Social Science & Medicine 93 (2013) 21–28

Contents lists available at SciVerse ScienceDirect

Social Science & Medicine

journal homepage: www.elsevier.com/locate/socscimed

Assessing health system performance: A model-based approach

John Gerring [a,*], Strom C. Thacker [a,b], Ruben Enikolopov [c], Julián Arévalo [d], Matthew Maguire [a]

[a] Department of Political Science, Boston University, 232 Bay State Road, Boston, MA 02215, United States
[b] Department of International Relations, Boston University, 152 Bay State Road, Boston, MA 02215, United States
[c] New Economic School, Nakhimovsky pr. 47, Moscow, 117418, Russia
[d] Facultad de Economía, Universidad Externado de Colombia, Calle 12 No. 1-17 Este, Bogotá, Colombia

ARTICLE INFO

Article history:
Available online 14 June 2013

Keywords:
Global health
Health policy performance
Health policy assessment

ABSTRACT

It is difficult to assess countries' relative success in addressing issues of public health because countries are subject to very different background conditions. To address this problem we offer a model-based approach for assessing health system performance. Specifically, an index of public health is regressed against a vector of variables intended to capture economic, educational, cultural, geographic, and epidemiological endowments. The residual from this model is regarded as a plausible measure of public health performance at the national level.

We argue that a model-based approach to performance is informative for policymakers and academics as it focuses attention on those aspects of a country's health profile that are not constrained by structural factors. This sharpens comparisons across countries and through time, and also allows one to evaluate the degree to which health systems have lived up to their potential.

Introduction

How should one assess country performance in the health sector? 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 efforts. "Improvement" and "deterioration" become matters of speculation, and accountability for policy choices is impossible to establish.

Here, we limit ourselves to a discussion of health sector effectiveness, leaving aside issues of equity and efficiency. Even so, one is at pains to evaluate performance across countries and through time. The problem is not an absence of indicators; there are a plethora of indicators of health performance to choose from (Backman et al., 2008; Kruk & Freedman, 2008; Murray & Evans, 2003). However, most policy-related indicators (e.g., immunization rates) are partial; they do not provide a total picture of health sector performance in a country. Moreover, few indicators provide sufficient country or historical coverage to judge performance on a global scale and over time.

Mortality-based outcome measures such as infant mortality or life expectancy compensate for these shortcomings. However, they are strongly affected by factors that lie outside the health sector. Indeed, a large portion of the variance in health outcomes may be explained as a correlate of economics, geography, education, or disease vectors that affect countries differently for reasons having nothing to do with the health sector in a particular country. 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 United States and Canada, it is virtually meaningless to compare this statistic across countries with vastly different 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 background factors.

This is the intuition behind most international comparisons. When writers point to the extraordinary achievements of countries like Costa Rica, Cuba, and Sri Lanka they are (implicitly) comparing human development achievements in these countries relative to certain baseline characteristics thought to lie outside—or at least be separable from—the social policy sector (Caldwell, 1986; Ghai, 2000; Halstead, Walsh, & Warren, 1985; McGuire, 2010; Mehrotra & Jolly, 1997; Riley, 2007). Unfortunately, this handicapping exercise is rarely conducted in a systematic and explicit fashion. One

* Corresponding author.
E-mail address: jgerring@bu.edu (J. Gerring).

0277-9536/$ – see front matter © 2013 Elsevier Ltd. All rights reserved.
http://dx.doi.org/10.1016/j.socscimed.2013.06.002

may limit the comparison to other developing countries. But this means that we can say nothing about the relative success of the health sectors in countries with very different levels of development. Moreover, it requires a categorical judgment about a quintessentially scalar phenomenon—development. Separating "developing" and "developed" countries is highly arbitrary at the breakpoint. Finally, it excludes other factors that also affect public health performance but are not the responsibility of the health sector.

To overcome these difficulties, model-based approaches have been developed for use in health policy (e.g., Wang, Jamison, Bos, Preker, & Peabody, 1999; WHO, 2000 [reviewed by Jamison & Sandbu, 2001]), human development (e.g., Kakwani, 1993), and other policy contexts (e.g., Ndulu & O'Connell, 2007). The approach laid out in this study differs from previous studies in four respects: (a) the spatial and temporal breadth of the dataset (including all country-years from 1960 to 2010), (b) the inclusion of an extensive set background factors, (c) close attention to problems of specification (including a large number of robustness tests), and (d) both cross-sectional and fixed-effect models.

In the first section of the paper we define a strategy for measuring health outcomes and show changes over time in the global distribution of public health. In the second section, we construct a cross-sectional model of health sector performance that includes background factors with expected impact on a country's health. In the third section, we use the residuals, or unexplained variance, from these models to assess countries' relative achievement and improvement across the 1960–2010 period. In the fourth section, we construct a fixed-effect model whose purpose is to track within-country change over time. In the fifth section, we conduct a series of sensitivity tests in order to probe the robustness of the results from these models. The concluding section reflects on the interpretation and possible uses of model-based measures of policy performance. Note that *country-years* are the relevant units of analysis in the following discussion.

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, transformed by the natural logarithm (to account for expected non-linearities). Of all possible 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) is sometimes regarded as a more valid statistic it is so highly correlated with IMR that the differences are of very little practical significance. We opt for IMR solely by reason of its more extensive coverage. With respect to life expectancy, one might prefer a more sensitive measure that takes account of 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 (unadjusted) life expectancy, as we shall see.

Although IMR (ln) is a component of life expectancy, our index combines both statistics in a single indicator of public health, understood as the sum of the (standardized) values for the two component variables. We adopt this approach 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 highly correlated, 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, the combination of life expectancy and IMR (ln) offers a more reliable, more sensitive, and more insightful measure of public health than either would provide on its own. Data sources are explained in Table A1.

Table 1
Correlation table: human development indicators.

	Life expectancy	IMR (ln)	Public health index
Life expectancy		−0.9025	0.9753
IMR (ln)	−0.9025		−0.9753
HALEs (males)	0.8157	−0.7810	0.8179
HALEs (females)	0.8225	−0.7972	0.8302
DALYs	−0.9416	0.7761	−0.8800
Child mortality rate (ln)	−0.9263	0.9958	−0.9853
Malnutrition (height for age)	−0.6630	0.7790	−0.7394
Poverty headcount ratio ($2/day)	−0.7560	0.8021	−0.8145
Human development index (HDI)	0.9181	−0.9320	0.9513

Pearson's r correlations, based on varying samples (no imputed data). All are significant at 99%. See Table A1 for variable definitions and sources.

Readers may be curious to know how the resulting index compares with other measures of public health such as health-adjusted life expectancy (HALE), disability-adjusted life years (DALYs), child mortality, malnutrition (as proxied by height for age), as well as broader quality-of-life measures such as the poverty headcount ratio (percentage of the population living on less than $2 a day), and the UNDP's Human Development Index. Table 1 shows inter-correlations among these alternate measures and (a) the components of our index and (b) the composite index itself. Not surprisingly, our public health index is highly correlated with other measures of public health and with other quality-of-life measures. Good (bad) things generally go together.

One of the benefits of an index with broad coverage is that one can employ it to compare global distributions at varying points in time. Fig. 1 displays a kernel density plot of the distribution of public health in 1960 and 2010, with the area under each portion of

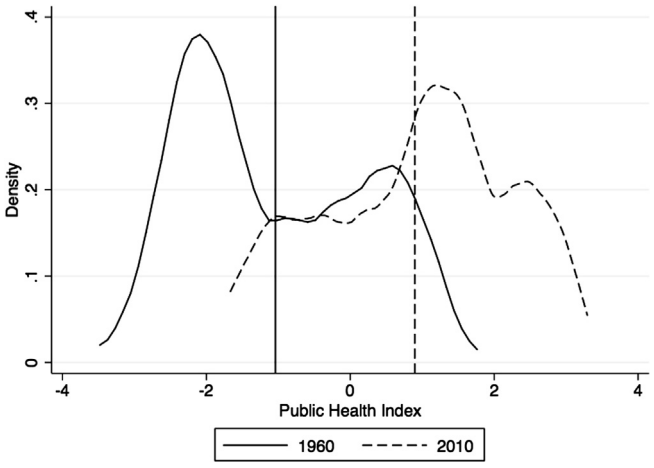

Fig. 1. Global distribution of public health by country. Kernel density plot of the distribution of public health in 1960 and 2010. Vertical line: mean value for that year's distribution. Unit of analysis: countries.