Global Infant Mortality: Initial results from a cross-country infant mortality comparison project

Rebecca Anthopolos
Children's Environmental Health Initiative, Nicholas School of the Environment and Earth Sciences
Duke University
Durham, NC 27708-0328

Charles M. Becker
Department of Economics, Duke University
Durham, NC 27708-0097

August 2008

Abstract: The United Nations Millennium Development Goals have highlighted the usefulness of the infant mortality rate as a measure of progress in improving neonatal health care services, and more broadly as an indicator of basic health care overall. However, prior research has shown that infant mortality rates can be underestimated dramatically, depending on a particular country's live birth criterion, vital registration system, and reporting practices. These problems are especially great for perinatal mortality. This study seeks to assess infant mortality undercounting for a global dataset using an approach popularized in economics some three decades ago, when researchers sought to create internationally comparable, purchasing power parity-adjusted per capita income measures. Using a one-sided error, frontier estimation technique, it is possible to recalculate rates based on estimated parameters to obtain a standardized infant mortality rate and concurrently to derive a measure of likely undercount for each nation.

(*JEL* classifications **J11**, I10, C13)

1. INTRODUCTION

Comparative measures of economic development or social welfare are difficult to devise. The proxies traditionally used suffer from severe imperfections, and in consequence new measures have emerged. In particular, the economists' use of (deflated) GDP per capita has met with severe and justified criticism, even when the exchange rates used to convert various currencies to a common unit are adjusted for differences in purchasing power. "Augmented" GDP measures have been devised that account for natural resources, subtracting social bads such as pollution costs, and treating certain expenditures (for example, on maintaining social order) as an intermediate rather than a final product that should not be included. Yet, these measures still miss the fundamental point that human welfare has many components, of which many are not economic. This awareness in turn has given rise to a set of eight "millennium development goals" (MDGs), promoted under the aegis of the United Nations, that are intended to capture the multi-dimensional aspect of economic and social development (http://www.un.org/millenniumgoals).

The fourth millennium development goal is to reduce child mortality. Each broad goal in turn is broken down into pieces and ancillary but related objectives. For example, for Botswana, the child mortality MDG contains further objectives of reducing infant mortality from 48 per thousand live births in 1991 to 27 in 2011, to reduce the under five years mortality rate (U5MR) by two-thirds over this same period, to reduce child protein energy malnutrition (PEM) from 18% in 1998 to 8% in 2011, and to immunize 80% of all one-year olds by 2009 (Republic of Botswana, 2004). In all, Botswana had 22 specific goals, some of which seem ideological (such as #21, "develop further an environment conducive for beneficial trade and foreign direct investment"), but which for the main part reflect aspects of social welfare far more clearly than GDP measures.

This broader concept of development also would appear to have the advantage of being easier to calculate, especially for components such as infant and child mortality. Researchers who specialize in poor and middle-income countries especially appreciate this feature, since economic indicators are fraught with a multitude of measurement errors. Since many of the social components of the MDGs are almost certain to be highly correlated with economic prosperity, tracking them is useful for assessing overall economic policy success as well, and measurement errors are likely to be less.

Or so it has long been assumed by development experts disinterested in data sources and quality. In this paper, we argue that infant mortality rates tend to be wildly and systematically inaccurate, but that it is possible to bring some order to comparative assessments by making systematic, consistent corrections across

countries. It is important to emphasize the systematic nature of the corrections. At present, a researcher either must use inconsistent data reported by national statistical services (and generally available on the WHO website at http://www.who.int/healthinfo/morttables/en/index.html), or must accept corrections made by United Nations' demographers.

The underlying problem with vital statistics data is that they do not provide universal coverage. Furthermore, the errors are not random: they tend to be larger in poor and less urbanized nations. The problem is particularly acute in terms of measuring deaths during the first day of life, and, to a lesser extent, days 2-6. Differences in what is *de facto* regarded as a live birth further weaken cross-country comparability, while varying quality of national statistical offices' (NSO) efforts can make time series comparisons problematic as well. Most critically, the errors are essentially one-sided: it seems overwhelmingly likely that there are far more unreported infant deaths than unreported live births to children who survive infancy. In response, we seek to derive estimates that are reasonably comparable, and that reflect systematic rather than somewhat idiosyncratic corrections to official NSO data.

We begin the narrative by discussing and documenting the problem. Section 3 then addresses estimation strategy, while the following section provides a first pass at estimating a "true" relationship between infant mortality and socioeconomic variables, using UN data. Section 5 then uses these results to derive an initial correction of WHO data. We emphasize that these results are preliminary and incomplete: Section 6 summarizes additional corrective steps needed.

2: THE DATA: UNDER-REPORTING CORRECTIONS AND INFANT MORTALITY PATTERNS

Broadly speaking, there are four sources of data on infant mortality across countries. First, the World Health Organization (http://www.who.int/whosis/mort/en) collects data from NSOs throughout the world, and reports them without correction, though terse assessments of quality are offered. The United Nations Statistics Division (http://unstats.un.org/unsd/demographic/products/vitstats) also collects data and assesses quality; efforts as well are made to correct for under-reporting. Finally, bodies such as the EU's Eurostat (http://epp.eurostat.ec.europa.eu/portal/page?_pageid=1090,30070682,1090_33076576&_dad=portal&_schema=PORTAL), WHO's regional Pan-American Health Organization (http://www.paho.org/english/dd/ais/coredata.htm), or CIS Stat (http://www.cisstat.com/rus) offer separate and in some cases independent assessments of mortality in particular regions. An excellent way to get a sense of credibility of a particular mortality value generated from national vital statistics registration is to compare it, if possible, with estimates from detailed, retrospective household

surveys. The most important of these are the Demographic and Health (DHS) surveys (http://www.measuredhs.com).

No data are flawless. Mortality rate estimates can be overstated if deaths are more likely to be reported than an undercounted base population. Error in age-specific mortality is likely to arise as well if there are systematic errors in reporting age of death. However, in the case of infant mortality, unreported deaths relative to reported deaths are likely to exceed unreported births relative to all births, at least in developing countries, leading to a systematic downward bias in infant mortality statistics. Indeed, given the difficulty in consistently counting live births in developing countries, Kramer *et al.* (2002) recommend that countries with weak monitoring systems report a combined measurement of stillbirths and neonatal mortality. One could also follow a strategy implied in Wegman (1996), subtracting first hour deaths when comparing infant mortality across nations. More conventionally, demographers such as Kingkade and Sawyer (2001) and Aleshina and Redmond (2005) employ data fitting techniques to correct for underreporting in the first months of life.

Unreported deaths are especially likely when the infant lives only a very short period, so that no registration has occurred. Indeed, midwives may announce to the mother and family that a stillbirth occurred, rather than a live birth followed shortly by death, regarding this report as an act of mercy to a grieving family. It seems plausible that unreported death will be more likely for births outside of hospitals both because risks are higher and reporting systems are weaker. Non-hospital births are more common in poorer countries and rural areas, and there is evidence of dramatic rural under-reporting in some countries (Anderson and Silver, 1986; Becker *et al.*, 1998). In former Soviet republics, live births were recorded as such only if gestation and weight conditions were met (Anderson and Silver, 1997; Kramer *et al.*, 2002; for a discussion of global practices, see Wegman, 1996). While most countries have officially changed this policy to conform to WHO practice, in practice the old conditions are often used, again especially in rural areas. Several former Soviet republics also serve as examples of large recorded improvements in infant mortality that almost certainly reflect deteriorating data collection rather than genuine health improvements (Anderson and Silver, 1997; Becker *et al.*, 1998).

These points have long been recognized, and several are discussed at greater length in Hill and Choi (2006). They use DHS surveys to assess neonatal mortality, focusing on death heaping (at day 7) and underreported early neonatal mortality rates (ENMRs, defined as day 0-6 mortality) relative to late neonatal mortality (LNMRs, day 7-27 mortality). They adjust data to correct for heaping, and then compare adjusted ENMR/LNMR ratios for developing countries relative to historic rates for England and Wales, controlling

for total infant mortality rate. They find little evidence of systematic bias in the ENMR/LNMR ratios over time, though the ratio does vary considerably across region. Thus, once day-7 death heaping has been corrected, there is little reason to believe in systematic relative undercounting from DHS data. However, the issue is not fully resolved, since DHS surveys are neither universal nor annual, and since it is not obvious that the historic comparison employed is appropriate. Most importantly, there are several reasons to suspect that even DHS infant mortality rate data suffer from some under-counting, even if ENMR/LNMR ratios do not (Hill and Choi, 2006:443-444; note in particular the comparison with a detailed site analysis from Maharashtra discussed in Bang *et al*., 2002).

The consequences of these various sources of under-reporting can be large. Wuhib *et al.* (2003) find that switching from Soviet to WHP live birth definitions raised the 1996 infant mortality rate in Kazakhstan's Zhambyl oblast (province) from 32 deaths per thousand live births to 58.7 deaths. The extent of underreporting in official data for transition nations is detailed in Aleshina and Redmond (2005), who contrast (still possibly underreported) DHS estimates with official tallies. The largest discrepancy occurred in Azerbaijan, where the official 2001 IMR, 17, contrasts with the survey estimate of 74. In a majority of cases, the survey IMR estimate was more than double the official estimate. Aleshina and Redmond (2005) also estimate that adjusting the live birth definition to WHO standards would raise recorded IMRs by 5 percent to 40 percent, depending on the country and year. Thus, while live birth definition matters, it hardly explains the entire discrepancy. Kingkade and Sawyer (2001) force transition nations' mortality patterns in the first three months of life relative to month 4-10 infant mortality to replicate US and German data from periods of similar overall mortality. Doing this raises 1987-2000 IMRs from a low of 0.3% in Slovakia (to 11.6 deaths/thousand) to highs of 167% in Azerbaijan (to 60.5) and 111% in Albania (59.8). Aleshina and Redmond (2005) use Trussel's (1975) version of the Brass method and use model life tables to convert survey survivorship data for older ages into infant mortality rates for Kazakhstan, Tajikistan, and Azerbaijan. While a wide range of possible IMRs result, they tend to be well above official estimates, especially for Tajikistan and Azerbaijan.

Comparison of official statistics and survey data also generate very different regional patterns. DHS and similar surveys almost always find considerably higher rural than urban infant mortality. For example, in their analysis of a fairly typical survey, Sullivan and Tureeva (2004) report rural IMR 74% greater than urban IMR in Uzbekistan. This pattern is confirmed for India as well (National Neonatology Foundation, 2004: 20). However, because of greater under-reporting, official data commonly find higher urban IMR, at least in transition nations (Becker *et al*., 1998)

A comparison of infant mortality rate estimates from four different sources – UN Statistical Division, UNICEF, WHO aggregate estimates, and the summation of total infant morality by four sub-periods and by specific causes of death from the WHO mortality database – makes it apparent that there is considerable divergence. In particular, the summed values tend to be lower than other estimates, even for countries with very high levels of coverage, though there are cases where the summed values are greater than other estimates. Furthermore, the detailed breakdown is not available for most very low-income countries, while it is generally present for high income countries.

It also can be seen that even for countries with very high rates of vital statistics coverage that huge differences in reported values may occur. In countries such as Thailand, Belize, or Mexico, the large range may reflect weaker reporting at the disaggregated (cause and sub-period of infant death) level. But countries such as Albania, Egypt, and Mongolia have very different aggregated numbers reported by different sources. Somewhat ironically, the level of conformity among IMR estimates is often greatest among some of the poorest (and likely worst enumerated countries), presumably because all sources report imputed values based on population structure and fertility estimates. Thus, for example, the UN and WHO figures are virtually identical for Niger, Myanmar, or Côte d'Ivoire, while they differ substantially for Turkey or South Africa. Paraguay appears to be in a category of its own in terms of having an astonishing level of disagreement.

What is the researcher to do? Economists tend to grab whichever data set is handiest without concern for the possibility that the IMR numbers reported may differ markedly from other reported values. To repeat our earlier point, we are most troubled by the apparent inconsistency in generating specific values, and by systematic biases that are likely to emerge. At present, global data sets use estimates from vital statistics (perhaps with a few, country-specific corrections in many cases) when these are of high quality and with good coverage. Where data are poorer, the estimates may be generated by retrospective surveys (for a discussion of problems in doing so, see Sullivan and Tureeva, 2004). Otherwise, the international bodies fall back on estimated imputed via a modified Brass method from population size and structure, and fertility estimates. However, as all demographers know, these imputation techniques make strong assumptions on population and mortality stability, and on low population movements (Aleshina and Redmond, 2005). These assumptions were reasonable for the Africa of the 1960s that Brass and Trussel had in mind as they developed imputation techniques. They are much less well suited for the more turbulent and mobile world of today.

The primary alternative to date has been to use data from DHS and similar surveys to find patterns for low and middle-income countries. In a detailed presentation on neonatal mortality rates, Hall (2002)

surveys what is known, presents detailed data, and discusses limitations to the surveys. The growing number of regular surveys makes this a valuable exercise. This is particularly the case now that several countries, and most importantly India, generate consistent regional surveys with reasonable frequency (for a detailed study of India, see National Neonatology Forum, 2004). Nonetheless, these advances do not address the need to generate a consistent set of estimates for all countries.

3: ESTIMATION STRATEGY

We approach the under-reporting problem differently, seeking to use reported mortality data rather than making standard Brass-Trussel corrections. Our rationale for doing so is driven in part by a desire to ultimately generate a consistent, comprehensive cross-country panel data set for an extended time series; it also reflects concern that the underlying Brass-Trussel model assumptions are less appropriate today than in the 1960s and 1970s, when the framework was first developed.

The first step is to develop a model of the determinants of infant mortality. A reasonable point of departure is standard Beckerian neoclassical household choice model in which health inputs are chosen subject to resource constraints. This framework leads to the prediction that infant health will improve and mortality risk will decline a household resources improve, as technology and general health knowledge improve, and as public health efforts (themselves dependent upon state resources and hence GDP) increase. In summary, we anticipate that IMR will decline with the level of economic and social development, effort devoted to public health, access to medical care, and quality of individual health practices. In practice, the challenging part of the exercise lies not in modeling infant mortality, but rather in estimating a plausible reduced form equation. Explanatory variables are highly correlated and to some extent causally related, yet data limitations make it impractical to estimate a multi-equation structural model. We instead explore a range of single-equation reduced form specifications in which we recognize that each included variable is likely to pick up a range of effects.

Within this theoretical framework, then, we explore several explanatory variables to identify consistent predictors of infant mortality. Related to economic and social development, we test adult literacy, secondary school enrollment, whether the country is a democracy, whether the country belonged to the former Soviet Union, and the percentage of females having their first baby before age 18. Moreover, we consider income inequality and overall level as indicated by World Bank series on Gini coefficients and gross national income per capita, adjusted for purchasing power. In the realm of health, we investigate private and government health expenditure as a percentage of GDP, the number of physicians per 1,000 persons, and the

maternal mortality rate. Additionally, we examine the percent of the population in urban areas, which reflects the level of development (to the extent that it is missed by GDP) as well as access to medical care. Finally, in an effort to directly capture underreporting, we use the WHO's measure of vital event coverage, which indicates the percent of vital events believed to be included in official counts. We transform this variable to indicate whether a country has coverage known to be at least 85 percent, as higher coverage reduces the opportunity for missing values.

After considering several alternate specifications exploring non-linear functional forms and interaction effects, we focus attention in the empirical analysis on three that provide the best fits and that are appealing from a theoretical perspective. The five empirical specifications considered included (1) GDP per capita in quadratic form, maternal mortality rate in quadratic form, and the former Soviet republic indicator; (2) GDP per capita in quadratic form, maternal mortality rate, percent urban, and the former Soviet Republic indicator; (3) GDP per capita, maternal mortality rate, percent urban, a multiplicative interaction of vital event coverage indicator and percent urban population, the vital event coverage indicator, and the former Soviet Union indicator; (4) GDP per capita in quadratic form, maternal mortality rate, adult literacy rate, and the former Soviet Republic indicator; and (5) GDP per capita in quadratic form, maternal mortality rate, Gini coefficient, and the former Soviet republic indicator. We discarded these last two specifications because of a large number of missing values, as well as consistent lack of significance of adult literacy and the inequality index across age-specific mortality rates. While this suggests varying explanatory variables of infant death by age group, we seek a set of explanatory variables that performs consistently and independent of age for reasons discussed in a later section. Due to space limitations, we report results for model (1) while noting whether our findings hold for the other two model specifications. Complete regression results are available upon request.

Our specification includes both linear and quadratic maternal mortality rate and per capita income terms because there is no reason to assume a linear relationship. Since higher incomes relax health spending constraints, per capita income is a key determinant of a country's infant mortality rate at all age groups (that is, first day, first week, first month…). However, if returns to spending diminish sharply, then marginal effects will decline with income. While public health measures are somewhat problematic due to endogeneity issues, the maternal mortality rate can be seen as representing good public health effort conditional on income levels. It also picks up some of the access and individual practice effects. Our preferred specification also includes a former Soviet republic indicator for two reasons. First, the USSR developed public health infrastructure far more advanced than is typical for middle income countries, and especially for the poorer republics that have not yet regained Soviet-level real income levels. Thus, one would expect lower IMRs in

former Soviet republics. In addition, weight-based definitions of live births (officially abandoned but still in use in many areas) will lead to IMR underreporting in the former Soviet Union. Although public health efforts have deteriorated drastically throughout the former USSR, the system of mandatory expectant mother check-ups and considerable if decayed infrastructure leads us to anticipate that being a former Soviet republic should be negatively associated with infant mortality rates, even correcting for underestimation (which should be higher in former Soviet republics).

That we limit reported regression results to this simple specification does not imply that the other models have less explanatory power. In practice, summary statistics vary little, and the remaining models have equal if not more explanatory power. However, two basic measures of social and economic development, per capita income and maternal mortality, explain a high proportion of infant mortality variation, and are the only variables to be consistently significant.

Considering alternate specifications enables us to examine the sensitivity of our results to model specification. To further investigate the robustness of our findings, we analyze three datasets. Not only does enhance our results, it also contributes to understanding for which countries and age groups underreporting is most severe. The datasets are based on UN IMRs that includes all countries, including those with figures based on sample estimates, UN IMRs for countries also having WHO raw infant death counts by age group, and a set that uses WHO raw infant death counts to calculate a WHO IMR as well as sub-period infant mortality rates. As mentioned, the first dataset does not exclude countries for which the United Nations has made corrections. The second dataset is limited to countries that on average have stronger vital event registration systems, along with higher levels of overall economic development, to have the capacity to report disaggregated rates: Seventy-seven percent of these countries have a vital event coverage statistic known to be greater than 85 percent. Lastly, the WHO IMR dataset pertains to the same set of countries as the second dataset; however, infant death rates have not been corrected.

The next issue concerns estimation. As long as we are dealing with aggregate IMR estimates, simultaneity problems seem minor. However, the nature of errors is that underestimates are almost certainly more likely than overestimates. The latter will occur to the extent that deaths are reported accurately while births are underreported; the former will be common if deaths are underreported relative to births. Underreporting of deaths is universally more common, possibly excluding tiny errors in a few highly developed countries. Therefore, we argue that errors will be one-sided, making standard OLS "average" infant mortality regressions inappropriate, since they assume that errors have zero mean, and in effect result

from random reporting error. Our estimation approach seeks to uncover whether these one-sided errors exist, and if so for which age groups.

This problem was first addressed in production and cost analysis with the aim of identifying firm inefficiency. Production and cost functions are properly regarded as being envelopes, and hence the frontier approach enabled both estimation of the envelope and measurement of the extent of inefficiency of particular firms (for example, Huang, 1984; Kumbhakar and Lovell, 2000 provide a detailed econometric presentation). While the stochastic frontier function technique has become standard in productivity and cost analysis, its application to other questions appears to be limited. Its application in mortality analysis seems natural. If situations as those depicted in Figure 1 prevail, in which the true relationship (the solid line) is obscured by under-reporting in many if not all cases, then an OLS estimate (dashed line) will produce biased coefficients. If the errors are negatively correlated with level of economic development, literacy, urbanization, and recorded maternal mortality – all of which seems plausible – then these coefficient estimates will be biased upward. That is, the true negative relationship will be understated. Furthermore, the predicted IMR and number of infant deaths in poor countries will be systematically understated. In practice, since IMR estimates in very poor counties are based on imputation techniques or small sample survey, and, while subject to considerable error, these estimates are not obviously systematically biased, it may well be that the greatest understatement occurs in middle-income countries with substantial but imperfect vital events registration systems.

More generally, the frontier estimation problem is posed by assuming some fixed number of observations and then forming the relevant log likelihood function, inheriting the usual maximum likelihood properties for the estimated coefficients and variance. The difference is that the error consists of two components.[1] These consist of one error term V_i that is symmetrically distributed i.i.d. as $N(0, \sigma_v^2)$, capturing the effects of random measurement error and random shocks to the observations. The other error component U_i is a one-sided term that is distributed i.i.d. as $N(\theta, \sigma_u^2)$, capturing the effects of non-random measurement error – that is, systematic underreporting. Then the observed mortality rate MR_i can be expressed as a function of non-stochastic determinants $\boldsymbol{X}$ and an error term ε_i as

$$MR_i = \boldsymbol{\beta}'\boldsymbol{X}_i + \varepsilon_i \text{ and } \varepsilon_i = V_i - U_i. \quad \textbf{(1)}$$

[1] This presentation follows Huang (1984) and Morrison (1993).

An expected-maximization (EM) algorithm is then used to estimate the parameter vector $\Theta' = (\boldsymbol{\beta}, \sigma_v^2, \sigma_u^2)$. Letting VMR_i represent the true mortality rate – more conventionally, the latent frontier – we can write:

$$VMR_i = \boldsymbol{\beta}'\boldsymbol{X}_i + V_i . \quad \textbf{(2)}$$

Hence,

$$MR_i = VMR_i - U_i. \quad \textbf{(3)}$$

The algorithm involves an iterative procedure that includes an expectation step that estimates sufficient statistics of VMR given the observed MR. A maximization procedure then estimates a new Θ' using a maximum likelihood procedure on (2) and (3). These new Θ' estimates then generate new sufficient statistics, and the procedure repeats until, if all goes well, the algorithm converges. The algorithm is available as an option in recent versions of STATA, which we use.

The choice to move from ordinary least squares to frontier regression is based on whether residuals are negatively skewed. In equation (1), if the one sided error is greater than zero, the residuals are negatively skewed, suggesting technical inefficiency in the data. Coelli (1995) suggests a simple test for negative skewness: since negative skewness occurs when the third sample moment of the residuals is less than zero, a test of the hypothesis that the third sample moment is greater than or equal to zero is appropriate (Khumbakar and Lovell 2000:73). From this point, researchers also must choose among the half normal, exponential, or truncated frontier models. In accordance with much of the literature, we focus on the relatively simple half-normal form (Khumbakar and Lovell, 2000:90).

Frontier estimation also allows for exploring the determinants of both error components. Estimation should begin under the assumption of heteroskedasticity and then test restrictions for whether the model is homoskedastic (Kumbhakar and Lovell 2000:122). Although guidance is offered in terms of production analysis, it can also be applied to the infant mortality problem. As traditionally conceived, the symmetric error component might be heteroskedastic if the sources of noise vary with the size of producers. Moreover, the technical inefficiency itself component is heteroskedastic if the sources of inefficiency vary with the size of the producers (116). Researchers also must justify why a variable is exogenous and a determinant of the error instead of being placed in the main function. Kumbhakar and Lovell define an exogenous variable as one that influences the structure of the production process beyond the control of involved managers (262). These variables are not inputs to the production process but still exert influence on

producer performance. In other words, we seek to delineate between the environment in which an infant death is recognized as occurring and the process that directly contributes to an infant death.

An obvious determinant of the stochastic error component is the total number of live births. For small sample reasons, the extent of random error in reported infant death varies with total births in a particular country. To capture inefficiency in the reporting process, we focus on the aforementioned dummy for whether a country has vital event coverage known to be greater than 85 percent. Extent of vital event coverage may be correlated with variables that determine recorded infant death but is itself exogenous to the infant death equation. The percent of the population residing in urban areas also relates to the strength and fluidity of the vital event registration system, at the same time that it helps determine quality of and access to medical care. While the former interpretation points to including it as a determinant of the one sided error term, it could also be placed in the main equation as a proxy for medical care quality. In our alternative models, we consider the variable in both roles, as well as interacting it with the vital event coverage indicator. Once our frontier regressions are estimated, "true" frontier value estimates can be calculated for each country as a function of its characteristics, and the level of error (corresponding to the estimate of firm inefficiency) can be determined.

While the frontier estimation process is useful in that it allows us to separately estimate determinants of inefficiency, a disadvantage is that these estimates are sensitive to functional form. Demographic and economic theory offer good insights into which variables should affect mortality rates, but there is little *a priori* restriction on functional form. The results reported below use log-log forms for the ordinary least squares regressions to allow for comparison with the frontier model. Quadratic terms are included in cases where we believe model fit would be improved in anticipation that elasticities may decline with increased inputs. The evidence for technical inefficiency (underreporting) presented below is fairly convincing; however, we acknowledge that statistically significant estimates of the one-sided error do not obtain in every specification. On the other hand, since we do not know the appropriate specification, a reasonable approach is to try alternates and hunt for the best fits, and then examine whether underreporting exists in those cases. We should note as well that poor specification of functional form can be expected to increase the V_i term relative both to the one-sided error and the non-stochastic component, so that we are more likely to miss than overestimate the extent of underreporting.

A complication emerges when we turn to disaggregated components of IMR; namely, birth day mortality, day 1-6 mortality (week 1 less day 0), day 7 – 27 mortality (weeks 2-4 mortality), and day 28 – 364 or post-neonatal mortality (*PNNMR*). Underreporting incidence declines with infant age, suggesting that

more accurate measures of aggregate IMR can be obtained if we divide the overall rate into its components. Because of the measurement error problems in the first three neonatal mortality (*NNMR*) components associated with birth heaping and increased random error earlier in life, for simplicity we focus on separate determinants of NNMR and PNNMR. However, we also take care to note how our analysis fares for these disaggregated components. Briefly, results for the disaggregated components do not detract from our hypothesis.[2]

The WHO database enables us to distinguish NNMR and PNNMR rates. Once again, we want to explore how these measures vary with socio-economic determinants of infant mortality. However, we cannot estimate the two equations independently, unless the explanatory variables are identical, since the error terms will be correlated. The standard procedure for dealing with this problem is to use seemingly unrelated regressions (SUR). Unfortunately, this technique is not at present integrated with frontier techniques, forcing us to choose between them. Our approach in this paper is to run independent frontier regressions using identical explanatory variables, in which case SUR collapses to ordinary regression. A future step will investigate simultaneity, with NNMR and shorter period mortality augmented with the estimated error from the first frontier regressions. These regressions then will be jointly estimated with PNNMR in an SUR framework. These sub-period regressions can then be used to calculate estimated components. The greater of estimated and reported values is taken at all times; one option is to scale up the subperiod estimates to yield the frontier IMR estimate. The advantage of the simultaneous equations approach is that we can permit PNNMR to depend on general health variables, while NNMR alone will depend on maternal mortality.

An alternative correction that has the potential to remove biases is to generate a panel data set, and then use a fixed effects model to capture country-specific biases. This approach is intended to be the topic of a companion paper, but we note here that it is not without problems. In particular, data quality systematically varies over time in many countries: it improves with overall economic development, and deteriorates with crises. This complicates time series analysis, and for simplicity we stick with a simple cross-country analysis here – while noting that refined estimates will need to introduce data from multiple periods.

[2] The authors will provide results for further disaggregated IMRs on request.

4: SOCIO-ECONOMIC DETERMINANTS OF INFANT MORTALITY

Before presenting re-estimated infant mortality rates along with OLS and frontier regression results, Figure 2 illustrates the negative skewness in OLS residuals for specification 1 (with squared GDP per capita and maternal mortality rate terms, and the FSU dummy) when estimating *NNMR*. As negative skewness is visually obvious, we clearly have good reason to proceed to frontier regression. Beyond this impressionistic evidence, we consider whether negative skewness exists based on two alternate examinations. First, a detailed summary of the residuals shows that the mean is to the left of the median, which is by definition negative skewness. Moreover, and of equal importance, STATA's frontier command stores a test for negative skewness[3] when running homoskedastic models. Thus, for the corresponding homoskedastic model without explanatory variables in the error functions, we reject the null hypothesis that the OLS residuals are not skewed. (p-value=.02). Findings of negative skewness are robust to model specification (at the 10 percent significance level).

To be clear that negative skewness indicates a special circumstance, we examine negative skewness for models using the UN IMR, the WHO calculated IMR, and PNNMR as dependent variables. For model 1 and our alternative specifications, using the UN IMR does not result in negative skewness across methods of examination (histogram of residuals, residual summary, and STATA's test for negative skewness in the corresponding homoskedastic models). Moreover, as anticipated, we do not find negative skewness in the PNNMR model for all three specifications. That is, skewness and hence undercounting are associated with neonatal mortality: once a child has survived for a month, both its birth and death are likely to be recorded.

In contrast, the long left-tailed residual distribution appears in the model using WHO calculated IMRs for two out of the three specifications. Since the WHO rate is the sum of the NNMR and PNNMR, it is not surprising that negative skewness is somewhat weaker. Even though the mean is to the left of the median in these 2 specifications, STATA's test for negative skewness is not in our favor. The reason for this seeming contradiction will become apparent in discussing the percent of technical inefficiency comprising the composite error in a later section. Frontier regression for each of our dependent variables confirms findings of negative skewness if we have a significant one-sided error term. We note that in some cases for mortality

[3]This test, developed by Coelli (1995), is based on the third moment of OLS residuals.

rates more finely disaggregated by time period, the mean and median are extremely close to zero and frontier regression will still result in technical inefficiency.[4]

The core regressions results we report in Tables 1 and 2 use the UN IMR, WHO uncorrected IMR, NNMR, and PNNMR as dependent variables. In Table 1, the first two regressions present results using UN data for all countries for which estimates are available. Our cautious approach in including quadratic terms is supported by joint significance tests, which are all highly significant. First, we note the strength of the maternal mortality rate and GDP per capita in explaining infant mortality rates: with these two variables, we explain 91 percent of variability in our regression model. While our alternative specifications have comparable R-squared statistics, we do not gain power by including independent variables other than per capita income and maternal mortality rate. In aiming to present the most parsimonious model, we focus on the first specification. We observe that in the global dataset there is no indication of systematic underreporting. The σ_u^2 terms is not significant, and in the corresponding homoskedastic model (results not shown) the frontier regression collapses to OLS. Alternative specifications concur with this result. Examining the coefficients, the overall infant mortality rate declines with ln GDP at an accelerating rate over the entire observed range (its maximum occurs at about USD 18). Not surprisingly, IMR rises with maternal mortality.

As noted, the former Soviet republic indicator may reflect both underreporting and the presence of an established welfare state with significant maternal and infant health care, even in low income settings. In frontier modeling, these two effects can be examined separately. In an OLS setting, they must be combined. As it turns out, when the entire United Nations IMR dataset is considered, the former Soviet Union dummy is not significant in either case (regressions 1 and 2). Quite simply, we do not find systematic inefficiency in the full data set, since the UN imputations do not make systematic errors, while our simple model with GDP per capita and maternal mortality rate captures an overwhelming share of infant mortality variation. We return to this point in discussing Table 2.

Table 2 reports the same model but for the WHO IMR, NNMR and PNNMR. While this smaller sample of 66 countries is anything but random, the non-randomness is not overly problematic for our purposes. A dataset with countries having stronger vital event registration systems and high economic development would lead us to underestimate rather than overestimate the reporting bias. Thus, we view this

[4] Inspection of OLS residuals for further disaggregated component reveals negative skewness in the following cases: Week 1, 3/3 specifications in all examinations; Week 2 to 4, 1/3, no negative skewness found by the Coelli test; Day 1 to 6, 2/3, results show negative skewness at 13 percent significance level; Less than day 1, 3/3, in all examinations.

as helping us conservatively estimate corrected rates, as we would rather fall slightly below the true rates than erroneously predict reporting error. The excluded group also includes many formerly socialist countries whose data practices are not fully reconciled with WHO. As these data are often available (see http://demoscope.ru/weekly/pril.php or http://www.mednet.ru/statistics/), in future versions we hope to extend this analysis to include merged data.

Table 2 confirms our hypotheses. Underreporting takes place early in infant life, as evidenced by the technical inefficiency term in the NNMR model. This result is independent of model specification. Furthermore, it is more difficult to explain variability in this early period. In contrast, the PNNMR model does not consistently detect inefficiency. In our alternative model with squared per capita income, maternal mortality, percent urban and the former Soviet Union dummy, there is mild evidence of technical inefficiency as the one-sided error term is significant at the 10 percent level; however, the corresponding homoskedastic model finds zero percent technical inefficiency. While we do not find inefficiency in PNNMR,[5] the inefficiency in the first month of life is strong enough to appear in the overall, uncorrected, WHO IMR. This result holds for model 2, however, in model 3 the percent urban and coverage variables are not jointly significant. Furthermore, in the corresponding homoskedastic models, we find from 36 – 51 percent of the composite error due to technical inefficiency despite the fact that the inefficiency terms (constants) are not significant.

Table 2 also reveals that the former Soviet Union dummy is important in explaining neonatal mortality but not post-neonatal mortality in the WHO dataset. This impact is independent of one-sided underreporting error; as the impact is negative, it implies that the residual effects of Soviet health care systems continue to have a salutary mortality risk effect.

Table 1 also reports results using UN IMRs for those countries having disaggregated data in the WHO mortality database. Interestingly, there does appear to be undercounting as the σ_u^2 term is significant, a result that is robust to model specification. Estimates of technical inefficiency in the corresponding homoskedastic models reveal an interesting result: although in the UN model for the global dataset the percent of the composite error due to technical inefficiency is invariably zero, in this limited dataset mostly comprised of better off countries for which few, if any, imputations are made by the UN, the percentage of the error due to inefficiency is as high as 41 percent. An obvious conclusion to draw is that the UN imputations made for countries with poor databases do not suffer from systematic underestimation.

[5] As expected, this model collapses to OLS when the homoskedastic version is estimated.

The next step is to ask whether one can gain additional information by further disaggregating neonatal mortality. In looking at our main and alternative specifications for week 1, weeks 2 to 4, days 1 to 6 and less than day 1 (results not shown), we see that in all cases there is evidence of underreporting., as evidenced by significant one sided error terms. One age group does not stand out relative to the rest. However, when we turn to the corresponding homoskedastic models to obtain STATA's estimates of technical inefficiency for each model, a somewhat different story emerges. Table 3 shows that undercounting is strongly driven by birth day mortality undercounting. In fact, undercounting is strong enough at this stage to dominate in the week 1 IMR, as undercounting for days 2 to 6 is comparatively much less. Moreover, considering that inefficiency manifests in all first month models (homoskedastic or heteroskedastic), the strength of undercounting at day 1 is persistent.

Table 3 also addresses the seeming contradiction between results for heteroskedastic and homoskedastic models, which first manifested in discussing STATA's test for negative skewness in the homoskedastic WHO model. In addition to altering findings of negative skewness, whether we estimate a homoskedastic or heteroskedastic model also affects the one-sided error. This discrepancy in results appears in modeling PNNMR, the WHO calculated rate, and in sub-periods of first month mortality. Often, inefficiency will appear in heteroskedastic models-- for example, in the cases of weeks 2 to 4 and days 2 to 6 mortality – however, in the homoskedastic version undercounting seems to disappear. Moreover, while running alternate specifications of homoskedastic or heteroskedastic models, results can vary widely. What is the cause? The presence of inefficiency, while to some degree contingent on model specification, also depends on the strength of undercounting in our dependent variable. For example, in the case of PNNMR, there is little, if any, undercounting. Furthermore, there are presumably no periods within post-neonatal mortality period where one-sided inefficiency exists; thus we do not find conflicting results. In contrast, when considering the WHO calculated IMR, which is the sum of the NNMR (where undercounting is unambiguous and predominantly driven by day 1 mortality) and PNNMR, mixed results emerge. In the homoskedastic model, we do not find negative skewness. However, technical inefficiency comprises a notable percent of the composite error. The obvious conclusion is that findings are less straightforward when dealing with a summed variable in which not all of the disaggregated rates have one-sided error. In effect, we are testing whether the one-sided error manifesting in a sub-period rate is strong enough to dominate in the summed variable.

In addition to telling us whether undercounting exists, as indicated by a significant one-sided error term and estimated percent inefficiency, our analysis explores the causes of undercounting. For NNMR,

whether a country has known coverage greater than 85 percent system significantly predicts the one-sided error in models 1 and 2. Moreover, in model 3, the coverage variable, percent urban population, and a coverage – percent urban interaction term are jointly significant (p=0.0001). Models for week 1 and days 1 to 6 mortality rates present similar results although with varying degrees of significance. In day of birth mortality, explanatory variables of the one-sided error are only significant in models 1 and 2. The divergent results in model 3 may be due to increased random error at this very early age. Interestingly, for weeks 2 to 4 mortality, in models 1 and 2, only the constant terms within the asymmetric error function are significant. However, in model 3, known coverage, percent urban population and their interaction are jointly significant at the five percent significance level.

It is necessary to ask if undercounting during the neonatal period is made up at a later age; in other words, does birth heaping, which leads to higher post-neonatal mortality, eventually eliminate any observed undercounting during the early stages of life? We tested this by using PNNMR as a predictor of the one-sided error term (results not shown) in models for NNMR and its disaggregated components. Furthermore, we tested whether undercounting occurring up to the first week of life is accounted for during weeks 2 to 4 (results also not shown). For each age category up to first month mortality and using each of the three models, heaping does not appear to be responsible for underestimated rates, as there is only one instance where PNNMR has mild significance ($p < 0.10$). Do elevated weeks 2 to 4 mortality rates compensate for undercounting in week 1 mortality? For day 1, days 1 to 6, and week 1 mortality rates, only model 1 shows a slightly significant weeks 2 to 4 mortality term; in the day 1 model, it enters at the 5 percent significant level, while in the latter two, at the 10 percent level. Thus, in light of overwhelming evidence against the notion of birth heaping, we cannot attribute large undercounting at early ages to increased rates at a later date.

5: CORRECTED INFANT MORTALITY ESTIMATES

Do these various corrections in fact matter? If so, where are they most important? A definitive answer must await many more specifications and quite likely the use of a panel data set with merged transition economy data. However, as a start, initial results for the 66 countries for which the WHO provides disaggregated rates are telling. In this paper, we purposely chose not to present re-estimated rates for all countries, as without having disaggregated rates for a truly global dataset, we are faced with the methodological challenge in making out of sample predictions. Until we overcome this problem, most likely by expanding our sample to a multi-year panel, we restrict ourselves to the 66 country sample. While, as mentioned, countries in this latter sample tend to be better-off with more advanced vital event management systems, it is pedagogically useful to divide the sample into two subsets. Tables 4 and 5 present corrected

infant mortality rates for the lesser-developed and industrialized subgroups, respectively. Such a split places results in context, as when considering our corrections as a percentage of UN reported rates, the proportions can be misleading. When a rich country begins with a very low infant mortality, even a small correction can be proportionally large. For example, our correction increases Switzerland's UN reported rate from 4 to approximately 6, a 50 percent rise. In contrast, poorer countries with higher UN reported rates may have larger corrections in absolute terms, but proportionately smaller increases. Why does our model predict any correction at all for wealthy countries like Switzerland? While there is a possibility that even rich countries may have slightly underestimated rates, our model is most likely yielding a correction in order to accommodate countries at the lower end of the economic development. That is, given the simple functional form that we have imposed, larger errors for poorer, high mortality countries also affect the predicted values and hence error estimates for wealthier countries. Fortunately, in absolute terms, corrections for this wealthy group are generally very small.

Tables 4 and 5 display corrected infant mortality rates based on models 1 and 2 (we note that model 3 fails to converge). After regressing the ratio of neonatal mortality and post-neonatal mortality on the UN reported rate, we solve for the newly estimated NNMRs and PNNMRs while constraining their sum to be at least equal the UN rate. Without being able to discount the possibility of one-sided error in this subsample, we correct these rates using frontier regression as previously described. Thus, now we have three values for NNMR and PNNMR: the WHO calculated rate, the prediction based on the initial OLS regression, and the frontier corrected rate. We calculate a corrected infant mortality rate by summing over the maximum values for each country.

The corrected rates generated by models 1 and 2 are equivalent in 14 cases. Of the remaining 52 countries, model 2 produces a larger estimate for 29 countries. Given that model 1 is more modest than model 2, we focus on model 1 while noting that the same analysis can be applied to model 2 estimates. Table 4 contains 42 of the 66 countries in this sample, of which approximately 60 percent have vital event coverage known to be greater than 85 percent. In other words, the picture of undercounting painted by these data will naturally underestimate the true extent globally. We begin by looking at countries where the WHO uncorrected rate is within one death per thousand of the UN reported rate, which includes some of the formerly socialist, poorer European countries (*e.g.*, Bulgaria, Hungary, Slovakia, and Latvia), as well as middle-income South American countries (*e.g.*, Argentina, Costa Rica, Chile and Uruguay). Despite the fact that the UN apparently does not make any imputations for these countries, our estimates suggest that infant mortality rates are underestimated by as much as 8 deaths per 1,000 live births. This is more surprising, albeit not unrealistic, for countries such as Argentina and less eye-opening when we turn to Mauritius, where frontier

estimation increases the UN reported rate of 15 to about 23. Moreover, a subset of these countries, including Argentina, Mauritius, Slovakia, and Uruguay, have UN infant mortality rates less than the WHO un-corrected values. Given the diverse range of countries excluded from UN corrections, it is interesting to consider the criteria applied to determine which countries require imputative measures.

Overall, the UN rates are significantly higher than the WHO values for the majority of countries in Table 4. Despite this, our approach generates even higher estimated rates. In Nicaragua, for example, model 1 estimated the infant mortality rate to be higher by about 23 deaths per 1,000 live births. This large discrepancy is repeated for an array of countries including Kyrgyzstan, Haiti, Peru, Philippines and Moldova. In the cases of Albania, Czech Republic, Estonia, Kuwait, Mexico, Poland, Lithuania, and Slovenia, frontier estimation increased UN reported rates by less than or equal to 2 deaths per 1,000 live births. Furthermore, frontier generated rates equal to the UN value for South Africa, Mongolia, and Bahrain.

When examining the size of correction for these countries in the context of their respective levels of vital event coverage, we do not find extreme corrections for countries with known coverage of at least 85 percent. Admittedly, there is a degree of subjectivity as for several of these countries, including Argentina, Bulgaria, and Mauritius, frontier estimation increases infant mortality rates by at least five. However, in considering countries that have corrections of at least 10, we do not find any with vital event coverage known to be greater than 85 percent. We do, of course, face instances where unexpected results occur: for example, according to the WHO, South Africa has vital event coverage statistic of less than 60 percent. However, our results suggestion no correction above and beyond that of the UN is necessary. This may simply point to the unbiasedness of UN imputations.

Although Table 5 lists industrialized countries in which we are relatively less interested, there is still information to be gained. Examining these forecast values helps assess the validity of our methodology, and we also note instances where our corrections are likely inaccurate. For 16 out of the 24 countries, our frontier correction forecast values are approximately within one death per thousand love births of the UN corrected rate. In many of the instances where frontier produces a corrected rate greater than two deaths per thousand live births than that of the UN, we observe that the UN reports an IMR below the WHO crude rates. Of special note is the peculiar case of Belgium, which has a UN infant death rate of 4, a WHO calculated rate of 8.893, and a frontier estimated rate of about 14. Without having a more complete panel data set and more countries in our sample, our estimations are bound to be subject to error, especially at the lower mortality end.

6: NEXT STEPS

Much work remains before it will be possible to generate a consistent panel data set with infant mortality estimates for nearly all countries on an annual basis. The first steps are obvious: it is necessary to examine alternative specifications and expand the number of observations to a multi-year panel in exploring determinants of sub-periods of infant mortality. The approach described above continues to be appropriate, though some complications are added by the time series.

There are also two related investigations that need to be conducted. One centers on the possibility that certain types of mortality are especially undercounted. We are not overly optimistic on this point, but the possibility should not be overlooked. WHO data also contain separate estimates of urban and rural mortality, and these clearly should be estimated separately, since rural infant mortality is likely to suffer from far greater undercounting than its urban counterpart. The second supplemental study is to econometrically investigate the determinants of idiosyncratic error. The approach is straightforward, but we have relatively little to guide us in terms of functional form and previous study.

These steps are conceptually simple. The most complex step will be to simultaneously estimate a vector of sub-period (and likely sub-region) mortality rates in a panel data, frontier analysis setting. To our knowledge, the combination of the three tasks has not been undertaken, but there is no obvious reason that it cannot be done.

Once the estimates are in hand, the remaining work is straightforward. For each country, and each year, the regressions will yield an estimated infant mortality rate, as well as day of birth, days 2-7, weeks 2-4, aggregate neonatal, and post-neonatal mortality rates. The sub-period estimates can then be summed to determine an alternate infant mortality estimate. The direct and indirect *IMR* estimates can then be compared to vital statistics data for the countries with high coverage rates, and possibly be replaced with directly counted numbers in a few cases. More commonly, the estimated equations will be used to "backcast" prior infant mortality rates, using previous estimates of GDP, urbanization, and maternal mortality – and, if we are successful, factors that are found to determine undercounting.

This work has not yet been completed, though Tables 4 and 5 hint at likely findings. Undercounts are greatest in low and middle-income countries with substantial but inaccurate vital statistics reporting. These countries' infant mortality rates come mainly from counts rather than imputations, but undercounting is a major problem. Most of the world's poorest countries do not have comprehensive vital statistics, and so

estimated infant mortality rates come from small to moderate surveys, or from imputational procedures. These approaches do not appear from our regressions to contain a systematic, or idiosyncratic, bias. However, these estimates are not constructed for the purpose of creating a consistent time series: rather, they tend to offer best guesses. The estimates generated from frontier panel regressions will provide the internal consistency needed. With luck, it will contribute to a better understanding of the actual picture of mortality at very young ages throughout the world.

REFERENCES

Aleshina, Nadezhda and Gerry Redmond, 2005 (March), "How high is infant and child mortality in Central and Eastern Europe and the CIS?" *Population Studies* **59**(1): 39-54.

Anderson, Barbara, Doo-Sub Kim, and John H Romani, 1997 (June), "Health Personnel, Son Preference, and Infant Mortality in China." University of Michigan PSC Research Report No. 97-401, http://www.psc.isr.umich.edu/pubs/abs.html?ID=1124.

Anderson, Barbara, and John H Romani, 1997 (November), "Socio-economic Characteristics and Excess Female Infant Mortality in Jilin Province, China." PSC Research Report No. 97-409, Http://www.psc.isr.umich.edu/pubs/abs.html?ID=1151.

Anderson, Barbara, John H Romani, Heston E Phillips, and Johan van Zyl, 2001 (May), "Safe Water, Access to Health Care, and Other Factors Affecting Infant and Child Survival among the African and Coloured Populations of South Africa: An Analysis Based on the 1994 October Household Survey." University of Michigan: PSC Research Report No. 01-472, http://www.psc.isr.umich.edu/pubs/abs.html?ID=1291.

Anderson, Barbara A., and Brian D. Silver, 1986 (December), "Infant Mortality in the Soviet Union: Regional Differences and Measurement Issues" *Population and Development Review* **12**(4):705-738.

Anderson, Barbara A., and Brian D. Silver, 1988 (June), "The Effects of the Registration System on the Seasonality of Births: The Case of the Soviet Union," *Population Studies* **42**(2): 303-320.

Anderson, Barbara, and Brian D Silver, 1997, "Issues of Data Quality in Assessing Mortality Trends and Levels in the New Independent States," in José Luis Bobadilla, Christine A Costello, and Faith Mitchell, Eds., **Premature Death in the New Independent States**. Washington, DC: National Academy Press. 1997.

Bang, Abhay, M.H. Reddy, and M.D. Deshmukh, 2002 (December 7), "Child mortality in Maharashtra," *Economic & Political Weekly* **37**(49): 4947-4965.

Becker, Charles M., Margarita Ibragimova, Grace Holmes and Damira Bibosunova), 1998 (November), "Maternal care *vs.* economic wealth and the health of newborns: Bishkek, Kyrgyz Republic and Kansas City, USA." *World Development*, **26**(11): 2057-2072.

Coelli, Tim, 1995, "Estimators and Hypothesis Tests for a Stochastic Frontier Function: A Monte Carlo Analysis." *Journal of Productivity Analysis* **6**(3): 247-268.

Hall, Sarah, 2002, "Neonatal mortality in developing countries: what can we learn from the DHS data?" Southampton, UK: University of Southampton, Statistical Sciences Research Institute, *Applications & Policy Working Paper A05/02*.

Hill, Kenneth, and Yoonjoung Choi, 2006 (May 23), "Neonatal mortality in the developing world," *Demographic Research*, **14**(18): 429-452.

Huang, Cliff J., 1984, "Estimation of stochastic frontier production functions and technical inefficiency via the EM algorithm," *Southern Economic Journal* **50**(3): 847-856.

Kingkade, W. and E. Arriaga, 1997, "Mortality in the new independent states: patterns and impacts," in José Luis Bobadilla, Christine A Costello, and Faith Mitchell, Eds., **Premature Death in the New Independent States**. Washington, DC: National Academy Press. 1997.

Kingkade, W. and C. Sawyer, 2001 (August), "Infant mortality in Eastern Europe and the former Soviet Union before and after the breakup," Salvador de Bahia, Brazil: paper prepared for the 2001 IUSSP meetings.

Kramer, Michael S., Shiliang Liu, Zhongcheng Luo, Hongbo Yuan, Robert Platt, and K.S. Joseph, 2002 (September 15), "Analysis of perinatal mortality and its components: time for a change?" *American Journal of Epidemiology* **156**(6): 493-497.

Kumbhakar, S.C. and C.A.K. Lovell, 2000. **Stochastic Frontier Analysis**. Cambridge, UK: Cambridge University Press.

Morrison, Andrew R., 1993, "Unproductive migration reconsidered in a stochastic frontier production function framework for analyzing internal migration," *Oxford Economic Papers* **45**(3): 501-518.

National Neonatology Forum, 2004 (November), **State of India's Newborns**. New Delhi: National Neonatology Forum.

Platt, Robert W., K.S. Joseph, Cande V. Ananth, Justin Grondines, Michael Abrahamowicz, and Michael S. Kramer, 2004, "A proportional hazards model with time-dependent covariates and time-varying effects for analysis of fetal and infant death," *American Journal of Epidemiology* **160**(3): 199-206.

Republic of Botswana and United Nations, 2004, **Botswana Millennium Development Goals: Status Report 2004**. Gaborone: Ministry of Finance and Development Planning.
http://unctdatabase.undg.org/index.cfm?module=CoordinationProfile&page=Country&CountryID=BOT

Richardus, J.H., 2003, "Suboptimal care and perinatal mortality in ten European regions: methodology and evaluation of an internal audit," *Journal of Maternal-Fetal & Neonatal Medicine* **14**: 267-276.

Romani, John, and Barbara Anderson, 2002 (August), "Development, Health and the Environment: Factors Influencing Infant and Child Survival in South Africa." University of Michigan PSC Research Report No. 02-525, http://www.psc.isr.umich.edu/pubs/abs.html?ID=1355.

Sullivan, J.M. and N.K. Tureeva, 2004 (April), "Infant and Child Mortality," Ch. 8 in Ministry of Health of the Republic of Uzbekistan *et al.*, **Uzbekistan Health Examination Survey 2002**. Tashkent: Ministry of Health, pp. 77-85. http://www.measuredhs.com/pubs/pdf/FR143/00FrontMatter.pdf

Trussel, T. James, 1975, "A re-estimation of the multiplying factors for the Brass technique for determining childhood survivorship rates," *Population Studies* **29**(1): 97-108.

Wegman, Myron E., 1996 (December), "Infant mortality: some international comparisons," *Pediatrics* **98**: 1020-1027.

Wuhib, Tadesse, Brian J. McCarthy, Terence L. Chorba, Tatiana A. Sinitsina, Ivan V. Ivasiv, and Scott J.N. McNabb, 2003, "Underestimation of infant mortality rates in one republic of the former Soviet Union," *Pediatrics* **111**(5): e596-e600.

Table 1
Infant Mortality Rate (IMR) Parameter Estimates from Frontier Function and OLS Regressions, UN data

Regression	*(1)*	*(2)*	*(3)*	*(4)*
Specification	Log-log	Log-log	Log-log	Log-log
Data source	UN	UN	UN	UN
Regression type	Frontier, normal/half normal	OLS	Frontier, normal/half normal	OLS
Dependent variable	IMR	IMR	IMR	IMR
Regressors:				
Constant	3.151[a]	3.12[a]	6.940[a]	5.512[a]
GDP per capita	0.175	0.140	-0.742 [a]	-0.490 [c]
(GDP per capita)^2	-0.030[b]	-0.029[a]	0.022	0.009
Maternal Mortality Rate	0.010	0.063	-0.019	0.125
(Maternal Mortality Rate)^2	0.030[b]	0.027[b]	0.043[c]	0.018
FSU	0.110	0.069	-0.198[b]	-0.188 [b]
ln σ_u^2				
Known Coverage > 85 %	29.973		-3.056	
Constant	-33.280		-2.138 [a]	
ln σ_v^2				
Total Live Births	-0.049		0.053	
Constant	-1.865 [c]		-4.105[a]	
Pr > $\bar{\chi}^2$ (*F*)	0.000	(533.49)	0.000	(137.39)
Log Likelihood (R^2)	-32.554	(.914)	12.024	(.926)
N	157	157	66	66

Notes: Standard errors in parentheses N = Number of observations[a] Significant at the .01 level [b] Significant at the .05 level [c] Significant at the .10 level[d] Significant at the .15 level

Table 2
Infant Mortality Rate (IMR) Parameter Estimates from Frontier Function and OLS Regressions

Regression	*(5)*	*(6)*	*(7)*	*(8)*	*(9)*	*(10)*
Specification	Log-log	Log-log	Log-log	Log-log	Log-log	Log-log
Data source	WHO	WHO	WHO	WHO	WHO	WHO
Regression type	Frontier, normal/half normal	OLS	Frontier, normal/half normal	OLS	Frontier, normal/half normal OLS	OLS
Dependent variable	Neonatal Mortality	Neonatal Mortality	Post Neonatal Mortality	Post Neonatal Mortality	IMR	IMR
Regressors:						
Constant	0.351	0.176	8.532 [a]	3.371	4.617[b]	2.435
GDP per capita	0.349	0.427	-1.422 [a]	-0.341	-0.367	0.060
(GDP per capita)^2	-0.028	-.034[c]	-0.065 [b]	0.005	0.006	-0.017
Maternal Mortality Rate	0.431[a]	0.273	-0.022	0.272 [c]	0.115	0.276
(Maternal Mortality Rate)^2	-.041[b]	-.026	0.046 [c]	-0.014	0.013	-0.020
FSU	-0.321[b]	-.232[b]	-0.159	-0.183	-0.266[b]	-0.243[b]
ln σ_u^2						
Known Coverage > 85 %	-0.922[b]		-37.107		-36.543	
Constant	-1.086 [a]		-0.478		-1.368[b]	
ln σ_v^2						
Total Live Births	-2.228[b]		-.359[c]		-0.030	

Constant	18.860 [b]		1.044		-2.602	
Pr > χ^2 (*F*)	0.000	(15.91)	0.000	(38.32)	0.000	(31.34)
Log Likelihood (R^2)	-0.984	(.546)	-4.063	(.761)	-3.937	(.727)
N	66	66	66	66	66	66

Notes: Standard errors in parentheses N = Number of observations [a] *Significant at the .01 level*
[b] *Significant at the .05 level* [c] *Significant at the .10 level*

Table 3: Summary of underreporting within first month of life for homoskedastic frontier models

Age Specific Infant Mortality Rate	Percent Inefficiency	Negative Skewness by STATA test (Yes if p < .10)
Week 1		
Model 1	89	Yes (p=.01)
Model 2	100	Yes (p=.02)
Model 3	88	Yes (p=.00)
Weeks 2-4		
Model 1	75	No (p=.45)
Model 2	0	No (p=.57)
Model 3	0	No (p=.79)
Days 2 to 6		
Model 1	58	No (p=.25)
Model 2	72	No (p=.13)
Model 3	62	No (p=.13)
Less than day 1		
Model 1	93	Yes (p=.00)
Model 2	89	Yes (p=.01)
Model 3	100	Yes (p=.00)

Table 4: Corrected Infant Mortality Rates for Less-Developed Subset of 66 Countries

Country	WHO	UN	IMR, Model 1	IMR, Model 2	Frontier as % UN, Model 1	Frontier as % UN, Model 2
Albania	12.681	25.000	26.046	29.598	104.183	118.391
Argentina	15.989	15.000	22.965	22.337	153.100	148.915
Bahrain	6.840	14.000	14.000	14.501	100.000	103.580
Belize	19.533	31.000	34.580	34.580	111.549	111.549
Brazil	16.931	27.000	34.613	30.642	128.196	113.488
Bulgaria	12.732	13.000	19.934	21.768	153.340	167.443
Chile	8.363	8.000	13.555	14.938	169.441	186.728
Colombia	14.419	26.000	33.032	32.316	127.044	124.294
Costa Rica	10.771	11.000	14.535	16.671	132.135	151.558
Croatia	7.060	7.000	8.940	9.408	127.708	134.400
Czech Republic	3.770	6.000	7.391	7.956	123.175	132.595
Dominican Republic	10.494	35.000	43.529	42.601	124.368	121.718
Ecuador	15.018	25.000	31.887	31.642	127.550	126.570
El Salvador	9.933	26.000	32.229	31.576	123.959	121.446
Estonia	5.502	10.000	11.470	11.419	114.698	114.186
Guatemala	30.201	39.000	43.048	43.456	110.380	111.425
Guyana	26.813	49.000	55.650	54.678	113.572	111.588
Haiti	17.883	62.000	81.680	73.728	131.742	118.916
Hungary	7.634	8.000	9.963	9.963	124.533	124.533
Jamaica	7.206	15.000	25.082	27.164	167.216	181.096
Kuwait	8.497	10.000	11.384	11.384	113.840	113.840
Kyrgyzstan	19.593	55.000	69.025	68.038	125.500	123.705
Latvia	9.517	10.000	12.673	13.279	126.725	132.786
Lithuania	7.744	9.000	9.469	9.469	105.209	105.209
Mauritius	15.386	15.000	22.533	23.690	150.217	157.933
Mexico	14.579	21.000	23.376	23.376	111.315	111.315
Mongolia	36.922	58.000	58.000	58.000	100.000	100.000
Nicaragua	11.198	30.000	53.586	50.553	178.620	168.510
Panama	14.582	21.000	25.920	26.107	123.427	124.320
Paraguay	12.851	37.000	46.800	45.927	126.485	124.127
Peru	12.115	33.000	43.201	39.220	130.911	118.847
Philippines	11.880	28.000	45.207	42.708	161.454	152.530
Poland	7.750	9.000	10.686	10.686	118.731	118.731

Moldova	12.554	26.000	37.097	38.490	142.682	148.040
Romania	16.880	18.000	24.639	27.466	136.884	152.591
Slovakia	8.789	8.000	9.814	10.488	122.671	131.105
Slovenia	3.726	6.000	7.986	9.613	133.104	160.215
South Africa	20.159	43.000	43.195	43.000	100.453	100.000
Suriname	12.038	26.000	30.943	31.205	119.011	120.020
Thailand	8.095	20.000	22.621	27.473	113.106	137.367
Uruguay	13.564	13.000	16.370	14.985	125.922	115.268
Venezuela	15.684	18.000	22.768	22.316	126.490	123.976

Note: Results are limited to countries for which the WHO provides disaggregated rates to avoid making out of sample predictions. For model 1, explanatory variables include a constant, GDP, GDP^2, maternal mortality, $(\text{maternal mortality})^2$ and FSU. For model 1, explanatory variables include a constant, GDP, GDP^2, maternal mortality, percent urban, and FSU. In both models, total live birth is used to explain the symmetric error and known coverage >= 85 for the one-sided term. Model 3 does not converge.

Table5
Corrected Infant Mortality Rates for Subset of Industrialized Countries

Country	WHO	UN	IMR, Model 1	IMR, Model 2	Frontier as % UN, Model 1	Frontier as % UN, Model 2
Australia	4.808	5.000	5.359	5.359	107.189	107.189
Austria	3.901	5.000	5.669	5.122	113.383	102.432
Belgium	8.893	4.000	14.078	14.078	351.958	351.958
Canada	5.390	5.000	5.395	5.390	107.901	107.805
Denmark	5.737	5.000	6.091	5.737	121.823	114.741
Finland	2.859	4.000	5.368	4.165	134.193	104.126
France	4.297	5.000	5.429	8.268	108.587	165.365
Germany	4.260	5.000	5.000	5.040	100.000	100.809
Greece	5.225	7.000	7.712	7.712	110.170	110.170
Ireland	5.015	6.000	6.492	6.492	108.196	108.196
Israel	6.293	5.000	6.412	6.412	128.245	128.245
Italy	4.695	5.000	5.814	5.814	116.272	116.272
Japan	3.042	3.000	5.185	7.894	172.826	263.122
Luxembourg	3.829	5.000	9.093	12.757	181.866	255.145
Netherlands	4.919	5.000	6.170	9.514	123.395	190.284
New Zealand	6.244	5.000	7.662	7.662	153.246	153.246
Norway	4.149	4.000	6.339	7.983	158.470	199.564
Portugal	4.986	6.000	6.234	6.287	103.904	104.776
Singapore	2.048	3.000	4.944	6.061	164.809	202.023
Spain	3.464	5.000	5.857	5.501	117.145	110.026
Sweden	3.373	3.000	4.922	4.916	164.074	163.858
Switzerland	4.915	4.000	6.026	7.614	150.645	190.360
Great Britain	4.851	5.000	5.029	5.942	100.587	118.834
United States	6.754	7.000	7.389	9.453	105.551	135.044

Note: Results are limited to countries for which the WHO provides disaggregated rates to avoid making out of sample predictions. For model 1, explanatory variables include a constant, GDP, GDP^2, maternal mortality, $(\text{maternal mortality})^2$ and FSU. For model 1, explanatory variables include a constant, GDP, GDP^2, maternal mortality, percent urban, and FSU. In both models, total live birth is used to explain the symmetric error and known coverage >= 85 for the one-sided term. Model 3 does not converge.

Figure 1

Average OLS vs. Frontier function estimates of infant mortality rates

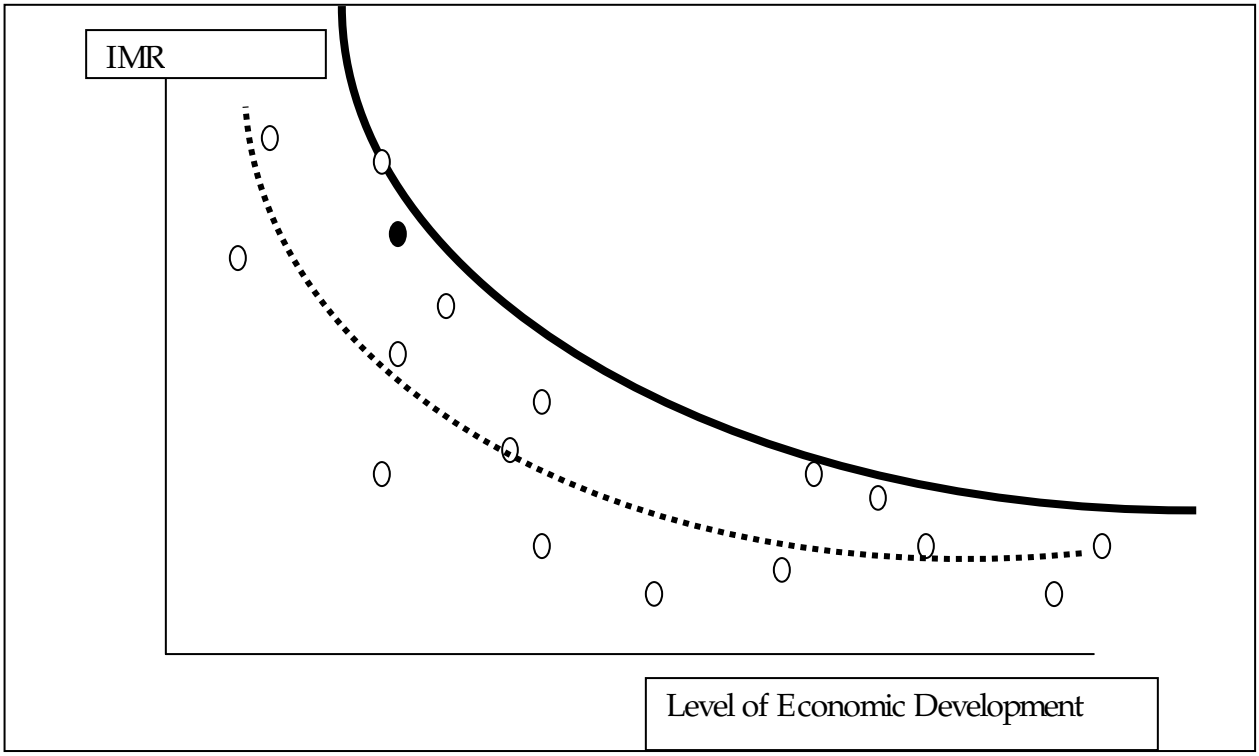

Figure 2

Negative skewness in neonatal mortality model for specification 1