Public Health Performance:

A Model-Based Approach

John Gerring
Professor
Department of Political Science
Boston University
232 Bay State Road
Boston MA 02215
jgerring@bu.edu
(corresponding author)

Strom C. Thacker
Associate Professor
Department of International Relations
Boston University
sthacker@bu.edu

Ruben Enikolopov
Assistant Professor
New Economic School
Moscow
REnikolopov@nes.ru

Julián Arévalo
PhD candidate
Department of Political Science
Boston University
arevalo@bu.edu

Comments Welcome!
Please do not cite without permission of authors.

June 11, 2008

ABSTRACT

It is difficult to assess countries' relative success in addressing issues of public health because countries' are subject to very different environmental and economic conditions. These conditions directly affect health outcomes, but are in no way the responsibility of a country's public health sector. To address this recurring problem, we suggest a model-based approach for assessing public health performance.

Our procedure begins with an outcome-based measure of public health, combining two mortality indicators (IMR and life expectancy) in a single index. This index is then regressed against a series of variables intended to capture geographic and economic endowments with a substantial impact on public health over the 1960-2004 period. The residual from this analysis is regarded as a measure of public health Achievement (at a particular point in time) or Improvement (relative to past performance). Next, we examine the over-performers and under-performers identified by these two models, which point to some surprises (notably in the MENA region). The models are then subjected to numerous robustness tests and compared with non-adjusted ("raw") public health statistics.

We conclude that a model-based approach is informative both for policymakers and for academics. Typically, what is of interest – both substantively and theoretically – are those aspects of a country's public health profile that are not geographically or economically constrained. A model-based procedure allows one to focus on the degree to which countries have lived up to their potential. Our contribution is to provide a more systematic and rigorous model for analyzing this intuitive idea.

How should one assess public health performance on a global scale? This is a critical issue for publics, policymakers, and policy specialists. Without reliable indicators of performance one cannot gauge the success or failure of private and public sector efforts. "Improvement" and "deterioration" become matters of speculation, and accountability for policy choices is impossible to establish.

Yet, one is at pains to benchmark country performance across countries and through time in this difficult policy area. The problem is not an absence of indicators. Indeed, there are many potential measures of public health – infant mortality, child mortality, life expectancy, disability-adjusted life years (DALY's), stature, and so forth. The problem is that each of these indicators is strongly affected by factors that lie outside the health sector. Indeed, a large portion of the variance in public health outcomes may be explained as a correlate of modernization or geography. This means that while it may be meaningful to view life expectancy as a measure of the performance of health sectors across similarly situated countries, such as the US and Canada, it is virtually meaningless to compare this statistic across countries with vastly different economic and geographic endowments. We do not learn much, if anything, about the relative success of health sectors in the US and Sri Lanka by comparing life expectancy in these two countries – unless, that is, we can find a way to partial out the causal effect of economic and geographic factors.

This is the intuition behind most international comparisons. When writers point to the extraordinary achievements of Costa Rica, Cuba, and Sri Lanka they are comparing human development achievements in these countries relative to certain baseline characteristics that are thought lie outside – or at least be separable from – the social policy sector (Caldwell 1986; Ghai 2000; Halstead, Walsh & Warren 1985; McGuire forthcoming; Mehrotra & Jolly 1997; Riley

2007). And yet, this handicapping exercise is rarely conducted in a systematic and explicit fashion.[1] One approach is to limit the comparison to other developing countries. But how should one define the concept of a "developing" country? Given that the most common measures – per capita GDP and urbanization – are matters of degree, the category seems highly imprecise, and necessarily includes countries with a wide range of scores on these attributes. And if the comparison is limited to countries with identical (or virtually identical) scores on, say, per capita GDP, then the comparisons are very limited. Sri Lanka's per capita income is virtually identical to the Maldives, but is quite a bit higher than other south Asian countries, thus precluding any comparison with India and Bangladesh. Including additional measures of development, such as urbanization, makes this "exact matching" technique even more complicated, and perhaps impossible. Inevitably, one faces questions about which factors ought to be considered part of the baseline.

How, then, might one identify, and model, baseline characteristics so that comparisons across health sectors can be made in a meaningful way? This is the question motivating the present study. In the first section of the paper we define a strategy for measuring public health outcomes and show changes over time in the global distribution of this composite index. In the second section, we present a simple model of public health performance that includes background factors intended to measure the effect of modernization and geography on a country's public health achievement. In the penultimate section, we conduct a series of sensitivity tests in order to probe the robustness of the findings drawn from these models. The

[1] A model-based approach has been used occasionally in public health (e.g., Jamison et al. 2004; Wang et al. 1999; World Health Organization 1999: Annex Table 5), in human development (e.g., Kakwani 1993), and other policy contexts (e.g., Ndulu & O'Connell 2007). Our approach differs from previous studies of public health in four respects: a) the spatial and temporal breadth of the dataset (including all country-years from 1960-2004 in a full, imputed sample), b) the inclusion of an extensive set of controls for geographic factors, c) greater attention to problems of specification and mis-specification (including a large number of robustness tests), and d) both latitudinal ("achievement") and longitudinal ("improvement") comparisons.

concluding section reflects on the interpretation and possible uses of model-based measures of policy achievement.

Measuring Public Health

To measure the health of societies we focus on mortality data, specifically life expectancy and the infant mortality rate (IMR), understood as the number of babies that do not survive to age one, per 1000 live births.[2] Of all possible public health indicators, these are probably the most reliable and the most widely available through time and across countries.

While child mortality (deaths before the age of five [U5MR]) is sometimes regarded as a more reliable statistic, it is so highly correlated with IMR that the difference between these measures must be considered negligible (see Table 1). We opt for IMR by reason of its more extensive coverage. With respect to life expectancy, one might prefer a more sensitive measure which takes account varying levels of morbidity such as disability-adjusted life years (DALY) or health-adjusted life expectancy (HALE). Unfortunately, these adjustments are possible only in recent years and tend in any case to be highly correlated with the unadjusted life expectancy statistics, as shown in Table 1.

-- Table 1 about here --

Although IMR is a component of life expectancy, our index combines both statistics in a single indicator of public health. We do so for several reasons. First, by incorporating data from two sources we are able to build a larger sample of observations, one that is also probably more typical of the total population of nation-states that we seek to represent. The two statistics are

[2] IMR is transformed by the natural logarithm to account for expected non-linearities in the causal relationships of interest.

highly correlated (see Table 1), so this statistical manipulation imposes little loss of information. Second, these two mortality-based statistics describe somewhat different components of the topic. Although life expectancy is the "summary" concept, it might be argued that loss of life at a very early age is a greater human tragedy since it represents the loss of nearly a whole life. Finally, because of the greater vulnerability of newborns, IMR tends to be sensitive to policy interventions and societal behavioral changes to a much greater degree than life expectancy, as evidenced by the greater variance of IMR. For all these reasons, we think the combination of life expectancy and IMR offers a more reliable, more sensitive, and more insightful measure of public health than either would provide on its own.

The data source for life expectancy and IMR in this study is the World Development Indicators (2007). As it happens, other mortality datasets (e.g., United Nations 1991; United Nations 2006), are very highly correlated with the mortality statistics collected in the WDI, which reflects the fact that the underlying sources of data are virtually identical.

Missing data for life expectancy and IMR is handled in two steps. First, in cases where data is missing within a country's time-series for one of these variables we interpolate missing data. Because of the highly trended quality of these variables, and because data are usually missing for only several years at a time (if at all), the interpolation of missing data within a single time-series poses little difficulty and allows for a complete annual time-series for all included countries. Second, in cases where data is available for only one of the two variables (after interpolating), we impute missing data (with Stata's linear imputation procedure) using the other indicator. This imputation procedure increases the sample size from 7775 to 8807 observations. (Each observation represents a country-year.)

Life expectancy and IMR are then combined into a single variable, relying on the first component of a principal component factor analysis. This comprises our index of public health, used in all subsequent analyses.

Readers may be curious to know how closely this index of public health accords with other measures of public health such as health-adjusted life expectancy (DALE), disability-adjusted life years (DALYs), child mortality, malnutrition (as proxied by height for age), as well as other quality-of-life measures such as literacy, school attainment, the poverty headcount ratio (percentage of the population living on less than $2 a day), and the UNDP's Human Development Index (HDI). Table 1 displays the correlations between these alternate measures and a) the components of our index and b) the composite index itself. Not surprisingly, the public health index is highly correlated with other measures of public health and with other quality-of-life measures. Good (bad) things go together. This means that our index of public health might also be regarded more broadly as an index of human development -- though we do not adopt that interpretation here.

One of the benefits of an index with broad coverage (across countries and through time) is that one can employ it to compare global distributions of a good at varying points in time. Figures 1 and 2 show kernel density plots of the distribution of public health in 1960 (solid lines) and 2004 (dotted lines). Note that 2004 is the last year for which reasonably complete data is available in the 2007 World Development Indicators database and thus comprises the end-point for all subsequent analyses. (The analysis is restricted to country-years for which real data exists for either life expectancy or IMR (or both), or where missing data within a single time-series can be interpolated; it does not include the larger dataset arrived at through multiple-imputation.)

Figure 1 treats countries as units of analysis, while Figure 2 weights each country by its population, thus adopting individuals as the unit of analysis. The area under each curve indicates the percentage of countries (Figure 1) or individuals (Figure 2) with a corresponding level of public health on our composite index. A higher score on the X axis indicates a higher level of public health. Means for each period are indicated by a solid vertical line for 1960 and a dotted line for 2004.[3]

-- Figures 1 and 2 about here --

Although these two figures provide somewhat different views of public health inequality at the global level, trends registered from 1960 to 2004 are broadly similar. The range between minimum and maximum values falls from 4.21 in 1960 to 3.45 in 2004. Likewise, distributions about the mean decrease substantially over this period -- from 1.042 to 0.864 when countries are considered as units of analysis and from 1.062 to 0.669 when individuals are the chosen units of analysis. Most important, the mean level of public health has improved dramatically over the past four decades—over one standard deviation relative to the 1960 distribution in both graphs.

It would appear that a global convergence in public health has occurred over the postwar era and that the current level is much higher than it once was, echoing the conclusions of previous studies (Gerring 2007b; Goesling, Firebaugh 2004; Neumayer 2003). Insofar as the chosen indicators capture important features of human wellbeing, immense progress has been made over the past four decades.

Even so, enormous differences in quality of life remain, as captured by the tails of the distributions in Figures 1 and 2. One's chances of survival still depend largely upon where one

[3] It should be noted that because the population of sovereign countries changes considerably over the observed time-period, with many new countries coming into formal existence, the sample in 1960 (111) is considerably smaller than the sample in 2004 (192). However, this has only minimal impact on the shape of the density function. An additional analysis focused on the 111 countries remaining in the sample from 1960 to 2004 reveals a similar set of graphs and very similar means. Therefore, we show only the full-sample results.

happens to be born. In Angola, 154 out of every 1,000 babies will die before they reach their first year of life and life expectancy is only 41 years – levels that approximate mortality rates among pre-modern populations. In Singapore, by contrast, the infant mortality rate is about 2.6 per 1,000, and life expectancy is 79 (estimates drawn from the World Development Indicators, 2007). How are we to understand these differences?

A Model-based Approach

Having arrived at an index of public health, it remains to model that outcome so that our comparisons – across countries and through time -- reflect something other than the simple and unrevealing fact that some countries are richer, and more economically developed, than others. Yet, in raising the prospect of "adjusted" scores, one must face the difficult question of all handicapping exercises. Which factors ought to be included in the model?

Our approach is to include only those factors a) that are measurable (across the postwar era and across most sovereign nation-states), b) that show a strong empirical relationship to public health in cross-country regression models, c) that are theoretically justifiable, and d) that lie outside the purview of short-term policies and politics (they are not political in nature). The final criterion reflects our principal theoretical interest. We are interested in gauging country performance in public health relative to each country's natural endowments, which for all intents and purposes may be reduced to two: *geography* and *modernization*.

Following these precepts, we propose a model of "expected" public health that includes several economic variables including urbanization, per capita GDP[4], and several interaction terms (GDPpc squared and GDPpc cubed) intended to capture non-linearities in the relationship

[4] In order to smooth out the effects of year-to-year variations in GDP this variable is measured with a five-year moving average.

between modernization and health. It also includes a variety of geographic variables that might impact the spread of disease and other aspects of the physical quality of life: latitude (distance from the equator), a dummy variable for islands, and a series of climatic zone variables (the percent of a country's territory classified as polar, boreal, temperate desert, sub-tropical, tropical, wet temperate, tropical desert, and water).

We also include a measure of disease exposure -- HIV/AIDS prevalence in neighboring countries[5] -- which we regard as geographic in nature since it hinges on the proximity of countries to the outbreak of the disease. Specifically, the HIV exposure for Country A is calculated as the mean value of HIV prevalence in all countries (not including A): (1) whose capitals lie within 1600 kilometers of A, or (2) whose borders are contiguous or nearly so (as bodies of land separated by small bodies of water). The intuition is that if a country is surrounded by other countries with high rates of infection its exposure rate is correspondingly high; for reasons related solely to proximity, HIV/AIDS is likely to be widespread. Accordingly, if a country maintains a lower (higher) rate of infection than its neighbors, it is judged a success (failure) relative to our model-based assessment.

Having arrived at a benchmark model of public health, we impute a full sample of sovereign and semisovereign countries from 1960 to 2004 with multiple-imputation procedures (Honaker et al. 2001). This avoids potential problems of sample bias that might result when countries are deleted by virtue of incomplete data. The "full" sample – actually five datasets,

[5] Data for HIV prevalence rates prior to 2001 are scarce-to-non-existent, so constructing historical scores required some assumptions. We assumed that if a country had an HIV prevalence rate of less than one percent in 2001, its previous score on this variable was probably so low as to have a negligible effect on its mortality rate. Thus, all such countries were assigned a score of zero in years prior to 2001. For countries with HIV prevalence rates of 1% or greater in 2001, we needed some way to estimate their history. A key issue was the year of onset, i.e., the year at which HIV reached epidemic proportions, and hence began to have a significant impact on IMR and life expectancy. We estimated this year of onset by looking closely at changes in trend of tuberculosis infection rates over the past three decades for countries currently experiencing high HIV/AIDS rates (data for TB are drawn from the WHO). TB rates for these countries inflected upward in the late 1980s, suggesting 1989 as the "take-off" year for HIV/AIDS. Thus, all countries with >1% HIV prevalence rates in 2001 were assigned a 0% score *prior to* 1989. We linearly interpolated data for these countries between 1989 and 2001 (the first year of reliable data).

intended to represent the variance of the estimates deriving from multiple imputation -- is employed in all regression analyses. Definitions, sources, and descriptive statistics for all variables are contained in Appendix A.[6]

In the first column in Table 2, we see the results of an OLS regression model (with Newey-West standard errors) in which the public health index is regressed against all the endowment variables along with year fixed-effects (annual dummies). The fit of this first model is quite good (R^2=.85), suggesting that baseline factors account for a substantial amount of the variance in public health across countries. This is referred to as an Achievement model because it indicates the public health achievement of countries at particular points in time.

-- Table 2 about here --

Among the various factors represented in the model, economic variables—GDP per capita, its polynomials, and urbanization—loom large. Of course, we do not mean to suggest that economic development *by itself* explains public health. Evidently, the economic terms in our equation are representing much more than income and demography. We presume that they are also playing a "proxy" role for various correlates of modernization that have a direct or indirect impact on public health, e.g., government-sponsored policies, infrastructure, and education (Wilensky 1975). To the extent that these factors co-vary empirically with economic development, they are correctly understood as integral to that secular process. For example, if countries tend to adopt more extensive and effective social policies as they develop economically, this fact should be reflected in our baseline model. By the same token, any deviations from the norm – perhaps by virtue of spending more or less than they "should" (given their level of development), or by virtue of allocating money more or less efficiently to public

[6] Several additional factors were also tested but were discarded if their performance was inexplicable in light of existing theory and research, or if theoretical priors were not clear.

health problems – will be reflected in the residual for this model. This is what we mean by gauging country performance in the health sector.

A second econometric issue is the possible endogeneity that may exist between GDP per capita and public health. Insofar as the latter influence the former (Sachs 2001), the coefficient for GDP per capita is evidently biased. However, there are good reasons to suppose that this bias is relatively small and -- more important -- that it is equal across countries. To the extent that public health influences economic development this causal relationship should hold globally. If the bias is constant, the residuals from this analysis are still good indicators of the *relative* success of countries around the world in addressing human wellbeing.

Our purpose, in any case, is not to test the relationship between right- and left-hand variables. Thus, whether the coefficients reported in Table 2 are precisely estimated is not of great concern (we assume that they are only approximations of some underlying data-generating process). We are interested specifically in what the baseline model does *not* explain. We regard the residuals, or the difference between the actual and predicted values of our public health index, as a broad measure of country performance in the health sector. That is, given a country's economic and geographic endowments at a particular point in time, as well as global trends in disease and technology (captured by year fixed effects), how impressive is a country's health performance?

Residuals generated by the Achievement model for one year – 2004 -- are displayed in Table 3, along with 90% confidence intervals.[7] A high positive residual indicates that a country over-performs in that year, while a large negative residual indicates under-performance relative

[7] The derivation of these confidence intervals are a bit complicated, due to the multiple-imputation procedure. Amelia generates five datasets. Regression analyses are conducted on each of the five data sets. The residuals and standard errors for each of these analyses are then combined, taking into account of the variance registered across the five datasets. Confidence intervals for each residual are computed with the standard formula $x \pm t_{0.05} \cdot se(x)$.

to the parameters of the model. Note that we less interested in the absolute size of the residuals (which are of course affected by any biases in the model) than in the placement of country residuals *relative to each other*.

-- Table 3 about here --

Note also that although the models in Table 2 are derived from the entire population of sovereign and semisovereign countries (including a fair bit of imputed data), the countries listed in Table 3 and in subsequent tables represent a subset of cases that meet the following data requirements. First, at least 30 years of real (non-imputed) data must exist for our public health index, and this data must include the final year in our analysis (2004). Second, reliable GDP per capita data must be available for the most recent decade. These two criteria reduce the potential sample of sovereign countries from 194 to 166.

Table 3 identifies over-performers -- defined for present purposes as the top thirty countries -- throughout the world: in Asia (Vietnam, Sri Lanka, Singapore, Thailand, Malaysia, Brunei), Europe and its off-shoots (Spain, Greece, Portugal, Moldova, Italy, Czech Republic, Slovenia), Latin America (Costa Rica, Ecuador, Paraguay, Nicaragua, Colombia, Honduras, Panama), the Middle East and North Africa (Syria, Egypt, Tunisia, Jordan, Israel, UAE, Oman, Kuwait), Australia and sub-Saharan Africa (Eritrea). Overall, the results seem sensible in light of standard impressions of these countries. Many of the top performers such as Sri Lanka, Thailand, and Costa Rica are familiar to policy experts. Other cases are more surprising, suggesting that further research may be warranted.

Under-performers (i.e., the bottom thirty countries listed in Table 3) are also found throughout the world. However, they are fairly scarce in Asia (Bhutan, Azerbaijan, Kazakhstan), the Middle East and North Africa (Lebanon), Latin America and the Caribbean (Dominican

Republic, Haiti, St. Kitts and Nevis), Australia (Papua New Guinea) and Europe and its off-shoots (Turkey). They are relatively common, by contrast, in sub-Saharan Africa (Niger, Guinea, Mali, Guinea-Bissau, Congo, Botswana, Central African Republic [CAR], Liberia, Djibouti, Cameroon, Nigeria, Gabon, Sierra Leone, Ivory Coast, Equatorial Guinea, Angola, Swaziland, Lesotho, South Africa, Chad, Zambia, Zimbabwe).

In some circumstances policymakers and publics are concerned primarily with a country's progress (or regress) over time, rather than its absolute level of policy performance. Where temporal comparisons are more important than spatial comparisons, we propose a fixed-effect approach to model specification. Here, dummy variables for each country (minus one) are inserted into the baseline model while time-invariant variables (i.e., geographic variables) are excluded. We refer to this as an Improvement model because it measures change over time (relative to the mean value for each country during the sample period).

Regression results for the entire 1960-2004 period are shown in Model 2, Table 2. Residuals produced by this model for 2004 are shown in Table 4. They represent a country's position in 2004 relative to its mean (average) value over the four-decade period. Note that although virtually all countries have higher public health scores in 2004 than in 1960, some have experienced greater improvements than others. Countries with high positive residuals have experienced the greatest improvements in the postwar era, taking background factors into account. Countries with high negative residuals have experienced the least improvement. Thus, the fixed-effect model retains country comparisons, but now the comparisons are understood relative to each country's average performance over the observed period.

-- Table 4 about here --

The historical record over the last four decades reveals the following regional picture. Over-performers (in the top 30) are found in Latin America (Chile, Peru, Nicaragua, El Salvador, Guatemala, Ecuador), Asia (Nepal, Bangladesh, Vietnam, Singapore, Sri Lanka, South Korea), Europe (Spain, Italy, Greece, Macedonia, Portugal, Iceland), sub-Saharan Africa (Comoros, Eritrea), and Middle East and North Africa (Egypt, Syria, Kuwait, UAE, Oman, Tunisia, Yemen, Turkey, Saudi Arabia, Algeria). These are the high Improvers.

Some of these countries are also high Achievers, such as Nicaragua, Ecuador, Vietnam, Singapore, Spain, Italy, Greece, Portugal, Eritrea, Egypt, Syria, Kuwait, UAE, Oman, and Tunisia. However, the lists of Achievers and Improvers (Tables 3 and 4) are by no means identical. Over the entire 1960-2004 period, residuals produced by the Achievement and Improvement models are correlated at 0.67 (Pearson's r) and 0.61 (Spearman's r). Countries like Costa Rica, a high Achiever in 2004 but a middling Improver (#53 in Table 4), evidently made their biggest gains in public health prior to 1960, and thus do not register substantial improvement over the observed time-period.

It should also be noted that the performance of individual countries is rarely monotonic or linear. Thus, the residuals generated by the Improvement model is perhaps most useful when examined over time. In order to get a sense of the temporal variation in performance realized by countries over this four-decade time-period we provide a graph of Improvement residuals for Egypt, one of the most-improved countries in our sample.[8] Of particular interest in Figure 3 are the changes in trend from decade to decade and the dramatic improvement in public health performance (relative to background factors) that began in the mid-to-late 1980s. Longitudinal

[8] The graph is smoothed by calculating a five-year moving average of the change in this country's residuals over time, with the first four years handled as averages of 1, 2, 3, and 4 years, respective.

graphs of this nature are extraordinarily informative insofar as they allow for a sensitive tracking of country performance over time, while holding constant certain factors.

-- Figure 3 about here --

Robustness Tests

Naturally, there will be questions about proper model specification in an exercise of this sort. Since country performance is evaluated against the backdrop of chosen background characteristics, the selection of variables is particularly fraught.

Before beginning, it is important to reiterate that the findings of interest in this study are contained in the residuals for each country, *not* in the coefficients and standard errors attached to specific variables. Of course, we must be concerned if a variable's performance confounds theoretical expectations and plausible assumptions. Yet, we are not concerned to arrive at precise parameter estimates. The purpose of this statistical exercise is therefore quite different from the usual employment of statistical models, which is to test a general causal relationship.

More particularly, we are concerned with the *relative ranking* of countries, as suggested by their residual in a given model, and secondarily with the robustness of the residuals attached to particular countries. To test the stability of these findings we run a series of sixteen robustness checks for the two benchmark models (Achievement and Improvement). Each robustness test involves a single change in the benchmark model. Results across the models are compared according to two metrics: a) the rank correlation (Spearman's r) of the residuals and b) the values correlation (Pearson's r) of the residuals, as shown in Table 5.

-- Table 5 about here --

First, we test how our choice of health indicators might affect the results. The first two models contained in Table 5 show how results vary when each one of these factors is removed from the dependent variable (seriatim).

The second set of models (3-6) return to the original dependent variable. This time, we play with the specification of the model, excluding the following variable(s), seriatim, from the models: HIV prevalence, GDP & GDP squared and cubed, urban population, and finally, the geographic controls.

The third set of models add other theoretically plausible variables into the benchmark models, as indicated in rows 7-15 in Table 5. All variables are defined in Appendix A. Finally, we make one key change in the sample, excluding OECD countries (row 16 in Table 5). This tests the possibility that different causal relationships exist in the developing and developed worlds.

The results of the robustness tests contained in Table 5 may be summarized as follows. Alterations in the dependent variable, as well as subtractions from the benchmark models, result in modest changes in the results. Clearly, it matters how the public health index is constructed and the kinds of variables included as regressors in the Achievement and Improvement models. Even so, the results of these specification tests are fairly highly correlated with the results generated by the benchmark models.

The more important point is that when variables are *added* to the benchmark models, or when the sample is redefined (excluding the OECD cases), these alternative specifications have negligible impact on the results. Residuals generated by these models are correlated with residuals from the benchmark models at .98 or greater. That is to say, the rankings of countries, and the actual residuals, are virtually identical.

The bottom line is this. If observers can agree on the construction of a set of outcomes to measure the status of public health (internationally and through time), and if agreement can be garnered on the utility of including both geographic and economic controls, a stable model for judging health performance is available. While other variables may always be added to this benchmark model, these additions are unlikely to change the results of the analysis by more than a few degrees. Subtractions from the model matter -- though not as much as one might have thought. Additions to the model, by contrast, have very little impact, presumably because they tend to be correlated with factors that are already included in the model. Accordingly, we are reasonably confident that the results presented in this study are not merely an artifact of vagaries in model specification.

Conclusions

An important step towards an effective solution to the myriad problems of international public health is a meaningful metric of country performance. In this paper we have suggested a methodology for measuring success and failure through time and across countries with vastly different endowments, which we understand as encompassing both geographic and economic components.

Our approach begins with two outcome indicators – life expectancy and infant mortality -- that are widely recognized as summary measures of public health. These are combined by factor analysis into a single composite indicator to compose a single index of public health. The second step is to identify and measure those natural and economic endowments which may affect the quality of public health in a country. The third step is the construction of a statistical model in which the composite indicator of public health is regressed against these endowment variables in a global sample of countries. This model generates residuals (unexplained variance) for each country-year, which we interpret as a measure of country performance during that period. The Achievement model measures countries' absolute level of performance against what might be expected, given their individual endowments. The Improvement model measures countries' improvement over some period (in this case, 1960-2004) relative to its mean value in that period and relative to any changes in status in economic development.

It is noteworthy that although many of the strong performers (as judged by both the Achievement and Improvement models) have had stagnant economies in the postwar period (e.g., Egypt, Syria, Namibia), there are also a number of fast-growth economies in the top rungs (e.g., Chile, Malaysia, Singapore, Thailand, UAE, Vietnam). Indeed, fast and slow-growing economies are interspersed throughout Tables 2 and 3. Reassuringly, when a growth term is included in these models the coefficient for that term is not significant, and its inclusion scarcely alters the findings depicted in Tables 2 and 3 (as indicated in Table 5). This is encouraging, for it suggests that there is no consistent tradeoff between growth and public health.

Consider that if achieving growth and public health were simply a matter of allocating necessary resources one would expect different societies to make different choices, thus ending up at different end-positions. Some would be "capitalist" states, in which growth is prioritized over other goods; others would be "socialist" states, in which social development is granted priority. Yet, there is little empirical support for this zero-sum view of the policy world. This non-relationship is comprehensible when one recalls that there is no correlation between aggregate taxing and spending or social policy spending and growth performance. Big spenders, and big welfare states, do not appear to grow more or less slowly.[9] And this, in turn, reinforces

[9] Granted, at *some* point increasing taxation would squelch all economic activity. Similarly, the absence of any revenue whatsoever would prevent the state from maintaining the peace, enforcing contracts, and providing other public goods; at this extreme, transaction costs would also become extremely high. In any case, when we say there is no correlation between state size and growth, we mean that there is no detectible correlation across the range

our sense that social policy and economic policy are *non-rival*. Both can be pursued successfully at the same time. Alternatively, both can be neglected, as the experience of many countries amply demonstrates.

How much difference do model-based adjustments make in our judgments of country performance relative to “raw” (unadjusted) indicators? In order to answer this question, Table 6 compares the results of our analysis with results based solely on countries’ scores on the public health index. That is, the residuals from the Achievement model are compared with the public health index in 2004 (column 1) and the residuals from the Improvement model for 2004 are compared with countries’ improvements in the public health index from 1960-2004 (column 2). Both country rankings (Spearman’s r) and a correlation of actual values (Pearson’s r) are provided. It will be seen that although correlations are positive and statistically significant, they are not especially strong.

-- Table 6 about here --

Clearly, a model-based analysis, taking account geographic and economic background factors, provides different answers than an analysis of raw (unadjusted) public health indicators. Our suggested approach to public health performance thus differs fairly dramatically from most extant approaches, which generally do not take into account the varying endowments faced by countries around the world. To be sure, writers often employ a back-of-the-envelope procedure, which usually amounts to a comparison between one country’s performance on a measure of human development and other, similarly situated, countries’ performance on that same indicator. Thus, Costa Rica may be compared with other developing countries and Sweden may be compared with other advanced industrial countries. Our approach may be understood as an effort to systematize the intuitions behind these simple cross-case comparisons.

---

of taxing and spending demonstrated by extant states.

In concluding, it is important to recall that the core motivation of a model-based approach to public health is to arrive at a viable basis for judging success and failure by reference to things that citizens and policymakers could affect without any change in a country's geographic or economic endowments. We do *not* purport to have provided a full causal model of public health, which would necessarily include many additional variables not found in our benchmark models. Our approach is thus properly classified is a descriptive model with strong prescriptive overtones, and in this respect mirrors the goals of the Human Development Index (HDI), the Physical Quality of Life Index (Morris 1979), and various measures of policy efficiency (where the interest of the model is in capturing the relative efficiency of social spending with respect to a corresponding policy goal).[10]

A residual-based approach to public health does not, of course, shed light on the *reasons* for the unexplained variance, i.e., the reasons why some countries have positive residuals and others have negative residuals. We suspect that the relative efficiency and effectiveness of health systems in countries around the world play a large role in this story. However, it would be incorrect to regard the residual simply as a measure of how well private and public health providers are performing. Life expectancy and infant mortality are also affected by many additional factors – most importantly, individual-level behavior. Do families utilize the medical facilities and technology available to them? Are dietary and exercise practices conducive to good health? Are family structures and social norms conducive to the empowerment of women and the preservation of children? Many factors contributing to a country's aggregate mortality rate presumably have little to do with its geographic or economic endowments. Unfortunately, these tend not to be factors that can be measured and tested in a crossnational format. Indeed, we

[10] Evans et al. (2001), Gupta & Verhoeven (2001), Moore (2003). For a review and critique see Ravallion (2003).

suspect that beyond natural and economic endowments the causal story of public health varies considerably from country to country.

It bears emphasis, therefore, that the interpretation of a country's residual in the Achievement and Performance models is a complex matter. In claiming that a country "over-performs" or "under-performs" we do not intend to point the finger at any one source. The reason for its over- or under-performance can only be understood through further analysis, presumably including in-depth country-level study. Our hope is that the benchmark models provided in this study will provide a suitable tool for case-selection where scholars wish to conduct case studies of specific countries.[11]

Even so, it seems reasonable to suppose that a country's residual indicates its likely potential for improving public health. Countries with very high residuals are probably bumping up against what it is possible to achieve, given current economic output. By the same token, countries with negative residuals in the Achievement model could probably do a lot more with the resources currently available to them. They are not geographically or economically constrained. Perhaps, the use of model-based analyses of public health will assist in identifying these targets of opportunity and in applying political pressure that might lead, ultimately, to improved performance.

Also implicit in our analysis is the idea that countries with high residuals provide exemplars of "best practices," practices which could be adapted for use in low-performing countries. In this respect, a model-based procedure may help to shed light on the contributing causes of public health and on concrete options for reform.

[11] For example, "deviant" cases -- identified here as those scoring high or low in the Achievement or Improvement models -- may illuminate new facets of public health that are not well understood. Other case-selection strategies are also feasible. In any case, model-based selection procedures allow researchers to mine the entire field of crossnational data in order to choose appropriate cases for in-depth analysis (Gerring 2007a).

References

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

Deininger , Klaus and Lyn Squire (1996) A New Data Set Measuring Income Inequality. *The World Banks Economic Review Vol. 10, No 3.*

Evans, David B., Ajay Tandon, Christopher J.L. Murray, and Jeremy A. Lauer. 2001. "Comparative Efficiency of National Health Systems in Producing Health: An Analysis of 191 Countries." World Health Organization: GPE Discussion Paper Series No. 29.

Gallup, John Luke; Jeffrey D. Sachs; Andrew Mellinger. 1999. "Geography and Economic Development." *International Regional Science Review* 22:2 (August) 179-232.

Gerring, John. 2007a. *Case Study Research: Principles and Practices*. Cambridge: Cambridge University Press.

Gerring, John. 2007b. "Global Justice as an Empirical Question." *PS: Political Science and Politics* 40:1, 67-78.

Ghai, Dharam (ed). 2000. *Social Development and Public Policy: Some Lessons from Successful Experiences*. London: St. Martin's.

Gleditsch, Nils Petter; Peter Wallensteen, Mikael Eriksson, Margareta Sollenberg & Håvard Strand, 2002. *'Armed Conflict 1946–2001: A New Dataset'*, *Journal of Peace Research* 39(5): 615–637.

Goesling, Brian; Glenn Firebaugh. 2004. "The Trend in International Health Inequality." *Population and Development Review* 30:1 (March) 131-46.

Gupta, Sanjeev; Marijn Verhoeven. 2001. "The Efficiency of Government Expenditure: Experiences from Africa." *Journal of Policy Modeling* 23, 433-67.

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

Honaker, James, Anne Joseph, Gary King, Kenneth Scheve, and Naunihal Singh. 2001. "AMELIA: A Program for Missing Data (Windows version)." Cambridge, MA: Harvard University. http://GKing.Harvard.edu/

Humphreys, Macartan. 2005. "Natural Resources, Conflict, and Conflict Resolution: Uncovering the Mechanisms." *Journal of Conflict Resolution* 49:4, 508-37.

Jamison, Dean T.; Martin Sandbu; Jia Wang. 2004. "Why Has Infant Mortality Rates Decreased at Such Different Rates in Different Countries?" Disease Control Priorities Project, Working Paper No. 21.

Kakwani, Nanak. 1993. "Performance in Living Standards: An International Comparison." *Journal of Development Economics* 41:2 (August) 307-36.

La Porta, Rafael, Florencio Lopez-de-Silanes, Andrei Shleifer, and Robert W. Vishny. 1999. "The Quality of Government." *Journal of Economics, Law and Organization* 15:1, 222-79.

Monty G. Marshall and Keith Jaggers. 2000. "Polity IV Dataset Project: Political Regime Characteristics and Transitions, 1800–1999." Page numbers cited refer to Dataset User's Manual.

McGuire, James W. Forthcoming. *Politics, Policy, and Mortality Decline in East Asia and Latin America*. Unpublished ms.

Mehrotra, Santosh; Richard Jolly (eds). 1997. *Development with a Human Face: Experiences in Social Achievement and Economic Growth*. Oxford: Clarendon.

Moore, Mick. 2003. "Polity Qualities: How Governance Affects Poverty." Peter P. Houtzager and Mick Moore (eds), *Changing Paths: International Development and the Politics of Inclusion* (Ann Arbor: University of Michigan Press) 167-203.

Morris, David. 1979. *Measuring the Condition of the World's Poor: The Physical Quality of Life Index*. New York: Pergamon.

Ndulu, Benno J.; Stephen A. O'Connell. 2007. "Policy Plus: African Growth Performance 1960-2000." Unpublished manuscript.

Neumayer, Eric. 2003. "Beyond Income: Convergence in Living Standards, Big Time." *Structural Change and Economic Dynamics* 14:3 (September) 275-96.

Ravallion, Martin. 2003. "On Measuring Aggregate 'Social Efficiency.'" World Bank Policy Research Paper 3166.

Riley, James C. 2007. *Low Income, Social Growth, and Good Health: A History of Twelve Countries*. Berkeley: University of California Press.

Sachs, Jeffrey. 2001. *Macroeconomics and Health: Investing in Health for Economic Development*. Geneva: World Health Organization.

Stonebruner, Rand L. and Daniel Low-Beer. 2004. "Population-level HIV Declines and Behavioral Risk Avoidance in Uganda." *Science* 304:5671 (April 30) 714-8.

United Nations. 1991. *Child Mortality in Developing Countries*. New York: United Nations.

United Nations. 2006. *World Mortality Report*. New York: Department of Economic and Social Affairs, Population Division, United Nations.

Wang, Jia, Dean T. Jamison, Eduard Bos, Alexander Preker, and John Peabody. 1999. *Measuring Country Performance on Health: Selected Indicators for 115 Countries*. Washington: World Bank.

Wilensky, Harold L. 1975. *The Welfare State and Equality: Structural and Ideological Roots of Public Expenditures*. Berkeley: University of California Press.

World Bank. 2006. *World Development Indicators*. Washington: World Bank.

World Bank. 2007. *World Development Indicators*. Washington: World Bank.

World Health Organization. 1999. *The World Health Report: Making a Difference*. Geneva: World Health Organization.

World Health Organization. 2008. *Statistical Information System*. http://www.who.int/whosis.

Appendix A:

Description of Variables

	Variables in Benchmark models	**Mean**	**SD**
Public health index	First component derived from principal components analysis of IMR (ln) and life expectancy. Composed by authors.	0.076	0.890
HIV/AIDS prevalence of neighbors	For Country A, the mean value of HIV prevalence in all countries (not including A): (1) whose capitals lie within 1600 kilometers of A, or (2) whose borders are contiguous or nearly so (as neighboring bodies of land separated by small bodies of water). Source: WHO Global Health Atlas (http://www.who.int/globalatlas/)	1.017	1.665
Island	A dummy variable that equals one if a country is an island. Coded by authors using CIA Factbook.	0.289	0.435
Latitude	Absolute distance from the equator in latitude degrees, transformed by the natural logarithm. Source: La Porta et al (1999).	-1.577	0.883
GDP per capita	GDP per capita (purchasing power parity) in constant 2000 US$, transformed by the natural logarithm. Source: WDI (2006). Missing data for several cases imputed from Penn World Tables.	7.496	1.449
Urban pop.	% of population leaving in urban areas. Source: WDI (2006).	48.058	24.098
Boreal	The proportion of the country's area in boreal zones. Source: Gallup et al (1999).	0.048	0.141
Temperate desert	The proportion of the country's area in temperate desert zones. Source: Gallup et al (1999).	0.030	0.113
Polar	The proportion of the country's area in polar zones. Source: Gallup et al (1999).	0.011	0.048
Subtropical	The proportion of the country's area in subtropical zones. Source: Gallup et al (1999).	0.302	0.299
Tropics	The proportion of the country's area in tropical zones. Source: Gallup et al (1999).	0.164	0.231
Wet temperate	The proportion of the country's area in wet temperate zones. Source: Gallup et al (1999).	0.192	0.290
Tropical desert	The proportion of the country's area in tropical desert zones. Source: Gallup et al (1999).	0.111	0.255
Water	The proportion of the country's area in water zones. Source: Gallup et al (1999).	0.061	0.093

Descriptive statistics are drawn from imputed datasets, including all sovereign and semi-sovereign states and all years from 1960-2004. N=13,616.

Peripheral Variables	
HALE	Healthy life expectancy, specifically, the "average number of years that a person can expect to live in 'full health' by taking into account years lived in less than full health due to disease and/or injury" (WHO web site: http://www.who.int/whosis/indicators/2007HALE0/en/). Source: WHO (2008)
DALY	Disability-adjusted life years, which "combines in one measure the time lived with disability and the time lost due to premature mortality. One DALY can be thought of as one lost year of 'healthy' life and the burden of disease as a measurement of the gap between current health status and an ideal situation where everyone lives into old age free of disease and disability" (WHO web site: http://www.who.int/whosis/indicators/2007HALE0/en/). Source: WHO (2008).
Child mortality	The number of children dying before reaching five years of age per 1,000 live births. Source: WDI (2007)
Poverty headcount ratio	The percentage of population living on less than $2.15 a day at 1993 international prices. Referred to in the text as "$2/day." Source: WDI (2007)
Malnutrition (height for age)	The percentage of children under five whose height for age is more than two standard deviations below the median for the international reference population ages 0 to 59 months. For children up to two years of age, height is measured by recumbent length. For older children, height is measured by stature while standing. Source: WDI (2007)
IMR	The number of infants dying before reaching one year of age per 1,000 live births, transformed by the natural logarithm. Source: WDI (2007)
Life expectancy	Life expectancy at birth indicates the number of years a newborn infant would live if prevailing patterns of mortality at the time of its birth were to stay the same throughout its life. Source: WDI (2006)
Growth	Annual percentage growth rate of GDP at market prices based on constant local currency. Aggregates are based on constant 2000 U.S. dollars. GDP is the sum of gross value added by all resident producers in the economy plus any product taxes and minus any subsidies not included in the value of the products. It is calculated without making deductions for depreciation of fabricated assets or for depletion and degradation of natural resources. Source: WDI (2007).
Health expenditure	Total health expenditure is the sum of public and private health expenditure. It covers the provision of health services (preventive and curative), family planning activities, nutrition activities, and emergency aid designated for health but does not include provision of water and sanitation. Source: WDI (2007).
Imports	Imports of goods and services represent the value of all goods and other market services received from the rest of the world. They include the value of merchandise, freight, insurance, transport, travel, royalties, license fees, and other services, such as communication, construction, financial, information, business, personal, and government services. They exclude labor and property income (formerly called factor services) as well as transfer payments. Source: WDI (2007).
Oil production	Millions of barrels per day per capita. Source: Humphreys (2005).
Democracy stock	Democracy stock for a country in a year is the sum of each country's Polity2 score from 1900 to the present year, with a 1 percent annual depreciation rate. Source: Marshall and Jaggers (2000)
Tax revenue	Tax revenue refers to compulsory transfers to the central government for public purposes. Certain compulsory transfers such as fines, penalties, and most social security contributions are excluded. Refunds and corrections of erroneously collected tax revenue are treated as negative revenue. Source: WDI (2007).
Telephones	Telephone mainlines are fixed telephone lines connecting a subscriber to the telephone exchange equipment.
Gini index	Index of inequality of income distribution. Source: Deininger and Squire (1996).
Conflicts	Includes data on Extra-State conflict, Insterstate conflict, Internal conflict, internationalized internal conflict and country listed as location of *any* kind of conflict. Source: Gleditsch et al (2002).

Table 1:

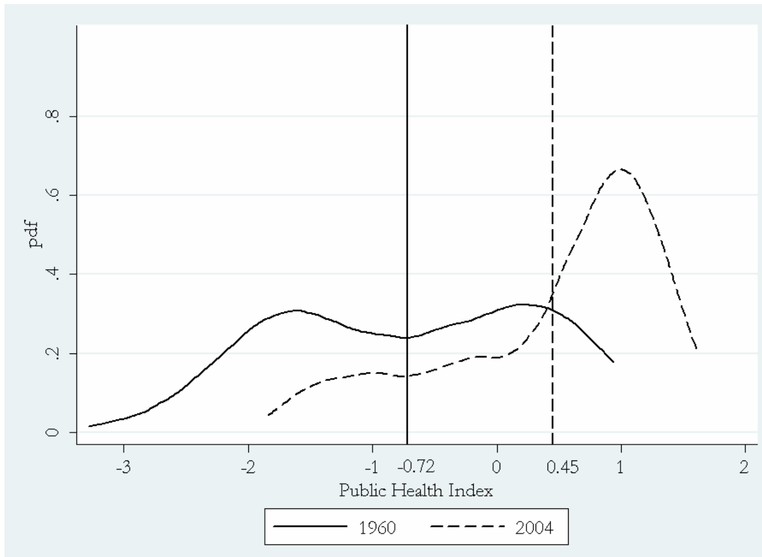
Correlation Table: Public Health and Human Development Indicators

	Life expectancy	**IMR (ln)**	**Public health index**
Life expectancy	--	-0.8187	0.9262
IMR (ln)	-0.8187	--	-0.9623
HALEs (males)	0.8028	-0.7604	0.8149
HALEs (females)	0.8133	-0.7754	0.8292
DALYs	-0.9356	0.7597	-0.8756
Child mortality rate (ln)	-0.8594	0.9100	-0.9081
Malnutrition (height for age)	-0.5575	-0.6040	-0.5978
Literacy	0.8295	-0.7239	0.7887
School attainment	0.8205	-0.7656	0.8055
Poverty headcount ratio ($2/day)	-0.7508	0.7333	-0.7781
Human Development Index (HDI)	0.8990	-0.9149	0.9404

Pearson's r correlations, based on varying samples. All are significant at 99%. See Appendix A for variable definitions and sources.

Figure 1:

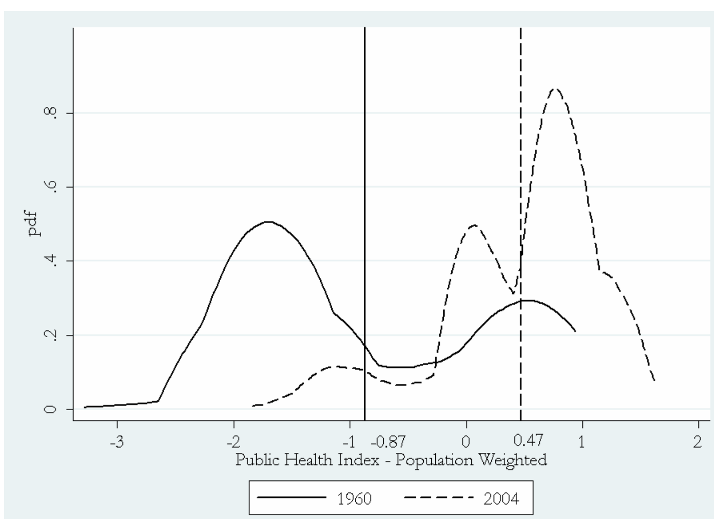
The Distribution of Public Health: By Country



Kernel density plot of the distribution of public health in 1960 and 2004. Vertical lines: mean value for that year's distribution. Unit of analysis: Countries.

Figure 2:

The Global Distribution of Public Health: By Individual



Kernel density plot of the distribution of public health in 1960 and 2004. Vertical lines: mean value for that year's distribution. Unit of analysis: Individual inhabitants (population-weighted function based on country data).

Table 2:

Benchmark Models of Public Health Performance

	1. Achievemen t	2. Improvement
HIV exposure	-0.0411***	-0.0274***
	[0.00161]	[0.00106]
GDP per capita	0.3899***	0.1318***
	[0.01609]	[0.01438]
GDP per capita squared	0.0067***	0.0037***
	[0.00093]	[0.00060]
GDP per capita cubed	-0.0010***	-0.0001
	[0.00008]	[0.00008]
Urban population	0.0057***	0.0068***
	[0.00030]	[0.00044]
Island	0.3038***	
	[0.01311]	
Latitude	0.0480***	
	[0.00678]	
Boreal	0.1909***	
	[0.03819]	
Temperate desert	-0.1968***	
	[0.04233]	
Sub-tropical	-0.3612***	
	[0.02546]	
Tropical	-0.3826***	
	[0.02805]	
Wet temperate	0.2566***	
	[0.02596]	
Tropical desert	-0.5711***	
	[0.02525]	
Water	0.3013***	
	[0.07725]	
Constant	-3.3240***	-1.9378***
	[0.08038]	[0.06952]
Year fixed effects	Yes	Yes
Country fixed effects	No	Yes
Years	1960-2004	1960-2004
Observations	8656	8656
Countries	194	194
R square	0.84	0.73

Dependent variable: life expectancy and IMR (composite index). The excluded climate variable is Polar. All variables defined in Appendix A. Newey-West standard errors in brackets. * sig. at 10%; ** sig. at 5%; *** sig. at 1%

Table 3:

Public Health Achievement in 2004

1	Vietnam	1.09 ±	0.1	57	France	0.163 ±	0.0	113	Uruguay	-0.10 ±	0.094
2	Sri Lanka	0.96 ±	0.3	58	China	0.158 ±	0.0	114	Luxembourg	-0.10 ±	0.084
3	Syria	0.94 ±	0.0	59	Ghana	0.155 ±	0.1	115	Benin	-0.10 ±	0.108
4	Costa Rica	0.77 ±	0.0	60	Switzerland	0.153 ±	0.0	116	Denmark	-0.10 ±	0.124
5	Singapore	0.74 ±	0.1	61	Sudan	0.152 ±	0.1	117	Argentina	-0.11 ±	0.068
6	Malaysia	0.67 ±	0.0	62	Malta	0.151 ±	0.3	118	Trinidad	-0.12 ±	0.104
7	Israel	0.66 ±	0.0	63	Bulgaria	0.150 ±	0.1	119	Jamaica	-0.12 ±	0.122
8	Spain	0.59 ±	0.1	64	Tajikistan	0.145 ±	0.1	120	Tanzania	-0.13 ±	0.091
9	UAE	0.58 ±	0.0	65	Ukraine	0.143 ±	0.0	121	UK	-0.13 ±	0.146
10	Thailand	0.56 ±	0.1	66	St Vincent	0.142 ±	0.2	122	Mongolia	-0.14 ±	0.106
11	Ecuador	0.55 ±	0.0	67	Austria	0.132 ±	0.0	123	Guyana	-0.15 ±	0.094
12	Oman	0.54 ±	0.0	68	Seychelles	0.131 ±	0.2	124	Mauritania	-0.15 ±	0.331
13	Paraguay	0.53 ±	0.1	69	Canada	0.129 ±	0.0	125	Kenya	-0.19 ±	0.092
14	Egypt	0.52 ±	0.0	70	Poland	0.128 ±	0.0	126	Georgia	-0.19 ±	0.119
15	Moldova	0.51 ±	0.0	71	India	0.110 ±	0.1	127	Ethiopia PDR	-0.19 ±	0.081
16	Colombia	0.51 ±	0.0	72	Peru	0.108 ±	0.0	128	Latvia	-0.19 ±	0.090
17	Greece	0.50 ±	0.1	73	Indonesia	0.105 ±	0.1	129	Madagascar	-0.20 ±	0.425
18	Portugal	0.49 ±	0.1	74	Suriname	0.098 ±	0.3	130	Bahrain	-0.21 ±	0.108
19	Nicaragua	0.45 ±	0.1	75	Japan	0.096 ±	0.1	131	Gambia	-0.22 ±	0.109
20	Australia	0.45 ±	0.2	76	Togo	0.088 ±	0.1	132	Bolivia	-0.23 ±	0.277
21	Brunei	0.44 ±	0.2	77	Mozambiqu	0.085 ±	0.2	133	Burkina Faso	-0.24 ±	0.116
22	Tunisia	0.43 ±	0.0	78	Philippines	0.081 ±	0.1	134	Rwanda	-0.25 ±	0.114
23	Eritrea	0.41 ±	0.0	79	Fiji	0.080 ±	0.4	135	Ireland	-0.27 ±	0.092
24	Jordan	0.40 ±	0.1	80	Hungary	0.076 ±	0.0	136	Guinea	-0.27 ±	0.082
25	Honduras	0.40 ±	0.1	81	Kyrgyzstan	0.072 ±	0.0	137	Niger	-0.31 ±	0.419
26	Italy	0.38 ±	0.1	82	Guatemala	0.071 ±	0.0	138	Mali	-0.31 ±	0.087
27	Kuwait	0.38 ±	0.1	83	Feda Stsa	0.068 ±	0.4	139	Azerbaijan	-0.31 ±	0.109
28	Slovenia	0.37 ±	0.0	84	Antigua &	0.064 ±	0.4	140	Dominican	-0.32 ±	0.086
29	Czech	0.35 ±	0.0	85	Saudi	0.063 ±	0.1	141	St Kitts	-0.35 ±	0.109
30	Panama	0.34 ±	0.0	86	Korea REP	0.062 ±	0.0	142	Bhutan	-0.35 ±	0.095
31	Finland	0.33 ±	0.1	87	Pakistan	0.062 ±	0.0	143	Turkey	-0.37 ±	0.118
32	Serbia/Mont	0.32 ±	0.2	88	Mexico	0.048 ±	0.0	144	Chad	-0.38 ±	0.096
33	El Salvador	0.32 ±	0.1	89	Estonia	0.044 ±	0.0	145	Zambia	-0.38 ±	0.388
34	Bangladesh	0.31 ±	0.1	90	Cambodia	0.036 ±	0.1	146	Zimbabwe	-0.42 ±	0.384
35	Cyprus	0.30 ±	0.2	91	Dominica	0.035 ±	0.3	147	Guinea-Bissau	-0.43 ±	0.094
36	Germany	0.29 ±	0.4	92	Armenia	0.034 ±	0.0	148	PNG	-0.45 ±	0.122
37	Algeria	0.27 ±	0.0	93	Uganda	0.025 ±	0.1	149	Congo, Rep.	-0.53 ±	0.090
38	Vanuatu	0.26 ±	0.3	94	Macedonia	0.023 ±	0.0	150	Lebanon	-0.54 ±	0.107
39	Tonga	0.26 ±	0.3	95	Uzbekistan	0.013 ±	0.1	151	South Africa	-0.62 ±	0.093
40	Chile	0.25 ±	0.0	96	Iran	0.004 ±	0.1	152	Kazakhstan	-0.63 ±	0.110
41	Bosnia-	0.25 ±	0.0	97	Kiribati	-0.00 ±	0.3	153	Liberia	-0.68 ±	0.092
42	Croatia	0.25 ±	0.0	98	Cape Verde	-0.01 ±	0.3	154	Central	-0.70 ±	0.089
43	Iceland	0.25 ±	0.4	99	Mauritius	-0.01 ±	0.1	155	Djibouti	-0.71 ±	0.117
44	Slovak	0.24 ±	0.0	100	Netherlands	-0.02 ±	0.1	156	Lesotho	-0.72 ±	0.103
45	Western	0.23 ±	0.3	101	Lithuania	-0.02 ±	0.0	157	Swaziland	-0.73 ±	0.494
46	Nepal	0.22 ±	0.0	102	Belgium	-0.03 ±	0.0	158	Haiti	-0.79 ±	0.071
47	Sweden	0.21 ±	0.0	103	Laos	-0.03 ±	0.1	159	Nigeria	-0.80 ±	0.113
48	Solomon Is	0.20 ±	0.2	104	Romania	-0.05 ±	0.0	160	Cameroon	-0.81 ±	0.102
49	Albania	0.20 ±	0.0	105	Malawi	-0.05 ±	0.1	161	Ivory Coast	-0.88 ±	0.111
50	St. Lucia	0.20 ±	0.3	106	United	-0.06 ±	0.1	162	Sierra Leone	-0.88 ±	0.099
51	Burundi	0.19 ±	0.1	107	Sao Tome	-0.06 ±	0.2	163	Gabon	-0.89 ±	0.126
52	Yemen REP	0.19 ±	0.0	108	Senegal	-0.06 ±	0.1	164	Angola	-1.13 ±	0.108
53	Comoros	0.18 ±	0.4	109	New	-0.07 ±	0.2	165	Botswana	-1.19 ±	0.074
54	Belize	0.18 ±	0.3	110	Namibia	-0.08 ±	0.2	166	Eq Guinea	-1.43 ±	0.236
55	Venezuela	0.18 ±	0.1	111	Belarus	-0.09 ±	0.0				
56	Norway	0.17 ±	0.0	112	Brazil	-0.09 ±	0.1				

Residuals from model 1, Table 2, in 2004, followed by 90% confidence intervals.

Table 4:

Residuals from the Improvement Model in 2004

1	UAE	0.62 ±	0.0	57	Australia	0.129 ±	0.2	113	Tonga	-0.09 ±	0.126
2	Oman	0.59 ±	0.0	58	Slovak	0.126 ±	0.1	114	Panama	-0.10 ±	0.106
3	Egypt	0.53 ±	0.0	59	Laos	0.121 ±	0.1	115	Antigua &	-0.10 ±	0.149
4	Syria	0.47 ±	0.0	60	Croatia	0.120 ±	0.1	116	Guinea-	-0.10 ±	0.137
5	Peru	0.43 ±	0.0	61	Moldova	0.119 ±	0.1	117	Lithuania	-0.11 ±	0.151
6	Chile	0.42 ±	0.1	62	Malta	0.118 ±	0.1	118	Feda Stsa	-0.11 ±	0.346
7	Portugal	0.41 ±	0.0	63	France	0.117 ±	0.1	119	China	-0.11 ±	0.094
8	Nicaragua	0.37 ±	0.1	64	Finland	0.115 ±	0.1	120	Djibouti	-0.13 ±	0.146
9	Vietnam	0.37 ±	0.1	65	Estonia	0.111 ±	0.1	121	Malawi	-0.13 ±	0.154
10	Comoros	0.36 ±	0.1	66	Western Samoa	0.110 ±	0.1	122	Benin	-0.13 ±	0.103
11	Singapore	0.36 ±	0.1	67	Senegal	0.094 ±	0.1	123	Trinidad	-0.15 ±	0.097
12	Tunisia	0.36 ±	0.0	68	Cape Verde	0.091 ±	0.1	124	Romania	-0.16 ±	0.122
13	El Salvador	0.35 ±	0.0	69	Hungary	0.083 ±	0.0	125	Latvia	-0.16 ±	0.099
14	Macedonia	0.35 ±	0.1	70	Bosnia-	0.078 ±	0.1	126	Haiti	-0.16 ±	0.133
15	Saudi Arabia	0.34 ±	0.0	71	Kiribati	0.076 ±	0.2	127	Guyana	-0.17 ±	0.115
16	Kuwait	0.32 ±	0.1	72	Argentina	0.074 ±	0.1	128	Ghana	-0.18 ±	0.140
17	Greece	0.31 ±	0.0	73	Brazil	0.068 ±	0.1	129	Netherlands	-0.18 ±	0.119
18	Guatemala	0.30 ±	0.0	74	Guinea	0.066 ±	0.1	130	Sao Tome	-0.18 ±	0.136
19	Spain	0.30 ±	0.1	75	Gambia	0.063 ±	0.1	131	Ethiopia PDR	-0.19 ±	0.115
20	Italy	0.30 ±	0.1	76	St Kitts	0.059 ±	0.2	132	Jamaica	-0.22 ±	0.138
21	Nepal	0.30 ±	0.1	77	Malaysia	0.044 ±	0.1	133	Uganda	-0.22 ±	0.124
22	Ecuador	0.29 ±	0.0	78	Thailand	0.040 ±	0.1	134	Ukraine	-0.23 ±	0.120
23	Yemen REP	0.28 ±	0.1	79	Madagascar	0.038 ±	0.1	135	Burkina Faso	-0.23 ±	0.125
24	Bangladesh	0.28 ±	0.1	80	Kyrgyzstan	0.036 ±	0.1	136	Togo	-0.23 ±	0.100
25	Turkey	0.28 ±	0.1	81	Venezuela	0.035 ±	0.1	137	Bulgaria	-0.23 ±	0.134
26	Korea REP	0.27 ±	0.1	82	Sweden	0.035 ±	0.1	138	Suriname	-0.24 ±	0.131
27	Algeria	0.26 ±	0.0	83	New Zealand	0.032 ±	0.1	139	Rwanda	-0.25 ±	0.122
28	Sri Lanka	0.26 ±	0.2	84	Luxembourg	0.019 ±	0.1	140	Burundi	-0.25 ±	0.161
29	Iceland	0.26 ±	0.1	85	Uruguay	0.015 ±	0.0	141	Lebanon	-0.28 ±	0.120
30	Eritrea	0.25 ±	0.1	86	Philippines	0.009 ±	0.1	142	Belarus	-0.28 ±	0.106
31	Iran	0.23 ±	0.1	87	Solomon Is	0.006 ±	0.0	143	Sierra Leone	-0.28 ±	0.089
32	Vanuatu	0.23 ±	0.1	88	Canada	0.003 ±	0.1	144	Uzbekistan	-0.29 ±	0.170
33	Czech	0.23 ±	0.1	89	PNG	0.001 ±	0.0	145	Georgia	-0.30 ±	0.124
34	Bhutan	0.22 ±	0.0	90	India	-0.00 ±	0.1	146	Chad	-0.34 ±	0.103
35	Austria	0.22 ±	0.0	91	Switzerland	-0.00 ±	0.0	147	Liberia	-0.35 ±	0.132
36	Indonesia	0.21 ±	0.1	92	UK	-0.00 ±	0.1	148	Tanzania	-0.35 ±	0.149
37	Mongolia	0.20 ±	0.1	93	Cambodia	-0.00 ±	0.1	149	Gabon	-0.37 ±	0.131
38	Bahrain	0.20 ±	0.1	94	Fiji	-0.01 ±	0.2	150	Namibia	-0.38 ±	0.176
39	Bolivia	0.19 ±	0.0	95	Dominican	-0.01 ±	0.0	151	Azerbaijan	-0.39 ±	0.138
40	Honduras	0.19 ±	0.0	96	Seychelles	-0.01 ±	0.1	152	Angola	-0.41 ±	0.130
41	St. Lucia	0.19 ±	0.1	97	Mozambique	-0.01 ±	0.1	153	South Africa	-0.42 ±	0.163
42	Brunei	0.18 ±	0.1	98	Pakistan	-0.01 ±	0.1	154	Congo, Rep.	-0.44 ±	0.079
43	Poland	0.18 ±	0.1	99	Norway	-0.02 ±	0.0	155	Nigeria	-0.44 ±	0.111
44	Germany	0.17 ±	0.1	100	Ireland	-0.02 ±	0.1	156	Zambia	-0.45 ±	0.175
45	Albania	0.17 ±	0.0	101	Dominica	-0.02 ±	0.1	157	Kazakhstan	-0.49 ±	0.189
46	Jordan	0.17 ±	0.1	102	Tajikistan	-0.04 ±	0.1	158	Cameroon	-0.52 ±	0.133
47	Mexico	0.17 ±	0.0	103	Belize	-0.04 ±	0.1	159	Central	-0.53 ±	0.103
48	Belgium	0.16 ±	0.1	104	Denmark	-0.04 ±	0.1	160	Swaziland	-0.54 ±	0.139
49	Japan	0.16 ±	0.1	105	Paraguay	-0.06 ±	0.0	161	Ivory Coast	-0.55 ±	0.122
50	Slovenia	0.14 ±	0.1	106	St Vincent	-0.06 ±	0.1	162	Kenya	-0.64 ±	0.127
51	Colombia	0.14 ±	0.1	107	United States	-0.07 ±	0.1	163	Eq Guinea	-0.71 ±	0.143
52	Cyprus	0.13 ±	0.1	108	Niger	-0.08 ±	0.1	164	Lesotho	-0.75 ±	0.190
53	Costa Rica	0.13 ±	0.1	109	Armenia	-0.08 ±	0.1	165	Zimbabwe	-0.90 ±	0.152
54	Mauritius	0.13 ±	0.1	110	Mauritania	-0.08 ±	0.0	166	Botswana	-1.40 ±	0.162
55	Israel	0.13 ±	0.1	111	Mali	-0.08 ±	0.1				
56	Serbia/Monte	0.13 ±	0.0	112	Sudan	-0.08 ±	0.1				

Residuals from model 2, Table 2, in 2004, followed by 90% confidence intervals.

Figure 3:

Residuals from the Improvement Model: An Example

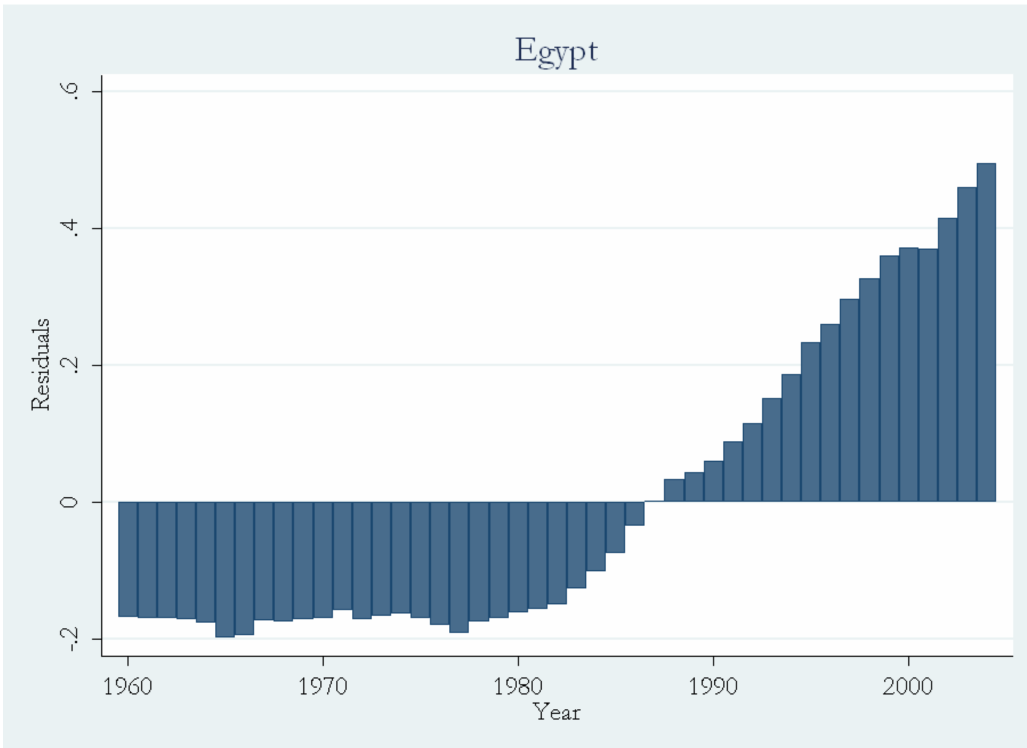


Table 5:

Robustness Tests

	Achievement Model		Improvement Model	
	Spearman's r	Pearson's r	Spearman's r	Pearson's r
I. Alternative outcome measures				
1. Life expectancy only	0.874	0.891	0.826	0.827
2. IMR (ln) only	-0.672	-0.661	-0.664	-0.609
II. Subtracting regressors				
3. - HIV prevalence	0.970	0.950	0.831	0.770
4. - ln(GDP) & its polynomials	0.882	0.867	0.971	0.971
5. - urban population	0.974	0.978	0.970	0.975
6. - geographic controls	0.775	0.842	--	--
III. Adding regressors				
7. + Growth (GDP per cap)	0.998	0.999	0.998	0.999
8. + Health expenditure per cap	0.999	0.999	0.999	0.999
9. + Imports (share of GDP)	0.998	0.991	0.998	0.999
10. + Oil production per cap	0.991	0.993	0.981	0.987
11. + Democracy stock	0.997	0.998	0.998	0.998
12. + Tax revenue	0.994	0.996	0.998	0.999
13. + Telephone mainlines	0.997	0.998	0.999	0.999
14. + Gini index	0.984	0.999	0.998	0.999
15. + Conflicts	0.998	0.996	0.993	0.995
IV. Changes to the sample				
16. - OECD countries	0.995	0.998	0.999	0.999

Table 6:

Comparisons between "Raw" and Model-based Measures of Public Health Performance

	Public health Index (raw)	
	Spearman's	Pearson's
Achievement (2004)		
Model 1 residual	0.522	0.491
Improvement (1960s-2004)		
Model 2 residual	0.761	0.722

All correlations significant at 99%.