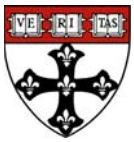

PROGRAM ON THE GLOBAL DEMOGRAPHY OF AGING

Working Paper Series

Short and long-term relationship between physician density on infant mortality: a longitudinal econometric analysis

Mansour Farahani S. V. Subramanian and David Canning

June 2009

PGDA Working Paper No. 49
http://www.hsph.harvard.edu/pgda/working.htm

The views expressed in this paper are those of the author(s) and not necessarily those of the Harvard Initiative for Global Health. The Program on the Global Demography of Aging receives funding from the National Institute on Aging, Grant No. 1 P30 AG024409-06.

Short and long-term relationship between physician density on infant mortality: a longitudinal econometric analysis

Mansour Farahani S. V. Subramanian and David Canning
Harvard School of Public Health, Boston, MA, USA

June 2009

Abstract: While countries with higher levels of human resources for health typically have better population health, the evidence that increases in the level of human resources for health leads to improvements in population health is limited. We provide estimates of short-run and long-term effects of physician density on infant mortality. We use a dynamic regression model that allows an estimation of both short- and long-run effects of physician density on infant mortality. We also used instrumental variables analysis to identify the causal effect of physician density on health. We estimate that increasing the number of physicians by one per 1,000 population decreases the infant mortality rate by 15% within five years and by 45% in the long-run. We find all countries are moving towards their own steady state at around 3% a year and are only half way there after 15 years. We conclude that the long-run effects of human resources for health are substantially larger than previously estimated. Our results suggest that health sector inputs can play a role in reducing infant mortality. However, meeting the Millennium Development Goal of reducing child mortality rate by two thirds from 1990 to 2015 would have required much earlier action.

Introduction

Building and strengthening capacity in human resources for health has been recognized as critical to alleviating health crises in less developed countries, besides contributing to the sustainable development of health systems in all countries (Chen et al., 2004). Multiple studies demonstrate that countries with higher levels of human resources for health typically have better population health (Flegg, 1982, Robinson and Wharrad, 2000, Anand and Baernighausen, 2004, Speybroeck et al., 2006, Or et al., 2005, Aakvik and Holmas, 2006, Jamison et al., 2004). The density of human resources for health, including the supply of physicians, nurses, and other health professionals, has been shown to be positively correlated with percentage of deliveries assisted by skilled birth attendants and the proportion of children fully immunized against measles (Anand and Barnighausen, 2007, Speybroeck et al., 2006) and negatively correlated with maternal, infant, and under-5 mortality (Anand and Baernighausen, 2004). In a recent cross-sectional analysis of 83 countries, Anand and Baernighausen (2004) report that of the different components associated with human resources for health, physician supply was significant in explaining country variations in maternal, infant, and under-5 mortality.

While prior studies have established that countries with higher levels of physician density typically have better population health, the evidence that *increases* in the level of physician density *lead* to improvements in population health is limited. Most of the previous studies examining the effect of physician density are cross-sectional in design (Anand and Baernighausen, 2004, Flegg, 1982, Robinson and Wharrad, 2000, Speybroeck et al., 2006), and are thus incapable of attributing improvements in

population health directly to increases in physician density. There are two problems with the cross-sectional approach. One is that the level of physicians may be correlated with country characteristics, such as climate, that affect health outcomes. The second is that studies assume that the effect of physician density on health is immediate and thereby may underestimate the full long-run impact of physician density on population health.

There are compelling reasons to anticipate a delay in the population health response when we increase the number of physicians. For instance, interventions that improve the health of young women have effects throughout their reproductive years that may also improve the health of their children. Physicians can also act as catalysts in motivating change in a patient's lifestyle, which most noticeably affects morbidity and mortality in the long run (Andersen and Blair, 1997, Bull and Jamrozik, 1998, Galuska et al., 1999). Physicians have a pivotal role in the implementation of new technologies, whether the in the form of new vaccines, drugs and medical procedures, which are, in turn, a major source of health improvements (Cutler et al., 2006). Besides having an impact on health in the short-run, increases in physician density are also likely to contribute to faster adoption of these technological innovations in the longer run (McClellan and Kessler, 1999, Booth-Clibborn et al., 2000, Packer et al., 2006).

Utilizing longitudinal panel data from 99 countries over the years 1960-2000, we investigate whether increases in physician density lead to reductions in infant mortality rates, and provide estimates of both the short- and long-term effects of increased physician density.

Data

We constructed a longitudinal panel data set with intervals of five years[1] from 1960 to 2000 from three different data sources including the World Bank's World Development Indicator, Penn World Table, and Barro-Lee's educational attainment dataset (World Bank, 2006, Heston et al., 2006, Barro and Lee, 2001). The infant mortality rate (IMR), defined as the number of deaths in children under one year of age per 1000 live births, and obtained from the World Development Indicators (World Bank, 2006), is the outcome of interest. There is, however, a concern with the quality of the infant mortality rate data (Hill and Pebley, 1989, Hill, 1999, Hill et al., 2007). In developed countries, the most common sources of data on infant mortality rates are vital registrations, while in developing countries the best data come from World Fertility Surveys in the early part of the time period and Demographic and Health Surveys in the later part. However, survey data for particular years are frequently missing, and IMR data are created by interpolation or extrapolation over time and even across countries (Hill and Pebley, 1989, Bos et al., 1992, Hill, 1999, Hill et al., 2007). Due to doubts about data quality, we performed an inspection of the infant mortality data, and removed the data deemed unreliable (United Nations, 1992, United Nations, 2006). Although pure measurement error in the dependent variable should not induce biases; it is possible that the extrapolation procedure used in producing the World Bank data induce systematic biases.

We also used under 5 mortality and life expectancy as dependent variables. However, since Life expectancy at birth estimates are derived from IMR and q5 estimates from surveys or censuses with model life tables, and Under 5 mortality rate is highly

[1] We prefer using IMR every fifth year starting in 1960 to averaging the five-yearly data as in United Nation data base, since averaging introduces serial correlation. Nevertheless, our results are robust to using five-year averages.

correlated with infant mortality (Bos et al., 1992). Our comparison of the data on child mortality from UNICEF's report with our data on infant mortality showed that they are highly correlated, both across countries (0.964 in 2000) as well as for changes over time (the correlation of the percentage changes from 1960 to 2000 is 0.913). Given this high correlation, it is not surprising that the regression results are substantially similar. (The results can be provided upon request).

The main reason for choosing IMR results over life expectancy at birth is a pure measurement issue. Data on adult mortality rates are often even less reliable (Hill, 2003, Bos et al., 1992) which is why we prefer to work with infant mortality as our health outcome measure. Life expectancy figures are generally derived from model life tables rather than observed directly from death registrations. The life expectancy estimates reported in are simply updated using an infant (or child) mortality figure applied to the model life table.

Our dataset is a panel of 99 countries at five-year intervals from 1960 to 2000. While the data for infant mortality are available at annual frequencies in the World Bank databases, the education and physician data are available only at five-year intervals. We use data from countries where we have at least three observations out of nine potential in this period. We have data on outcome variable and explanatory variables for 602 observations out of a possible 891, giving an unbalanced panel where each country has between three and nine observations. If the potential observations are missing in some systematic way, this may bias our results. We discuss the pattern of the missing data and the sensitivity of our results to the sample size in the appendix.

We use data on physician density, defined as the number of physicians per 1,000 population obtained from the World Development Indicators (World Bank, 2006). We have restricted this analysis to physician supply, since longitudinal data on other human resources for health is lacking. Furthermore, physician density has been shown to be strongest predictor of health outcomes among the various components of human resources for health (Anand and Baernighausen, 2004).

Ideally, we would prefer to add data on nurses and midwives, and health infrastructure variables, such as number of the beds. However, our dataset would be restricted to some of the OECD countries and mainly from the 1980s onward. Therefore, paucity of data on other health system resources, especially human resources for health, makes us use physicians as a proxy measure of the other health system inputs. For other factors out of the health system that can affect health, we could not use time-invariant variables, such as the country coastal area, whether it is landlocked, or located in tropical, etc. because our final model would remove them through first differencing. For the Time variant variables like average temperature, altitude and rainfall, just like other studies we did not find any significance in neither the static nor the dynamic models. Other variables which were related to development level such as length of paved roads, railways, and capacity for electricity generation were the cause of multicollinearity with GDP. The data of percentage of population who had access to water and sanitation could also be useful. But the available data reduce the sample size to 63 countries over 20 year period. We also tried to include the Gini coefficient as a measure of inequality of income distribution or inequality of wealth distribution. However, the data available in the World Bank data base for the Gini were sparse. We collected all the articles pertaining to Gini

measurement over the period of study (1960-2000). However, because of inconsistency due to the variation in data collection (for example, total population vs. employed adult population) or in method (Total consumption vs. gross income), we were not able to make a usable data set.

We now turn to the functional form of the relationship we estimate. It is common practice to use the log of IMR, as opposed to the level of IMR (Anand and Baernighausen, 2004, Robinson and Wharrad, 2000, Flegg, 1982, Jamison et al., 2004) as the variable to be explained. The level of the IMR is bounded below by zero and many countries are near this bound, and can, in consequence, show only very small improvements in the level of IMR. Rather than model the outcome variable as bounded, we take the log of IMR as our dependent variable. Using the log of IMR means we can interpret the effects of our explanatory variables in terms of the percentage change they generate in the IMR.

While prior studies (Anand and Baernighausen, 2004, Robinson and Wharrad, 2000, Flegg, 1982, Jamison et al., 2004) use the log of physician density as an explanatory variable, we undertook an analysis of which functional form gives the best fit to the data and found that the level of physician density has better explanatory power than the log of physician density and even outperforms a model that includes both the log of physician density and its square. The fit of the different functional forms was evaluated by using non-nested hypotheses tests (Davidson, 1993, Mizon and Richard, 1986). The details of this analysis are reported in the appendix.

We also use income levels, measured by the log of per capita Gross Domestic Product (GDP) at year 2000 purchasing power parity (PPP) adjusted dollars from the

Penn World Tables (PWT 6.2) (Heston et al., 2006) and the average years of schooling in the total population aged 15 years and above, (Barro and Lee, 2001) as covariates in this analysis. Since we account for country fixed effects in our analysis, we have controlled for any confounding variables that are fixed in a country over time.

In any analysis of observational rather than experimental data confounders are a fundamental problem which is difficult to overcome. Our panel data approach is more robust than cross-sectional studies in that we control for country-specific factors that are fixed over time. However, this leaves open the issue of time-varying confounders. We do control for education, income, and a worldwide time trend. While this takes account of some of the time-varying confounders, there are clearly other potential confounders not included in the model.

One major potential confounder is technical progress in health care that can improve health outcomes even with a steady level of health inputs; this has been a major source of long-term advancement in health. We control for worldwide technical progress by using time dummies, but there may be country-specific technical progress in health as well (see Jamison et al 2004). Our dynamic model allows for diffusion of health technologies from advanced to less advanced countries, so that countries converge towards a common technological frontier. However, since we cannot claim to have eliminated all confounders, our results must be interpreted with caution.

Statistical Analysis

Cross-sectional specification: we first estimated a cross-sectional regression model, similar to those utilized in prior studies given as:

$$LIMR_i = \alpha PHY_i + \beta X_i + e_i \quad (\mathbf{1})$$

where log infant mortality rate ($LIMR_i$) in country i at one time-period (we use the year 2000) is a function of physician density in the country (PHY_i), the effect of which is given by α; and a vector of covariates X_i, the effect of which is given by β; and an error term (e_i)[2].

Longitudinal specification: Utilizing the longitudinal panel data, we specified the following model:

$$LIMR_{it} = \alpha PHY_{it} + \beta X_{it} + \tau_t + e_{it} \quad (\mathbf{2})$$

where *LIMR* in country i at time t is a linear function of physician density (αPHY_{it}) and covariates (βX_{it}) in that period. Additionally, we also model a time-specific effect (τ_t) that estimates the global level of health technology available at time *t*; and allow for an error term (e_{it}). Model (**2**) is similar to model (**1**), except that we run the ordinary least squares estimator on the entire sample over time as well as across countries. An alternate and more robust specification for longitudinal panel data is:

$$LIMR_{it} = \alpha PHY_{it} + \beta X_{it} + \tau_t + \eta_i + e_{it} \quad (\mathbf{3}).$$

[2] Assumed to be random for each observation but independently drawn from a common distribution.

The main improvement from model (**2**) to (**3**) is the specification of a time-invariant country-specific fixed effect (η_i) that reflects unobserved country variables that affect infant mortality but do not change over time.

All these specifications suffer from reliance on the assumption that the effect of physician density on health outcomes is instantaneous. If the physician density increases, the current IMR will reflect the short-run effect but the long-run effects will not yet have come into operation. We therefore expect these approaches to underestimate the long-run effects of physician density. Estimation of long-run effects requires a dynamic structure where the full effect of changes in physician density may only come about with a lag.

Long-run dynamic specification: To allow for the possibility of long-run effects of physician density on infant mortality, we propose the following specification:

$$LIMR_{it} = \alpha_0 PHY_{it} + \alpha_1 PHY_{i,t-1} + \beta_0 X_{it} + \beta_1 X_{i,t-1} + \lambda LIMR_{i,t-1} + \tau_t + \eta_i + e_{it} \quad (\mathbf{4}).$$

Model (**4**) posits that the log of infant mortality rate $LIMR_{it}$ depends not only on the current level of physician density PHY_{it} and other covariates X_{it}, but also on the lagged physician density and other covariates measured in the previous period five years before, $PHY_{i,t-1}, X_{i,t-1}$. We also include the lag of the log of IMR, $LIMR_{i,t-1}$ as an explanatory variable. This measures the persistence in the IMR; in most cases, IMR responds to health sector inputs, but is slow to move from its previous levels.

In this framework, α_0 is the immediate short-run impact of an increase in physician density on the log of IMR. In addition to this short-run effect, there is also a long-run effect. After a five year lag, physician density will start to have an effect.

While the lag in this framework is only five years, the response of the infant mortality rate to the model unleashes a process of dynamic adjustment to a new steady state that can take a long time. In the long run, an increase in physician density will raise the log of infant mortality rate by $\frac{\alpha_0 + \alpha_1}{1-\lambda}$.

Model (**4**) has the advantage of including specification (**3**) as a special case. If $\lambda = 0, \alpha_1 = 0$ and $\beta_1 = 0$, the lagged effects disappear, short-run and long-run effects are the same, and we have exactly specification (**3**). However, in all cases other than this extreme one, we have both the short-run impact of changes on health inputs, and a new level of inputs that gradually converge towards the new long-run steady state.

A major issue in this kind of study is identifying a causal effect as opposed to a merely an association between physician density and health. There are two issues of concern. First, there is the threat that both good health and high density of physicians are associated with some other country-level factor, such as "good government," that is not included in the model. With variables omitted from the model, a statistical relationship between health and physician density may be merely an incidental association. Using panel data and allowing for country fixed effects can alleviate this concern. The second possibility is that countries with good health could have higher demand for health care professionals, which subsequently results in higher number of physician per capita. In this case, a statistical relationship between health and physician density does not establish causality between physician density and better health (a case of reverse causation). This study uses the instrumental variables estimation method to address both of these issues (Wooldridge, 2002).

We estimate model (**4**) using the Generalized Method of Moment (GMM) estimator which corrects for (a) the potential endogeneity[3] of contemporaneous changes in the independent variables, (b) unobserved heterogeneity at the country level (the country fixed effects), and (c) the endogeneity of lagged level of IMR in the dynamic specification. It should be noted that conventional estimators, such as ordinary least squares or generalized least squares, will be biased and inconsistent in this framework (Wooldridge, 2002, Cameron, 2005, Blundell et al., 2000).

To address these problems, the GMM estimator uses the suitably lagged levels of the variables as instruments, after the equation has been first-differenced to eliminate country-specific effects.[4] However, when the marginal processes are close to random walk process, the lagged levels of the series are weakly correlated with the subsequent first differences, and as a result the instruments available for the first-differenced equations are weak (Blundell and Bond, 1998). Arellano and Bover (1995) describe that if the original equations in levels were added to the system, additional moment conditions could be used to increase efficiency. In these equations, predetermined and endogenous variables in levels are instrumented with suitable lags of their own first differences. Strictly exogenous regressors enter the instrument matrix in the conventional instrumental variables fashion: in first differences, with one column per instrument. In this estimation, we relax the assumption of exogeneity on all the variables and treat them all as endogenous. The only exogenous variable to enter the instrumental matrix is time. We use this approach to estimate model (**4**).

[3] In econometrics endogeneity in a regression model refers to the correlation of the independent variable with the error term.

[4] The within transformation is not useful in this context, since it introduces the shocks from all time periods into the transformed error term.

It is clear that the GMM framework deals consistently and efficiently with the estimation problems such as endogeneity and unobservable country-specific heterogeneity. This consistency, however, critically hinges upon the identifying assumption that the lagged values of IMR and the other explanatory variables are valid instruments in the regression. A crucial necessary condition in this respect is the lack of serial correlation in the errors, e_{it}. To address these concerns, a battery of specification tests complements the estimation results. In particular, we performed a Sargan test of overidentifying restrictions (Cameron, 2005). This test is based on the sample analog of the moment conditions exploited in the estimation process, and evaluates the overall validity of the set of instruments.[5] The validity of lagged levels dated t-2 as instruments in the first-differenced equations is rejected by the Sargan and Hansent tests of overidentifying restriction (for Hansent test results not reported)(Sargan, 1958, Hansen, 1982). This is consistent with the presence of measurement errors. Instruments dated t-3 and earlier are not rejected, therefore, used as instrument.

The Sargan or Hansen tests are prone to weakness, therefore should be interpreted cautiously. Sargan test is not robust, but not weakened by many instruments, while Hansen test is robust, but can be weakened by many instruments. The test actually grows weaker the more moment conditions there are. That is the error in this test is proportional to the number of instrumental variables, so if the asymptotic approximations are to be used, this number must be small. While these two tests are the most commonly used methods to detect serial correlation of the error term in a dynamic model based on panel

[5] In robust one-step GMM, non-sphericity in the errors is suspected, therefore the Sargan statistic is inconsistent. In that case, a theoretically superior overidentication test for the one-step estimator is that based on the Hansen statistic from a two-step estimate. We performed the two-steps GMM in order to obtain a consistent Hansen statistic.

data, their application is limited to uncorrelated disturbances under the null. Therefore, if there are reasons to expect autoregressive errors in a panel regression model, or if one suspect that the dynamics of the model have been incorrectly specified, there is a strong possibility of autocorrelation being present in the residuals. Hence, it is natural that we may consider a test of uncorrelated errors as a null against an AR(1) error as an alternative. If the disturbance has an AR(1) structure, the usual instruments of lagged values of the dependent variables in the differenced equations are no longer valid. Furthermore, our estimator that uses lags as instruments under the assumption of white noise errors loses its consistency if in fact the disturbances are autocorrelated. Thus, Sargan and Hansen tests would be no longer applicable because they use inconsistently estimated residuals based on even optimal two-step estimation which also use invalid instruments. In order to remedy this problem, we used a t-test developed by Jung. The t-test utilizes consistently estimated residuals based on IV estimation which uses the lags of exogenous variable as instruments for the lagged dependent variables (Jung, 2005). The assumption of strict exogeneity of an explanatory variable is rather strong. However, it is safer to assume this than to restrict the serial correlation structure of the errors where it is suspected that the error has an autoregressive structure. For all GMM regressions, our tests of overidentifying restrictions indicate that we cannot reject the hypothesis that our identification assumptions are valid.

We report the results of two tests of the null hypothesis of no serial correlation in the errors in levels. One test is based on the difference between the Sargan test statistics described above and another one obtained after dropping all the moment conditions that would be invalid if the errors in levels were first-order serially correlated (see Arellano

and Bond, 1991). This test is labeled AR (1) in Table 4, which cannot reject the null hypothesis of no serial correlation in the errors in levels. The second is a test of the hypothesis that the errors in the differenced equation are not second order serially correlated. This test is labeled AR (2) in Table 4. The AR (2) is normally distributed, while the Sargan and the difference-Sargan statistics are chi-square distributed under the null hypothesis. The AR (2) test, reported at the bottom of the column, indicates that there is no further serial correlation and the overidentifying restriction is not rejected.

Results

Table 1 reports the global average of the different variables for each time period. IMR more than halved between 1960 and 2000, while physician density almost quadrupled during the same period. Table 2 reports the mean, standard deviation, minimum and maximum of each variable over the whole sample period. Table 3 presents the results for the static models (**1**), (**2**), and (**3**)[6]. Estimates for model (**1**), for the 2000 cross-section, suggest a strong inverse relationship between physician density and the log of IMR. Adding one physician per 1000 persons appears to reduce IMR by almost 24%. Estimates for model (**2**) based on longitudinal data also show a strong inverse association between physician density and the log of IMR, and are consistent with a 21% reduction in IMR for every physician added per 1000 population. Finally, including country fixed effects (Model (3)) strengthens the association suggesting that adding one physician for every 1000 population is consistent with a reduction in infant mortality by about 30%, which is opposite to what one expects to see when the country fixed-effect is added. Generally, when country fixed-effect is added to the equation, the coefficient of interest

[6] All the estimations are robust and clustered at the country level.

tends toward zero. However, in this case one possible explanation could be that the rate of the expansion of physician workforce is much slower than other health system inputs, due to the long and expensive training process; we see that adding the fixed effect shows stronger effects. This result implies that the association between IMR and physician density found in cross-section is robust and not the result of confounding variables that are country-specific and fixed over time.

Table 4 reports the results estimating dynamic model (4), where one lag for both dependent and independent variables is included. Theoretically, we expected the lag effects; however, we needed to first rule out the unit root. In order to do that we used augmented Dickey-Fuller test which assumes that all series are non-stationary under the null hypothesis against the alternative that at least one series in the panel is stationary. The unit root was rejected at 0.001 level of statistical significance. To determine the number of lags, we used the Akaike's information criterion (AIC). As one can see in the table the lag of the independent variables are individually statistically insignificant. Nonetheless, the F test indicates that they are jointly statistically significant. Therefore, dropping them might cause omitted variable bias. We find significant contemporaneous effects of physician density: an increase of one physician per 1000 population reduces IMR by about 14% immediately. However, we also find significant long-run effects. While the lagged values of our explanatory variables are not individually statistically significant, the lagged value of the log of IMR is highly significant.[7] This difference indicates that the static models are mis-specified. The large coefficient on the lag of IMR means that IMR is slow to adjust to changes in health inputs and tends to persist close to

[7] An F-test for joint significance of the lagged and the contemporary coefficients shows that they are highly significant.

its previous level. In addition, we find significant effects of income on the IMR but do not find statistically significant effects for education.

Table 5 shows the estimated long run effects of changes in our explanatory variables on IMR. The long-run effect of physician density is three times as large as the short-run effect; an increase of one physician per 1000 population reduces IMR by about 45% in the long run. Given the mean IMR in this time period is 63, a 45% reduction in IMR means 2.8 deaths averted per 1000 live births. Since the sample mean is about one physician per 1000 population, to achieve this level of reduction, on average, a doubling of the physician density around the world is required.

The speed of convergence to the long-run steady state is around 17% over a five-year period (the speed of convergence is 1- λ), or about 3% convergence per year. There is a quick reduction of IMR of around 14% and after about 15 years, every country is about half way to the long run steady state reduction of 45%.

Interpretation

One of the United Nations Millennium Development Goals (MDGs) is to reduce IMR by 2015. Health system inputs are among the factors that may influence the achievement of the MDGs. One of these inputs is healthcare workers. As far as infant mortality is concerned, healthcare workers are required to provide the medication, prenatal care as well as pediatric services. Adequate number of committed motivated health workforce with required public health and clinical competencies is a prerequisite for achieving MDG.

Using longitudinal data from 1960 to 2000 for 99 countries, we report substantial long-term effects of physician density on reductions in IMR. Our static models, using

specifications (**1**), (**2**), or (**3**), largely concur with the previous results in terms of their effect size (after adjusting for the fact that we use the level of physician density, while previous studies usually use the log of physician density). However, the dynamic model specification (**4**) produces a much larger long-run effect, though this effect is slow in coming. This long run effect may be three times as large as the contemporaneous effect usually estimated.

The Millennium Development Goal is to reduce child mortality by two-thirds, from 93 children of every 1,000 dying before age five in 1990 to 31 of every 1,000 in 2015. These deaths occur mainly in the developing world. Sub-Saharan Africa has the highest rates. While many developed and developing countries are on their track towards their goal, as one can see in Table 6, some of the Sub-Saharan countries are way off their target, partly because of the AIDS epidemics. This projection is based on the trend in 1960-2000. In some of these countries the downward trend of IMR that started in the 1960s reversed in the 1990s. However, with more access to antiretroviral treatment and prevention of mother-to-child HIV transmission (PMTCT) services, hopefully the outcome would be better than what we projected.

Several caveats should be taken into account when interpreting the results. The first is measurement error in the data and the extent of missing data in our sample. Second, we only considered one aspect of human resources for health (physician density). While previous studies found physician density to be amongst the most important components of human resources for health, we do not wish to imply that other components such as supply of nurses, or hospital beds, are unimportant. Considering additional dimensions of human resources for health in international longitudinal data,

however, would require information that is not presently available. Lastly, additional time-varying confounders within countries have the potential to bias these results.

There are areas for further research which will have to be addressed. Having physicians as a proxy for health system resources, even for human resources for health, limits the external validity of the results. Had we been able to run a more detailed analysis aided by accurate data of the nursing workforce, results would have allowed for more precision concerning the impact of human resources on infant mortality. Another limitation was the lack of ability to account for the distribution of the physicians. There are large variations in the physician distribution; in some countries access to physicians are limited to the urban area. We are also assuming that physicians in all the countries providing similar practices, which is due to the paucity of data on quality of training for medical professionals. Country-level information on how physician are practicing within the health care system of that particular country may be informative in identifying best medical practices for health care structure, licensure/registration, scope of practice, education and continuous education, and collaboration with other health care professionals, and the use of traditional care providers. New models are needed to assess care, and treatment and their impact on patients and their communities, all of which are largely influenced by culture and setting.

In conclusion, we find that the increase in physician density is an important determinant of infant mortality, particularly in the long-run. However, note should be taken that paucity of data on other health system resources, especially human resources for health, makes us use physicians as a proxy measure of the other health system inputs. Therefore, we should caution on the interpretation of the results, since physician density

is reflective of a more general commitment of resources to health care which is not accounted for here.

Prior studies, with their focus on contemporaneous effects, may have underestimated the full impact of physician supply on reducing IMR. Over time, investment in programs that increase the number of health care professionals can play an important role in helping to achieve the Millennium Development Goal of reducing infant mortality by two thirds in 2015 relative to 1990. However, our results suggest that the efforts to achieve this goal should have been more substantial early in the period, in order to reap the long-term benefits of early investment. It may be much more difficult to achieve the goal by increasing health sector inputs at this point in time, since we can expect to see only the short-term effects of increases in current health sector inputs before 2015.

Table 1 Sample Means 1960-2000

	Infant Mortality Rate	Physicians per 1000 population	GDP per capita	Years of Schooling
1960	105	0.39	3821	3.6
1965	89	0.54	4781	3.7
1970	87	0.49	5253	4.0
1975	76	0.61	6248	4.2
1980	71	0.68	6342	4.5
1985	58	0.94	7148	5.1
1990	52	1.10	7986	5.5
1995	46	1.17	8779	5.9
2000	40	1.38	10339	6.3
Average 1960-2000	63	0.91	7370	5.0

Table 2 Descriptive Statistics for the Whole Sample

Variable	Mean	Std. Dev.	Min	Max
IMR	63	50	3	204
Physician density	0.91	0.97	0.01	4.3
GDP	7370	7544	314	34365
Total years of education	5.04	2.86	0.17	12.05
Log IMR	4.14	1.08	1.1	5.32
Log GDP	8.9	1.12	5.75	10.45

Data from 602 observations used in the model

Log IMR: log of infant mortality rate, defined as the number of deaths of children under one year of age per 1000 live births in the same year. Physician density: number of physicians in 1000 population. GDP: the gross domestic product per capita at 2000 real purchasing power (PPP) adjusted dollars from the Penn World Tables 6.2 (PWT6.2) of Heston and Summers (2002). Years of schooling is the average years of schooling in the total population, aged 15 years and above, from the Barro-Lee educational attainment dataset.

Table 3 Regression Results for the Static Models

	Cross-section 2000	Panel (1960-2000)	
	1	2	3
Estimator	OLS	OLS	Fixed-effect
Dependent Var.	Log IMR	Log IMR	Log IMR
Physician density	-0.239***	-0.212***	-0.302***
	(0.08)	(0.03)	(0.03)
Log GDP per capita	-0.522***	-0.39***	-0.261***
	(0.09)	(0.07)	(0.03)
Total years of education	-0.0963***	-0.137***	-0.013
	(0.03)	(0.02)	(0.02)
Constant	8.549***	8.009***	6.594***
	(0.63)	(0.46)	(0.27)
R^2	0.888	0.881	0.976
# countries	92	99	99
# observation	92	602	602

Cluster robust standard errors in parentheses
* significant at 10%; ** significant at 5%; *** significant at 1%
Column one is the cross-section estimation of data in year 2000; columns 2 and 3 are the estimation of the panel 1960-2000 without and with fixed-effects respectively.
Time dummies included in the panel data models (not reported).
Fixed effects are included in the model in column 3 (not reported).
For the description of the variables see note to Table 2.
All models estimated with ordinary least squares (OLS).

Table 4 Regression Results for the Dynamic Model

Dependent Variable		Log IMR
Explanatory Variables	Parameter	GMM Estimate
Lag Log IMR	λ	0.830***
		(0.057)
Physician density	α_0	-0.143**
		(0.057)
Lag Physician density	α_1	0.0688
		(0.053)
Log of GDP	β_0	-0.338**
		(0.16)
Lag log of GDP	β_1	0.216
		(0.162)
Years of Schooling	γ_0	-0.011
		(0.02)
Lag Years of Schooling	γ_1	0.008
		(0.017)
Constant		1.632***
		(0.61)
Number of observations		480
Number of countries		99
AR(1) (z-statistic)	Pr > z = 0.014	-2.47
AR(2) (z-statistic)	Pr > z = 0.048	-1.98
Sargan test (χ^2)	Pr > chi2 = 0.000	135.54
Sargan difference test		0.44

Robust standard errors in parentheses
* significant at 10%; ** significant at 5%; *** significant at 1%
Estimates are GMM (system estimator) with two-step estimates and heteroskedasticity-consistent standard errors and test statistics.
AR(1) and AR(2) are tests for first-order and second-order serial correlation, asymptotically N(0,1).
Time dummies are included (not reported).

Table 5: The Long-run Effects from the a Dynamic Model Specification

Explanatory Variables	Parameter Calculation	Long Run Effect Estimate
Long-run Physician density	$(\alpha_0 + \alpha_1)/ (1 - \lambda_1)$	-0.446**
		(0.077)
Long-run log of GDP	$(\beta_0 + \beta_1)/(1 - \lambda_1)$	-0.714**
		(0.097)
Long-run Years of Schooling	$(\gamma_0 + \gamma_1)/(1 - \lambda_1)$	-0.016
		(0.031)

Robust standard errors in parentheses
* significant at 10%; ** significant at 5%; *** significant at 1%

Table 6: Changes in IMR in Selected Sub-Saharan countries

country	IMR 1990	IMR 2000	Projection IMR 2015	MDG IMR 2015	% change 1990-2000 Observed	% change 2000-2015 Needed	Physician Density in 100,000 pop. In 2000
Botswana	45	74	68	15	64	-80	29
Zimbabwe	53	73	64	17	38	-76	16
Swaziland	78	98	85	26	26	-74	18
Kenya	64	77	67	21	20	-73	14
Rwanda	103	118	99	34	15	-71	5
South Africa	45	50	58	15	11	-70	77
Lesotho	75	75	64	25	1	-67	5
Zambia	101	102	87	33	1	-67	12
Congo, Dem. Rep.	129	129	109	43	0	-67	7
Sierra Leone	175	167	140	58	-5	-65	7
Uganda	93	85	72	31	-9	-64	8
Togo	88	80	68	29	-9	-64	7
Ghana	75	68	58	25	-9	-64	9
Senegal	90	80	68	30	-11	-63	6
Sudan	74	65	57	24	-12	-62	16
Guinea-Bissau	153	132	113	50	-14	-62	18
Tanzania	102	88	74	34	-14	-62	5
Benin	111	95	80	37	-14	-61	6
Niger	191	159	132	63	-17	-60	2
Malawi	146	117	98	48	-20	-59	3
Mozambique	158	122	102	52	-23	-57	2

Countries ordered according to the percentage of change between 1990 and 2000
Projections are calculated from the regression estimates using the IMR 2000
Negative number denotes reduction

References

AAKVIK, A. & HOLMAS, T. H. (2006) Access to primary health care and health outcomes: The relationships between GP characteristics and mortality rates. *Journal of Health Economics,* 25**,** 1139-1153.

ANAND, S. & BAERNIGHAUSEN, T. (2004) Human resources and health outcomes: cross-country econometric study. *Lancet,* 364**,** 1603-1609.

ANAND, S. & BARNIGHAUSEN, T. (2007) Health workers and vaccination coverage in developing countries: an econometric analysis. *The Lancet,* 369**,** 1277-1285.

ANDERSEN, R. E. & BLAIR, S. N. (1997) Encouraging Patients To Become More Physically Active: The Physician's Role. *Annals of Internal Medicine,* 127**,** 395-400.

ARELLANO, M. & BOND, S. (1991) Some Tests of Specification for Panel Data: Monte Carlo Evidence and an Application to Employment Equations. *Review of Economic Studies,* 58**,** 277-97.

ARELLANO, M. & BOVER, O. (1995) Another Look at the Instrumental Variable Estimation of Error-Components Models *Journal of Econometrics,* 68**,** 29-51.

BARRO, R. J. & LEE, J.-W. (2001) International data on educational attainment: updates and implications. *Oxf. Econ. Pap.,* 53**,** 541-563.

BLUNDELL, R. & BOND, S. (1998) Initial Conditions and Moment Restrictions in Dynamic Panel Data Models. *Journal of Econometrics,* 87**,** 115-43.

BLUNDELL, R., BOND, S. & WINDMEIJER, F. (Eds.) (2000) *Estimation in Dynamic Panel Data Models: Improving on the Performance of the Standard GMM Estimator,* Amsterdam New York and Tokyo, Elsevier Science, JAI

BOOTH-CLIBBORN, N., PACKER, C. & STEVENS, A. (2000) Health Technology Diffusion Rates Statins, Coronary Stents, and MRI in England. *International Journal of Technology Assessment in Health Care,* 16**,** 781-786

BOS, E., VU, M. & STEPHENS, P. W. (1992) Sources of World Bank Estimates of Current Mortality Rates. *Policy Research Working Papers.* Washington, DC, World Bank.

BULL, F. C. & JAMROZIK, K. (1998) Advice on exercise from a family physician can help sedentary patients to become active. *American Journal of Preventive Medicine,* 15**,** 85-94.

CAMERON, A. C. T. P. K. (2005) Microeconometrics : methods and applications. Cambridge University Press.

CHEN, L., EVANS, T., ANAND, S., BOUFFORD, J. I., BROWN, H., CHOWDHURY, M., CUETO, M., DARE, L., DUSSAULT, G., ELZINGA, G., FEE, E., HABTE, D., HANVORAVONGCHAI, P., JACOBS, M., KUROWSKI, C., MICHAEL, S., PABLOS-MENDEZ, A., SEWANKAMBO, N., SOLIMANO, G., STILWELL, B., DE WAAL, A. & WIBULPOLPRASERT, S. (2004) Human resources for health: overcoming the crisis. *The Lancet,* 364**,** 1984-1990.

CUTLER, D., DEATON, A. & LLERAS-MUNEY, A. (2006) The Determinants of Mortality. *Journal of Economic Perspectives,* 20**,** 97-120.

DAVIDSON, R. M. J. G. (1993) Estimation and inference in econometrics. Oxford University Press.

FLEGG, A. T. (1982) Inequality of Income, Illiteracy and Medical Care as Determinants of Infant Mortality in Underdeveloped Countries. *Population Studies,* 36**,** 441-458.
GALUSKA, D. A., WILL, J. C., SERDULA, M. K. & FORD, E. S. (1999) Are Health Care Professionals Advising Obese Patients to Lose Weight? *JAMA,* 282**,** 1576-1578.
GREENE, W. H. (2003) *Econometric analysis,* Upper Saddle River, N.J., Prentice Hall.
HANSEN, L. P. (1982) Large sample properties of generalised method of moments estimators. . *Econometrica* 1029-1054.
HESTON, A., SUMMERS, R. & ATEN, B. (2006) Penn World Table Version 6.2. Center for International Comparisons of Production, Income and Prices at the University of Pennsylvania.
HILL, K. (1999) *Trends in child mortality in the developing world : 1960 to 1996,* New York, NY, UNICEF.
HILL, K. (2003) Adult Mortality in the Developing World: What We Know and How We Know It. New York, Population Division, Department of Economic and Social Affairs, United Nations.
HILL, K., LOPEZ, A. D., SHIBUYA, K. & JHA, P. (2007) Interim measures for meeting needs for health sector data: births, deaths, and causes of death. *The Lancet,* 370**,** 1726-1735.
HILL, K. & PEBLEY, A. R. (1989) Child mortality in the developing world. *Population and development review,* 15**,** 657-687.
JAMISON, D., SANDBU, M. E. & J., W. (2004) Why Has Infant Mortality Decreased at Such Different Rates in Different Countries? *Disease Control Priorities Project,* 2**,** 21.
JUDSON, R. A. & OWEN, A. L. (1999) Estimating dynamic panel data models: a guide for macroeconomists. *Economics Letters,* 65**,** 9-15.
JUNG, H. (2005) A Test for Autocorrelation in Dynamic Panel Data Models. Tokyo, Hitotsubashi University Research Unit for Statistical Analysis in Social Sciences.
MCCLELLAN, M. & KESSLER, D. (1999) A Global Analysis Of Technological Change In Health Care: The Case Of Heart Attacks. *Health Affairs,* 18**,** 250-255.
MIZON, G. E. & RICHARD, J.-F. (1986) The Encompassing Principle and its Application to Testing Non-Nested Hypotheses. *Econometrica,* 54**,** 657-678.
OR, Z., WANG, J. & JAMISON, D. (2005) International differences in the impact of doctors on health: a multilevel analysis of OECD countries. *Journal of Health Economics,* 24**,** 531-560.
PACKER, C., SIMPSON, S. & STEVENS, A. (2006) International diffusion of new health technologies: a ten-country analysis of six health technologies. *International Journal of Technology Assessment in Health Care,* 22**,** 419-428.
ROBINSON, J. & WHARRAD, H. (2000) Invisible nursing: exploring health outcomes at a global level. Relationships between infant and under-5 mortality rates and the distribution of health professionals, GNP per capita, and female literacy. *Journal of Advanced Nursing,* 32**,** 28.
SARGAN, J. (1958) The estimation of economic relationships using instrumental variables. *Econometrica,* 26**,** 393-415.

SPEYBROECK, N., KINFU, Y., POZ, M. R. D. & EVANS, D. B. (2006) Reassessing the relationship between human resources for health, intervention coverage and health outcomes. *Background paper prepared for the World Health Report 2006 - Working Together for Health.* Geneva, World Health Organization.
UNITED NATIONS (1992) Child Mortality Since the 1960s- A Database for Developing Countries. New York, United Nation. Department of Economic and Social Development.
UNITED NATIONS (2006) World population prospects: the 2006 revision.
WOOLDRIDGE, J. M. (2002) *Econometric analysis of cross section and panel data,* Cambridge, Mass., MIT Press.
WORLD BANK (2006) World Development Indicators. The World Bank, Washington: D.C.

Appendix:

This appendix gives more details on three issues mentioned in the paper: missing data, the functional form, and the estimation of the dynamic model.

(a) Missing Data

Given that about one third of our data is missing, there is concern that the data is not missing at random but is purposeful in some way, which would cause a bias in our estimates. Table A1 reports the pattern of missing data by region. We have a slightly higher proportion of missing data in the Middle East and slightly lower proportion of missing data in Europe and the Americas. Africa, the poorest region in our sample, has about an average proportion of missing data. An examination of the data showed it is not the level of development that was driving missing data, but rather being a formerly Eastern bloc country. These transition countries frequently lack data for the early years of our sample.

Missing observations is not a problem for the consistency, or unbiasedness, of the estimator if the missingness is at random or if being missing depends on the value of an explanatory variable. The Arellano & Bond estimator (GMM) is an unbiased estimator, even with an unbalanced panel as in this case (Judson and Owen, 1999).

However, missing data will be a problem if being missing depends on the value of the outcome variable, the infant mortality rate. Table A2 shows the number of observations and missing data for each variable in our entire potential data set. While every variable has missing data, the years of schooling variable has the most missing data points (about half are missing) and effectively limits our usable sample. For our whole potential sample, there are missing data on the infant mortality rate about 25% of the time. However, if the sample is limited to one with three explanatory variables, there are missing data for infant mortality only 6% of the time. This means the lack of data on infant mortality is the major cause of missing observations in our sample.

While the missing data has not been driven by the level of infant mortality and is unlikely to introduce bias to our results, we could potentially improve the efficiency of our estimates (i.e. reduce the standard errors) by imputing values of the missing data and including the imputed data in our analysis. However, this would add an extra level of complication to an already complex methodology.

One way of enlarging the sample size is to drop years of schooling from the model, which was found to be insignificant anyway (see table 4). We examined how sensitive the results are to changing the sample in this way, and concluded that it does not significantly change the effect of physicians per capita (see Table A3).

Table A1

Region	Countries	Observations	Missing Data Points	% Data Missing
World	99	602	289	32.4
Africa	30	184	86	31.8
Asia and Oceania	17	100	53	34.6
Central and South America	22	142	56	28.3
Middle East	20	110	70	38.8
Europe and North America	14	93	33	26.2

Table A2

Variable	Observations	Missing Data Points	% Data Missing
GDP per capita	1345	549	29.0
Years of Schooling	921	973	51.4
Physicians per Capita	1342	552	29.1
Infant Mortality Rate	1427	467	24.7
Conditional on observing GDP per capita, Years of Schooling and Physicians per Capita			
Infant Mortality Rate	746	46	5.8

(b) Functional Form

Figure 2 plots the log of physicians per 1000 population against the log of IMR and plots the level of physicians per 1000 population against the log of IMR in 1960. The graphs suggest that the log of IMR varies linearly with the level of physician density but non-linearly with the log of physician density. We tested three models for the function form, using the level of physician density, the log of physician density and finally, both the log of physician density and its square (to allow for a non-linear effect), as explanatory variables. We estimate the static panel data model with fixed effects with these three different specifications for how physician density affects the log of IMR, keeping the other specifications fixed.

The three models are:

M1 : $LIMR_{it} = \alpha PHY_{it} + \beta X_{it} + \tau_t + \eta_i + e_{it}$

M2 : $LIMR_{it} = \alpha \log(PHY_{it}) + \beta X_{it} + \tau_t + \eta_i + e_{it}$

M3 : $LIMR_{it} = \alpha \log(PHY_{it}) + \alpha'[\log(PHY_{it})]^2 + \beta X_{it} + \tau_t + \eta_i + e_{it}$

We then conducted non-nested hypothesis tests to select between these models.
The tests used include Davidson-MacKinnon test (J test), Cox's and Pesaran's for non-nested models (Greene, 2003) :

Test 1. Null Hypothesis M1 is correct. Alternative is model M2.
t(741) -2.84510 p-val 0.056 accept

Test 2. Null Hypothesis M2 is correct. Alternative is model M1.
t(741) 10.80994 p-val 0.00000 reject

Test 3. Null Hypothesis M1 is correct. Alternative is model M3.
t(741) 0.95698 p-val 0.33889 accept

Test 4. Null Hypothesis M3 is correct. Alternative is model M1.
t(740) 4.99587 p-val 0.00000 reject

In each case, we accept the level specification for physician density and reject the specifications using the log of physician density at the 5% significance level.