

**Determinants of the Number of Deaths from COVID-19: Differences between Low-Income and High-Income Countries in the Initial Stages of the Pandemic.**

Magali Valero<sup>1</sup>

University of Michigan at Dearborn

mvalero@umich.edu

Jorge Valero

Facultad de Economía, Universidad Autónoma de Nuevo León

jorge.valerogl@uanl.edu.mx

*International Journal of Social Economics* Vol. 48 No. 9, pp. 1229-1244. DOI: 10.1108/IJSE-11-2020-0752

Licensing statement: This Author Accepted Manuscript is deposited under the Creative Commons Attribution Non-commercial International Licence 4.0 (CC BY-NC 4.0), and any reuse is allowed in accordance with the terms outlined by the licence.

---

<sup>1</sup> Corresponding Author.

# **Determinants of the Number of Deaths from COVID-19: Differences between Low-Income and High-Income Countries in the Initial Stages of the Pandemic.**

## **Abstract**

The purpose of this study is to understand the factors that contribute to the number of reported coronavirus (COVID-19) deaths among low-income and high-income countries, and to understand the sources of differences between these two groups of countries. Multiple linear regression models evaluate the socio-economic factors that determine COVID-19 deaths in the two groups of countries. The Oaxaca-Blinder decomposition is used to examine sources of differences between these two groups. Low-income countries report a significantly lower average number of COVID-19 deaths compared to high-income countries. Community mobility and the ease of carrying the virus from one place to another are significant factors affecting the number of deaths, while life expectancy is only significant in high-income countries. Higher health expenditure is associated with more reported deaths in both high and low-income countries. Factors such as the transport infrastructure system, life expectancy, and the percent of expenditure on health lead to the differences in the number of deaths between high and low-income countries. This paper contributes to the emerging literature on COVID-19 and its relation to socio-economic factors by examining the differences in reported deaths between low-income and high-income countries. Our study shows that mobility measures taken by individuals to limit the spread of the virus are important to prevent deaths in both high- and low-income countries. In addition, our results suggest that countries with weak health institutions underestimate the number of deaths from COVID-19, especially low-income countries. The underestimation of COVID-19 deaths could be affecting a great number of people in poverty in low-income economies.

**Keywords:** Coronavirus, COVID-19, Deaths, Mobility, Oaxaca-Blinder Decomposition, High-Low-income Countries.

## 1. Introduction

With the onset of the 2019 coronavirus (COVID-19) epidemic, the need to understand the disease arises as it is spreading throughout the world. The detrimental effects of this pandemic need not be of the same magnitude in different countries. For instance, Barro, R. J. et al. (2020) show the mortality rates of the Great Influenza Pandemic of 1918-1920 varied widely across countries and years. It is important to understand the main factors influencing deaths from COVID-19 and how they relate to differences in COVID-19 deaths across countries.

In this paper, we seek to understand the determinants of the number of deaths from the coronavirus disease. Some recent studies aim to understand differences in COVID-19 mortality among countries. Sornette et al. (2020) analyze COVID-19 mortality statistics, and find a higher number of deaths in countries with a higher life expectancy and lower stringency of confinement. Chaudry et al. (2020) study COVID-19 mortality by focusing on socioeconomic factors as well as country preparedness and government actions. They find higher mortality among countries with a higher rate of obesity, less smoking prevalence, a fewer amount of nurses, more income dispersion, and a higher GDP. They also find no relation between mortality and border closings or full lockdowns. One might question why higher GDP countries exhibit higher mortality from COVID-19. As we seek to understand the differences in the number of reported deaths among countries, the focus in this paper is on how mortality differs in high-income countries versus low-income countries, and how these differences relate to socioeconomic factors.

It is difficult to measure COVID-19 fatality rates in the early phases of the epidemic given the lack of accurate information on the number of infections or resolved cases (Atkins, 2020). This could be more severe for developing economies. However, the number of deaths is easier to observe, so counting the number of deaths is more reliable than counting the number of infections

(Barro et al., 2020). Moreover, the number of infections or the number of tests will vary with requirements and availability in different countries, which makes focusing on deaths more appropriate (Fernández-Villaverde and Jones, 2020).

It is important to note that COVID-19 deaths are said to be underestimated by at least 50% in the United States and by more than 60% in the rest of the world. Generally, an individual would first be identified as infected, and if they later die from the virus then the death would be recorded as a COVID-19 death. For individuals not identified as infected first, the correct accounting as a COVID-19 death will depend on the effectiveness of the government and of the health system. Using the available data, we study the determinants of the number of deaths that are reported and captured in the data repository maintained by Johns Hopkins University (JHU CSSE, 2020). We therefore study the determinants of reported deaths from COVID-19, which could differ from actual deaths.

Our study examines how deaths relate to factors at the onset of the pandemic while focusing on the number of deaths at a relative point in time. Figure 1 shows the log average number of deaths in high-income versus low-income countries<sup>2</sup>. Day zero corresponds to the day the country reached the first five deaths. High-income countries experience more deaths from COVID-19 compared to low-income countries. The difference between the two increases as the disease progresses. We study the number of deaths 20 days and 40 days after the first reported 5 deaths in the country.

[Insert Figure 1 here]

---

<sup>2</sup> Countries in our sample are split based on median 2018 per capita GDP.

The number of deaths depends on both how much the disease spreads, and on the probability that an infected person dies. With the arrival of this pandemic, many countries have implemented a set of social distancing rules. These rules reduce contact between individuals and reduce the probability of infection, for instance, quarantine and social distancing procedures, and restrictions on international travel. The probability that an infected person dies is a function of their age, wealth level, and quality of health care. We start by studying the factors that influence the spread of the infection and the proportion of infected that die. We find both commonalities and differences between low-income and high-income countries. The regression analysis shows that the factors that influence the rate at which the disease spreads (mobility and the transport infrastructure measured with number of arrivals) are important determinants of the number of deaths for both groups of countries at both the 20-day and the 40-day mark. When we analyze factors that affect the probability of going from infected to dead, important differences emerge. Life expectancy, a measure of the age of the population, is only significant in high-income countries, where it is positively associated with the number of deaths. Additionally, a lower percentage of expenditure on health is associated with fewer deaths in low-income countries. We interpret this as signaling the strength of the health system, where a weaker health system is less able to correctly classify deaths as COVID-19.

Understanding the factors associated with a higher number of deaths in low-income and high-income countries is important. We also question what leads to the differences in the average number of deaths between the two groups of countries. Utilizing the ideas in the Oaxaca-Blinder model, we examine to what extent the differences between the two groups are due to differences in endowments. For instance, differences in death rates could be due to differences in life expectancies (a difference in endowments), or they could be due to individuals receiving a higher

level of medical attention in countries with the same amount of health spending (a difference in the beta coefficients or the estimation errors).

We show that the differences in deaths in low-income and high-income countries 20 days after the first five deaths stem from differences in the endowments of the number of arrivals (a proxy for the transport infrastructure) and life expectancy, and to differences in how life expectancy and health expenditure affect the number of deaths. At 40 days however, the number of arrivals into the country, which should influence the fraction of individuals that are infected, and to a lesser degree health expenditure, explain the differences in the number of deaths between the two groups.

Our study contributes to the understanding of the effect that containment, social distancing policies, and economic variables can have on death outcomes from the coronavirus epidemic. The paper is organized as follows. In the next section, we review the epidemiological background needed to study COVID-19. Section 3 details the factors that we consider in our study. Section 4 discusses the data. Section 5 presents our results. Section 6 concludes.

## **2. Background**

Variations of the classic epidemiological model of compartments (Kermack and McKendrick, 1927; Kermack and McKendrick, 1932) are being used to study COVID-19. These epidemiological models study the spread of the disease through several compartments, - susceptible population, infected population, recovered, or dead. When a susceptible individual has contact with an infected individual, there is a certain probability that the susceptible person will be infected. Once infected the individual has a certain probability,  $\delta$ , of dying and a certain probability,  $(1-\delta)$ , of recovering. The model shows there is a parameter  $R_0$  that determines the

growth in the number of infected at time 0.  $R_0$  is the ‘basic reproduction number’ and represents the number of infections caused by one infected person in a population where everyone is susceptible.

Epidemiological models help us understand the course of the epidemic, but there is also a need to understand how economic decisions affect rates of COVID-19 infection and outcomes. As such, economists have started bridging this gap. Avery et al. (2020), Stock (2020), and Eichenbaum et al. (2020) study the epidemiological model of compartments and how it can fit into economic models. Eichenbaum et al. (2020) model the number of deaths, and show that they fall with a reduction in work and consumption. We use as a basis epidemiological models with two purposes. First, the models provide a basis for pinpointing the factors that are important in the spread of the disease, and hence affect the number of deaths. We consider factors that affect both the number of infections an individual can generate and the probability that an infected individual dies. Second, the models show that parameters vary through time and there is time component to the growth and progression of the disease. They highlight the importance of analyzing countries in a similar stage of the epidemic, which suggests taking a relative point in time where countries would be in the same stage.

We assume that some model parameters that study mortality vary depending on individual behavior. For instance, the probability of infection falls with the degree of contact and hence the social distancing measures. Wang et al. (2020) establish the reproduction parameter  $R_0$  to initially be 3.1 in Wuhan (December 2019-January 2020) and lower the parameter to 2.6 once social distancing rules are implemented (January 24, 2020). The value of the parameter further falls as the epidemic is controlled. Fernández-Villaverde and Jones (2020) consider the reproduction

parameter to be first a function of the disease itself, and hence the same for all regions, but then also a function of regional characteristics such as density and culture.

To study the number of deaths forces our focus on the determinants of the likelihood of going from susceptible to infected, determinants that influence the value of the reproduction parameter  $R_0$ , and also those determinants of  $\delta$ , the probability of dying once infected. Although the parameters can vary through time depending on specific policies and containment measures implemented, we study the same relative point in time for all countries by selecting a starting point in the evolution of the disease. In particular, our starting point is the day in which the country reaches the first five deaths from COVID-19. We then analyze the number of cumulative deaths at a later date. The median time from the onset of symptoms to recovery is estimated to be 2 weeks for mild cases and 3-6 weeks for severe cases; while looking at patients who died from COVID-19, the median number of days between symptoms and outcome is of 2-8 weeks (WHO, 2020)<sup>3</sup>. We therefore select a period of 20 days after the first five deaths for our analysis to compare high-income versus low-income countries; and we additionally study the number of deaths at 40 days after the first five deaths.

### **3. Factors influencing the number of COVID-19 deaths**

The number of deaths will depend on the average number of infections that an infected individual can generate, measured through  $R_0$ , and on the probability  $\delta$  that an infected individual dies. We consider factors that influence both the number of infections and the number of deaths. In particular, the rate at which the disease spreads will depend on social distancing measures,

---

<sup>3</sup> Wang et al. (2020) report an average hospitalization period of  $12.39 \pm 4.77$  days.



which vary across countries. Social distancing measures reduce the contact between individuals, reducing the probability that an infected individual has contact with a susceptible individual. We use community mobility and the number of arrivals into the country in this category of factors. We consider the probability that an infected individual dies as a function of the age of the individual, their wealth level, and the quality of the health system. We measure these factors at the country level by separating countries according to GDP per capita, and using life expectancy, health expenditure as a percentage of GDP, and the death rate in the country. We also consider the size of the shadow economy, as it can influence both the number of infections and the probability that an infected individual dies. We now discuss each of these variables in more detail.

*Mobility.* The spread of the virus should be lower if the community is engaging in social distancing measures, with less people infected and hence a lower number of individuals dying, other things equal. Governmental authorities have imposed different rules intended to slow the spread of the virus, such as stay at home orders and gatherings that limit the number of people, among others. The rules are not effective unless they can be enforced, and as implementation and monitoring could differ among populations, it can affect outcomes across nations. Additionally, it is possible that even when the government does not implement social distancing measures the population isolates themselves (or they change their behavior before any government guidance is given). Our approach is to focus on the actions that individuals take, rather than the restrictions that the governments impose. We thus look at mobility of the community as a representation of distancing, instead of considering governmental distancing rules<sup>4</sup>.

---

<sup>4</sup> When we considered governmental distancing efforts, using the Stringency index of Hale, Petherick, Phillips and Webster (2020), we find that such lockdown policies reduced deaths only in high-income countries. The Stringency Index is not included as it is highly correlated with Mobility.

We use Google Community Mobility Reports to measure social distancing efforts during the COVID-19 pandemic (Google LLC, 2020). The report measures the percentage change in visits to specific places within a geographic area, compared to a baseline for that day of the week. The baseline used is the same day of the week during the 5-week period of January 3 to February 6, 2020. In particular, we use the mobility measure to transit stations, which reports the percentage change in visit to places of public transport hubs such as subways, busses, and train stations. Mobility then represents the percentage change in visits to transit stations from the baseline days to the day when the country reported their first five deaths. The average mobility in high-income and low-income countries is shown in Figure 2. Notice a stronger retraction in mobility on the day of the first five deaths in low-income countries compared to high-income countries<sup>5</sup>. These differences in mobility seem to disappear as time goes on.

[Insert Figure 2 here]

*Number of arrivals.* Generally, more advanced economies have stronger transport infrastructure, and hence more interactions between people, which should help foster the spread of COVID-19. To measure this, we use the number of arrivals, which refers to the number of inbound tourists who travel to a country (overnight visitors). Number of arrivals proxies for the intensity with which the virus is carried between countries, while it also proxying for better transportation infrastructure within the country. We use the number of arrivals into a country for 2018 (in thousands), collected from the World Development Indicators database (The World Bank, 2020). The number of arrivals is about 1 million in Bolivia, 39 million in Germany and 80 million in the United States. While it is hard to gauge with the number of arrivals of 2018 whether the

---

<sup>5</sup> The date of the first five deaths is on average sooner (in calendar terms) for high-income countries than for low-income countries. The population from low-income countries then reacts stronger in terms of mobility on the date of the first five deaths.

closing of borders or the reduction in the number of flights was effective in reducing COVID-19 deaths, we should generally expect the transmission of the virus to be more severe for more transited countries and those with better transport networks, which would imply a greater number of deaths.

*Size of the shadow economy.* The consequences of COVID-19 can be more severe for those working in the informal economy (Egger et. al., 2021). There are two channels in which the shadow economy can influence COVID-19 deaths. Firstly, the informal sector includes cases where neither the employee nor the employer contribute to social security (ILO 2018) and therefore individuals lack social safety nets which gives them the possibility of being seen in hospitals, therefore affecting mortality and its measurement. In particular, we would expect a larger shadow economy to lead to more deaths as it captures individuals possibly lacking access to medical care. On the other hand, the shadow economy can also measure the erosion in the power of the state, and proxy for governability of the population and therefore compliance. If citizens do not abide by the public institutions regulations and recommendations, this can affect the number of infected individuals and therefore lead to more deaths. There are multiple ways in which state institutions can slow down the number of infections, for instance the recommendation or mandate on the use of masks, social distancing measures, restrictions on social gatherings, etc. However, these rely on the influence of the state on the population and hence on compliance<sup>6</sup>. Through this effect, we also expect a larger shadow economy to lead to more deaths. We take the estimations of the shadow economy for 2015, provided by Medina and Schneider (2018) as a measure of access and strength of institutions in the country. The shadow economy (also called informal economy, black

---

<sup>6</sup> For instance, the use of facemasks reduces the risk of transmission and mitigates the spread of COVID-19 (Lyu and Wehby, 2020), countries with high compliance of mask wearing have lower incidence of COVID compared to non-mask wearing countries (Cheng et. al, 2020).

economy, etc.) includes economic activities that are hidden from the government authorities for monetary, regulatory, and institutional reasons. The shadow economy measure of Medina and Schneider (2018) reflects legal economic and productive activities that would have contributed to GDP had they been recorded, and it measures the size of the shadow economy as a percent of GDP.

*Life expectancy.* Age can also be a determinant of the number of COVID-19 deaths. According to the Center for Disease Control and Prevention (NCHS 2020), in the United States using data up to October 31, 2020, 79% of COVID-19 deaths were of individuals aged 65 and above, and 92% of individuals aged 55 and over. Similarly, for the province of Wuhan in China, Wu et al. (2020) find that compared to 30-59 year olds, those under 30 years of age are 0.6 times more likely and those over 60 years of age are 5.1 times more likely to die if symptomatic. In a related study, Stojkoski et al. (2020) find life expectancy is the most precise predictor of mortality from COVID-19. We include life expectancy in our analysis (similar to Sornette et al., 2020) as it is a broader measure than age<sup>7</sup>. For instance, it is possible that in countries where life expectancy is higher there are more individuals with conditions, beyond age, that make them more susceptible to the effects of the virus. We expect that in countries where life expectancy is lower we will observe fewer deaths from COVID-19. Life expectancy at birth in 2018 is taken from the WDI database (The World Bank, 2020).

*Death Rate and Health Expenditure.* Countries with more effective health systems should experience less COVID-19 related deaths. We measure the effectiveness of the health system using the percentage of GDP that the country spends on health. Countries with a larger fraction of the

---

<sup>7</sup> The correlation between life expectancy and the percentage of the population aged 65 and above is 0.6898. The correlation between the percentage of the population aged 65 and above and death rate is 0.69.

population in poverty generally also suffer from health systems that are more inefficient. A direct measure of poverty is not widely available for our sample countries and including it would imply the loss of many observations. We use death rate as a proxy variable for poverty and for the weakness of the health system. A country with high death rates has deficiencies in its health system. For instance, Hartz et al. (1989) find that even in American hospitals the mortality rate increases for lower quality hospitals. At the country level, Center et al. (2012) study prostate cancer and find that in more developed countries the incidence rates are higher and the mortality rates lower compared to developing countries such as parts of South America, the Caribbean and Sub-Saharan. Death rate can also proxy for the general health-level of the population. We use as variables the 2017 current health expenditure as a percentage of GDP, and the 2018 death rate per 1000 people, both from the WDI database.

#### **4. Data**

Data on the number of deaths is from the COVID data repository maintained by Johns Hopkins University (JHU CSSE, 2020); the COVID data we use is as of October 19, 2020. Our focus is on reported COVID-19 deaths that are captured in the data repository, which may differ from actual COVID-19 deaths. In order to consider differences in the number of COVID-19 deaths, our first data requirement is that as of October 19, 2020, at least 20 days have passed since the country surpassed the five deaths figure. Our starting sample is then 162 countries. Further restricting the sample to countries with available mobility data (as of the five deaths date) our sample falls to 116 countries. Furthermore, we require that there is data availability on GDP per capita, life expectancy, size of the shadow economy, health expenditure, death rate and the number of arrivals into the country. Our final full sample consists of 92 countries. We subdivide the sample into two, based on median per capita GDP. The low-income group consists of 46 countries with

average GDP per capita of \$4,668, and the high-income group consists of 46 countries with average GDP per capita of \$38,999.

The earliest date of COVID-19 data from Johns Hopkins University is 1/22/20. For accurate dates of the first reported confirmed case we double check with news sources. The country with the first confirmed case in our sample is Thailand, with a confirmed case on January 13, 2020, which was also the first confirmed case of COVID-19 outside of China<sup>8</sup>. Thailand was followed by Japan on January 16, and the United States on January 22. The first country to report a death was Philippines on February 2, while the first countries to reach five deaths were Italy (Feb 24) followed by Japan (Feb 29) and the United States (March 2). Between the time of the first confirmed case and the time when the country had five confirmed deaths, there are a minimum of 10 days (Turkey) and a maximum of 124 (Nepal). Similarly, twenty days after those five first deaths there are a minimum of 6 deaths (Costa Rica and Malta) and a maximum of 4365 (Spain)<sup>9</sup>.

Average and median differences of the variables between low-income and high-income countries are shown in Table 1. In high-income countries, there are significantly more deaths 20 days after recording the fifth death. In particular, 20 days after the fifth death there are on average 104 deaths in low-income countries versus 391 deaths in high-income countries (the median number of deaths is 23 in low-income versus 82 in high-income countries). In high-income countries there is also higher life expectancy at birth (80 years versus 71 for low-income countries), a smaller shadow economy (16% versus 33%), and higher health expenditure as a percentage of GDP. We also show that for both high-income and low-income countries, mobility significantly

---

<sup>8</sup> China is excluded from our sample since by 1/22/20 it had already reported more than five deaths.

<sup>9</sup> A table with the countries in our sample, dates on which the country first reported five deaths, number of deaths recorded 20 days and 40 days later, and the factors we consider in our analysis is available from the authors upon request.

drops. On the day when the country reported five deaths mobility to transit stations had dropped by an average of 52% in low-income and by 48% in high-income countries. Differences in average mobility and death rates between the two groups are not significant. High-income countries have more than double the number of arrivals compared to low-income countries, indicating they have both higher international mobility and better transport infrastructure, which facilitate the spread of COVID-19.

[Insert Table 1 here]

## 5. Results

### *Number of deaths 20 days after first five deaths*

Table 2 shows the results of OLS regressions where the dependent variable is the number of reported deaths in the country 20 days after the first five reported deaths. The first column shows results for the full sample while in the following two columns the sample is split into low- and high-income countries. The full sample results show more deaths are associated with higher mobility to transit stations and with a greater number of arrivals. However, by splitting the sample differences emerge between the two groups. Using Chow's test (Chow, 1960) we find the coefficients of the two regressions are significantly different (test statistics at the foot of the table), highlighting the importance of separating the two groups.

[Insert Table 2 here]

We can focus first on the parameters of the variables that affect the growth in the number of infected ( $R_0$ ). For both low-income and high-income countries, Mobility is positive and statistically significant. In countries with higher mobility where people are less strict at following social distancing rules (if such rules are in place) more deaths result. Higher mobility increases the

spread of the disease and with more people infected one would expect a greater number of deaths. Notice however that the significance and coefficients are higher in high-income countries. Similarly, in both low-income and high-income countries the number of international arrivals is positively related to the number of deaths. This might happen through two channels. First, more international arrivals proxies for more socially open economies, which could, at the onset of the pandemic, have led to a stronger spread of the disease. Second, countries with more international arrivals can be associated with a more advanced transport infrastructure. One would expect that countries with stronger transport infrastructures experience a wider-spread of the disease, as mobility of individuals between cities and within cities is more prominent and easier. Shadow economy, which measures the institutional factor, is positive in high-income countries; a larger shadow economy is associated with more COVID-19 deaths. A larger shadow economy can mean a population with lower social safety nets and access to medical care, which leads to more deaths. Additionally, the larger shadow economy also proxies for the lack of compliance of the population to public recommendations and restrictions. A larger shadow economy proxies for individuals unwilling to follow rules, which results in more deaths perhaps because it leads to a faster spread of the virus.

We focus next on the parameters that affect the probability of death once infected,  $\delta$ , which are Life Expectancy, Death Rate, and Health Expenditure. Within high-income countries, more deaths occur in countries with higher life expectancy, although only marginally significant the coefficient is large. As mentioned earlier, higher life expectancy is associated with an older population, and preliminary data suggests that COVID-19 is more deadly for such population. However, in the low-income regression life expectancy is not significantly different from zero.



One possibility is that in low-income countries other factors such as the strength of the health system play an important role in the number of deaths<sup>10</sup>.

Death Rate and Health Expenditure are not significant in the high-income regression but they have the expected signs; in countries with higher death rates and those with lower health expenditure, a weaker health system, there are more COVID deaths. Interestingly, in low-income countries a higher percentage of health expenditure and a lower death rate are associated with more deaths. It is not necessarily true that the stronger health system will cause more COVID deaths, but rather, it is possible that the stronger health system is better able to identify COVID-related deaths. Both death rate and health expenditure can measure poverty in the national health systems, and a poorer health system is less likely to detect a COVID-19 death. Burn-Murdoch et al. (2020) from the Financial Times compare the number of deaths from all causes between March and April 2020 to the average deaths for the same period between 2015 and 2019. The difference, excess deaths, outnumber official coronavirus death tolls, deaths rising above normal by up to 60%. As an example, in the province of Gayas in Ecuador, 245 coronavirus deaths were reported, but the number of excess deaths compared to the baseline is 10,200. The degree of underestimation of COVID-19 deaths might vary across countries. For instance, while in Sweden and France 92% of excess deaths are attributed to COVID-19, in Turkey only 44% and Jakarta only 14% of excess deaths are attributed to COVID-19 (The Economist, 2020), suggesting underestimation of deaths that is more severe in lower-income countries. In the case of the United States Katz et al. (2020) find coronavirus related deaths could be underestimated close to 50%. We posit that in lower-income countries and countries with deficient health systems, a fraction of those COVID-19 deaths

---

<sup>10</sup> For instance, using data of the official death counts of Mexico as of July 16, 2020 (Gobierno de México, 2020), 59% of the COVID-19 deaths are of people aged under 65; while in the United States only 25% of deaths are of individuals under 65, CDC (2020).

are not detected as infected first, and upon their death, they are not classified as a COVID-19 death. The coefficient of health expenditure for low-income countries suggests that those countries with a stronger health system are more efficient at identifying coronavirus deaths. Undetected COVID-19 deaths are a result of failures in the health system and of poverty, which are more prominent in low-income countries.

#### *Number of deaths 40 days after first five deaths*

We have looked at the number of deaths at a particular point in time but one might question how things change as time passes. As mentioned earlier the infection rate and death rate from infection could be time varying. We can study the number of deaths 40 days after the first five deaths, the regression results are shown in Table 3.

[Insert Table 3 here]

Mobility and Number of arrivals, which influence the initial spread of the infection, are significantly different from zero for both groups. Using the estimated coefficients from Table 3 and the mean values from Table 1, the results show that if mobility increases by 1% the number of deaths would be 1.9% higher in high-income countries and 3.5% higher in low-income countries. The high figure for low-income countries is due to the lower amount of detected deaths at 40 days for this group. The variables associated with the probability of death once infected,  $\delta$ , influence low-income countries, and continue to have the opposite sign: health systems in worse conditions register less COVID-19 deaths, possibly because of deficiencies in the systems and the underestimation of COVID-19 deaths. The coefficient of health expenditure is also positive and significant for high-income countries, suggesting countries with stronger health systems are better at identifying COVID-19 deaths. Interestingly, the life expectancy coefficient is negative and

insignificant in the case of high-income countries. One possibility is that high-income countries take better care of the elderly by the 40-day mark. Additionally, the  $R^2$  for low-income countries falls slightly when compared to the 20-day regression, while in high-income countries  $R^2$  increases by more than 40%.

### *Oaxaca-Blinder Decomposition*

After analyzing the factors that determine the number of deaths, our next step is to examine the determinants of the differences in the number of deaths in low- and high-income countries. We study whether the differences are due to differences in the endowments of the variables in the regressions. We use the Oaxaca-Blinder decomposition to learn whether the differences between low-income and high-income countries are a result of differences in endowments (differences in  $X$ ), or if given an endowment low-income and high-income countries experience a different impact from such endowments. We follow the Oaxaca-Blinder decomposition method (Oaxaca, 1973; Blinder, 1973) and the implementations of the model in Jann (2008) and Hlavec (2014). Referring to the high-income countries' regression with  $H$  and to the low-income countries' regression with  $L$ , we can write the differences in the mean number of deaths between the two groups as:

$$\Delta\bar{Y} = \bar{Y}_H - \bar{Y}_L = \bar{X}_H' \hat{\beta}_H - \bar{X}_L' \hat{\beta}_L$$

We assume that COVID-19 is random and does not distinguish between low-income and high-income countries, yet we observe the coefficients of both groups to be different. Assuming the reference coefficients are those of the high-income group and that the coefficients of both groups are independent, we can write the equation as:

$$\Delta\bar{Y} = (\bar{X}_H - \bar{X}_L)' \hat{\beta}_H + \bar{X}_L' (\hat{\beta}_H - \hat{\beta}_L)$$

Where the first term is the portion of differences in the average number of deaths that can be explained by differences in the endowments and the second portion the one that cannot be explained (as is due to difference in the coefficients and in the error term). The model is discussed and implemented by Jann (2008) in Stata and by Hlavac (2014) in R. We follow the measurement in Jann.

The results are shown in Table 4, for 20 days after the first five deaths in the left panel and 40 days in the right panel.

[Insert Table 4 here]

The decomposition for deaths at 20 days shows differences in the average number of deaths between the two groups can be explained by differences in the number of arrivals and life expectancy<sup>11</sup>. At 40 days, differences in the endowments of number of arrivals and health expenditure explain differences in deaths between the two groups. Mobility is not relevant in this decomposition, and although it influences the number of deaths in our regression analysis, it does not explain differences in the number of deaths between low- and high-income countries. The unexplained portion in the decomposition at 20 days shows life expectancy and health expenditure as significant, while at 40 days the difference lies in life expectancy and number of arrivals.

Overall, results indicate that differences in the number of deaths 20 days after the first five deaths can be explained by differences in the endowments of number of arrivals and life expectancy, and to differences in the impact of life expectancy and health expenditure on deaths. However, 40 days after the first five deaths differences in the number of deaths is strongly linked to differences in the transport infrastructure, measured with number of arrivals, and to a smaller

---

<sup>11</sup> Differences in mean endowments are shown in Table 1.

degree to differences in the endowments of health expenditure. There could also be other factors not studied which could explain more deaths through differences in endowments. For instance, using data from Our World in Data (Roser et al., 2020), the difference in the average number of COVID-19 tests per million in population in high versus low-income countries is of 20,725 at the time of the first five deaths and of 4,244 40-days after<sup>12</sup>. As testing rates improve countries should become more effective at identifying COVID-19 deaths, but we still see testing rates widening between low-income and high-income countries.

Our findings also show a possible underestimation of COVID-19 deaths, especially among low-income countries. Low-income countries with more expenditure on health are better able to detect COVID-19 deaths in the initial stages of the pandemic, at the 20-day mark, compared to low-income countries with a lower percentage of expenditure on health. At the 40-day mark, this effect is more pronounced and we can see it among high-income countries as well. This suggests that countries with less well-developed health systems underestimate COVID-19 deaths, and that this is most important among low-income countries. The decomposition results further suggest that these differences in the strength of the health system can partially explain the differences in the number of deaths between low and high-income countries. Preliminary evidence on this is starting to emerge. The Economist (2020) estimates excess mortality for 11 regions or countries once 50 coronavirus deaths have been reached, out of which 9 are in our high-income and 2 in our low-income sample. Their numbers imply the underestimation of deaths is of 25% in high-income and 71% in low-income countries. For the case of Mexico City, Romero-Zavala and Despeghel

---

<sup>12</sup> Using testing data in our analysis would restrict our sample to 23 low-income and 29 high-income countries.

(2020b) compare observed mortality against expected mortality between April 1 and May 20, and estimate deaths are underestimated by 70%.

## **6. Conclusion**

Important differences are emerging between low-income and high-income economies as they relate to the COVID-19 pandemic. Low-income countries have less reported COVID-19 deaths compared to high-income countries, at 20 and 40 days after their first 5 deaths. However, the factors that affect the number of individuals infected and the spread of the infection are important to deaths in both types of countries. In particular, more community mobility and a stronger transport infrastructure system are associated with a higher number of deaths. We also find higher health expenditure is associated with more deaths in low-income countries. Factors such as the transport infrastructure for the populations' mobility and life expectancy lead to the differences between low- and high-income countries in the number of deaths 20 days after the first five deaths. At 40 days the number of arrivals and hence the transportation infrastructure as well as the country's expenditure on health are able to explain differences in the number of deaths.

There might be other factors that are important when studying differences in the number of Covid-19 deaths between low-income and high-income countries that are not considered here. For instance, it is possible that differences in deaths are related to the social structure of the countries, where individualism could play a major role. Further, geographical factors could also be relevant (Sornette et al. 2020), as the virus is not equally distributed among geographic areas. Temperature and humidity can possibly play a role and lead to higher transmission risk of the virus (Araujo and Naimi, 2020; Sajadi and others, 2020). More research is needed to fully understand all the factors leading to COVID-19 deaths and how they correlate with each other.

Our results suggest that low-income countries could have a more prominent issue with identifying COVID-19 deaths. The strength of the health institutions in the country are of utmost importance in the identification of such deaths. It will be difficult to determine how COVID-19 affected deaths in low-income countries until more data on excess deaths becomes available. A public policy implication is that the measures implemented to contain the spread of the virus which limit the spread of the disease, in particular the social distancing restrictions (mobility) and other restrictions on the movement of people (number of arrivals) are important in preventing deaths. Further, citizens in low-income economies need to be aware of the underestimation of deaths and most likely the underestimation of COVID-19 cases. The underestimation of COVID-19 cases and deaths could be affecting a great number of people in poverty. Further research is needed to understand how the virus is moving into areas of poverty, since it is possible that people in such areas are unaware and unprotected.

## References

- Atkeson, Andrew, "How Deadly Is COVID-19? Understanding The Difficulties With Estimation Of Its Fatality Rate," Working Paper 26965, National Bureau of Economic Research April 2020.
- Araujo, Miguel B., and Babak Naimi. 2020. "Spread of SARS-CoV-2 Coronavirus likely to be constrained by climate," *medRxiv*.
- Avery, Christopher, William Bossert, Adam Clark, Glenn Ellison, and Sara Fisher Ellison, "Policy Implications of Models of the Spread of Coronavirus: Perspectives and Opportunities for Economists," Working Paper 27007, National Bureau of Economic Research April 2020
- Barro, R. J., J. F. Ursua, and J. Weng (2020, April). The coronavirus and the great influenza pandemic: Lessons from the "spanish flu" for the coronavirus's potential effects on mortality and economic activity. Working Paper 26866, National Bureau of Economic Research.
- Blinder, A. S. (1973). Wage Discrimination: Reduced Form and Structural Estimates. *The Journal of Human Resources*, 8(4), 436-455. doi:10.2307/144855
- Burn-Murdoch, J., Romei, V., & Giles, C. (2020, April 26, 2020). Global coronavirus death toll could be 60% higher than reported. *Financial Times. International Edition*. Retrieved from [https://www.ft.com/content/6bd88b7d-3386-4543-b2e9-0d5c6fac846c?utm\\_source=Nature+Briefing&utm\\_campaign=527dbb4637-briefing-dy-20200428&utm\\_medium=email&utm\\_term=0\\_c9dfd39373-527dbb4637-44471613](https://www.ft.com/content/6bd88b7d-3386-4543-b2e9-0d5c6fac846c?utm_source=Nature+Briefing&utm_campaign=527dbb4637-briefing-dy-20200428&utm_medium=email&utm_term=0_c9dfd39373-527dbb4637-44471613)
- Center, M. M., Jemal, A., Lortet-Tieulent, J., Ward, E., Ferlay, J., Brawley, O., & Bray, F. (2012). International variation in prostate cancer incidence and mortality rates. *European urology*, 61(6), 1079-1092.
- Chaudhry, R., Dranitsaris, G., Mubashir, T., Bartoszko, J., & Riazi, S. (2020). A country level analysis measuring the impact of government actions, country preparedness and socioeconomic factors on COVID-19 mortality and related health outcomes. *EClinicalMedicine*, 25, 100464.
- Cheng, V.C.-C.; Wong, S.-C.; Chuang, V.W.-M.; So, S.Y.-C.; Chen, J.H.-K.; Sridhar, S.; To, K.K.-W.; Chan, J.F.-W.; Hung, I.F.-N.; and Ho, P.-L., 2020. The role of community-wide wearing of face mask for control of coronavirus disease 2019 (COVID-19) epidemic due to SARS-CoV-2. *Journal of Infection*, Vol 81 issue 1.
- Chow, Gregory C. (1960). "Tests of Equality Between Sets of Coefficients in Two Linear Regressions". *Econometrica* 28 (3): 591–605.
- CIA. (2020). Country Comparison: Infant Mortality Rate (2017). Retrieved May 14, 2020, from Central Intelligence Agency. U.S. Department of State <https://www.cia.gov/library/publications/the-world-factbook/rankorder/2091rank.html>



- Egger, D., Miguel, E., Warren, S. S., Shenoy, A., Collins, E., Karlan, D., . . . Udry, C. (2021). Falling living standards during the COVID-19 crisis: Quantitative evidence from nine developing countries. *Science Advances*, 7(6), eabe0997.
- Eichenbaum, Martin S, Sergio Rebelo, and Mathias Trabandt, “The Macroeconomics of Epidemics,” Working Paper 26882, National Bureau of Economic Research March 2020.
- Fernández-Villaverde, J., & Jones, C. (2020). *Estimating and Simulating a SIRD Model of COVID-19*. Manuscript, University of Pennsylvania.
- Gobierno de México. (2020). Covid-19 México. Rangos de Edad y Sexo (Histograma de Defunciones Nacional). Retrieved May 30, 2020, from Gobierno de México <https://coronavirus.gob.mx/datos/#DOView>
- Goldberg, P. K., & Reed, T. (2020). The effects of the coronavirus pandemic in emerging market and developing economies An optimistic preliminary account. *Brooking Papers on Economic Activity*, 6, 1.
- Google LLC, 2020. "Google COVID-19 Community Mobility Reports." <https://www.google.com/COVID19/mobility/> Accessed: October 19, 2020.
- Hale, T., Petherick, A., Phillips, T., Webster, S.: Oxford COVID-19 Government Response Tracker, Blavatnik School of Government, Online data set, <https://www.bsg.ox.ac.uk/research/research-projects/coronavirus-government-response-tracker> (2020).
- Hartz, A. J., Krakauer, H., Kuhn, E. M., Young, M., Jacobsen, S. J., Gay, G., . . . Rimm, A. A. (1989). Hospital characteristics and mortality rates. *New England journal of medicine*, 321(25), 1720-1725.
- Hlavac, M. (2014). *oaxaca: Blinder-Oaxaca decomposition in R*. Available at SSRN 2528391.
- ILO. (2018). *Women and men in the informal economy: A statistical picture* (Third ed.). Geneva, Switzerland: International Labor Organization: International Labour Office.
- Jann, B. (2008). The Blinder–Oaxaca decomposition for linear regression models. *The Stata Journal*, 8(4), 453–479.
- JHU CSSE. (2020). *COVID-19 Data Repository by the Center for Systems Science and Engineering (CSSE) at Johns Hopkins University*. Retrieved from: [https://github.com/CSSEGISandData/COVID-19/tree/master/csse\\_COVID\\_19\\_data/csse\\_COVID\\_19\\_time\\_series](https://github.com/CSSEGISandData/COVID-19/tree/master/csse_COVID_19_data/csse_COVID_19_time_series)
- Katz, J., Lu, D., & Sanger-Katz, M. (2020, April 29, 2020). U.S. Coronavirus Death Toll Is Far Higher Than Reported, C.D.C. Data Suggests. *The New York Times*. Retrieved from <https://www.nytimes.com/interactive/2020/04/28/us/coronavirus-death-toll-total.html>
- Kermack, William Ogilvy and A. G. McKendrick, “A contribution to the mathematical theory of epidemics, part I,” Proceedings of the Royal Society of London. Series A, 1927, 115 (772), 700– 721.

- Kermack, William Ogilvy and A. G. McKendrick,, “Contributions to the mathematical theory of epidemics. II – The problem of endemicity,” *Proceedings of the Royal Society of London. Series A*, 1932, 138 (834), 55–83
- Lyu, Wei and Wehby, George, 2020. Community use of face masks and Covid-19: Evidence from a natural experiment of state mandates in the U.S. *Health Affairs*. Vol 39 No. 8.
- Medina, L., & Schneider, F. (2018). *Shadow economies around the world: what did we learn over the last 20 years?* IMF Working Paper. African Department. International Monetary Fund. Retrieved from <http://pinguet.free.fr/fmi1817.pdf>
- NCHS, N. V. S. S. (2020). *Provisional Death Counts for Coronavirus Disease (COVID-19)*. CDC Centers for Disease Control and Prevention Retrieved from [https://www.cdc.gov/nchs/nvss/vsrr/covid\\_weekly/index.htm#AgeAndSex](https://www.cdc.gov/nchs/nvss/vsrr/covid_weekly/index.htm#AgeAndSex).
- Oaxaca, R. (1973). "Male-Female Wage Differentials in Urban Labor Markets". *"International Economic Review"*, 14, "693-709".
- CDC. (2020). Provisional COVID-19 Death Counts by Sex, Age, and State. In p. Deaths involving coronavirus disease 2019 (COVID-19), and influenza reported to NCHS by sex and age group and state. (Ed.): Centers for Disease Control and Prevention.
- Fernández-Villaverde, J., & Jones, C. (2020). *Estimating and Simulating a SIRD Model of COVID-19*. Manuscript, University of Pensilvania.
- Romero-Zavala, M., & Despeghel, L. (2020a, May 25, 2020). ¿Qué nos dicen las actas de defunción de la CDMX? *Nexos. Taller de Datos, XL*.
- Romero-Zavala, M., & Despeghel, L. (2020b). ¿Qué nos dicen las actas de defunción de la CDMX? Actualización al 5 de julio 2020. *Nexos. Taller de Datos*.
- Sajadi, Mohammad M., Parham Habibzadeh, Augustin Vintzileos, Shervin Shokouhi, Fernando Miralles-Wilhelm, and Anthony Amoroso. 2020. “Temperature and latitude analysis to predict potential spread and seasonality for COVID-19,” Available at SSRN 3550308.
- Sornette, D., Mearns, E., Schatz, M., Wu, K., & Darcet, D. (2020). Interpreting, analysing and modelling COVID-19 mortality data. *Nonlinear dynamics*, 101(3), 1751-1776.
- Stock, James H, “Data Gaps and the Policy Response to the Novel Coronavirus,” Working Paper 26902, National Bureau of Economic Research March 2020.
- Stojkoski, V., Utkovski, Z., Jolakoski, P., Tevdovski, D., & Kocarev, L. (2020). The socio-economic determinants of the coronavirus disease (COVID-19) pandemic. *arXiv preprint arXiv:2004.07947*.
- The Economist (Producer). (2020, May 14, 2020). Tracking COVID-19 excess deaths across countries. *COVID-19 data*. Retrieved from [https://www.economist.com/graphic-detail/2020/04/16/tracking-COVID-19-excess-deaths-across-countries?utm\\_campaign=coronavirus-special-edition&utm\\_medium=newsletter&utm\\_source=salesforce-marketing-cloud&utm\\_term=2020-05-02&utm\\_content=cover\\_text\\_url\\_2](https://www.economist.com/graphic-detail/2020/04/16/tracking-COVID-19-excess-deaths-across-countries?utm_campaign=coronavirus-special-edition&utm_medium=newsletter&utm_source=salesforce-marketing-cloud&utm_term=2020-05-02&utm_content=cover_text_url_2)
- The World Bank, 2020. “World Development Indicators.” URL: <http://data.worldbank.org/data-catalog/world-development-indicators>

- Wang, H., Wang, Z., Dong, Y., Chang, R., Xu, C., Yu, X., . . . Huang, J. (2020). Phase-adjusted estimation of the number of coronavirus disease 2019 cases in Wuhan, China. *Cell discovery*, 6(1), 1-8.
- WHO (2020). Report of the WHO-China Joint Mission on Coronavirus Disease 2019 (COVID-2019). World Health Organization.
- Wu, J. T., Leung, K., Bushman, M., Kishore, N., Niehus, R., de Salazar, P. M., Leung, G. M. (2020). Estimating clinical severity of COVID-19 from the transmission dynamics in Wuhan, China. *Nature Medicine*, 1-5.

**Table 1. Differences of Means and Medians between Low-Income and High-Income Countries.**

	Means				Medians			
	Low-Income	High-Income	t-Stat		Low-Income	High-Income	Z	
Dead after 20 days	104.61	390.74	-2.42	**	23.50	82.00	-2.31	**
Dead after 40 days	420.74	2682.52	-2.54	**	53.00	217.00	-2.13	**
GDP per Capita	4667.90	38999.15	-10.06	***	3975.92	34238.81	-8.26	***
Mobility	-51.78	-47.98	-0.73		-47.00	-55.00	-0.69	
Number Arrivals	8015.32	17244.68	-2.57	**	2661.50	9804.00	-3.22	***
Shadow Economy	33.29	16.32	8.97	***	31.58	15.85	6.97	***
Life Expectancy	70.93	79.92	-9.32	***	73.12	81.16	-7.39	***
Death Rate	7.55	8.31	-1.23		6.71	9.00	-1.97	**
% Health Expenditure	5.87	8.16	-4.88	***	5.87	8.51	-4.34	***
Number Obs.	46	46						

\*, \*\*, and \*\*\* indicate statistical significance of the differences at the 10%, 5% and 1% levels, respectively.

**Table 2. Regression Results for Number of Deaths 20 Days after First Five Deaths.**

	Full Sample	Low-Income Countries	High-Income Countries
Constant	-1292.64 (804.12)	-276.27 (386.79)	-4874.59 (2,893.55)
Mobility	5.02** (2.15)	2.13* (1.21)	8.94* (4.73)
Number Arrivals	0.02*** (0.003)	0.01*** (0.003)	0.02*** (0.005)
Shadow Economy	6.19 (4.84)	0.35 (2.32)	23.35* (13.73)
Life Expectancy	17.44* (10.39)	4.14 (5.60)	64.87* (34.83)
Death Rate	-0.84 (15.82)	-14.14 (9.38)	20.52 (27.20)
% Health Expenditure	15.05 (21.88)	35.79** (15.72)	-46.59 (41.53)
R <sup>2</sup>	0.5213	0.4463	0.5901
Adj. R <sup>2</sup>	0.4876	0.3612	0.5271
Num. Obs.	92	46	46

Chow's test for whether coefficients of low-income and high-income countries' regressions are different:  $F(7,78) = 2.44$ ,  $\text{Prob} > F = 0.0260$ . Standard errors in Parenthesis. \*, \*\*, and \*\*\* indicate statistical significance of the coefficients at the 10%, 5% and 1% levels, respectively.

**Table 3. Regression Results for Number of Deaths 40 Days after First Five Deaths.**

	<b>Full Sample</b>	<b>Low-Income Countries</b>	<b>High-Income Countries</b>
Constant	3067.40 (4,516.71)	-2051.76 (1,936.93)	22302.00 (15,805.00)
Mobility	24.22** (12.08)	14.69** (6.05)	52.30** (25.86)
Number Arrivals	0.17*** (0.018)	0.04*** (0.013)	0.19*** (0.027)
Shadow Economy	11.50 (27.20)	4.34 (11.63)	33.27 (74.99)
Life Expectancy	-73.32 (58.35)	28.29 (28.04)	-289.59 (190.26)
Death Rate	-155.97* (88.87)	-68.07 (46.99)	-175.24 (148.55)
% Health Expenditure	573.34*** (122.89)	216.78*** (78.70)	452.27** (226.83)
R <sup>2</sup>	0.7349	0.406	0.7916
Adj. R <sup>2</sup>	0.7162	0.3146	0.7596
Num. Obs.	92	46	46

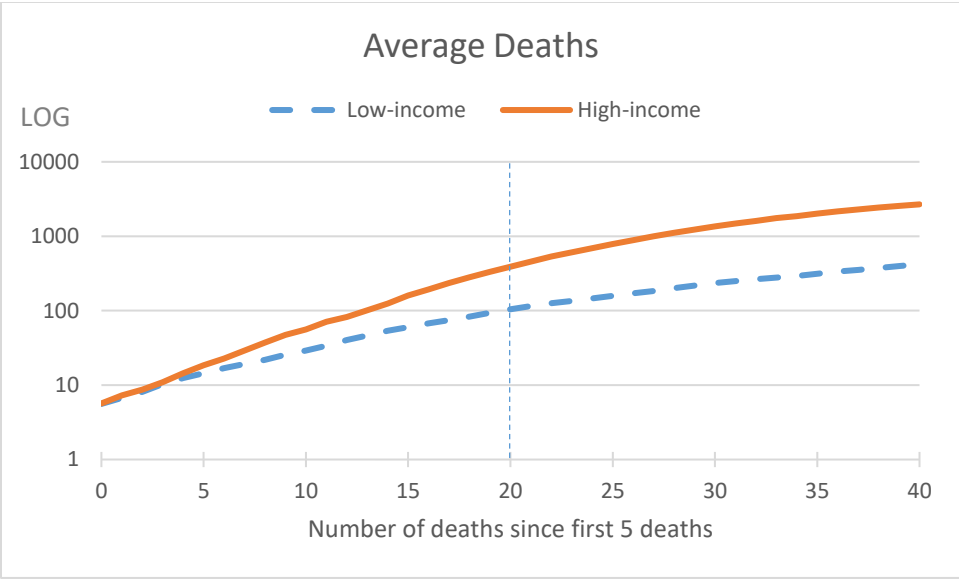
Chow's test for whether coefficients of low-income and high-income countries' regressions are different:  $F(7,78)=3.40$ .  $\text{Prob}>F = 0.0032$ . Standard errors in Parenthesis. \*, \*\*, and \*\*\* indicate statistical significance of the coefficients at the 10%, 5% and 1% levels, respectively

**Table 4. Oaxaca-Blinder Decomposition for the Number of Deaths 20 Days after First Five Deaths**

	20 days		40 days	
	Coef.	Std. Err.	Coef.	Std. Err.
<b>Differential</b>				
Prediction High-GDP	390.74***	118.41	2682.52***	893.74
Prediction Low-GDP	104.61***	29.83	420.74***	144.65
Difference	286.13**	122.11	2,261.78**	905.37
<b>Explained</b>				
Mobility	34.00	50.10	198.95	290.71
Number Arrivals	181.05**	84.25	1,743.40**	722.99
Shadow Economy	-376.74	231.78	-564.64	1,274.30
Life Expectancy	583.34*	319.44	-2,604.21	1,733.64
Death Rate	15.70	24.43	-134.08	157.71
% Health Expenditure	-106.86	97.74	1,037.37*	562.04
Total	310.98	326.05	-323.20	1,882.71
<b>Unexplained</b>				
Mobility	-352.49	254.24	-1,947.52	1,382.14
Number Arrivals	74.89	48.05	1,190.28***	344.31
Shadow Economy	765.42	464.92	963.13	2,526.63
Life Expectancy	4,307.56*	2503.06	-22,547.83*	13,644.10
Death Rate	261.45	217.51	-808.64	1,176.37
% Health Expenditure	-483.36*	261.39	1,217.86	1517.13
Constant	-4,598.315	2919.28	24,353.74	15,922.91
Total	-24.85	322.5002	2,584.99	1,761.88

\*, \*\*, and \*\*\* indicate statistical significance of the coefficients at the 10%, 5% and 1% levels, respectively.

**Figure 1. Log Average Number of Deaths for Low-Income vs High-Income Countries, Relative to the Day When the Countries Reported Their First Five Deaths.**





**Figure 2. Average Mobility to Transit Stations for Low and High-Income Countries**

