over 65 and -0.043 for life expectancy). Their effects on the efficiency gap differs: adding life expectancy, indicating a healthy society, leads to an increasing efficiency gap (coefficient of $PubHex$ increases by 30% from -0.142 to -0.184), while the inclusion of the proportion of the population above 65 leads to a lower difference in the effect of public and private expenditure (the coefficient of $PubHex$ drops by 40% from -0.139 to -0.080). The difference in the effect of public and private funds decreases by only 15% when we add both variables at the same time. The efficiency gap remains statistically significant in all cases. The results support Hitiris and Posnett (1992) finding: population aging is statistically significant in explaining public health expenditure. It is interesting to note that the effect of private health care expenditure on infant mortality is invariant to the inclusion of the population share aged 65 and over, i.e. while public expenditure on health care is systematically higher in aging societies, this is not true for private expenditure.[9] However, using aging and life expectancy as explanatory variables in regressions on infant mortality causes an endogeneity problem: infant mortality directly enters into the calculation of life expectancy; a change in infant mortality, everything else equal, alters the share of the population aged 65 and above.

As an additional demographic factor we test population growth. Population growth is positively associated with infant mortality, but looses significance if we include other variables in the regression. Population growth is also correlated with low levels of public expenditure and GDP per capita. Its inclusion reduces the efficiency difference between public and private expenditure by roughly 15%.

[9] An interpretation of these results for the efficiency of public health care expenditure is difficult. A possible explanation for the observed effects might be the following: low infant mortality is correlated with high life expectancy, as the latter is an indicator for the "general health" of a population. Let us suppose that "general health" is not due to health care expenditure, but external influences such as geography or water quality (this assumption holds in our data set: in a regression of income per capita, access to improved sanitation, clear water access, and public health care expenditure on life expectancy, all factors but health expenditure (which is not significant) are associated with higher life expectancy). In a society with better "general health" less health care expenditure is needed than in a society with worse "general health". In case two countries spent the same total amount for health care, the one with the better "general heath" level can thus use a higher fraction of the expenditure to reduce infant mortality. This causes the coefficient on public expenditure to increase in absolute terms. The share of the population above 65 is at the same time a proxy for the general health of the society (that is why the elderly share is negatively correlated with infant mortality and positively correlated with life expectancy) and for high health care costs. I.e. in a healthy society public expenditure is not only high to keep infant mortality low, but also because it allows for a large share of the population to become 65 or older. This might cause the coefficient on public expenditure to shrink in absolute terms when population aging is included.

However, the difference remains statistically significant.

Corruption Gupta et al. (2001) find a statistically significant association between higher corruption and higher infant mortality due to an adverse effect of corruption on the efficient supply of publicly provided services. Rajkumar and Swaroop (2008) measure the quality of governance using a corruption index. They show empirically that public health care expenditure has a negative association with child mortality and the efficacy of public spending in lowering mortality has a positive association with the quality of governance. We use Transparency International's Corruption Perception Index as a proxy for corruption; the regression estimates confirm a statistically significant association of corruption and higher infant mortality.[10] The results are stable even if we include further explanatory variables in the regression. Given the results of Gupta et al. (2001) and Rajkumar and Swaroop (2008), we would expect an increase in the efficiency of public health care expenditure when we include corruption as an additional explanatory variable. In our regressions this is the case for private expenditure (the coefficient improves by approximately 10%), but not for public expenditure (the effect indicated by the coefficient of public expenditure decreases by almost 30%). The efficiency gap between public and private health care expenditure remains statistically significant on a 1% level. One reason for the decrease in the coefficient of public health care expenditure is the statistically very significant correlation between lower corruption and higher public expenditure, which might exceed the efficiency reduction due to corruption.[11]

Transmission variables

Assume that our findings are robust, consistent, and not biased: public health care expenditure is more efficient than private health care expenditure. We would then like to understand *why* this is the case. We find indications for a transmission of the higher efficiency of public health care expenditure via health infrastructure investment.

Filmer and Pritchett (1999) propose a logical chain that links health care spending with health outcome. The chain is divided in three transitory steps: firstly, sector efficacy measures the extent to which health care spending is trans-

[10] Transparency International gathers data from surveys about corruption mainly in politics and business. They use a 0 to 10 scale, where 0 stands for a perception of very high corruption. In our regressions we invert the scale, such that a higher score indicates higher corruption.

[11] In a robust regression of $lnGDP$ and the corruption index on public health care expenditure as a share of GDP, $lnGDP$ has a coefficient of +0.55 (t-statistic of 3.62), while the coefficient of corruption is -0.34 (t-statistic of -4.12).

formed into effective services; secondly, net sector efficiency addresses the extent to which the provision of public services leads to a crowding-out of private services; finally, clinical effectiveness of services transforms the consumption of effective health services into health improvements, it addresses potential systematic divergence in the cost effectiveness of services offered in the public or private sector. Unfortunately we have no global cross-national data about cost efficiency of health care services. We therefore shorten the process to two transitory steps: (a) from health care spending to health care service provision; (b) from health care service provision to health care service consumption.

We approximate the transition from expenditure to service provisions by the effect of public and private health care expenditure on the number of hospital beds per 1,000 people in public or private hospitals and on the number of physicians per 1,000 people. Regression results suggest that while public and private health care expenditure are positively associated with the number of physicians (Table 2.5, "Physicians"), public expenditure is correlated with an increased number of hospital beds and private expenditure with a decreased number of hospital beds (Table 2.5, "Hospital beds").[12] Corruption and crowding-out of public and private health care expenditure could be disturbing factors in the transmission from health care expenditure to health care provision. A corrupt government might try to embezzle parts of the public budget, thereby lowering the funds that can be effectively used for health care. Surprisingly, the estimates for the association of public and private funds with the number of hospital beds and physicians do not change when we control for corruption. Private health care might crowd out public health care: firstly, higher absolute public or private expenditure might reduce the absolute amount spent from the other resource. We implicitly control for such problem in any regression, public and private expenditure are not significantly correlated. Secondly, if physicians get high incomes in private practice, this might spill over into higher salaries in the public sector, too. Even for a constant absolute expenditure, the higher labor costs would in this case lead to lower service provision. Finally, the coexistence of a public and private system could cause publicly paid physicians to moonlight in the private sector to increase their salaries. Moonlighting would not alter public expenditure, but the service provision in the public sector would decrease as physicians would spent part of their working time in the private sector. Unfortunately, our data does not allow detecting the latter two effects.

[12] Note that the low coefficient and insignificance of the estimates for $PubHex$ in the regression "Physicians" indicates that there is no systematic difference between the effects of public and private health care expenditure, i.e. an increase of each of the two leads to a similar number of additional physicians.

The second transitory step goes from health care provision (number of physicians and hospital beds) to health care service consumption. We use the share of births attended by skilled staff as index of health care service consumption. A higher number of hospital beds is associated with an increase in the share of births attended by skilled staff (Table 2.5, "Attended births (I)"). The number of physicians and the share of births attended by skilled staff are surprisingly not significantly correlated. Note that the inclusion of corruption does not change the estimated correlation between attended births and physicians/hospital beds. However, corruption does explain the explanatory power of GDP per capita for the share of attended birth in a correlation in which we control for the number of hospital beds and physicians per 1,000 people (Table 2.5, "Attended births (II)").

Given these results, infrastructure investment might be the transmission mechanism from health care expenditure to health outcomes: (a) higher public health care expenditure leads to higher health infrastructure investment (e.g. hospital beds) than private expenditure; (b) a higher number of hospital beds allows broader health care access and thus better health outcomes. The results presented in Table 2.5 should be interpreted as a first indication for the outlined transmission mechanism. Further research should aim at assessing the validity and robustness of these results.

Table 2.5: Health Care Expenditure to Health Care Service Consumption

	Hospital beds	Physicians	Attended births (I)	Attended births (II)
lnGDP	0.950*	0.661***	0.670**	0.068
	(2.33)	(6.59)	(2.82)	(0.18)
PubHex	0.806*	0.017		
	(2.59)	(0.21)		
TotalHex	-0.458*	0.170**		
	(-2.22)	(2.99)		
Hospital beds			0.161*	0.159*
			(2.12)	(2.08)
Physicians			0.286	0.304
			(1.31)	(1.36)
Corruption				-0.272
				(-1.63)
_cons	-4.330	-5.252***	90.771***	94.836***
	(-1.38)	(-6.83)	(47.87)	(37.06)
R-squared	0.272	0.528	0.340	0.308
N	94	125	83	83

Dependent variable as outlined in column header
* $p<0.05$, ** $p<0.01$, *** $p<0.001$
Estimates from robust regression (t-statistics in parentheses)

2.3 Conclusion

Health care expenditure is globally on the rise; the way to expand national health care systems was not only at the core of the debate for the US presidential elections in 2008. However, the efficiency of public and private expenditure in improving a population's health is not well understood. Two findings from an investigation of health care expenditure in 163 countries argue in favor of a publicly governed expansion of the system. First, we provide strong evidence that public health care spending is more effective in reducing infant mortality than private health care spending. Second, we cannot find any indication of an association of higher private expenditure and lower infant mortality. Our estimates imply a high sensitivity of health outcomes to the allocation of health care funds, even if we control for other variables such as geographies and socioeconomic factors. We find indications that the higher efficiency of public health care expenditure is channeled through higher health care infrastructure investment as compared to private expenditure.

A fundamental extension of our investigation is to analyze countries by development status. To increase the sample size, panel data should be used rather than cross-section regressions.

Bibliography

[1] Berger, Mark C., and Jodi Messer (2002). Public financing of health expenditures, insurance, and health outcomes, *Applied Economics*, Vol. 34, pp. 2105-2113.

[2] Cochrane, Archie L., Selwyn A. St Leger, and Fred Moore (1978). Health service 'input' and mortality 'output' in developed countries, *Journal of Epidemiology and Community Health*, Vol. 32, pp. 200-205.

[3] Cook, R. Dennis, and Sanford Weisberg (1983). Diagnostics for heteroscedasticity in regression, *Biometrika*, 70(1), pp. 1–10.

[4] Crémieux, Pierre-Yvey, Pierre Ouellette, and Caroline Pilon (1999). Health care spending as determinants of health outcomes, *Health Economics*, Vol. 8, pp. 627-639.

[5] Deaton, Angus (2003). Health, Inequality, and Economic Development, *Journal of Economic Literature,* Vol. XLI, pp. 113-158.

[6] Filmer, Deon, and Lant Pritchet (1999). The impact of public spending on health: does money matter?, *Social Science & Medicine,* Vol. 49, pp. 1309-1323.

[7] Grosskopf, Shawna, Sharmistha Self, and Osman Zaim (2006). Estimating the efficiency of the system of healthcare financing in achieving better health, *Applied Economics*, Vol. 38:13, pp. 1477-1488.

[8] Grossman, Michael (1975). The Correlation Between Health and Schooling, in N. Terleckyj, ed., *Household Production and Consumption*, Columbia University Press, New York, pp. 147-211.

[9] Grossman, Michael and Robert Kaestner (1997). Effects of Education on Health, in J.R. Berhman and N. Stacey, eds., *The Social Benefits of education*, University of Michigan Press, Ann Arbor, pp. 69-122.

[10] Gupta, Sanjeev., Hamid Davoodi, and Erwin R. Tiongson (2001). Corruption and the Provision of Health Care and Education Services, in A. K. Jain, ed., *The Political Economy of Corruption*, Routledge, London, pp. 111-141.

[11] Hitiris, Theo, and John Posnett (1992). The determinants and effects of health expenditure in developed countries, *Journal of Health Economics*, Vol. 11, pp. 173-181.

[12] Kim, K., and P. M. Moody (1992). More Resources Better Health? A Cross-National Perspective, *Social Science and Medicine*, Vol. 34(8), pp. 837-842.

[13] Le Grand, Julian (1987). Inequalities in health: some international comparisons, *European Economic Review*, Vol. 31, pp. 182-191.

[14] Leu, Robert E. (1986). The public-private mix and international health care costs, in Culyer A. J., Jönsson B., eds., *Public and private health services*, Basil Blackwell, Oxford, pp. 41–63.

[15] Musgrove, Philip (1996). Public and Private Roles in Health: Theory and Financing Patterns, *World Bank Discussion Paper*, No. 339, World Bank, Washington, DC.

[16] Newhouse, Joseph (1977). Medical care expenditure: A cross-national survey, *Journal of Human Resources,* Vol. 12, pp. 115-125.

[17] Pritchett, Lant, and Lawrence H. Summers (1996). Wealthier is healthier, *Journal of Human Resources*, Vol. 31, pp. 841-868.

[18] Rajkumar, Andrew S., and Vinaya Swaroop (2008). Public spending and outcomes: Does governance matter?, *Journal of Development Economics*, Vol. 86, No. 1, pp. 96-111.

[19] Self, Sharmistha, and Richard Grabowski (2003). How effective is public health expenditure in improving overall health? A cross-country analysis, *Applied Economics*, Vol. 35:7, pp. 835 -845.

[20] Waldmann, Robert J. (1992). Income Distribution and Infant Mortality, *Quarterly Journal of Economics*, Vol. 107, No. 4., pp. 1283-1302.

2.4 Appendix

2.4.1 Data

Our primary source of data are the World Bank's World Development Indicators 2005. The 163 countries contained in our data set include all countries for which both infant mortality and health care funding data is provided between 1999 and 2003. The following countries are included: Albania, Algeria, Angola, Antigua and Barbuda, Argentina, Armenia, Australia, Austria, Azerbaijan, Bahamas, Bahrain, Bangladesh, Barbados, Belarus, Belgium, Belize, Benin, Bolivia, Bosnia and Herzegovina, Botswana, Brazil, Bulgaria, Burkina Faso, Burundi, Cambodia, Cameroon, Canada, Cape Verde, Central African Republic, Chad, Chile, China, Colombia, Comoros, Congo (Dem. Rep.), Congo (Rep.), Costa Rica, Cote d'Ivoire, Croatia, Cyprus, Czech Republic, Denmark, Djibouti, Dominica, Dominican Republic, Ecuador, Egypt, Arab Rep., El Salvador, Equatorial Guinea, Eritrea, Estonia, Ethiopia, Fiji, Finland, France, Gabon, Gambia, Georgia, Germany, Ghana, Greece, Grenada, Guatemala, Guinea, Guinea-Bissau, Guyana, Haiti, Honduras, Hungary, Iceland, India, Indonesia, Iran (Islamic Rep.), Ireland, Israel, Italy, Jamaica, Japan, Jordan, Kazakhstan, Kenya, Korea (Rep.), Kuwait, Kyrgyz Republic, Lao PDR, Latvia, Lebanon, Lesotho, Lithuania, Luxembourg, Macedonia (FYR), Madagascar, Malawi, Malaysia, Mali, Malta, Mauritania, Mauritius, Mexico, Moldova, Mongolia, Morocco, Mozambique, Namibia, Nepal, Netherlands, New Zealand, Nicaragua, Niger, Nigeria, Norway, Oman, Pakistan, Panama, Papua New Guinea, Paraguay, Peru, Philippines, Poland, Portugal, Romania, Russian Federation, Rwanda, Samoa, Saudi Arabia, Senegal, Sierra Leone, Singapore, Slovak Republic, Slovenia, Solomon Islands, South Africa, Spain, Sri Lanka, St. Kitts and Nevis, St. Lucia, St. Vincent and the Grenadines, Sudan, Swaziland, Sweden, Switzerland, Syrian Arab Republic, Tajikistan, Tanzania, Thailand, Togo, Tonga, Trinidad and Tobago, Tunisia, Turkey, Turkmenistan, Uganda, Ukraine, United Kingdom, United States, Uruguay, Uzbekistan, Vanuatu, Venezuela (RB), Vietnam, Yemen (Rep.), Zambia, and Zimbabwe.

The following variables are taken from the World Bank's World Development Indicators: infant mortality; child mortality; public, private, and total health care expenditure; GDP per capita; Gini index; HIV prevalence; smoking prevalence; access to sanitation in rural areas; undernourishment; physicians; hospital beds; attended births. Most data is for the year 2000, if not available we used the data from 1999 to 2003. The number of hospital beds used are from the period 1995-2004. All data taken from the Human Development Indicators are from the 2007-2008 tables. They provide life expectancy data for 2000-2005, the share of the population aged 65 and older for 2005, and the growth rates of the population and GDP per capita as averages of the period from 1975 to 2005. Transparency International's Corruption Perception Index is based on an aggregation of subjective answers to polls from the last three years. Female obesity (BMI>=30 kg/m^2, age group: 15-100 years) is provided by the World Health Organization for 2005. Data for female literacy is taken from the World Bank, if not available from the UNESCO or, as a last resort, the CIA World Factbook 2007.

Table 2.6: Variable definitions and summary statistics

Variable	Definition	Obs	Mean	Std. Dev.	Min	Max	Source***
Infant mortality	Mortality rate, under-1 (per 1,000 live births)	163	43.6	39.9	2.9	167.0	WB
Child mortality	Mortality rate, under-5 (per 1,000)	163	63.8	66.7	3.9	286.0	WB
PubHex	Health expenditure, public (% of GDP)	163	3.4	1.8	0.2	8.4	WB
PrivHex	Health expenditure, private (% of GDP)	163	2.5	1.4	0.3	10.1	WB
TotalHex	Health expenditure, total (% of GDP)	163	5.9	2.1	2.0	13.1	WB
GDP per capita	GDP per capita, PPP (constant 2000 international $)	163	8417.6	9315.0	460.1	56267.5	WB
Gini	Gini Index*	125	40.8	10.0	24.7	74.3	WB
Eco growth	Annual growth rate of GDP per capita 1975-2005 (%)	157	1.2	2.3	-6.3	11.7	HDI
HIV	Prevalence of HIV, total (% of population aged 15-49)	140	2.8	6.4	0.1	38.2	WB
Female obesity	Share of adults (>=15 years) who are obese (%), female	163	17.3	13.2	0.0	76.1	WHO
Female smoking	Share of adults (>=15/18 years) who smoke cigarettes (%), female	106	13.7	10.0	0.3	33.6	WB
Female literacy	Literacy rate, adult female (% of females ages 15 and above)	147	77.0	24.9	8.5	100.0	****
Sanitation rural	Improved sanitation facilities, (% of rural population with access)	142	55.2	31.7	2.0	100.0	WB
Undernourishment	Prevalence of undernourishment (% of population)	113	19.2	16.2	2.5	73.0	WB
Population 65+	Share of population aged 65 and over (% of population)	160	7.5	5.0	1.8	19.7	HDI
Life expectancy	Life expectancy at birth, total (years)	163	67.1	11.1	40.5	82.3	HDI
Pop growth	Annual population growth 1975-2005 (%)	160	1.7	1.1	-0.4	4.3	HDI
Corruption	Corruption Perception Index**	158	4.1	2.1	1.4	9.3	TI
Physicians	Physicians (per 1,000 people)	125	1.8	1.4	0.0	6.0	WB
Hospital beds	Hospital beds (per 1,000 people)	94	4.6	3.6	0.1	17.9	WB
Attended births	Births attended by skilled health staff (% of total)	160	78.2	27.0	5.6	100.0	WB

*Area between Lorenz curve and hypothetical line of absolute equality, as % of maximum area under line

**Corruption Perception Index: perceptions of the degree of corruption as seen by business people and country analysts

*** HDI: Human Development Index (United Nations Development Programme); TI: Transparency International; WB: World Bank

**** WB, UNESCO, and CIA World Factbook 2007

2.4.2 Additional regression table

Table 2.7: Health Care Expenditure on Infant Mortality - Outliers, Continents, and additional explanatory variables[a]

	(A)	(B)	(C)	(D)	(E)
lnGDP	-0.508***	-0.504***	-0.353***	-0.637***	-0.372***
	(-7.42)	(-9.53)	(-7.76)	(-8.82)	(-5.43)
PubHex	-0.103**	-0.128***	-0.121***	-0.138*	-0.106**
	(-2.66)	(-3.58)	(-3.99)	(-2.59)	(-3.07)
TotalHex	0.035	0.044	0.045	0.054	0.045
	(1.11)	(1.54)	(1.78)	(1.36)	(1.66)
Americas	-0.301	0.225	0.111	0.454**	-0.028
	(-1.92)	(1.14)	(0.96)	(2.79)	(-0.22)
Europe&CA	-0.365*	0.022	0.180	0.153	0.000
	(-2.58)	(0.11)	(1.31)	(0.95)	.
Middle East&NA	0.000	0.190	0.000	0.286	0.152
	.	(0.98)	.	(1.58)	(1.02)
S-S Africa	0.283	0.158	-0.100	0.416	0.181
	(1.83)	(0.86)	(-0.77)	(1.57)	(1.08)
Gini	0.015**				0.004
	(3.25)				(0.91)
Eco growth	-0.020				
	(-1.04)				
Corruption	0.077**				0.095***
	(2.64)				(3.47)
HIV		0.029***			0.023***
		(4.51)			(3.57)
Female literacy		-0.004			-0.004
		(-1.55)			(-1.62)
Sanitation rural		-0.008***			-0.007***
		(-4.32)			(-4.02)
Population 65+			-0.062***		
			(-3.95)		
Life expectancy			-0.039***		
			(-7.01)		
Pop growth			0.009		
			(0.15)		
Physicians				0.013	
				(0.23)	
Hospital beds				-0.001	
				(-0.07)	
Attended births				-0.006	-0.000
				(-1.86)	(-0.17)
_cons	-157.114	8.212***	9.391***	193.287	-194.857
	(-0.43)	(19.88)	(22.62)	(0.55)	(-0.64)
R-squared	0.916	0.907	0.922	0.895	0.935
N	119	120	157	81	104

Dependent variable: ln infant mortality
* p<0.05, ** p<0.01, *** p<0.001
Estimates from robust regression (t-statistics in parentheses)
[a] Estimates for year, East Asia&Pacific, South Asia, and outliers not displayed