Life Expectancy During the Covid-19 Pandemic: A Semi-Parametric Difference-in-Differences Analysis

Polemis, Michael and Stengos, Thanasis

University of Piraeus, University of Guelph

22 September 2020

Life Expectancy During the Covid-19 Pandemic:

A Semi-Parametric Difference-in-Differences Analysis

Michael L. Polemis[a,b*], Thanasis Stengos[c]

[a] *University of Piraeus, Department of Economics, Piraeus, Greece (corresponding author)*

[b] *Hellenic Competition Commission, Athens, Greece.*

[c] *University of Guelph, Department of Economics, Ontario, Canada*

September 2020

Abstract

The scope of this note is to investigate the causal effects of the COVID-19 pandemic on life expectancy over a sample of 47 countries split into two groups. The first one includes countries that have adopted general lockdown measures (treatment group), while the second one consists of countries that have imposed social distancing measures other than a national lockdown (control group). The investigated period starts from the first confirmed European case back (25.01.2020) until 28.07.2020 and covers the first wave of the pandemic for the sample countries. The empirical results based on a Semi-Parametric Difference-in-Differences framework, suggest a decline in life expectancy at birth estimated to 1.38 years on average even though the countries who did implement the lockdown measures were motivated and willing to do so. However, the decrease in life expectancy would have been double (2.9 years) in the absence of such willingness to adopt the policy. Lastly, the findings support the argument that national lockdown would be an effective policy tool to the hands of regulators and health practitioners to mitigate the negative effects of the pandemic infection it is pursued by motivated and willing participant countries.

Keywords: Covid-19; Life expectancy; Lockdown; Social distancing; Semi-parametric Difference-in-Differences

1. Introduction

The COVID-19 pandemic has severely stressed health systems all over the world in an unforeseen way, potentially leading to increases in morbidity and mortality rate (see for example Connor, et al, 2020; Vandoros 2020; Trias-Llimós and Bilal 2020). These increases, however, seem to negatively affect life expectancy triggering significant consequences in all the aspects of social life (Marois et al, 2020). To give but an example of the rapid penetration of the pandemic to humanity, it is estimated that over 30,5 million Covid-19 cases have been reported globally, leading to nearly 952,000 deaths as of 19 September 2020 (Johns Hopkins University, 2020).

It is noteworthy that the impact of this unprecedented situation on public health would have been even worse by the absence of restrictive measures to curb the spread of the pandemic (e.g. lockdowns, physical distancing, compulsory masking, flights cancellation, etc). This paper contributes to the existing literature since it is the first study so far that examines the impact of the national lockdown adopted by 33 countries on life expectancy at birth. In other words, this study investigates the links of life expectancy, which is a long-term issue with the effect of a short-run policy of a national lockdown.

For this reason, the flexible Semiparametric Difference-in-Differences (SDID) approach developed in Abadie (2005) is employed to estimate how the effect of treatment (national lockdown) varies with changes in individual characteristics. The reason for using this method over the traditional linear DID estimator is that the former is mostly suited for longitudinal surveys with a baseline and follow-up rounds (e.g. lockdown periods) as in this case (see Houngbedji, 2016 and Abadie 2005).

Despite the profound interest of the scientific community in studying the consequences of COVID-19 on public health (see among others Connor et al, 2020; Elgar et al, 2020; Vandoros, 2020; Nivette et al, 2020; Preis et al, 2020), little attention has been paid on the casual effects of the restrictive measures adopted to mitigate the spread of the pandemic infection on life expectancy at birth.

The most related work to ours is the recent study of Llimós and Bilal (2020), where they examine the impact of COVID-19 on the annual life expectancy in the region of Madrid. This study estimates the expected changes in annual life expectancy accounting for the excess mortality over the period from 9.3.2020 to 10.05.2020. The empirical findings support a decline in life expectancy at birth of 1.9 and 1.6 years for men and women, respectively. The authors argue that the relevant figures could be even bigger since their estimates are based on the existing mortality pattern. Moreover, they argue that a decline mortality rate is also possible due to the existence of a "*harvesting*" effect, as happened in other flu pandemics.

In a subsequent study, Marois et al (2020) provide first estimates of the potential direct impact of the COVID-19 pandemic on life expectancy for several countries and regions (North America and Europe, Latin America, and the Caribbean, Southeastern Asia, and Sub-Saharan Africa). They employ a microsimulation model based on several assumptions (e.g. mortality rate, prevalence rate, fertility, etc) that calibrates the life histories of 100,000 individuals by five-year age groups over a period of 52 weeks (one year). The empirical findings support that in all prevalence scenarios, life expectancy would drop even by 9 years (North America and Europe). However, life expectancy will not be substantially affected when the prevalence COVID-19 rate does not exceed 2%.

2. Data and Methodology

The sample consists of an unbalanced panel data set comprising of 47 countries, over the period 25.01.2020 to 28.07.2020 (T = 186). All the variables were obtained from Roser et al (2020) and extracted on 30 July 2020.[1] We must mention though that values included in this dataset may be changed in later releases, as is sometimes the case. The starting date for the national lockdown measure adopted by each sample country is shown in Table 1. As it is evident, the majority of the (treatment) countries implemented this restrictive measure shortly after the declaration of the pandemic by the World Health Organization (11.03.2020).[2]

<Table 1 about here>

Table 2 presents the summary statistics for the sample variables used in this study. It appears that the diabetes prevalence exhibits the lowest standard deviation among the sample variables equal to 1.997, while the GDP per capita the highest. The relevant variable is positively skewed (0.863) and the (excess) kurtosis value suggests a platykurtic distribution (>3). On the contrary, the dependent variable (life expectancy) is negatively skewed (-1.209) and does not follow the normal distribution, which is also the case for all the rest variables. It is noteworthy that the rest variables are positively skewed (heavy-tailed) revealing a leptokurtic distribution.

<Table 2 about here>

One of the assumptions in the interpretation of using difference-in-differences estimators to measure the effect of policy intervention as a causal effect is that, in the absence of the treatment, the outcome variable follows the same trend in treated and untreated groups (Vandoros, 2020; Card and Kruger, 1994). This assumption may not be credible if selection for treatment is correlated with characteristics that affect the dynamic of the outcome variable (Imbens and Wooldridge, 2009; Hirano et al, 2003).

[1] Retrieved from: https://ourworldindata.org/coronavirus.

[2] However, Slovakia is the only country that adopted national lockdown measures nearly one month after the declaration of the pandemic (08.04.2020).

Abadie (2005) introduces the SDID estimator addressing the imbalance of characteristics between treated and untreated groups through a reweighting scheme that allows differences in observed characteristics to create nonparallel outcome dynamics for the treated and untreated groups. Since the characteristics are treated nonparametrically any estimation error caused by functional misspecification is avoided (Chang, 2018). Furthermore, the SDID estimator allows for the use of covariates to describe how the average effect of the treatment varies for different groups of the treated population. The estimator is based on reweighing the trend for the untreated participants using a semiparametric estimator of their propensity score using a polynomial series of the covariates (Abadie 2005).

Following the SDID methodology, the effect of national lockdown can be disentangled by identifying two groups. The first group includes 34 countries (treatment group), that pursued the measure of the national lockdown. The second group (control group), comprises of the rest 13 countries (Belarus, Estonia, Iceland, Malta, Sweden, Switzerland, Japan, Malawi, Nicaragua, South Korea, Taiwan, Timor, and Uruguay) that did not impose a general (national) lockdown but only transitory protective measures (e.g. school closures, social distancing, COVID testing, health checks, etc).[3]

3. Results and discussion

We begin our analysis by estimating with OLS the following linear benchmark model

$$Life\ expectancy_{it} = b_o + b_1 Treated_i + b_2 GDP_{it} + b_2 Population_{it} + b_3 Median\ age_{it} + b_4 Male\ smokers_{it} + b_5 Female\ smokers_{it} + b_6 Female\ smokers + b_7 Poverty + b_8 Hospital + b_9 Diabetes + \varepsilon_{it} \quad (1)$$

[3] For the empirical application of the SDID estimator, we used the "*absdid*" command in Stata, developed in Houngbedji (2016). The relevant command estimates the average treatment of the treated (ATT) by comparing change over time of the outcome of interest across treatment groups while adjusting for differences between treatment groups on the observable characteristics at baseline that are correlated to the propensity score (Houngbedji, 2016).

The dependent variable in Eq. 1 is the life expectancy at birth in country *i* at day *t* and remains constant throughout the sample period. $Treated_i$ is a dummy variable equal to one for the countries adopted the national lockdown and zero otherwise. In other words, the policy intervention is simply a dummy variable in this benchmark linear OLS model without any Difference-in-Difference correction. GDP denotes the per capita gross domestic product at purchasing power parity in constant 2011 USD dollars. Population denotes the number of people divided by land area (density), measured in square kilometers. Median age denotes the median age of the population. Male and female smokers denote the share of males and females who smoke. Poverty denotes the share of the population living in extreme poverty. Hospital denotes the hospital beds available per 1,000 people and can be used as a proxy of the health infrastructure. Diabetes denotes diabetes prevalence as a percentage of the population aged 20 to 79 in 2017. Finally, ε_{it} is the error term.

<Table 3 about here>

The empirical findings of the benchmark model are presented in Table 3. As it is evident, the COVID-19 pandemic reduces life expectancy by about 2.9 years (95% *CI* from -3.107 to -2.776) since the relevant dummy variable (*Treated*) is negative and statistically significant. This means that the absence of willingness and motivation to implement the policy intervention measure of the national lockdown (counterfactual), the pandemic infection would have an inversely related effect on life expectancy. Most of the variables have the anticipated signs and are statistically significant at a 1% level of significance.

Having estimated the benchmark model, the analysis focuses on the interpretation of the empirical findings of the SDID estimator. Table 4 presents the empirical results of the SDID econometric analysis. Specifically, Model (1) reports estimates of the average life expectancy at birth. Model (2) shows how the life expectancy varies with economic (GDP per capita, extreme poverty), demographic (e.g. population, hospital beds), and medical characteristics (e.g. male and female smokers, diabetes prevalence). The coefficients reported in (1) and (2) are estimated using a logit specification of degree 4 to estimate the propensity score (SLE).[4]

<Table 4 about here>

As it is evident, from the relevant table, life expectancy is estimated to be decreased by 1.38 years on average in all the models. A careful inspection of the relevant table reveals some important findings. First, if we isolate all the additional covariates affecting life expectancy at birth (e.g. income effect, health conditions, population density, etc), we notice that there is a decrease in the post-treatment period compared to the control group (see column 1). The relevant SDID coefficient is negative (-1.158) and statistically significant, while the 95% confidence interval ranges from -1.588 to -0.727 during the first wave of the pandemic. This finding indicates that the policy intervention change (national lockdown) adopted by the 33 sample countries reduces annual life expectancy at birth by 1.9 years since these countries were willing, motivated by health safety considerations and ready to do so. The relevant estimate is in alignment with the study of Llimós and Bilal (2020).

Second, similar findings are reported when several independent variables enter the model (see column 2). In this case, the SDID estimator remains negative and statistically significant, though its magnitude is slightly larger than the previous estimate (1.606 compared to 1.158). Specifically, annual life expectancy at birth will be reduced by 1.606 years, 95% *CI* (-2.78, 5.99), while the rest of the variables have a positive and statistically significant correlation with life

[4] We have also used a linear polynomial function (LPM) of degree 4 to approximate the propensity score and the results were quite similar. To preserve space, the empirical findings are available from the authors on request.

expectancy as it is documented by previous studies (see among others Miladinov, 2020; Shafi and Fatima, 2019).

All in all, this study argues that the national lockdown measure coincides with reductions in life expectancy. However, the magnitude of the reduction rate would have been larger in the case of an absence of willingness, motivation and readiness of the participant countries to implement this intervention. In that case, the estimated reduction in life expectancy would have been by almost three years from the current 2019 level. The relevant finding justifies the effectiveness of the national lockdown as a restrictive policy tool to combat the novel coronavirus disease. Combining the two important findings, this study supports the notion that short-run health policies (e.g. national lockdown) incur long term implications (e.g. reduction in life expectancy growth). A similar result is also documented in the study of Verstraeten, et al (2016), where it is argued that a short-run policy change (decolonization) incurs negative effects at life expectancy growth in the Caribbean states. A subsequent study by Montez et al, (2020) for the US seems also to support this claim.[5] In this case, the link between life expectancy which constitutes a long-term issue with the short-run policy of a national lockdown can be explained through the channel that people's health may be compromised at birth. This could be partly justified by the fact that health resources have moved to take care of the reported COVID-19 infection cases, or early premature births.

[5] This study supports that shifts in state policy measures (e.g. liberal or conservative) predict life expectancy and US longevity.

4. Conclusions

This study employs a flexible SDID approach, to investigate the effect of a short-run health policy (e.g. national lockdown) on a long-run health issue (life expectancy). The econometric analysis suggests a structural link between the adoption of lockdown measures by the national governments to face the pandemic crisis and life expectancy at birth. The results indicate that the most stringent policy intervention such as a national lockdown incurs a decrease in life expectancy by 1.38 years on average. The negative relationship remains robust under several SDID specifications. In the absence of a willingness to adopt such a policy by the countries involved would have resulted in life expectancy at birth to be reduced by a larger amount equal to 2.9 years (e.g. more than the double). Lastly, this study has significant implications for policymakers and health practitioners, since a decrease in life expectancy due to the pandemic could imply a higher death toll as a result of the imminent second wave of the pandemic, pushing the national healthcare systems to their limits. Therefore, future research may shed light on this issue by extending the sample and including more covariates (spatial characteristics, demographic conditions, education level, etc).

Credit author statement

Michael Polemis, Conceptualization, Methodology, Formal Analysis, Investigation, Data Curation, Writing, Visualization.

Thanasis Stengos, Methodology, Writing, Supervision, Formal Analysis, Investigation, Data Curation, Writing

Disclosure

The authors report no conflicts of interest.

Table 1: National lockdown coverage: Dates and Countries

Date	AL	AT	BE	BG	HR	CY	CZ	DK	FI	FR	DE	GR	HU	IR	IT	LV	LI	LT	LU	MD	MC	ME
09.03.2020																						
10.03.2020																						
11.03.2020 Pandemic																						
12.03.2020																						
13.03.2020																						
14.03.2020																						
15.03.2020																						
16.03.2020																						
17.03.2020																						
18.03.2020																						
19.03.2020																						
20.03.2020																						
21.03.2020																						
22.03.2020																						
23.03.2020																						
24.03.2020																						
25.03.2020																						
26.03.2020																						
27.03.2020																						
28.03.2020																						
29.03.2020																						
30.03.2020																						
31.03.2020																						

Date	NL	MK	NO	PL	PT	RO	RS	SK	SI	ES	UA	UK
09.03.2020												
10.03.2020												
11.03.2020 Pandemic												
12.03.2020												
13.03.2020												
14.03.2020												
15.03.2020												
16.03.2020												
17.03.2020												
18.03.2020												
19.03.2020												
20.03.2020												
21.03.2020												
22.03.2020												
23.03.2020												
24.03.2020												
25.03.2020												
26.03.2020												
27.03.2020												
28.03.2020												
29.03.2020												
30.03.2020												
31.03.2020												
01.04.2020												
02.04.2020												
03.04.2020												
04.04.2020												
05.04.2020												
06.04.2020												
07.04.2020												
08.04.2020												
09.04.2020												

Notes: Light shaded areas denote the pre-lockdown period, while dark shaded areas represent the lockdown period. AL = Albania, AT = Austria, BE = Belgium, BG = Bulgaria, HR = Croatia, CY = Cyprus, CZ = Czech Republic, DK = Denmark, FI = Finland, FR = France, DE = Germany, GR = Greece, HU = Hungary, IR = Ireland, IT = Italy, LV = Latvia, LI = Liechtenstein, LT = Lithuania, LU = Luxembourg, MD = Moldova, MC = Monaco, ME = Montenegro, NL = Netherlands, MK = Northern Macedonia, NO = Norway, PL = Poland, PT = Portugal, RO = Romania, RS = Serbia, SK = Slovakia, SI = Slovenia, ES = Spain, UA = Ukraine, UK = United Kingdom.
Source: Authors' elaboration from Wikipedia and various national websites.

Table 2: Summary statistics

Variable	Observations	Mean	Median	Min	Max	Standard deviation	Skewness	Kurtosis
GDP	7,372	33,504	32,606	1,095	94,278	17,875	0.863	4.672
Population	7,686	602.5	106.7	3.404	19,348	2,862	6.356	41.65
Median age	7,558	41.21	42.30	18	48.20	5.415	-2.701	11.90
Male smokers	7,063	34.52	33.10	15.20	78.10	11.39	0.987	5.107
Female smokers	7,063	21.57	21.30	4.400	44	8.423	0.034	3.056
Poverty	5,300	3.259	0.500	0.100	71.40	11.27	5.206	30.33
Hospital	7,686	5.355	4.530	0.900	13.80	2.918	1.255	4.262
Diabetes	7,686	6.414	5.760	3.280	11.47	1.997	0.701	2.581
Life expectancy	7,872	79.68	81.32	64.26	86.75	4.047	-1.209	4.887

Notes: GDP denotes the per capita gross domestic product at purchasing power parity in constant 2011 USD dollars, Population denotes the number of people divided by land area (density), measured in square kilometers, Median age denotes the median age of the population, Male smokers denotes the share of males who smoke, Female smokers represents the percentage of female smokers, Poverty denotes the share of the population living in extreme poverty, Hospital, denotes the hospital beds available per 1,000 people, Diabetes, denotes the diabetes prevalence as a percentage of the population aged 20 to 79 in 2017 and finally Life expectancy denotes the life expectancy at birth in 2019.

Table 3: Linear regression results (benchmark model)

Dependent variable: *Life expectancy*	**OLS**
Treated	-2.942***
	(0.0845)
GDP	0.000131***
	(2.62e-06)
Population	0.000841***
	(0.000109)
Median age	0.460***
	(0.00892)
Male smokers	0.0442***
	(0.00368)
Female smokers	-0.0298***
	(0.00286)
Poverty	-0.0468***
	(0.00315)
Hospital	-0.557***
	(0.0163)
Diabetes	0.250***
	(0.0139)
Constant	58.41***
	(0.438)
Observations	4,991
R-squared	0.835

Notes: Treated is a dummy variable taking value one for countries that imposed national lockdown and zero otherwise, GDP denotes the per capita gross domestic product at purchasing power parity in constant 2011 USD dollars, Population denotes the number of people divided by land area (density), measured in square kilometers, Median age denotes the median age of the population, Male smokers denotes the share of males who smoke, Female smokers represents the percentage of female smokers, Poverty denotes the share of the population living in extreme poverty, Hospital, denotes the hospital beds available per 1,000 people, Diabetes, denotes the diabetes prevalence as a percentage of the population aged 20 to 79 in 2017, Life expectancy denotes the life expectancy at birth in 2019. Robust standard errors in parentheses. Significant at *** $p<0.01$.

Table 4: SDID regression results (control & treatment sample)

	SLE	
Dependent variable: *Life expectancy*	**(1)**	**(2)**
SDID estimator		
Constant	-1.158***	-1.606***
	[-1.588 to 0.727]	[-2.78 to 5.99]
Covariates		
GDP	-	0.0000485**
		[9.74e-06 to 0.00008]
Population	-	0.0117***
		[0.008 to 0.014]
Male smokers	-	-0.103
		[-0.250 to 0.043]
Female smokers	-	-0.360***
		[-0.472 to -0.247]
Poverty	-	-0.492***
		[-0.756 to -0.227]
Hospital	-	-0.296
		[-0.795 to 0.204]
Diabetes	-	-1.134***
		[-1.684 to -0.584]
Observations	4,991	4,991

Notes: GDP denotes the per capita gross domestic product at purchasing power parity in constant 2011 USD dollars, Population denotes the number of people divided by land area (density), measured in square kilometers, Male smokers denotes the share of males who smoke, Female smokers represents the percentage of female smokers, Poverty denotes the share of the population living in extreme poverty, Hospital, denotes the hospital beds available per 1,000 people, Diabetes, denotes the diabetes prevalence as a percentage of the population aged 20 to 79 in 2017, Life expectancy denotes the life expectancy at birth in 2019. Time and country dummies are included but not reported. Robust confidence interval in brackets. Significant at *** p<0.01.

References

Abadie, A. (2005). Semiparametric Difference-in-Differences Estimators, *Review of Economic Studies*, 72(1): 1-19.

Card, D., and Krueger, A. (1994). Minimum wages and employment: a case study of the fast-food industry in new jersey and Pennsylvania. *American Economic Review*, 84 (4): 772–793.

Chang, N-C. (2018). Semiparametric Difference-in-Differences with Potentially Many Control Variables, Papers 1812.10846, arXiv.org, revised Jan 2019.

Connor, J., Madhavan, S., Mokashi, M., Amanuel, H., Johnson, N., Pace, L., and Bartz, D. (2020). Health risks and outcomes that disproportionately affect women during the Covid-19 pandemic: A review, *Social Science & Medicine*, Volume 266, https://doi.org/10.1016/j.socscimed.2020.113364.

Elgar, F., Stefaniak, A., and Wohl, M. (2020). The trouble with trust: Time-series analysis of social capital, income inequality, and COVID-19 deaths in 84 countries, *Social Science & Medicine*, https://doi.org/10.1016/j.socscimed.2020.113365.

Hirano, K., Imbens, G. W. and Ridder, G. (2003). Efficient estimation of average treatment effects using the estimated propensity score. *Econometrica* 71: 1161‑1189.

Houngbedji, K. (2016). Abadie's semiparametric difference-in-differences estimator, *Stata Journal*, 16(2): 482-490.

Imbens, G.W. and Wooldridge, J. M. (2009). Recent developments in the econometrics of program evaluation. *Journal of Economic Literature*, 47: 5‑86.

Johns Hopkins University, (2020). Covid-19 Dashboard by the Center for Systems Science and Engineering. https://coronavirus.jhu.edu/map.html, Accessed date: 19 September 2020.

Marois G, Muttarak R, Scherbov S. (2020). Assessing the potential impact of COVID-19 on life expectancy. *PLoS ONE* 15(9): e0238678. https://doi.org/10.1371/journal.pone.0238678.

Miladinov, G. (2020). Socioeconomic development and life expectancy relationship: evidence from the EU accession candidate countries. *Genus* 76, 2 (2020). https://doi.org/10.1186/s41118-019-0071-0.

Montez, J.K., Beckfield, J., Cooney, J.K., Grumbach, J.M., Hayward, M.D., Koytak, H.Z., Woolf, S.H. and Zajacova, A. (2020). US State Policies, Politics, and Life Expectancy. *The Milbank Quarterly*, 98: 668-699. doi:10.1111/1468-0009.12469

Nivette, A., Ribeaud, D., Murray, A., Steinhoff, A., Bechtiger, L., Hepp, U., Shanahan, L., Eisner, M. (2020). Non-compliance with COVID-19-related public health measures among young adults in Switzerland: Insights from a longitudinal cohort study, *Social Science & Medicine*, https://doi.org/10.1016/j.socscimed.2020.113370.

Preis, H., Mahaffey, B., Heiselman, C., Lobel, M. (2020). Vulnerability and resilience to pandemic-related stress among U.S. women pregnant at the start of the COVID-19 pandemic, *Social Science & Medicine*, 266, https://doi.org/10.1016/j.socscimed.2020.113348.

Roser, M., Ritchie, H., Ortiz-Ospina, E., and Hasell, J. (2020). Coronavirus Pandemic (COVID-19). Published online at OurWorldInData.org.

Shafi, R., and Fatima, S. (2019). Relationship between GDP, Life Expectancy, and Growth Rate of G7 Countries, *International Journal of Sciences*, 8(06): 74-79.

Trias-Llimós, S., and Bilal, U. (2020). Impact of the COVID-19 pandemic on life expectancy in Madrid (Spain), *Journal of Public Health*, 42(3): 635–636.

Vandoros, S. (2020). Excess mortality during the Covid-19 pandemic: Early evidence from England and Wales, *Social Science & Medicine*, 258, https://doi.org/10.1016/j.socscimed.2020.113101

Verstraeten, S., Oers, H., Mackenbach, J. (2016). Decolonization and life expectancy in the Caribbean, *Social Science & Medicine*, 170: 87-96.