Child Mortality and Public Spending on Health: How Much Does Money Matter?

Deon Filmer
The World Bank

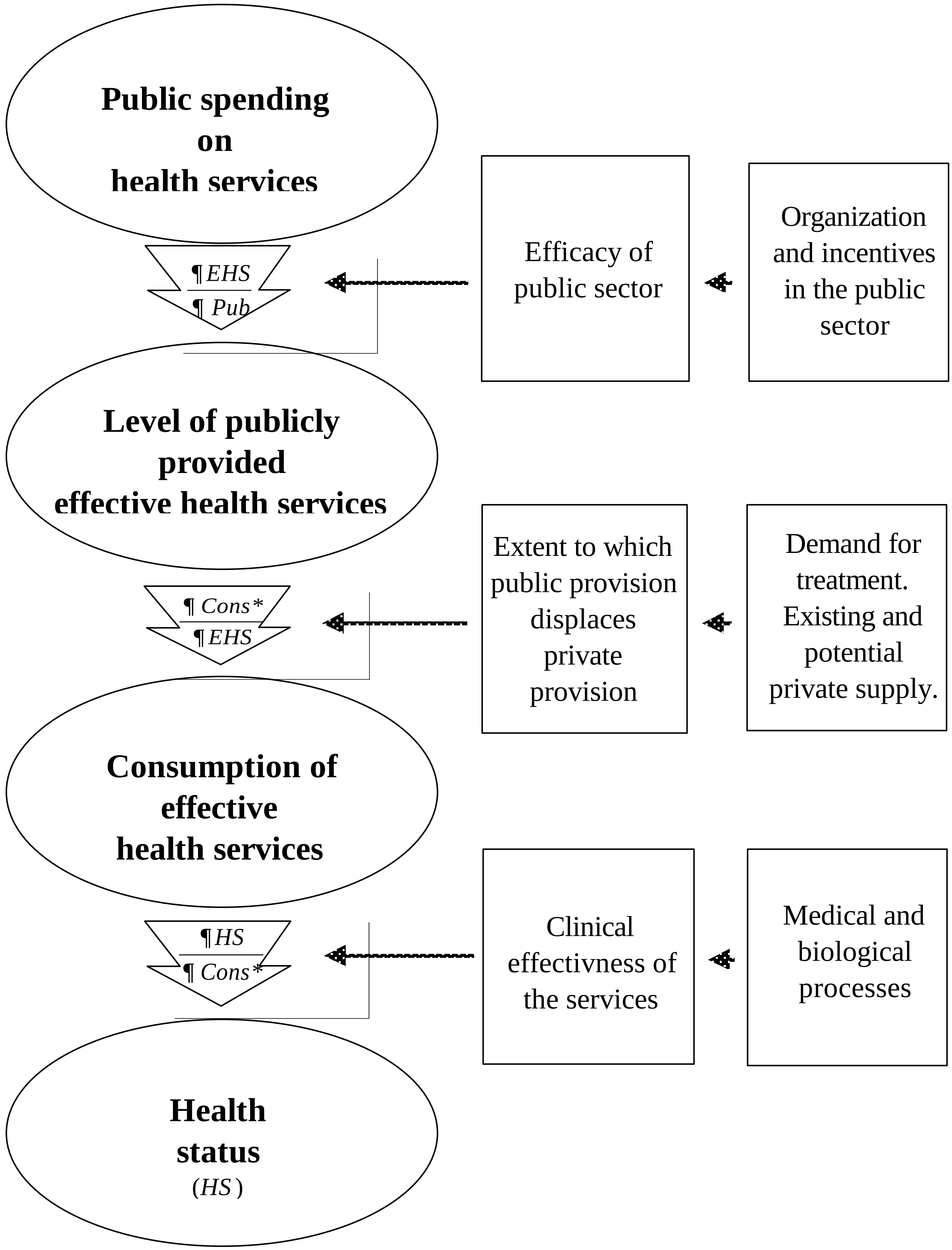
Lant Pritchett
The World Bank

October 17, 1997

Abstract
We use cross-national data to examine the impact of both non-health factors (economic, educational, cultural) and public spending on health in determining child (under-5) and infant mortality. There are two striking findings. First, ninety-five percent of cross-national variation in mortality can be explained by a country's income per capita, the distribution of income, extent of female education, level of ethnic fragmentation, and predominant religion. Second, the impact of public spending on health is quite small, with a coefficient that is numerically small and statistically insignificant at conventional levels. Independent variations in public spending explain less than one-tenth of one percent of the observed differences in mortality across countries. The estimates imply that for a developing country at average income levels the actual public spending per child death averted is $50,000 to $100,000. This stands in marked contrast to the typical range of estimates of the cost effectiveness of medical interventions to avert the largest causes of child mortality, which is $10 to $4,000. We outline three possible explanations for this divergence of the actual and apparent potential of public spending.

The authors can be reached at dfilmer@worldbank.org and lpritchett@worldbank.org

Figure 1: From Public Spending to Better Health



October 17, 1997

Child Mortality and Public Spending on Health: How Much Does Money Matter?[1]

In 1995 over 9 million children under five in developing countries died avoidable deaths. This staggering figure is more than the entire population of Sweden or of Zambia.[2] The cumulative human suffering in the individual and familial tragedies behind these statistics is overwhelming and creates a powerful impetus to action, to do *something*. This is a laudable impulse but if the desire to do something is allowed to be the enemy of good policy the result can be wasted, and possibly counter-productive, efforts. In this paper we examine cross national

[1] We would like to thank Nancy Birdsall, Jeffrey Hammer, Maureen Lewis, Samuel Lieberman, and Martin Ravaillon for helpful discussions. The findings, interpretations, and conclusions expressed in this paper are entirely those of the authors. They do not necessarily represent the views of the World Bank, its Executive Directors, or the countries they represent. The paper should not be cited without the permission of the authors.

[2] "Avoidable" deaths are defined as the excess of the average death rate for the 0-5 age group in the low- and middle-income countries of 88 per 1000 versus the level in the high-income countries, 9. Using a similar approach Gwatkin (1980) calculated the total number of deaths of under fives to be about 15 million.

differences in the widest and best measured indicators of health status: child (under-5) and infant mortality. We establish two major points about the cross national relationship between health status and public spending on health.[3]

First, the differences across countries in infant and child mortality are overwhelmingly explained by economic and social factors, that is, "development" broadly taken. The finding that "development" is strongly associated with improvements in mortality is neither surprising nor new (Caldwell, 1986, World Bank 1993). What is perhaps surprising is the *strength* of the relationship as essentially *all* (95 percent) of the cross national variations in either under-5 or infant mortality can be explained by five factors: the level of income and its distribution, the extent of female education, the extent of ethnolinguistic differences within a country, and whether it is predominately Muslim. While there are poor countries with exceptionally good health status, properly accounting for determinants besides income reduces the unexplained differences in health outcomes and leaves little to be explained by independent variations in health policy.

Second, there is an enormous gap between the apparent *potential* of public spending to improve health status and the actual *performance.* Reviews of the cost effectiveness of preventive and primary curative interventions suggest that a significant fraction of under five deaths could be avoided for as little as $10, and in many cases under $1000, per death averted (Jamison and others, 1993). However, differences in public spending on health account for essentially *none*

---

[3] We use the perhaps awkward phrase "public spending on health" throughout to avoid the ambiguity in the phrase "public health expenditures" which could mean either "expenditures on those items classified as 'public health'" or "all health expenditures by the public sector."

(0.15 percent) of the cross-national differences in health status. The extremely small actual impact of public spending that we estimate from the cross-national data implies that the typical public spending on health per child death averted is $50,000 to $100,000, a striking discrepancy between the apparent potential and actual performance.

Why is public spending on health ineffective at improving health status even though relatively cheap and effective medical interventions exist? Addressing that complicated question in detail is left to a companion paper (Filmer, Hammer, and Pritchett, 1997) and here we only suggest three likely explanations: (1) cross-national differences in the efficacy of the public sector mean that public spending on health does not always translate into a larger supply of effective health services, (2) the impact of a greater supply of effective health services in the public sector on health status depends on individual demand and market supply, and (3) public monies are spent on expensive, but ineffective, curative services.

I) Explaining cross-national variation in health status

Much of the intuitive appeal behind many proposed strategies to improve health status, such as Primary Health Care (PHC) or a "basic package" of cost-effective services, comes from the simple but powerful observation that there are countries with exceptionally good health status for their level of income (World Bank, 1997a). The relatively good health of Sri Lanka, China, Costa Rica, and Kerala, India is frequently cited as an indication of the potential benefits from PHC.[4] However, it is impossible to jump from some countries' good health outcomes to the conclusion that all (or even that *any*) of the unexplained differences in mortality are due to health policy. While it is possible that these countries' good health outcomes is due to health sector strategy, it is equally plausible that they share non-health characteristics like high levels of female education (King and Hill, 1992), better nutrition, more equal income distribution (Bidani and Ravallion, 1997) that explain their better outcomes.

To assess the *maximum* amount of the variation in health outcomes that can be explained by *independent* variations in health sector expenditures or policy, and so properly identify the "outliers," we estimate a multivariate regression that explains health outcomes without any health sector variables:

$$\textit{Health Status}_i = f\,(\textit{Income}_i\ ,\ \textit{Female Education}_i\ ,\ \textit{Income Distribution}_i\ ,\ Z_i\ )$$

where *Z* refers to a vector of additional country specific non-health related variables.

---

[4] These countries (regions) were highlighted at a conference sponsored by the Rockefeller Foundation (Halstead et al, 1985).

At this stage we resolve the problem of attribution of effects in a multivariate regression by assuming that our non-health factors may cause better health policy but that better health policy does not independently affect the included factors. That is, we assume more female education might lead to better health policy but not vice versa. This means we attribute to female education, for example, both whatever health improvements it may cause directly through behavioral changes *and* whatever indirect impact it may have through better health policy. This assumption about the causal ordering of health policy and non-health factors is seems reasonable for nearly all of the variables and we return to the only problematic case of the joint determination of health status and income below.[5]

[5] While there are a number of difficult details about the data and the empirical estimation of this equation we will not deal in detail with them here, as its main purpose is to place a (possibly quite strict) upper bound on the magnitude of the health variation across countries that can be attributed to differences in health care strategies (see appendix 1 for the details on the data).

The child mortality figures used are from a UNICEF publication (UNICEF, 1992).[6] While there are difficulties with measuring child mortality, it is arguably superior to alternative measures. Life expectancy is not reliably measured in many countries and many of the figures reported in official sources are not based actual data, but are extrapolations from child mortality and assumed life tables. More comprehensive measures of health status that go beyond mortality (such as QALYs or DALYs) are even less solidly based for cross-national comparisons. Infant mortality is perhaps more reliably measured, but fails to capture mortality from many of the health conditions of concern which are responsive to health care, such as diarrhea and respiratory infections. Especially at moderate to low levels, infant mortality is dominated by perinatal mortality.

The first empirical result is the variation in mortality associated with income. Column 1 of Table 1 shows that a large part of the variation in (the natural log of) under-5 mortality can be "explained" by (the natural log of) GDP per capita and a set of region dummy variables only.[7,8] Even excluding the regional dummy variables, 84 percent of mortality differences can be "explained" by income alone.

The second, even more striking, result in column 2 of Table 1, is that when a few other variables are included over 94 percent of the variation in the under-5 mortality rate is explained.

---

[6] Based on background work done by demographer Ken Hill (see United Nations, 1992).

[7] Dummy variables are based on World Bank regions: East Asia and Pacific, Latin America and Caribbean, Middle East and North Africa, South Asia, Sub-Saharan Africa, and (excluded from the regression) "Rest of World."

[8] In order to ensure the robustness of these, and subsequent regressions (except the median regression), the two observations with the largest impact on the parameter vector are dropped from the sample.

The level of female education, an estimate of the income distribution, a binary variable for whether the population of the country is predominantly Muslim, and the degree of "ethnolinguistic fractionalization" are each statistically significant and of plausible sign and magnitude.[9]

Table 1: Dependent variables: Under-5 mortality rate (natural log), 1990.

Column:	1	2	3	4
Dependent variable	Under-5 Mortality Rate			Infant M.R.
Method	OLS	OLS	Two-stage least squares	Two-stage least squares
GDP per capita (ln)	-.661 (10.04)	-.598 (9.28)	-.645+ (3.48)	-.462+ (2.78)
Female education		-.102 (4.03)	-.098 (3.73)	-.067 (2.46)
Income inequality		.012 (2.26)	.013 (2.08)	.003 (.505)
Percent urban		.003 (1.12)	.004 (.990)	.001 (.292)
Predominantly Muslim		.408 (3.13)	.399 (2.85)	.199 (1.53)
Ethnolinguistic fractionalization		.625 (4.14)	.626 (4.20)	.314 (2.12)
Tropical country		-.006 (.063)	-.004 (.043)	-.046 (.478)
Access to safe water		-.003 (1.16)	-.002 (-.777)	-.003 (1.44)

9 When female education, income inequality, ethnolinguistic fractionalization, or access to safe water are missing, the variable is set to zero and a dummy variable equal to one is included in the regression. Of the 104 countries included in the regression in column 2 of Table 1, 61 are not missing any of these variables, and 16 are missing only the inequality variable.

Additional variables	Dummy variables for area and a constant term.	Dummy variables for area and a constant term. Dummy variables for when female education, income inequality, ethnolinguistic fractionalization, or access to safe water are missing.		
R-Squared	.9023*	.9454	.9452	.9369
Number of observations	109	104	104	104

Notes: White heteroskedasticity-corrected t-statistics are in parentheses. Countries with largest influence on parameter vector on OLS case and hence dropped are for column (1) Hong Kong and Turkey, for columns (2) and (3) Rwanda and Korea (South), and for column (4) Sri Lanka and Turkey.
*When dummy variables for area are excluded, the R-squared for column (1) is equal to .8534.
[+]Instruments are, whether or not the country's main export is oil, and years since 1776 that the country has been independent.

The high explanatory power of these regressions is even more impressive when one considers the role of measurement error, both in mortality and in the independent variables. Imagine the hypothetical case that one was "explaining" an individual's height with a measure of height taken one minute later. If both measures were completely accurate to the recorded significant digits the regression R-squared would be 1, but if there were random error in the first measurement with variance s_v^2 then the R-squared would be $1 - s_v^2 / (s_y^2 + s_v^2)$ where s_y^2 is the true variance of height in the sampled population.

In the case of child mortality we can use a recent compilation of mortality estimates to gauge the degree of measurement error as there are multiple estimates for the same country for the same year (United Nations, 1992). The differences across estimates using different surveys and different methods for the same periods are substantial: the four estimates of child mortality per 1000 in Egypt in 1980 are 203, 167, 171, and 97, the four estimates for Ghana in 1975 are 187, 171, 130 and 125. The average coefficient of variation of these different measurements for a sample of countries with repeated measurements is 0.129 compared to a coefficient of variation for the sample as a whole of 0.880, which would suggest the *maximum* achievable R-squared, if all the true variation were explained, of 0.978.[10] The effect of measurement error in the independent variables has a similar effect in lowering the feasible R-squared.[11]

Many researchers do not trust cross-national comparisons because they doubt the data are sufficiently reliable (Srinivasan, 1994). While it is true that there are difficulties in accurately measuring child mortality, incomes, and educational levels across countries, this cannot be an explanation for a high degree of explanatory power; the more strongly one believes the data is "bad" the more puzzling is the high explanatory power of the regression.

[10] We use the coefficient of variation since the mean of the small sample for which we have repeated measurements is much larger (142) than the mean across all countries (87).

[11] At low levels, measurement error in the dependent and independent varibles are roughly additive in lowering the R-squared.

In addition to the high explanatory power, the regression is impressive as the direction and magnitude of the estimates on the variables are consistent with aggregate and household results elsewhere. The estimated elasticity of mortality with respect to income of around -0.6 is consistent with what has been found elsewhere, either using cross-sectional or time series evidence. Kakwani (1993) uses functional forms that allow for varying income elasticity in cross-national data and finds a range of elasticities between -0.5 and -0.6. Pritchett and Summers (1996) use time series on changes in income and under-5 mortality from 1960 to 1980 and find the long-run elasticity to be between -0.43 and -0.76 (depending on the instruments used in the instrumental variables estimation). Pritchett (1997) uses time-series of 22 countries with data going back to 1870 to do fixed effects estimation and finds an infant mortality elasticity with income of -0.59. Jamison, Wang, Hill and Londono (1996) combine cross section and time series data and find an income elasticity of -.65 in 1990[12]. Anand and Ravallion (1993) find that average income does have an important impact on health status, but that it operates only through its effect on the share of the population in poverty (less than a --1985 PPP-- dollar a day) while we find that adding an estimate of the proportion of population in poverty leaves our income estimate unaffected.[13]

In interpreting estimates of the impact of income on health there is a potentially serious econometric problem of reverse causation, as better health status might cause higher average

[12] They allow the income elasticity to vary across periods and find that the estimate increases from -.40 in 1960 to -.65 in 1990.

[13] In any case, the difference in the two results is not about whether income affects health status but about the particular functional form of the specification for estimating the income effect.

income. This is related to the problem of the attribution of effect between income and health policy, as better policy might cause better health which might lead to higher income.   While it is almost certainly true that better health leads to higher income at the individual level (Strauss and Thomas, 1995), the effect is less clear at the aggregate level. Moreover, it is unlikely that the mortality rate of children under five would effect higher incomes contemporaneously.   Pritchett and Summers (1996), using instrumental variables and fixed effects estimation on a panel of data show that wealthier is *causally* healthier for the cases of infant mortality, under-5 mortality, and life expectancy. While not going to the same level of detail in the estimation, we use two-stage least squares estimation (column 3 of Table 1) using as instruments for income whether or not a country's primary export is oil, and the percentage of years since 1776 that a country has been independent. Other than the much larger standard errors the results are largely unaffected, suggesting that whatever role health might have in causing higher incomes it does not affect the estimation of the cross-national impact of income on child mortality.

The results on female education are consistent with both aggregate and household level studies (King and Hill, 1993, Subbarao and Raney, 1995, Caldwell 1986, 1990, and Hobcraft, 1983). Table 2 presents the differences in under-5 mortality by educational status of the mother derived from forty-five Demographic and Health Surveys which imply that, for the average of this sample, mothers who have secondary schooling in addition to primary schooling (with usually about four years between levels) have child mortality rates 35.8 percent lower. Our cross country results, where mean female schooling is 4.97 years, imply increasing female schooling by 4 years would lead to 39.2 percent fall in under-5 mortality.

Table 2: Mean (standard deviations) under-5 mortality by education level of the mother: DHS results

	Mothers with no schooling	Mothers with primary schooling	Change (percent)	Mothers with secondary schooling	Change (percent)	Number of countries
East Asia and Pacific	119.0 (40.3)	69.9 (18.1)	41.3	39.2 (16.0)	43.9	3
Latin America and Caribbean	110.9 (52.5)	79.3 (43.1)	28.5	49.4 (26.1)	37.7	10
Middle East and North Africa	94.5 (34.3)	62.1 (14.7)	34.3	41.4 (15.6)	33.3	6
South Asia	127.7 (41.5)	84.8 (29.9)	33.6	70.9 (22.9)	16.4	4
Sub-Saharan Africa	179.8 (60.8)	139.9 (43.0)	22.2	87.5 (24.0)	37.5	22
All	144.4 (62.8)	106.5 (49.9)	26.2	68.4 (30.1)	35.8	45

Notes: Compiled from a set of DHS Final Reports from 1987 to 1995. Where there was a choice, the value for “completed primary” and/or “completed secondary” was chosen. In other cases the group may include mothers who have attended but not completed the educational level.

Income inequality and ethnolinguistic fractionalization are each associated with worse under-5 mortality. Flegg (1982) finds that the elasticity of infant mortality with respect to the Gini coefficient is 0.77 when female illiteracy is controlled for. Our results suggest that the elasticity is equal to 0.51 at mean income inequality. Our estimates imply that a country with the high inequality of Brazil (Gini of .596) compared to that of Sri Lanka (Gini of .301) could expect mortality to be 38 percent lower.

While it might appear that the cumulative empirical evidence of a redistributive impact on mortality is weaker than that of the level of income or female education, this is merely because reliable data on income distribution are scarce and recent. Nearly all estimates of the relationship between mortality and income use non-linear functional forms which imply a redistribution from rich to poor would increase the average level of mortality. For instance, Bidani and Ravallion

(1997) estimate the level (not the log) of life expectancy and infant mortality on a non-linear functional form for income that allows different elasticities for the poor and the non-poor and find strong confirmation of different impacts. The rejection of a linear functional form constitutes powerful evidence that income distribution matters.

Second, in estimating a functional form that is non-linear in income with aggregate data the specification of the relationship itself implies that an income inequality term must be included or else the regression is mis-specified. For instance, our estimates are typical in estimating a relationship between the log of average mortality and the log of average income.[14] But the aggregate data are averages of individual data and the log of average income is not the same as the average of log incomes and the difference between these two grows as the variance of income grows.[15] Therefore, if the non-linear functional form assumed for aggregates also holds in the individual level data (and there is no reason to assume that the form of the relationship between

---

[14] A variety of non-linear specifications have been used in this literature, sometimes a logistic form is imposed on mortality as the dependent variable with income in logs, sometimes just income is in logs, sometimes other non-linear specifications in income are used (e.g. quadratic).

[15] Mathematically, if one has a monotonically increasing (*f* ' > 0) but concave function *f*"<0) function then Jensen's inequality implies that Install Equation Editor and double-click here to view equation. but there is some value *x** such that Install Equation Editor and double-click here to view equation. . The difference between the two depends on the distribution of *x*, *g(x)* such that mean preserving spreads in *g(x)* increase the distance between the two. In the particular case of using the log of income the second order Taylor series expansion gives the difference between the average of ln incomes and the ln of average incomes as: Install Equation Editor and double-click here to view equation. The difference between the average of the logs and the log of the averages is sometimes *defined* as an index of inequality.

income and mortality within countries is different than the form across countries) then this implies that a measure of the distribution should necessarily be included in the regression.[16]

The fact that a country is predominantly Muslim is significantly positively associated with higher under-5 mortality, a result that has been suggested previously in the literature. Caldwell (1986) finds that no exceptionally good infant mortality health performers relative to income are predominantly Islamic, whereas many of the poor performers are. In our results, the coefficient on "predominantly Muslim" while strong for child mortality, is only insignificantly associated with higher infant mortality. This pattern of higher child mortality is consistent with the pattern of higher mortality for girls aged 1-4 in some Muslim countries such as Pakistan and Egypt (Filmer, King, and Pritchett, 1997). The exact causal mechanisms behind either of these relationships have yet to be fully investigated.

Ethnolinguistic fractionalization, which is defined as the probability any two individuals are *not* from the same ethnolinguistic group, is associated with higher mortality. For instance, moving from the level in Costa Rica (0.07) to that of Bolivia (0.70) is associated with a rise in mortality of about 40 percent. This result is new to the literature, but perhaps not surprising. It is well known that in many, although certainly not all, instances ethnolinguistic minorities have worse socioeconomic outcomes than the majority group. With a disadvantaged minority, the larger the group the higher the fractionalization and the higher the average mortality levels. Moreover, there is accumulating evidence that political fractionalization of all kinds, including that

[16] For a formal discussion of this issue, as well as a proof that an inequality measure adequately corrects for the mis-specification, see Heerink and Folmer (1994).

occasioned by ethnic conflict, makes it more difficult to achieve desirable outcomes (Alesina, Baqir, and Easterly 1997).

Perhaps surprisingly, some variables thought to be important, such as the percent of the population that is urban, whether the country is "tropical" (where populations are more exposed to certain diseases) and most surprising, the percent of the population with access to safe water are not found to have significant explanatory power for under-5 mortality. However, these estimates are conditional on the other variables in the regression (particularly income) and hence there may be only small amounts of independent variation to identify these effects.[17]

Outliers. The results in Table 1 can be used to define the "outliers," countries with under-5 mortality very different from the predicted level based on observed non-health policy characteristics. The examination of outliers has been a popular mode of argumentation for the importance of health policy, and is still used to motivate discussions of health policy strategy (Jamison and Murray, 1997, World Bank, 1997a). There are two interesting points.

First, the ranking of countries depends heavily on factors besides income. Table 3 presents the "ten best" and "ten worst" countries from the "income alone" (I) and the "income plus other variables" (II) regressions in Table 1. Since the variables in addition to income add considerable explanatory power, it is not surprising the regressions identify different outliers. Among the "ten best" countries from (I), only five reappear in the "ten best" list for (II) (Jamaica, Sri Lanka, Costa Rica, Singapore, and Trinidad and Tobago). Similarly only five of the "ten

[17] The correlations between each variable and (the natural log of) GDP per capita are in Appendix Table A2-1.

worst" countries identified with income alone reappear in the "ten worst" list for the full regression (Indonesia, Mozambique, Peru, Brazil, and Bolivia). Income alone clearly does not allow the identification of countries that are doing well (or poorly) with respect to mortality in a way that can be used to identify good or bad health policies.

Second, the difference in public spending on health between the "ten best" and the "ten worst" is very small, while the differences in public spending among countries within the "ten best" or "ten worst" groups are very large. The average fraction of GDP devoted to public spending on health among the "ten best" is 2.04 percent, versus an only slightly lower average of 1.81 percent among the "ten worst". Jamaica, with mortality 80 percent better than predicted spends 2.89 percent of GDP while Brazil, with mortality almost 50 percent worse than predicted, spends 2.97 percent. Sri Lanka, the second best, spends 1.67 percent compared to South Korea, which spends 1.88 percent . These are stark examples of the similarities. Within the "ten best" and "ten worst" groups there are large variations in the public sector health spending as a share of GDP. Costa Rica does well and has public spending of 7.5 percent of GDP while Sri Lanka and Jamaica do well spending only 2.89 and 1.67 percent of GDP respectively.

Table 3: Ten best and ten worst from residuals of under-5 mortality regressions

Specification	I	II			
Description of specification	Income only (Table 1, column 1)	Full regression, excluding public health expenditures as a share of GDP (Table 1, column 2)			
	Country	Country	Residual	Under-5 mortality rate	Pub hlth exp %GDP
Best ten	Israel	Jamaica	-.795	20	2.896
	Jamaica	Sri Lanka	-.519	35	1.673
	Sri Lanka	Costa Rica	-.463	22	7.494
	Hong Kong	Singapore	-.451	9	0.985
	Costa Rica	CAR.	-.419	169	0.954

	Singapore China Panama Mauritius Trinidad and Tobago	Kenya Zaire[+] Trinidad and Tobago Ethiopia[+~] Myanmar[+]	-.416 -.375 -.370 -.364 -.357	108 200 17 220 88	1.647 0.180 2.735 0.839 1.045
Mean (unweighted)					2.045
Worst ten	Sierra Leone Indonesia Mozambique Peru Brazil Laos Bolivia Saudi Arabia Namibia Turkey	Philippines Mozambique[+] Bangladesh* Syria[+] Bolivia Peru Brazil Indonesia Rwanda Korea (South)	.421 .432 .443 .461 .461 .502 .509 .634 .807 1.12	69 297 180 59 160 116 83 97 198 30	0.760 4.385 1.072 n/a 1.599 1.200 2.971 0.571 1.891 1.883
Mean (unweighted)					1.815
Notes: Missing values (with dummy variable flags) in the regression for (+) inequality (~) female education (*) ethnolinguistic fractionalization.					

II) Public spending on health

How much of the cross-national difference in under-5 mortality can by explained by differences in public sector health expenditures? In order to answer this question, we use a recently compiled data set of country level health expenditures. The data are the latest available update of those which appear in World Bank (1993) and were prepared for World Bank (1997a).[18] Table 4, column 1 shows the results of adding the (natural log of) public sector health expenditures as a share of GDP into the regression. Public expenditures on health do reduce under-5 mortality, although the effect is empirically small and imprecisely estimated, and is statistically significant only at the 10 percent level.[19] Public spending on health explains *very* little

[18] The sources for the updates are primarily country sources, e.g. ministries of health.

[19] We have included the log of "public spending on health as a share of GDP" as derived

of the variation in under-5 mortality over and above that which can be explained by non health-sector variables, the incremental R-squared is equal to 0.0015.[20] This means that only 0.15 percent of all mortality variations are explained by differences in spending. Introducing public sector health expenditures does not affect the other coefficients in the regression by much even though it is positively correlated with some, particularly GDP per capita.[21]

Table 4: Dependent variables: Under-5 or infant mortality rate (natural log)				
Column:	1	2	3	4
Dependent variable	Under-5 Mortality Rate			Infant M.R
Method	OLS	Median˜ regression	Two-stage least squares	Two-stage least squares
GDP per capita (ln)	-.611	-.570	-.596+	-.511+

from the aggregate "health production function" discussed below. Using the log of "public spending on health per capita" yields *the same point estimate and significance* for the effect of public spending on health (as a mathematical re-shuffling would show). The only difference in the latter specification is the coefficient on income.

[20] The incremental R-squared is generally not appropriate as a *statistical* procedure for resolving questions about the relative importance of variables, but if one has prior non-statistical knowledge about the causal ordering among the independent variables it is appropriate.

[21] In the data set the bivariate correlation coefficient between GDP per capita and public sector health expenditures as a share of GDP is 0.63.

	(9.71)	(4.58)	(3.67)	(3.39)
Public health expend. (ln of share of GDP)	-.135 (1.78)	-.090 (0.677)	-.192⁺ (.742)	-.078⁺ (.243)
Female education	-.093 (3.54)	-.076 (1.65)	-.091 (2.90)	-.061 (1.63)
Income inequality	.008 (1.28)	.010 (.911)	.008 (1.17)	.004 (.603)
Percent urban	.001 (.459)	.001 (.190)	.001 (.219)	.001 (.359)
Predominantly Muslim	.450 (3.18)	.265 (.803)	.446 (2.91)	.104 (.644)
Ethnolinguistic fractionalization	.549 (3.43)	.303 (1.04)	.534 (2.90)	.343 (1.64)
Tropical country	-.051 (.549)	-.038 (.183)	-.059 (.619)	-.165 (1.42)
Access to safe water	-.001 (.606)	-.001 (.140)	-.001 (.390)	-.003 (.753)
Additional variables	Dummy variables for area and a constant term.	Dummy variables for area and a constant term. Dummy variables for when female education, income inequality, ethnolinguistic fractionalization, or access to safe water are missing.		
R-Squared	.9469	.7839*	.9465	.9361
Number of obs.	98	100	98	98

Notes. Countries with largest influence on parameter vector in OLS case and hence dropped are for columns (1) and (3) Zaire and Korea (South), and for column (4) Israel and Turkey.
*The R-squared in column 2 is the pseudo R-squared calculated as 1 minus the ratio of the sum of absolute deviations from the predicted median to those from the actual median.
⁺Instruments are, neighbors' public health spending, neighbors' military spending, whether or not the country's main export is oil, years since 1776 that the country has been independent, neighbors' average local health services (column 4), and neighbors' average access to local health services (column 5).
~T-statistics derived from bootstrapping with 100 iterations.

Are these results on public spending on health reliable? The results on public spending are not unexpected as the small difference in public sector health spending between the best and worst "outliers" in Table 3 strongly hinted that it was unlikely to have empirically large effects. Empirical studies that include health resource variables, such as physicians, nurses, or hospital beds per capita rarely find large and significant impacts (Kim and Moody, 1992). There have been very few previous estimates of the impact of public spending in developing countries, mostly because of the lack of data, but what there is Musgrove (1996) summarizes: "multivariate estimates of the determinants of child mortality give much the same answer: income is always significant, but the health share in GDP, the public share in health spending, and the share of public spending on health in GDP never are." [22]

[22] There is a larger literature in developed countries, which also tends to find little impact of health spending, with some exceptions (Wolfe, 1986).

Three studies find a significant impact of public spending (these are Anand and Ravallion, 1993, Bidani and Ravallion, 1997, and Jamison et al, 1996). Anand and Ravallion (1993) themselves warn that "[a] number of caveats should be noted about these results. 1) They are based only on the patterns observed in this sample of 22 countries." As we emphasize below the impact of public spending on health is not an immutable parameter, but is likely to vary widely from country to country and hence the results may be sensitive to the sample used. Bidani and Ravallion (1997) use a particular functional form to desegregate the impact of various variables into its impact on the poor and non-poor. They consistently find an impact of public spending for the poor but not for the non-poor and their findings are consistent with a small or insignificant impact of spending on aggregate health status.[23] But their findings highlight the importance of considering the incidence of the health benefits: they benefit (at best) only the poor so cuts without reallocation would fall on the poor. As explored below the impact on the poor versus the non-poor is not a constant or immutable parameter, but depends on the composition and efficacy of public spending. Jamison et al (1996) use an unusual econometric procedure and find that public spending on health lowers mortality in their sample of Latin American countries.[24] When we attempt to replicate their finding using our regressions and limiting them to only Latin

[23] For infant mortality they find that the elasticity at the population weighted means is -0.213 for the poor and -0.056 for the non-poor, which is roughly comparable to our -0.078 in the aggregate.

[24] Jamison et al (1996) use a panel of data from 1960 to 1990 to first estimate (the natural log of) under 5 mortality as a function of income and year (and interactions) and whether the country was Latin American for a broad cross section of countries. They then use the anti-log of the predicted values to create a percentage deviation of the level of the actual and predicted for each country and then regress those percentage deviations on the public and private share of health in GDP for the sub-sample of Latin American countries.

America, we find a positive and significant impact of public spending on health if only income is included. However, adding our non-income control variables reduces the magnitude of the estimate by half (and it is becomes statistically insignificant).[25]

Before we move on to interpretation, there are several questions about the econometric validity of our (and previous) results that need to be addressed: robustness, measurement error, and reverse causation.

[25] The regression with just income has an R-squared of .6474 and estimates (t-statistics) for the sample of 21 countries:

$$\underset{}{\ln(u5)} = -\underset{(5.3)}{.708}\ \ln(GNPpc) - \underset{(3.7)}{.464}\ \ln PHS$$

The regression which includes our non-income control variables has an R-squared of .8467 and estimates (t-statistics) for the sample of 21 countries:

$$\ln(u5) = -\underset{(3.2)}{.759}\ \ln(GNPpc) - \underset{(1.2)}{.247}\ \ln PHS + \underset{(2.0)}{.144}\ FemEd + \underset{(.87)}{.018}\ Gini - \underset{(1.9)}{.247}\ Urb + \underset{(1.6)}{.967}\ Ethno - \underset{(.04)}{.026}\ Trop - \underset{(.76)}{.004}\ AcSaWa$$

First, are the results on public sector health spending driven by a few outlying observations or are they robust? First, in the OLS regressions the two observations with the largest impact on the parameter vector are deleted to avoid this problem. Second, the results of using a median regression estimation, which is much less sensitive to influential observations than OLS, are shown in column 2 of table 4.[26] The estimated impact of health spending is similar (but smaller -0.09 versus -0.13) and the biggest change is in the statistical significance of the estimates (the t-statistic is only .67 versus 1.78) but this is perhaps not too surprising as OLS is the more efficient estimator.

Second, reverse causation and measurement error are serious sources of bias in estimating the impact of public sector expenditures. If countries that would otherwise have high under-5 mortality devote larger amounts to public health sector spending in order to reduce mortality then the regression coefficient could be small, or even negative, and still be consistent with a powerful positive impact of public spending. Measurement error certainly exists because the accounting systems which track public spending on health are not coordinated across countries and therefore different countries are measuring different concepts. An instrumental variables estimation

[26] Median regression is equivalent to choosing the parameter vector in order to minimize the sum of absolute deviations (as opposed to the sum of squared deviations as is the case in OLS). Median regression is less sensitive than OLS to influential observations as any deviation from the regression line is "weighted" by the absolute value of the deviation, not the square of the deviation.

procedure potentially solves both of these problems. This is the case if the instruments are correlated with the public sector health spending, but not correlated with under-5 mortality except through spending, and the measurement error in the instruments is uncorrelated with that in public sector spending on health. As instruments, we use the average public sector health spending as a share of GDP and the average defense spending as a share of GDP of a country's geographic neighbors.

The two-stage least squares results presented in Table 4, column 3, are similar to the OLS results. The coefficient on public sector health spending is approximately 40 percent larger (in absolute terms), -.192 versus -.135, which is consistent with the OLS estimate being biased towards zero due to measurement error. However, the coefficient is now much more imprecisely measured and statistically insignificant from zero (t-statistic of only .742). This is a typical "good news-bad news" result from instrumental variables estimation as those who would want to argue for a large effect of public spending have a higher point estimate to support their case, whereas skeptics can point to the very low significance level.

A third test for the "robustness" of our results is to use an alternative outcome measure for health status. The results in Table 4, column 4 show that when infant mortality is used (and with our "preferred" two-stage least squares estimation) the results are again similar, although the coefficient on public sector health spending is smaller, which is consistent with the higher "genetic" component of neo-natal deaths (Rutstein, 1984), and statistically not significantly different from zero.

As an additional check we reproduced our results on a different data set constructed directly from the World Bank's Social Indicators of Development 1997 database (World Bank, 1997b). For the 173 countries that have data on under-5 mortality and World Bank atlas method GNP per capita in US dollars (which is not purchasing-power-adjusted), the explanatory power of the regression which includes variables similar to those presented above, is about 90 percent. The coefficient on income in this auxiliary regression is -0.39 which is a smaller effect than we found. This is partly explained by the fact that non purchasing-power-adjusted estimates systematically underestimate the incomes of poorer countries which would bias the estimate towards zero. In addition, random measurement error may cause a bias towards zero. When using purchasing-power-adjusted GDP per capita (from the Penn World Tables database) as an instrument for GNP per capita with the set of countries included in our main analysis, over 92 percent of the variation is explained, and the coefficient rises (in absolute value) to -0.49. Including the reported public spending on health as a share of GDP produces a point estimate of -0.07 (t-statistic of .92) for the 119 countries for which the relationship can be estimated.[27]

---

[27] The OLS regression has an R-squared on .9095 and estimates (t-statistics) for the sample of 175 countries:

ln(u5) = -.387 ln(GNPpc) + .008 FemIllit + .012 Gini - .001 Urb - .003 AcSaWa
(13.2) (3.9) (2.4) (.64) (1.9)

The regression which includes public sector health spending as a share of GDP has an R-squared of .9234 and estimates (t-statistics) for the sample of 119 countries:

ln(u5) = -.412 ln(GNPpc) + .008 FemIllit + .014 Gini - .001 Urb - .001 AcSaWa - .073 lnPHS

(12.0) (3.4) (2.8) (.23) (.54) (.92)

The two-stage least squares regression has an R-squared of .9215 and estimates (t-statistics) for the sample of 103 countries that are in our analysis:

ln(u5) = -.492 ln(GNPpc) + .008 FemIllit + .014 Gini + .002 Urb - .002 AcSaWa

(6.6) (2.6) (2.2) (.73) (.65)

All of these include dummy variables for when urban and access to safe water are missing, as well as dummies for area.

Cost of averting a death: Macro and medical estimates. Overall, these regressions show that differences in public sector spending on health do not go far in explaining why some countries have high, and others low, child mortality. Doubling the share of GDP from the mean of 2.96 to 5.92 percent would improve mortality by only between 9 (OLS) to 13 (2SLS) percent. These results can be used to calculate the effect on health status of an increase in public expenditures. In order to motivate this, think of the specification of the regressions in Table 4 columns 1 through 3 as being generated from the following simple model of the determination of aggregate health

Install Equation Editor and double-click here to view equation.

status

where H_i is public expenditures in the health sector of country *i*, *NH* is the "rest of GDP" (and includes all non-public sector health spending), *N* is the population, and *A* is a country specific

Install Equation Editor and double-click here to view equation.

factor. Dividing the numerators and denominators through by GDP and taking logs implies that is the log of health status is a function of the log of the public health expenditures as a share of GDP, non-public health sector spending as a share of GDP (i.e. with H+NH=GDP), GDP per capita, and the country specific factor. Allowing A_i to depend on the a set of observable and unobservable characteristics leads to the regression estimations reported in Table 4.[28]

[28] Because *H* and *NH* spending add to *GDP*, ß in this equation would be identified solely

through the non-linearity induced by taking logs. We therefore exclude the share of non-health spending from the regressions. When we include it, and impose the restriction that the coefficient on health and non-health sum to the coefficient on GDP per capita, the results are extremely similar. For simplicity however, we exclude it, and infer the value of the coefficient on the share of non-health spending ex-post.

Equation (1) implies an expression for the relationship between the public expenditures on

Install Equation Editor and double-click here to view equation.

health as a share of GDP and health status, that is:

The above relationship can be expressed in terms of the number of deaths averted and, in turn, can be inverted to derive the public sector health spending per additional death averted. An equivalent expression can be derived for the amount of GDP that is not public expenditures on health. From the difference one can calculate the amount that would need to be transferred from GDP to public sector health expenditures in order to avert an additional death. The values derived from these calculations, estimated at the median characteristics of countries with under-5 mortality rates greater than 20 are presented in Table 5.[29]

[29] The calculations are made for a country with an under-5 mortality rate of 126, a population of about 7 million, a GDP per capita of $1,816 (in 1985 international dollars), a crude birth rate of 34.15, and whose public health expenditures are 2.1 percent of GDP.

These results imply that the cost of averting the death of a child under-5, in terms of public sector health spending, is equal to $47,112 using the two-stage least squares point estimate. The range of the three different estimation techniques (two-stage least squares, OLS, median regression) is $47,112 to $100,927. The amount for non-health spending is between $863,000 and $1 million.[30]

Table 5: Cost of averting a death derived from different specifications of health status regressions.				
Source of estimate of a [coefficient on ln(share of public health in GDP)]	Value of a (ß)	Increasing public expenditures on health	Transferring expenditures from non-health to health	Increasing non-health GDP
Two-stage least squares	-.1917 (-.4044)	47,112	49,381	1,025,398

[30] Our confidence in these method for producing reasonable magnitudes is bolstered by the fact that Vicusi (1993) reports the estimated individual "willingness to pay" to avert a statistical death for the United States. The range of estimates is very wide, between $2.5 and $7.3 million which is between 124 and 375 times the average workers' income. There are few studies like this for developing countries, but a hedonic wage regression for Taiwan, China, gives an estimate of $650,000 which is 54 times average worker income (assuming GDP per capita is half average worker income). If we view the "non-public health share" part of our estimates to be related to private expenditures on health, then our estimate from the two-stage least squares estimate is $1 million, or 265 times the average worker income of the countries under study (again assuming GDP per capita is half average worker income) which is lower than the US range but higher than the Taiwanese estimates, but in both cases within a small multiple.

OLS	-.1354 (-.4756)	66,680	72,202	871,894
Median regression	-.0895 (-.4805)	100,850	114,193	863,107

The highest fraction of deaths, and in particular of child deaths, in developing countries are due to infectious and parasitic diseases (such as diarrheal diseases, TB, or malaria). As reported in Table 6, about 28 percent *of all deaths* in developing countries are due to infectious and parasitic diseases among children under-5. The two most prominent among those are children under-5 dying of diarrheal diseases and Acute Respiratory Infections (ARI) which each account for slightly over 10 percent of all deaths. An additional 8 percent of all deaths in developing countries are due to perinatal conditions. Since these are the largest causes of death, a substantial reduction in overall mortality must involve a fall in mortality from these causes.

Table 6: Percent of deaths by cause for population subgroups and cost per death averted

	Percent of deaths by cause (Estimates for 1985)			Estimates of cost per death averted of treatment
	Indust. mkt. economies	Developing countries		Developing countries
	All ages	All Ages	Under-5	
Infectious and parasitic	4.6	44.9	27.7	
Diarrheal Diseases	0.0	13.2	10.6	$ 1000 - $ 10,000[a]
Tuberculosis	0.2	7.9	0.8	$ 20 - $ 76[b]
ARI	3.6	16.6	11.3	$ 379 - $ 1,610[a]
Non-ARI measles and whooping cough	0.0	1.8	1.8	$ 10 - $ 561[c]
Malaria	0.0	2.6	2.0	$ 78 - $ 990[d]
Other	0.0	2.6	1.2	

Compl. of pregnancy	0.0	1.3	0.0	$ 836 - $ 3,967[e]
Perinatal conditions	0.6	8.4	8.4	
Non-communicable	78.6	29.8	0.0	
External (incl. injuries)	17.0	15.6	2.4	
Total	100	100	38.5	

Source: Causes of death calculated from Lopez (1993), sources for costs are Diarrheal-Martines and others (1993), TB-Murray and others (1993), ARI-Stansfield and Shepard (1993), Measles-Foster and others (1993), Malaria-Najera and others (1993).
Notes: a) Current US dollars, from model estimates, b) 1989 US dollars, from 3 Sub-Saharan African countries (Malawi, Mozambique, and Tanzania), c) Current US dollars, from model estimates, measles only, d) Malaria control, 1987 US dollars, from 2 studies, 1 in Sri Lanka ($78) using insecticides, one for Developing countries ($990) using vector control, e) Current US dollars, from model estimates, maternal and no-natal deaths are not distinguished in this study.

A selected group of estimates of the cost per death averted for a variety of treatments that address these conditions are presented in Table 6 (derived from Jamison and others, 1993). The cost of averting a death by treating each of the conditions are estimated using different methods (e.g. model estimates versus observed ranges) and should be used with extreme caution. The ranges for the three biggest killers are $1,000-$10,000 for diarrheal diseases, and a much more modest $379-$1,610 for ARI and $836-$3,967 for the complications of pregnancy. From these numbers it would appear that public sector intervention to treat each of these conditions could be done relatively cheaply and could save many lives.

III) How to reconcile the apparent and actual performance of public spending

The results discussed so far imply that between the "effectiveness" of public spending on health and the "cost effectiveness" of available medical interventions there is a several orders of magnitude gap in the cost per averting a child death. The cross-national regression results suggest a range between about $50,000 and $100,000 whereas the medical intervention cost effectiveness figures suggest a range between $10 and $4000 (excluding the upper limit of $10,000 on diarrheal diseases). Understanding why public spending on health has not had a strong effect on reducing mortality is crucial to designing public policy to reduce excess mortality and morbidity in developing countries. In this section we present only an outline of the analytical possibilities. A detailed exploration, including a review of the existing empirical literature, is contained in a companion paper (Filmer, Hammer, and Pritchett, 1997).

What is of interest is the impact of spending an additional public dollar on health status, or the total derivative of health status with respect to public spending. This total derivative can be analytically broken into three components, through a chain of partial derivatives, summarized in Figure 1:

* *Health production function.* The change in health status is affected as a proximate matter by changes in the consumption of various health services some of which are more, and some less, effective in improving health.

* *Net public sector impact.* However, even of particular services that are "cost effective" in improving health, this does not mean that *public spending* on those goods would be cost effective in improving health, as additional consumption in the public sector occasioned by public supply will crowd out, to differing extents, services that would have been consumed anyway. So the second term is the change in the consumption of health services (of various kinds) occasioned by an expansion in effective public supply of those services.

* *Public sector efficacy*. Moreover, the impact of public spending will depend first on the degree to which public spending is able to create effective public services. That is, in some countries a dollar spent in the public sector on health will create facilities and services which are effective at improving health status while in other countries it will create no services at all.

Since the total impact of public sector spending is the product of these three terms, if any one of them is low the total impact will be low. This framework provides a useful way of summarizing the differing conclusions and policy reform implications that various schools of thought and analysts draw from the typically low effectiveness of public spending.

Medical intervention cost-effectiveness (MICE) of publicly financed health services. Far and away the most widely held interpretation is that the composition of public spending across types of medical services and levels of medical facilities accounts for the typical low efficacy of public spending and a better allocation of spending accounts for the successes. In this interpretation, the reason why public spending on health has not produced health gains is that the bulk of public monies are spent on high cost curative services at facilities that are more complex than necessary. This, to a large extent, is the common intellectual impetus behind Primary Health Care (PHC) and the "basic package" based on MICE as approaches to health strategy, which are mainly relevant to low income countries, and to reforms to limit cost escalation (again based on MICE), which is relevant to the higher income countries. While this is logically possible, and while it is easy to demonstrate that evaluated on MICE grounds primary treatments like ORT or vaccinations are better than some types of heart surgery, the evidence for this strategy is still more deduced than demonstrated. While some of the "good outliers" certainly have something like a

PHC strategy, it has not been demonstrated that any large significant fraction of the differences in either the health performance or efficiency in health spending across countries is due to the composition of health spending.

Markets, market failures, and public interventions in health. A second school of thought suggests that what is relevant to the efficacy of public spending on health is not the composition of public spending on health services evaluated on MICE grounds, but evaluated by their economic characteristics. Various health interventions have different economic properties: some are public goods (vector control, traditional public health activities), some are private goods with little or no externalities (most curative services), while others are mostly private goods but with substantial externalities (infectious diseases activities, vaccinations). The actual impact of the public provision of any given service on health status is not measured by its effectiveness evaluated as a *medical* intervention, with and without the treatment, but by the difference in outcomes with and without the *public sector* intervention. Taking an aspirin when you have a headache might be a health intervention with high MICE, but public spending on aspirin is unlikely to be effective because it mostly displaces private spending and hence would not actually reduce the number of headaches. Vector control, on the other hand, might be much less effective than aspirin evaluated by MICE, but since vector control is a public good it is possible there would have been little or none in the absence of public sector action so this intervention might score high in Public Sector Cost Effectiveness (PSCE). The difference between MICE and PSCE depends on both the demand and the supply for particular health services, which will certainly vary not only across services but also from location to location, even within countries.

Once it is acknowledged that the consumption of health services is based on individual choices, it is possible that mortality is primarily determined by individuals' incomes and education, and that it is therefore not very sensitive to the "price" (by which we mean the total cost, including access, travel, and waiting) of health services, which is what might be influenced by the availability of clinical services through the public sector. Moreover, if private (profit and non-profit) entities are allowed to supply health services, then public sector spending is even less likely to have a large effect on aggregate mortality. In this case, public spending determines the composition of spending between the public and private sectors, but not health outcomes (Filmer, Hammer, and Pritchett, 1997).

One might question why we do not therefore estimate health status as a function of private expenditures on health. There are two reasons. First, when private expenditures decisions are voluntarily made then, in choosing to spend, people at least believe they will receive value for money. But this active choice makes private expenditures certainly, and perhaps irredeemably, endogenous and it is very difficult to separate out the impact of expenditures on health status from health status on expenditures.[31] Second, the best the data can give is total *expenditures* on health, but cannot decompose total expenditures into prices and quantities so there is no way of identifying the impact of increased quantities of privately purchased health services. Suppose for instance that the only private health expenditure was on penicillin, that the use of penicillin had a beneficial effect on stopping child deaths, and that the demand for penicillin was very unresponsive to price. Then suppose one country imported the drug at a low price while the

[31] Policy decisions regarding health spending are more exogenous and hence these factors can be separated out for public sector health spending.

other imposed a large tariff that raised the domestic prices. In this case private *expenditures* would have negative correlation with health status as higher prices would lead to a lower quantity consumed and hence worse mortality outcomes, even though total expenditures could be higher. Therefore any cross-national regression between private expenditures and health status has no obvious interpretation.

Public sector reform. Finally, there is a third school of thought. This school argues that it is not so much the *announced* strategy of the government, but how well the government actually performs that determines the impact of public spending. This interpretation points to the poor quality, low utilization, high unit costs, and medical ineffectiveness of public sector health facilities in many countries. If whatever services the facilities were intended to provide are not available then neither their *potential* MICE or their *potential* PSCE matter. When the penicillin intended for a rural clinic has been resold on the black market then the question of the relative effectiveness of public spending on various kinds of treatments seems academic, at best. The estimated impact of health spending might be low because the efficacy of the public sector varies so widely across countries.

Of course these three schools of thought are not mutually exclusive and, by construction of the framework, all three matter. The hard question for public policy is their relative empirical importance in any given situation. In some cases the government may be quite effective at supplying services and the private market weak (for example in inaccessible, poor rural areas). In others cases the government may be quite ineffective and the private sector strong. In some cases the market for inexpensive curative services (of a PHC) type may work quite well while elsewhere insurance markets are weak so public sector supply of tertiary services makes the most sense.

Conclusion

The results presented in this paper show that a remarkable amount of the cross-country variation in health status, as measured by the under-5 mortality rate, can be explained by variation in factors not related to non-health sector policy. Approximately 95 percent of the variation in under-5 mortality is explained with income, its distribution, female education, and other "cultural" factors. Moreover, the results show that, although income alone is a powerful determinant, other factors are significant determinants of under-5 mortality.

In addition, higher public spending on health as a share of GDP is shown to be very tenuously related to improved health status. The observed efficacy of public spending is several orders of magnitude lower than the apparent potential.

The correct interpretation of the empirical results and their policy implications depend on three factors, *cost effectiveness of public spending, the net impact of additional public supply, and public sector efficacy*. Each can explain the observed results and almost certainly each contributes to explaining the low typical efficacy of actual public spending. Each has difference implications for reform and which is the most important depends on the particular situation.

[Figure 1]

Appendix 1: Cross national data availability and reliability

This appendix provides details on four aspects of the cross national regressions: data on health status, data on independent variables, data on health sector strategy, and functional form.

a) Data on health status

The data used here are as reported by UNICEF (1992) except for the mortality rates for Zaire which are unbelievable: under-5 mortality rate for 1990 is reported as 130 and infant mortality rate as 79. These are replaced with the, more reliable, estimates reported in United Nations (1992) for 1984. These are 200 for under-5 mortality and 126 for infant mortality.

b) Data on non-health sector variables

Income	Real GDP per capita in 1995 international dollars (i.e. adjusted for Purchasing Power Parity) are from the Penn World Tables 5.6.
Education	Average education levels for men and women over 15 are from Barro and Lee (1996).
Income inequality	Gini coefficient as calculated by Deininger and Squire (1996) multiplied by 100.
Percent urban	Percent of the country's population that lives in urban areas. From the World Bank's Social Indicators of Development database.
Predominantly Muslim	Dummy equal to one if over 90 percent of the country's population is Muslim.
Ethnolinguistic fractionalization	Index of ethnolinguistic fractionalization for 1960. Measures the probability that two randomly selected people from a given country will not belong to the same ethnolinguistic group as reported in Easterly and Levine (1996).
Tropical country	Dummy equal to one if part of the country's territory lies within 20 degrees of the equator.
Access to safe water	Percent of population with access to safe water. From the World Bank's Social Indicators of Development database.
Oil exporter	Dummy equal to one if the country primary export is fuels (mainly oil) as classified by the World Bank's World Development Indicators (1996) plus Kuwait.

Years independent	The percentage of years since 1776 that a country has been independent, as reported in Easterly and Levine (1996).
Defense spending	Defense spending as a share of GDP, as reported in CIA (1994)

c) Data on health sector variables

Health Expenditures	As part of the World Development Report 1993, Investing in Health, figures on the magnitude of health expenditures were generated (Murray, Govindaraj, and Musgrove, 1995). These have undergone various updates, based on country level reports. We use the latest data available in our analysis.
Percentage of national health expenditures devoted to local health services	As reported in the WHO's Health for All Database. The observation closest to 1990 in the 1986-1993 period is used.
Percentage of the population with local health services, including availability of essential drugs, within one hour's walk or travel	As reported in the WHO's Health for All Database. The observation closest to 1990 in the 1986-1993 period is used.

Table A2-1: Summary statistics for estimation sample (Number of observations = 100)			
	Mean	Std. Dev.	Correlation with GDP per capita (ln)
Under-5 mortality rate (ln)	3.928	1.183	-.9274
Under-1 mortality rate (ln)	3.605	1.053	-.9200
GDP per capita (ln)	7.914	1.151	1.000
Public health expend. (ln of share of GDP)	-3.769	.7591	.6228
Female education	4.971	2.753	.8115
Income inequality	40.93	8.706	-.2432
Percent urban	49.96	25.01	.8075
Predominantly Muslim	.120	.327	.0096
Ethnolinguistic fractionalization	.4219	.2871	-.5112
Tropical country	.620	.488	-.6445
Access to safe water	67.13	25.73	.6909
East Asia and Pacific	.10	.302	.0377
Latin America and Caribbean	.21	.409	.0090
Middle East and North Africa	.09	.288	.1236
South Asia	.05	.219	-.1346
Sub-Saharan Africa	.31	.465	-.6812
Female education missing	.11	.314	-.2463
Income inequality missing	.23	.423	-.2649
Ethnolinguistic fractionalization missing	.08	.273	.0390
Access to safe water missing	.08	.273	.3405
Under-5 mortality rate	87.75	77.25	
Under-1 mortality rate	56.96	44.66	
GDP per capita	5004	5300	
Public health expend. (Share of GDP)	.02967	.01998	

References:

Alesina, Alberto, Reza Baqir, and William Easterly, 1997. "Public Goods and Ethnic Divisions," NBER Working Paper No.6009. Cambridge, MA.

Anand, Sudhir, and Martin Ravaillon, 1993. "Human Development in Poor Countries: On the Role of Private Incomes and Public Services," Journal of Economic Perspectives, 7(1):133-50.

Barro, Robert J. and Jong-Wha Lee. 1996. "International Measures of Schooling Years and Schooling Quality." American Economic Review, Papers and Proceedings 86 (2): 218-23.

Bidani, Benu, and Martin Ravaillon, 1997. "Decomposing Social Indicators using Distributional Data," Journal of Econometrics, 77:125-139.

Caldwell, John C., 1986. "Routes to Low Mortality in Poor Countries," Population and Development Review, 12(2):171-220.

Caldwell, John C., 1990. "Cultural and Social Factors Influencing Mortality Levels in Developing Countries," The Annals of the American Academy, 510:44-59.

CIA, 1994. "The World Factbook 1994" CIA.

Deininger, Klaus and Lyn Squire, 1996. "A New Data Set Measuring Income Inequality," The World Bank Economic Review, 10(3):565-591.

Easterly, William and Ross Levine, 1996. "Africa's Growth Tragedy: Policies and Ethnic Divisions," mimeo (September). Washington, DC: The World Bank.

Filmer, Deon, Jeffrey Hammer, and Lant Pritchett, 1997. "Health Policy: Weak Links in the Chain?" mimeo, Development Research Group, The World Bank. Washington, DC.

Foster, Stanley O., Deborah A. McFarland, and A. Meredith John, 1993. "Measles," in Disease Control Priorities in Developing Countries, Jamison et al eds.

Halstead, Scott B., Julia A. Walsh, and Kenneth S. Warren sed., 1985. Good Health at Low Cost (Conference Report), The Rockefeller Foundation. New York: The Rockefeller Foundation.

Hammer, Jeffrey S., 1997. "Economic Analysis for Health Projects," World Bank Research Observer, 12(1):47-71.

Heerink, Nico, and Henk Folmer, 1994. "Income Distribution and the Fulfilment of Basic Needs: Theory and Empirical Evidence," Journal of Policy Modeling, 16:625-652.

Hobcraft, John, 1993. "Women's Education, Child Welfare and Child Survival : A Review of the Evidence," Health Transition Review: The Cultural, Social and Behavioral Determinants of Health, 3:159-75.

Jamison, Dean T., W. Henry Mosley, Anthony R. Measham, Jose-Louis Bobadilla eds., 1993. Disease Control Priorities in Developing Countries, Oxford Medical Publication.

Jamison, Dean T., Jia Wang, Kenneth Hill, and Juan-Luis Londoño, 1996. "Income, Mortality and Fertility Control in Latin America: Country-Level Performance 1960-90," mimeo, LAC Technical Department, The World Bank. Washington, DC.

Kakwani, Nana, 1993. "Performance in Living Standards: an International Comparison," Journal of Development Economics. 41:307-336.

Kim, Kwangkee and Philip M. Moody, 1992. "More Resources Better Health? A Cross-National Perspective," Social Science and Medicine, 34(8):837-842.

Kunstadter, Peter, 1985. "Remarks [on Good Health at Low Cost]," in Haltead et al eds.(1985)

Gwatkin, Davidson, R., 1980. "How Many Die? A Set of Demographic Estimates of the Annual Number of Infant and Child Deaths in the World," American Journal of Public Health, 70:1286-1289.

Lopez, Alan D., 1993. "Causes of Death in Industrial and Developing Counties: Estimates for 1985-1990," in Disease Control Priorities in Developing Countries, Jamison et al eds.

Martines, Jose, Margaret Phillips, and Richard G. Feachem, 1993. "Diarrheal Diseases," in Disease Control Priorities in Developing Countries, Jamison et al eds.

Murray, Christopher, Karel Stybolo, and Annik Rouillon, 1993. "Tuberculosis," in Disease Control Priorities in Developing Countries, Jamison et al eds.

Musgrove, Philip, 1996. "Public and Private Roles in Health: Theory and Financing Patterns," World Bank Discussion Paper No. 339. Washington, DC.

Najera, Jose A., Bernhard Liese, and Jeffrey Hammer, 1993. "Malaria," in Disease Control Priorities in Developing Countries, Jamison et al eds.

Pritchett, Lant, 1997, "Divergence, Big Time," Journal of Economic Perspectives. (Summer)

Pritchett, Lant and Lawrence Summers, 1996. "Wealthier is Healthier," Journal of Human Resources, 31(4):841-868.

Rutstein, Shea Oscar, 1984. "Infant and Child Mortality: Levels, Trends, and Demographic Differentials," World Fertility Surveys Comparative Studies No. 43. Voorburg, Netherlands: International Statistical Institute.

Srinivasan, T. N., 1994. "Data base for development analysis: an overview," Journal of Development Economics, 44:3-27.

Stansfield, Sally K., and Donald S. Shepard, 1993. "Acute Respiratory Infection," in Disease Control Priorities in Developing Countries, Jamison et al eds.

Strauss, John, and Duncan Thomas, 1995. "Human Resources: Empirical Modeling of Household and Family Decisions," in J. Behrman and T.N. Srinivasan,, eds., Handbook of Development Economics, Vol 3.: Amsterdam: North Holland Press: 1883-2023.

Subbarao, K.Y. and Laura Raney, 1995. "Social Gains from Female Education: A Cross-National Study," Economic Development and Cultural Change, 44:105-128.

UNICEF, 1992. The State of the World's Children: 1992 New York: Oxford University Press.

United Nations, 1992. Child Mortality since the 1960s: A Database for Developing Countries New York: United Nations.

Vicusi, Kip, 1993. "The Value of Risks to Life and Death," Journal of Economic Literature, 31(4):1912-46.

Walsh, Judith A., Chris N. Feifer, Anthony R. Measham, and Paul J. Gertler, 1993. "Maternal and Perinatal Health," in Disease Control Priorities in Developing Countries, Jamison et al eds.

Wolfe, Barbara, 1986, "Health Status and Medical Expenditures: Is there a Link?" Social Science and Medicine 22(10), pp 993-999.

World Bank, 1993. World Development Report 1993: Investing in Health. New York: Oxford University Press.

World Bank, 1997a. Sector Strategy: Health, Nutrition, and Population. Human Development Network, The World Bank, Washington, DC.

World Bank, 1997b. Social Indicators of Development. New York: Oxford University Press.