DETERMINANTS OF INFANT MORTALITY.
Predicting infant mortality at global level

Alexandra Barbus
Babes-Bolyai University

Infant mortality is a very popular topic which has been frequently associated with high poverty, low level of education and a deficient health system. Starting from these premises, I further explore the topic from a new perspective, in order to determine the factors that influence the most the infant mortality rate. I use an OLS linear multiple regression, in STATA, on a sample of 89 countries in the world, based on data collected from the World Bank website.

Keywords: infant mortality, literacy, health system, Africa, GDP.

Introduction and background

The main goal of the countries considered as being part of the Third World is fighting against poverty, mortality, hunger and diseases. But poverty is no longer only the problem of the third world countries; it is rather a general social problem, perceived as an obstacle and a threat for development in communities throughout the world. The high rate of infant mortality is directly linked to poverty as a general phenomenon, but it is worthwhile exploring what components of poverty lie at the heart of infant mortality. Therefore, the research question that drives this study is: What are the factors that bear the highest influence on infant mortality?

The infant mortality was defined by many authors, who have offered various definitions. A widely accepted definition claims that "the infant mortality rate represents the probability that a child will die before reaching the age of 1 year in a specific period of time (usually a year) and in a specific geographical area (usually a country)" (Aleshina and Redmond, 2005). We have to add that, starting from this general consideration, there are several comments that emphasize the importance of the precise moment when the infant death occurs. Therefore the literature provides three types of infant mortality: perinatal or early neonatal mortality (infant deaths within six days of birth), neonatal mortality (infant deaths at ages under four weeks, or more precisely infant deaths at ages under 27 days) (Congdon and Southall, 2005) and post-neonatal mortality (infant deaths registered between the first and the 11th month after birth). In addition, another indicator related to infant deaths is

the infant mortality under the age of five years, which according to UNICEF represents the probability that a child dies between the age of 1 and 5 years.

It should also be mentioned that the concept of "life birth" is a crucial concept related to infant mortality rate. Thus, when the infant mortality rate is calculated, it should exclude the children who are not considered to have been born alive, because in this situation they could have not died after their birth (Aleshina and Redmond, 2005). The total number of children who have not been born alive is defined as the mortinatality rate (Mare, 2011). In this paper I will take into consideration the general definition (the probability that a child will die before reaching the age of 1 year).

The general causes of infant mortality (applicable to all the countries in the world) are quite hard to be identified because specific conditions have a major influence over this rate. Different authors have studied the infant mortality in different areas and their conclusions circle around several causal factors: living conditions and nutrition, bad sanitary conditions, social class and income, level of urbanization, ethnicity (Wolleswinkel-van den Bosch, 2000), mother's education level (Kravdal, 2004) or age at first birth (Rahman and Abidin). Trying to include these determinants in some general categories, we obtain several large categories: social conditions (poverty), demographic aspects (ethnicity and age), education and health system. Each of them is tackled in the following sections.

Infant mortality and poverty

The socioeconomic status determines the lifestyle. Together they influence the population's health status, the level of urbanization, the mortality rate, the literacy rate, the measure in which the population is interested in culture or in academic activities. Often we wonder what the cause is and what is the effect or the consequence: the degree of development influences the population's status (economic, social, cultural or health status) or the lifestyle determines the degree of development. In both cases we notice that development is a paramount factor and when we say development, inevitably we must discuss about poverty.

Some sources indicate that, even if the infant mortality rate is a demographic indicator, it can also be used to measure the success of an entire economy, based on the fact that it is determined by various measures and policies implemented by the governments (measures that involve

social, educational or sanitary aspects), which are correlated with the economical system adopted by those who govern the state (Sen, 1998). Therefore, infant mortality can also be linked to governmental performance.

This idea brings us closer to the African context. In 2011, the world's population passed the threshold of 7 billion people (www.worldometers.info), of which 15% live in Africa. It is relevant for our topic to note that, while the African population has an upwards tendency – "from 1960 to 2000, the Sub-Saharan African population tripled from 223.6 milion to 658.9 milion, representing approximately a 195 percent increse. This increase represents a jump from a 7 percent to an 11 percent contribution to the world's population, making it one of the fastest growing regions in the world" (Ukpolo, 2002) – the same area is the most affected by poverty. A study realized by Oxford University in collaboration with the United Nations indicates that the first 10 countries from the ranking of the poorest countries from the entire world are from Sub-Saharan Africa. The same study shows that in some cases the level of poverty reaches up to 90%.[1] This ranking is explained by several aspects such as the geographical region, the lack of natural resources, the natural calamities. A different ranking having the same goal (identifying the world's poorest countries) presents the same 10 countries but in a different order, placing Congo at the top[2], based on an analysis of the data provided by the World Bank regarding the GDP per capita in 2010.

Despite these rankings, the important aspect is represented by the similarities among these countries regarding several indicators: low life expectancy at birth (the lowest value is reached in Malawi: 47 years and the mean value of the majority of these countries is below 50 years[3]), high infant mortality rate, low literacy rate. A look at the natality rate, based on data provided by the World Bank, reveals that these 10 countries have big natality rates (Niger has the highest rate, 48.9 per 1.000 inhabitants, followed by Mali, Uganda and Chad, while the lowest value is reached in Germany with a value of 8.10).[4]

---

[1] These 10 poorest countries are: Niger, Etiopia, Mali, Burkina Faso, Burundi, Somalia, Centrafrican Republic, Liberia, Guinea and Sierra Leone

[2] [http://www.rediff.com/business/slide-show/slide-show-1-the-20-poorest-nations-in-the-world/20110311.htm], accessed on 25 October 2011.

[3] [http://apps.who.int/ghodata/?vid=710#], accessed on 30 September 2011.

[4] [http://data.worldbank.org/indicator/SP.DYN.CBRT.IN], accessed on 30 September 2011.

The influence of poverty over the infant mortality rate can be proved also by comparing these states. A UN Panel Report, published in 2005, discusses the case of Africa, especially the Sub-Saharan region, claiming that „whereas in the developed world less than one children die before age five, in most of sub Saharan Africa that number is one in 10, and in 14 countries it is one in five" (United Nations, 2005). The importance of poverty in the region is emphasized also by figures indicating that "only about 70 to 80% of children are still alive after their fifth birthday" (Strulick, 2004).

Moreover, nowadays poverty is even regarded as more dangerous than terrorism, global warming, pollution or war. A study realized by BBC World Service which analyzes the respondents' answers from 23 countries shows that 71% believe that extreme poverty is the main reason for concern and should have priority in front of other problems.[5] As an attempt to fight poverty, the United Nations, the World Bank and the International Monetary Fund have signed a document called „A Better World For All," which states that the poverty is a challenge for the international community and that they are ready to adopt various measures in order to fight against poverty (Minujin and Delamonica, 2002).

Infant mortality and education

Coming back to the general categories distinguished above, we see that education seems to be an important predictor for infant mortality rate. Apparently there is a link between the mothers' level of education and the infant mortality rate explained by the fact that pregnant woman who have a higher level of education are less predisposed to lose their children in the first year of life. But education has its own predictors: the parents' attitude toward education, the mean income, the community where they live, the level of urbanization, the policies promoted by the government (Kravdal, 2004). A higher level of education of the mother influences the attitude during pregnancy and afterwards and different choices the mother makes (the way of feeding the infant, the hygiene of the place where the infant lives, the periodical medical examinations or the permanent supervision of the child's health) (Levandowski, 2006). In addition, the education has an influence over the age at first birth, meaning that women with a higher level of education give birth for the first time at a higher age. Thus, the relationship between the mother's age at birth (we take into discussion the

[5] [http://www.bbc.co.uk/pressoffice/pressreleases/stories/2010/01_january/17/poll.shtml], accessed on 1 September 2011.

women aged between 12 and 20 years) and the survival's probability of the infant are negatively correlated. Children whose mothers are under the age of 20 years at birth have higher chances to deal with serious health problems before reaching the age of 1 year (Sahel, 2006). In Bangladesh the situation is quite severe, meaning that children whose mothers are under the age of 20 years have 1.5 more chances to die in the neonatal period than those who have mothers aged between 20 and 29 years (Rahman and Abidin, 2010). Kravdal (2004) also links level of education and infant mortality rate. In the communities where the number of highly-educated women is high there is a big interest in preventing the diseases which can cause infant deaths. In many cases this behavior is spread to the entire community, so that other mothers become concerned about this problem and they act cautiously.

Concomitantly, Kravdal offers other evidence to claim that the education is a determinant of the infant mortality rate. He asserts that a higher level of education is positively correlated with a smaller number of ill children; thus, if there are no diseases, there are no chances for the children to become ill. Not only is the mothers' level of education important. The fathers have an important role because they can influence their wives or partners to adopt a healthy behavior during and after the pregnancy, which is favorable for the child's health. In addition, more educated men tend to have better paid jobs, which generate a higher income, meaning that their children can benefit from better medical services (Kravdal, 2004).

A study conducted in China discusses the connection between the women's education and the infant mortality rate, analyzing the 1970-2001 period. The study emphasizes a negative correlation between women with a higher level of education and the infant mortality rate. Significant changes have taken place in this period: the mean number of school years among the female population has doubled from 3.8 years to 7.3 years, the number of medically attended births has increased significantly (in 1970 only one in four births was medically attended, while in 1990 the number rose to 2 out of 3), leading to a decrease of the infant mortality rate (Song and Burgard, 2001).

Even though during the analyzed period the measures implemented by Mao or other leaders regarding the sanitary system were quite different and the policies did not have continuity, the following tendency is obvious: more educated women visit the doctor more often and benefit more from specialized medical services (Song and Burgard, 2001).

A frequently used indicator for the level of education in a country is the literacy rate, which is also negatively correlated with the infant mortality rate; countries that have high literacy rate tend to have low infant mortality rate (Zakir and Wunnava, 1999). The capacity of the population (especially women) to read the instructions of the medicines is very important, as well as the capacity of understanding the necessary doze of medicine or of understanding the instructions received from the pediatricians (Kravdal, 2004).

A study designed and realized in Onondaga, New York, concludes that in 2006, almost a half of the American population was having problems with understanding written information on health, which could be extremely dangerous for their health (Levandowski, 2006). The numbers showed that between 1992 and 1994 the infant mortality rate was 11.3 at 1.000 inhabitants for mothers who did not have a secondary school degree, while for mothers with higher education (secondary school degree or university degree) the rate was 7.4 at 1.000 inhabitants.

Infant mortality and health

Another set of explanations for infant mortality rate refers to the health system. The link between the community's health condition and the life expectancy at birth is well documented. Important organizations, such as the World Health Organization or the Organization for Economic Co-operation and Development, often use the life expectancy at birth to describe the health condition and the quality of life in a country or community. Within the United Nations Development Programme (UNDP), the UN put together in 1990 the Human Development Index (HDI). This index "is a summary measure of human development and provides an alternative to the common practice of evaluating a country's progress in development based on per capita Gross Domestic Product (GDP)".6 It includes three dimensions: knowledge, a long and healthy life and a decent standard of living. The second dimension is measured, among others, by the life expectancy at birth. Moreover, the life expectancy at birth is considered as an indicator of the effectiveness of medical care and various authors use it in order to evaluate the performance of the health system (Smith and Bradshaw, 2006).

[6] [http://www.eoearth.org/article/Human_Development_Index], accessed on 4 October 2011.

Method

As the literature above shows, the link between the infant mortality rate and poverty, education and health system respectively is well documented. Being aware that all the three factors mentioned above can be measured in various ways, in this paper I define the following independent variables: GDP per capita for poverty, the literacy rate for youth female aged between 15 and 24 years for education and the life expectancy at birth for the health system. All these indicators were chosen based on the literature mentioned in the first part of this paper.

In order to answer the research question that drives this study – What are the factors that influence the most the infant mortality rate? – I test the following hypotheses:

The lower the GDP per capita, the higher the infant mortality rate in a country.
The higher the literacy rate for youth female, the higher the infant mortality rate.
The higher the life expectancy at birth, the lower the infant mortality rate.

In order to test these hypotheses, I use an OLS multiple linear regression analysis. The dependent variable is the infant mortality rate and the explanatory variables (the independent variables) are GDP per capita, the literacy rate for youth female and the life expectancy at birth. No dummy variables were included in the model and all the variables are scale. Data on infant mortality rate, GDP per capita, literacy rate and life expectancy at birth is obtained from World Bank website for 2009. I analyze the data using the statistical software STATA.

Two aspects distinguish this article from other studies: (1) the inclusion of all the independent variables in a single model and (2) the analyzed sample of 89 countries (I included all the countries in the world for which data was available). These two elements allow us to make general statements based on the results of the analysis.

Before moving on, I should mention that the main limitation of this study is related to the sample, which has excluded some countries based on the lack of data for several variables; thus, important cases could be neglected.

Statistical analysis

The analysis for the infant mortality rate shows that the mean of the infant mortality rate is 37.06, meaning that, at global level (based on this sample),

we have 37 infant deaths at 1.000 inhabitants; however, the standard deviation is 30.71. The lowest infant mortality rate in 2009 was in Singapore and the highest was in Sierra Leone (117 infant deaths at 1.000 inhabitants) (Annex 1).

For GDP per capita we notice that the distribution of the values is very asymmetrical. Only 27 countries have a GDP per capita above the mean value. The significant difference is also confirmed by the standard deviation of 9931.563. The minimum value of the GDP per capita in 2009 was in Burundi (163$), while the maximum value in Qatar (61.532$) – that is 378 times higher. The mean value of GDP per capita was approximately 6.841$ (Annex 1). The high variation is visible in the figure 2 below.

It is not surprising that the GDP per capita tends to be lower in the African countries, with a mean value of 256$. What is nevertheless striking is that the mean value for these states is almost 27 times smaller than the general mean (Figure 3).

Analyzing the literacy rate for youth female we see that the mean value is very close to 100% (almost all the values are situated in the proximity of 100%). The lowest value was in Chad, where only 39% of the youth female population can read and write at a basic level (Figure 4 and Annex 1).

Figure 1. Distribution of the infant mortality rate

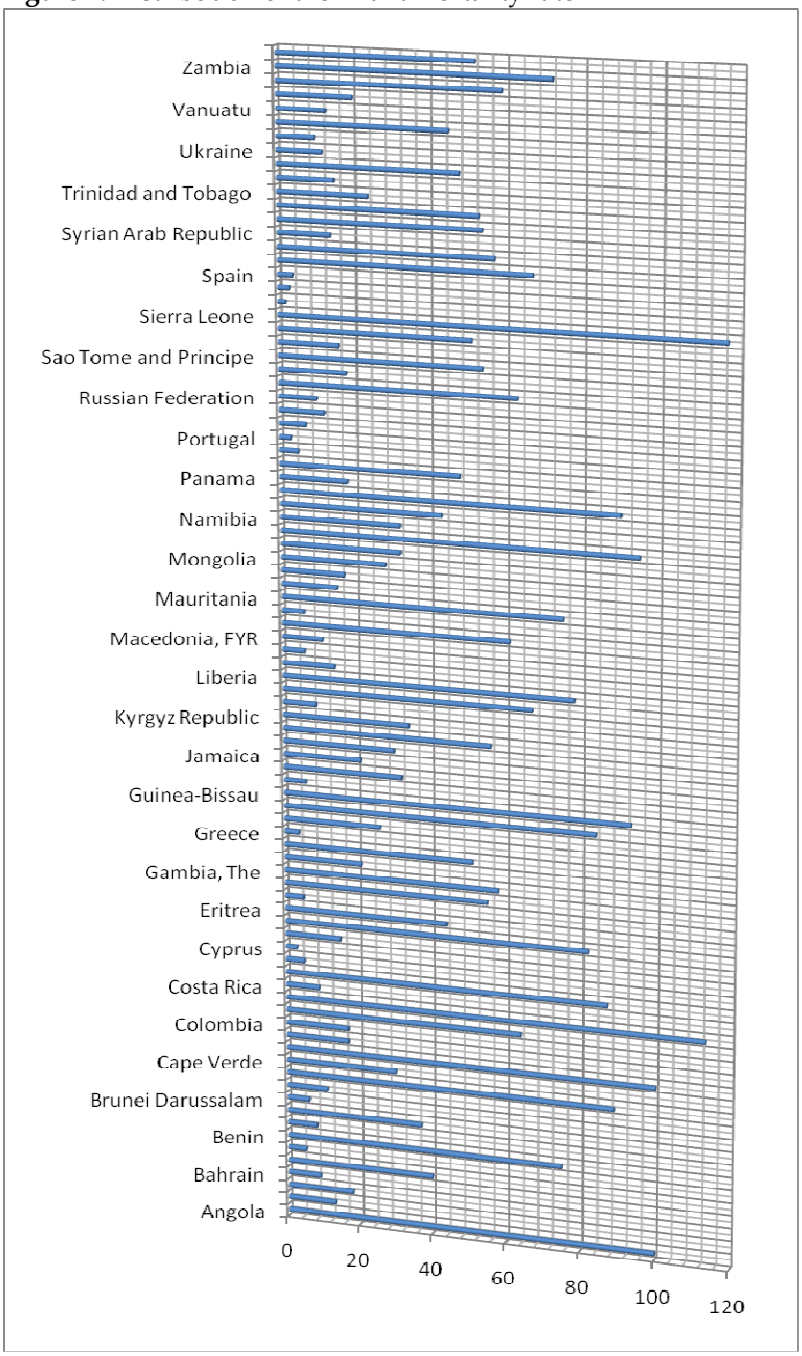


Figure 2. The distribution of GDP per capita

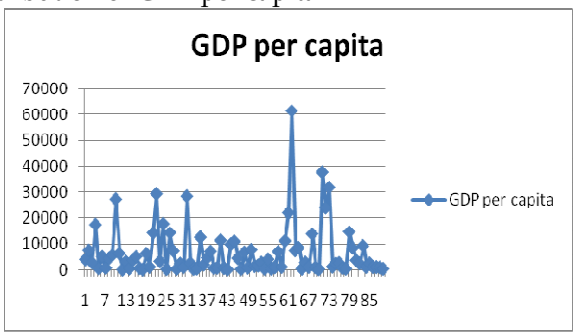


Figure 3. GDP per capita in the African countries

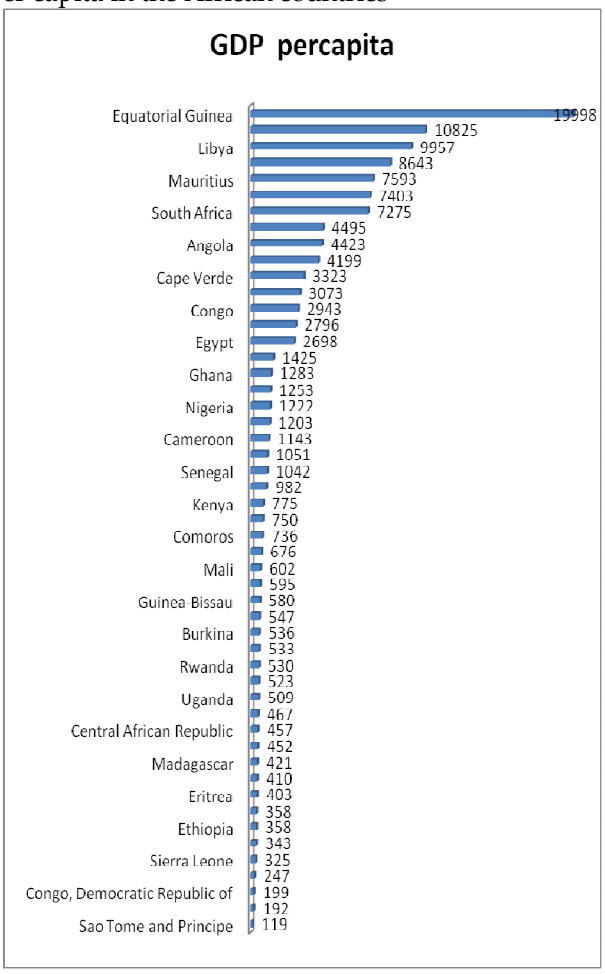
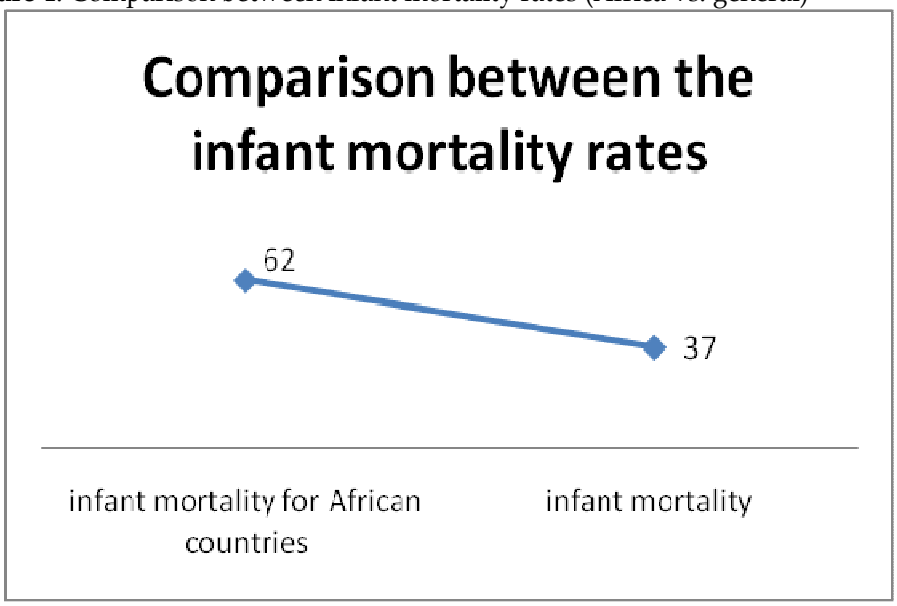


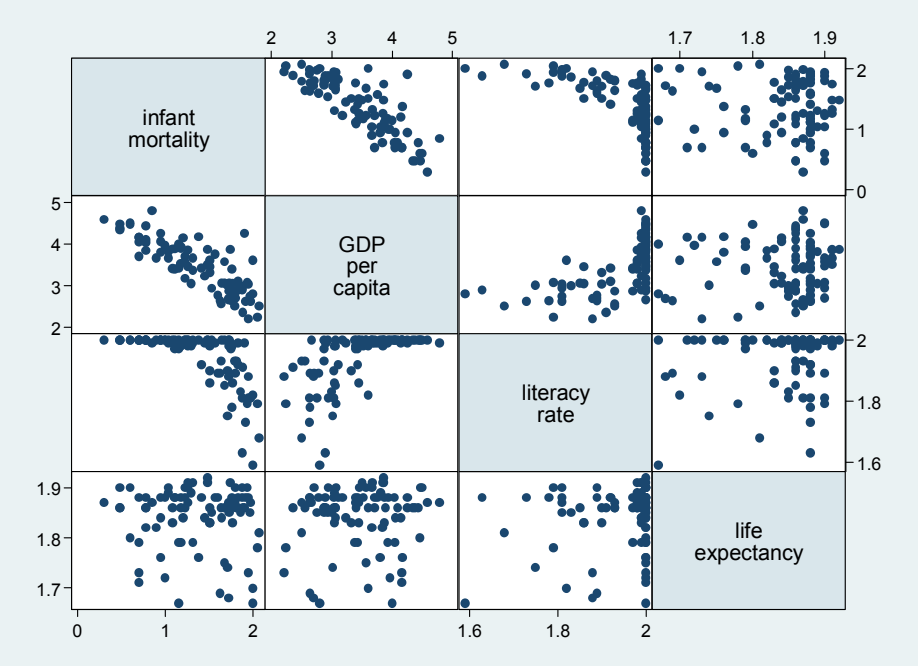
Figure 4. Comparison between infant mortality rates (Africa vs. general)



For the last explanatory variable, the life expectancy at birth, we have a more symmetrical distribution with a mean value of 70.04 years and a standard deviation of 8.94. The minimum value is registered in Chad and Liberia (47 years) and the highest life expectancy at birth is recorded in Kazakhstan (83 years), 13 years above the mean value. (Annex 1). An interesting aspect is the mean value for the African countries (61.96 dead children per 1.000); almost twice as high than the general rate (Annex 2).

All the independent variables resulted to be negative correlated with the infant mortality rate. The strongest correlation was found between the GDP per capita and the infant mortality rate (-0.82), followed by the literacy rate (-0.65) and the life expectancy at birth (-0.04) (Annex 2). The correlation can be seen in the scatter plot below.

Figure 5. The correlation matrix between the infant mortality rate: GDP per capita, the litarecy rate and the life expectancy at birth



The scatter plot above shows that the countries which have high infant mortality rates display small values of GDP per capita, meaning that these countries are generally poor (high values of infant mortality rate are correlated with small values of GDP per capita) and a large percentage of the youth female population is illiterate (high values of infant mortality rate are correlated with small values of literacy rate).

After analyzing the correlation, we have generated a multiple regression. From the three explanatory variables, only two resulted to be statistically significant (the life expectancy at birth resulted to be insignificant, with: the p value of 0.518 and the t value of 0.657). The R2 has a value of 0.7157, meaning that a percentage of approximately 72% of the variance of the infant mortality rate is explained by our model. (Annex 3)

In order to confirm that the life expectancy at birth is statistically insignificant, we generated a simple regression where the dependent variable was the infant mortality rate and the independent variable the life

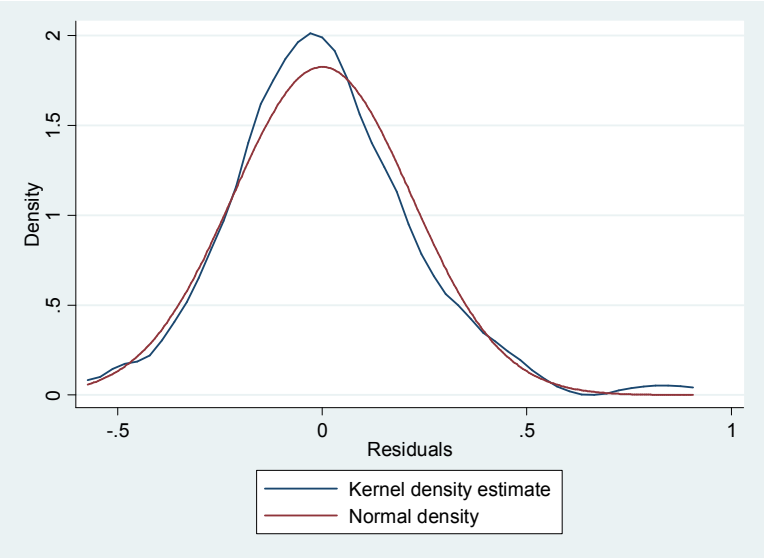
[7] The t value is the correspondent coefficient of the Beta coefficient in SPSS.

expectancy at birth. The results (presented in the Annex 4) show that the independent variable cannot be considered as a predictor for the infant mortality rate (the t coefficient is -0.41 and the p value is 0.683). Also, the R2 is only 0.0019.

Therefore we eliminate the life expectancy from the model and we generate another regression with one dependent variable and two explanatory variables (Annex 5). The software indicates that all the explanatory variables are now statistically significant. As we already saw, the most important predictor for the infant mortality rate is the GDP per capita (p 0.000, t -9.35), followed by the literacy rate of youth female population (p 0.000, t -3.21).

The residuals are normally distributed. There are several ways (graphs, functions and tests) to study the errors' distribution, like Kdensity function, Pnorm plot, Qnorm plot, Shapiro Wilk test (Annexes 7 and 8). I tested the model through all these methods.

Figure 6. The distribution of errors (residuals)



In addition, the White test and the Breusch-Pagan test reveal that the errors are homogenous. Moreover, the Cook-Weisberg test indicates that multicollinearity is absent in this model (there is one value of 8.67, still below 10), thus confirming the validity of the model (Annex 9).

Conclusion

The key predictor for infant mortality according to our analysis is the GDP per capita, followed by the literacy rate of the mothers. The life expectancy at birth cannot be considered a predictor for infant mortality. Thus, two of our three hypotheses are confirmed. The study emphasizes the importance of education as a mean to reduce the number of infant deaths, especially in the African countries, where infant mortality rates are at their highest. Besides the academic perspective, this should represent a clear direction for the international organizations that aim to reduce infant mortality or for various international programs that aim to support the development of the countries from the third world.

A potential direction for further research can be the inclusion of the father's education as an independent variable. Several studies cited above indicated that this particular element could have a major impact on the behavior of mothers.

Bibliography

Aleshina, Nadezhda, and Redmond, Gerry. 2005. How high is infant mortality in Central and EasternEurope and the commonwealth of Independent States?. Population Studies, 59: 1

Congdon. P., and H. Southall. 2005. Trends in inequality in infant mortality in the North of England, 1921-1973 and their association with urban and social structure. Journal of the Royal Statistical Society. Series A (Statistics in Society), 168: 4, 679-700

Kravdal, Oystein. 2004. Child Mortality in India: The Community-Level Effect of Education, Population Studies 58:2, 177-192

Levandowski. A. Brooke et all. 2006. Parental Literacy and Infant Mortality: An Evidence-Based Healthy Start Intervention. Health Promotion Practice

Mare, Codruta. 2011. Suport de curs- Demografie. Babes Bolyai University

Minujin, Alberto. Delamonica. and Enrique Jan Vandemoortelem. 2002. Economic growth, poverty and children. Environment and Urbanization. 14: 2

Murphy, Shaun, and Egger, Matthias. 2002. Studies of the social causes of tuberculosis in Germany before the First World War. International Journal of Epidemiology 31

Rahman, Md. Mosiur, and Abidin, Sumaiya. 2010. Factors Affecting Neonatal Mortality in Bangladesh. Journal of Health Management

Sahel, West African Club. OECD. 2006. The socio-economic and regional context of West African migrations.

Sen, Amartya. 1998. Mortality as an Indicator of Economic Success and Failure. The Economic Journal 108: 446

Smith, W. David, and Bradshaw, Benjamin. 2006. Variation in life expectancy during the Twentieth Century in the United States. Demography 43: 4

Song, Shige, and Burgard. A. Sarah. 2001. Dynamics of Inequality: Mother's Education and Infant Mortality in China, 1970-2001. Journal of Health and Social Behavior 52:349

Strulick. Holger. 2004. Child Mortality, Child Labour and Economic Development, The Economic Journal 114: 497

Ukpolo, Victor. 2002. Population Growth and Economic Growth in Africa. Journal of Developing Societies 18

United Nations. 2005. Poverty. Infectious Disease and Environment Degradation as Threats to Collective Security: A UN Panel Report. Population and Development Review 31:3

Wolleswinkel-van den Bosch. Judith H et all. 2000. Determinants of infant mortality and early childhood mortality levels and their decline in The Netherlands in the late ninetheenth century. International Journal of Epidemiology 29: 1031-1040

Zakir, Mohammed and Wunnava, Phanindra. 1999. Factors affecting infant mortality rates: evidence from cross-sectional data, Applied Economics Letters, 6: 271-273

Ximenes. Ricardo Arraes de Alencar et all.2009. Is it better to be rich in a poor area or poor in a rich area? A multilevel analysis of a case-control study of social determinants of tuberculosis, International Journal of Epidemiology 38: 1285-1296

Websites

www.capital.ro
www.eoearth.org
www.who.int
www.oecd.org
www.rediff.com
www.ucla.edu
www.worldometers.info

Annexes

Annex 1

Infant mortality rate	
Mean value	37.06
Minimum value	2
Maximum value	117
Std. Deviation	30.71
Skweness	0.768882
Kurtosis	2.516731

GDP per capita	
Mean value	6841.685
Minimum value	163
Maximum value	61.532
Std. Deviation	9931.563
Skweness	2.841897
Kurtosis	13.17773

Literacy rate	
Mean value	88.64
Minimum value	39
Maximum value	100
Std. Deviation	15.82
Skweness	-1.34975
Kurtosis	3.750701

Life expectancy birth	
Mean value	70.04
Minimum value	47
Maximum value	83
Std. Deviation	8.94
Skweness	-1.10279
Kurtosis	3.357627

Annex 2

Infant mortality in the African Countries

The African countries analyzed are: Algeria, Angola, Benin, Botswana, Burkina, Burundi, Cameroon, Cape Verde, Central African Republic, Chad, Comoros, Congo, Congo Democratic Republic of, Djibouti, Egypt, Equatorial Guinea, Eritrea, Ethiopia, Gabon, Gambia, Ghana, Guinea, Guinea-Bissau, Kenya, Lesotho Liberia, Libya, Madagascar, Malawi, Mali, Mauritania, Mauritius, Morocco, Mozambique, Namibia ,Niger, Nigeria, Rwanda, Sao Tome and Principe, Senegal, Seychelles, Sierra Leone, South Africa, Sudan, Swaziland, Tanzania, Togo, Tunisia, Uganda, Zambia, Zimbabwe.

Variable	Obs	Mean	Std. Dev.	Min	Max
infantmort	51	61.96078	26.69829	12	114

GDP per capita in the African countries

Variable	Obs	Mean	Std. Dev.	Min	Max
gdpcapita	51	256.6429	282.0244	1.042	982

Annex 3

Correlation	
	infant mortality rate
infant mortality rate	1
GDP per capita	-0.8247
literacy rate	-0.651
life expectancy	-0.0438

Annex 4

Multiple regression, output from STATA

```
    Source |       SS       df       MS              Number of obs =
89
-------------+------------------------------           F(  3,    85) =   71.32
       Model |  13.0441737     3  4.34805792           Prob > F      =
0.0000
    Residual |  5.18230285    85  .060968269           R-squared     =
0.7157
-------------+------------------------------           Adj R-squared =
0.7056
       Total |  18.2264766    88  .207119052           Root MSE      =
.24692

------------------------------------------------------------------------------
infantmort~y |      Coef.   Std. Err.      t    P>|t|     [95% Conf.
Interval]
-------------+----------------------------------------------------------------
gdppercapita |  -.5028702   .0543995    -9.24   0.000    -.611031
-.3947094
literacyrate |   -1.19827   .3673972    -3.26   0.002    -1.928755   -
.4677863
lifeexpect~y |   .2854683   .4400632     0.65   0.518    -.5894953
1.160432
       _cons |     4.91054   .9152639     5.37   0.000     3.09075
6.73033
```

The regression equation
Y= -0.503 X1 -1.19 X2 +0.28 X3 +4.91

Annex 5

Simple regression, output from STATA

The dependent variable: the infant mortality rate
The independent variable: the life expectancy at birth

```
    Source |       SS       df       MS              Number of obs =      89
-------------+------------------------------           F(  1,    87) =    0.17
       Model |  .035015697     1  .035015697           Prob > F      =  0.6834
    Residual |  18.1914609    87  .209097252           R-squared     =  0.0019
-------------+------------------------------           Adj R-squared = -0.0096
       Total |  18.2264766    88  .207119052           Root MSE      =  .45727

------------------------------------------------------------------------------
infantmort~y |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
lifeexpect~y |  -.3268103   .7986168    -0.41   0.683    -1.914148    1.260527
       _cons |   1.978174   1.471599     1.34   0.182    -.9467893    4.903137
```

Annex 6

Regression including the statistically significant variables

```
    Source |       SS       df       MS              Number of obs =      89
-------------+------------------------------           F(  2,    86) =  107.49
       Model |  13.0185177     2  6.50925886           Prob > F      =  0.0000
    Residual |  5.20795888    86  .060557661           R-squared     =  0.7143
-------------+------------------------------           Adj R-squared =  0.7076
       Total |  18.2264766    88  .207119052           Root MSE      =  .24608

------------------------------------------------------------------------------
infantmort~y |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
gdppercapita |  -.5055258   .0540623    -9.35   0.000    -.612998   -.3980535
literacyrate |  -1.152539   .3593538    -3.21   0.002    -1.86691   -.4381668
       _cons |   5.356744   .6017646     8.90   0.000     4.160476    6.553013
```

The regression equation

Y= -0.505 X1 -1.152 X3 +5.35

Annex 7

The errors and predicted values, output from STATA

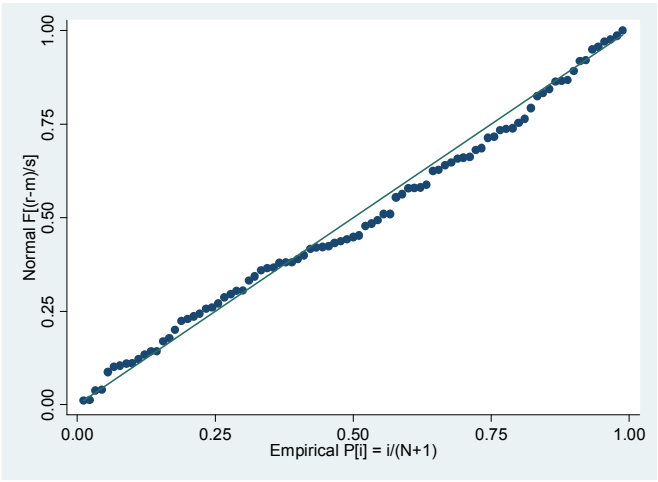
	inf mort~f	p	r
1.	2	1.586317	.4136828
2.	1.11	1.027961	.0820391
3.	1.26	1.294899	-.0348988
4.	.95	.9015182	.0484817
5.	1.6	1.783775	-.1837746
6.	.7	1.197664	-.4976639
7.	1.88	1.908597	-.0285965
8.	.9	1.175385	-.2753855
9.	1.57	1.42982	.1401804
10.	.78	.9214421	-.1414421
11.	1.04	1.106851	-.0668507
12.	1.95	1.944669	.0053305
13.	1.48	1.356697	.123303
14.	2	2.111914	-.1119141
15.	1.23	1.262055	-.032055
16.	1.23	1.124331	.1056695
17.	1.81	1.504934	.3050657
18.	2.05	2.126091	-.0760912
19.	.95	.9761321	-.0261321
20.	1.94	1.728955	.2110447
21.	.7	.901876	-.201876
22.	.48	.592259	-.112259
23.	1.18	1.183878	-.0038776
24.	1.91	1.073634	.8363657
25.	1.64	1.728239	-.0882393
26.	.7	.9342429	-.2342429
27.	1.74	1.379034	.360966
28.	1.76	1.969352	-.2093519
29.	1.32	1.366204	-.0462043
30.	1.71	1.573763	.136237
31.	.6	.6120785	-.0120784
32.	1.41	1.365541	.0444593
33.	1.92	2.043275	-.1232752
34.	1.97	1.879631	.0903695
35.	.78	.8751419	-.0951419
36.	1.51	1.431494	.0785055
37.	1.32	.9446731	.375327
38.	1.48	1.240224	.2397763
39.	1.75	1.706561	.043439
40.	1.53	1.572044	-.0420445
41.	.95	1.016457	-.0664567
42.	1.83	1.758705	.0712953
43.	1.89	1.945597	-.0555965
44.	1.15	1.024787	.1252131
45.	.78	1.077674	-.2976738
46.	1.04	1.083679	-.0436792
47.	1.79	1.851612	-.0616122
48.	.78	1.169276	-.3892758
49.	1.88	1.850065	.0299348
50.	1.18	.9752668	.2047332
51.	1.23	1.498265	-.2682651
52.	1.45	1.517339	-.0673385
53.	1.51	1.46581	.0441905
54.	1.98	2.023061	-.0430614
55.	1.51	1.518917	-.0089172
56.	1.63	1.796918	-.1669181
57.	1.95	1.706753	.2432465
58.	1.26	1.121468	.1385316
59.	1.68	1.754518	-.0745182
60.	.7	.9418725	-.2418725
61.	.48	.8633991	-.3833991
62.	.85	.7613052	.0886948
63.	1.08	1.213563	-.1335631
64.	1	1.162585	-.1625848
65.	1.8	1.731007	.068993
66.	1.26	1.080878	.1791224
67.	1.73	1.488696	.2413039
68.	1.2	.8924569	.3075431
69.	1.71	1.86742	-.1574198
70.	2.07	2.221929	-.1519292
71.	.3	.8031061	-.5031061
72.	.48	.7355705	-.2555706
73.	.6	.7431855	-.1431855
74.	1.83	1.560115	.2698851
75.	1.76	1.6681	.0918999
76.	1.15	1.187364	-.0373636

77.	1.73	1.627167	.1028333	
78.	1.72	1.798869	-.0788692	
79.	1.38	.8982398	.4817602	
80.	1.18	1.022759	.1572408	
81.	1.68	1.243174	.4368256	
82.	1.08	1.347204	-.2672044	
83.	1	.9655666	.0344334	
84.	1.65	1.5005	.1494996	
85.	1.11	1.343923	-.233923	
86.	1.3	1.579583	-.2795829	
87.	1.77	1.549216	.2207844	
88.	1.86	1.854478	.0055219	
89.	1.72	1.837447	-.1174474	

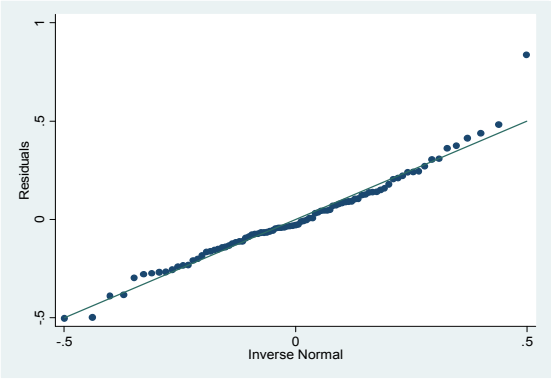
Annex 8

The distribution of errors

Normal Probability Plot



Normal Quantile Plot



Annex 9

Shapiro-Wilk test

Shapiro-Wilk W test for normal data

Variable	Obs	W	V	z	Prob>z
r	89	0.97333	1.999	1.527	0.06342

Annex 10

Cameron & Trivedi's decomposition of IM-test

Source	chi2	df	p
Heteroskedasticity	16.55	9	0.0563
Skewness	4.46	3	0.2159
Kurtosis	1.09	1	0.2958
Total	22.10	13	0.0538

The Breusch-Pagan test for multicollinearity
estat hettest

Breusch-Pagan / Cook-Weisberg
Ho: Constant variance
Variables: fitted values of logmortinf

chi2(1) = 8.67
Prob > chi2 = 0.0032

Annex 11

Test for ommited values

Ramsey RESET test using powers of the fitted values of logmortinf
Ho: model has no omitted variables
F(3, 82) = 2.41
Prob > F = 0.0729