Phase3

Karla Otoude and Chloé Marquis

Link to our github project

Our Github Project

I. Research question

How are changes in health and social protection expenditures associated with avoidable mortality across OECD countries over time?

One key factor explaining differences in population health across countries is the strength of social and health protection systems. In particular, investment in healthcare and social protection is expected to reduce mortality from causes that are considered treatable through effective public policies.

This report seeks to understand the relationship between social and health protection spending and avoidable mortality, used as an indicator of the effectiveness of a country's health system. Avoidable mortality captures premature deaths that could be prevented through healthcare, making it a relevant proxy for population health.

Using panel data for 41 OECD countries between 2000 and 2023, this report aims to analyse the global evolution of avoidable mortality across OECD countries and assess if high levels of social and health spending are correlated with lower avoidable mortality rates. A second part focuses on the Covid-19 crisis, to understand if countries with a stronger health protection pre-Covid experienced lower avoidable mortality rates from 2020 to 2021.

Method

To conduct our study we will be using a **linear mixed model (LMM)**. Lagged variables will be used as investments in both the social and health systems likely impact the population's general health with a delay.

For explanatory variables, in order to estimate the strength and quality of a country's health protection scheme, we will use the following variables:

- Health expenditure
- Social expenditure aggregates
- Number of hospitals beds

We also have more general variables such as :

- GDP per capita
- Unemployment rate
- Human development index
- Life expectancy

They will be used as control variables, since richer countries tend to have better living conditions leading to a healthier population.

The explained variable is :

- Avoidable mortality

II. Description of the datasets

This report relies on 8 datasets. More details on the datasets and their descriptive statistics can be found in Phase 2.

Health expenditure - OECD (2025)

This dataset details the annual health expenditure and financing as a percentage of GDP for 52 countries (2000-2023). It includes consumption of healthcare goods and services from different health providers.

Avoidable mortality - OECD (2025)

This dataset describes the number of "avoidable deaths" per 100,000 inhabitants annually for the period 2000-2023 for 46 countries. It contains 2 indicators referring to premature mortality (under age 75) : "preventable mortality" and "treatable mortality".

GDP per Capita - World Bank (2024)

The GDP per capita dataset presents the GDP per capita (constant 2015 US$) from 1960 to 2023 for 70 countries.

Hospital beds - OECD (2025)

This dataset provides data on the number of total hospitals beds per 1,000 inhabitants by function of healthcare and by type of care (ie. somatic or psychiatric care) for the period 2000-2023 for 48 different countries.

Life expectancy - World Bank (2023)

This dataset provides the life expectancy in years for women and men for a period ranging from 1960 to 2023, for 69 different countries or economies.

Social Expenditure Aggregate - OECD (2025)

This dataset includes statistics on public and private social expenditure at program level (e.g. Old age, Survivors, etc) as a percentage of GDP. It covers 38 OECD countries for the period 1980-2023.

Unemployement Rate - IMF (2025)

This dataset describes the percentage of the labor force that is unemployed and actively seeking employment. It is available from 1980 to 2025 and features 120 geographic regions.

HDI (Human Development Index) - UNDP (2025)

This dataset describes the human development index, per country (195), annually (1990-2023). It is a summary measure of average achievement in 3 dimensions of human development: a long and healthy life, being knowledgeable and having a decent standard of living.

Table 1: Brief summary of the datasets used in this report

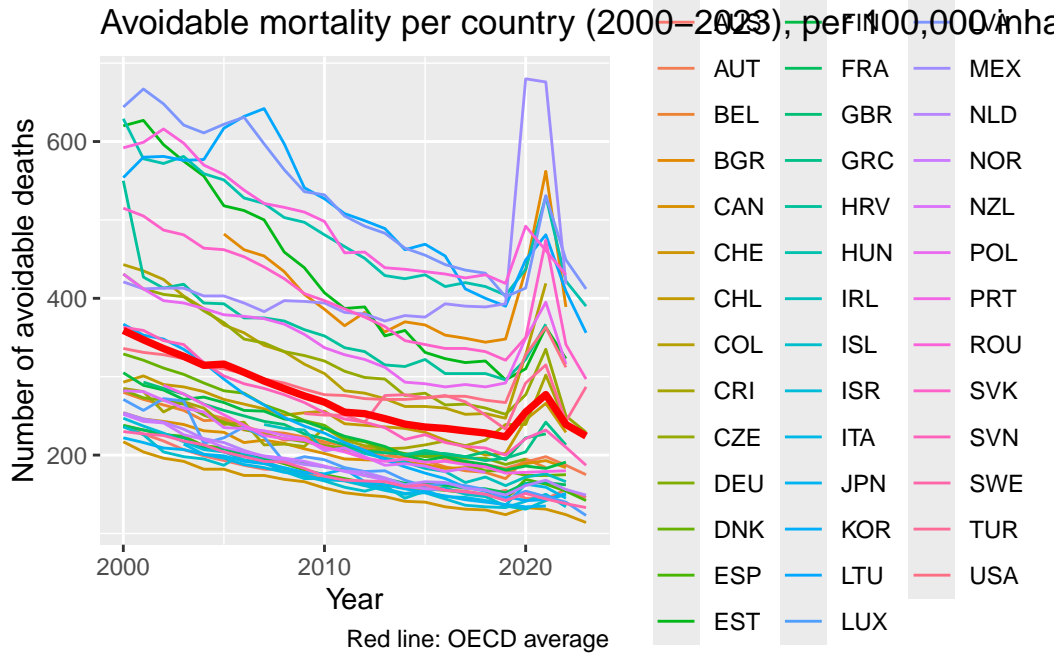| Dataset	Observations	Variables
Health Expenditure	1246	46
Avoidable Mortality	997	44
GDP per Capita	266	70
Number of Hospital Beds	1034	38
Life Expectancy	266	70
Social Expenditure	1046	34
Unemployement	122	32
HDI	4484	10

Description of the merged and cleaned dataset

Number of rows	Number of columns	Percentage of NA	Number of countries studied	Time period
888	10	1.05%	41	2000-2023

3. Data Analysis

We first start by a visual exploration of our main variable : avoidable mortality (*OBS_VALUE_mortality*) and its link to explanatory variables.
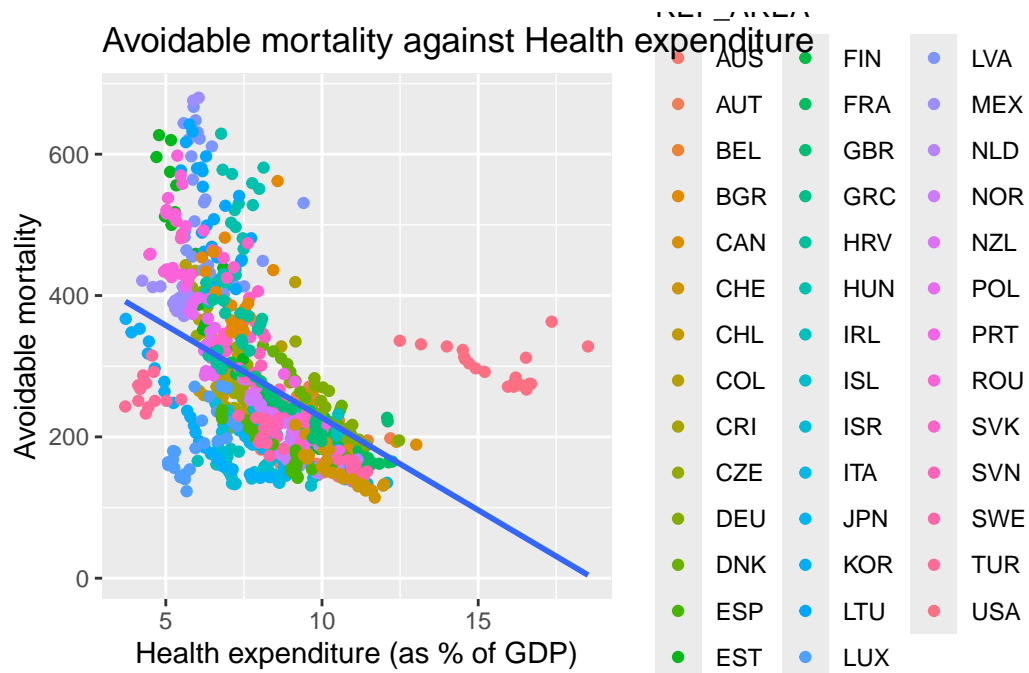
Evolution of avoidable mortality per country

Red line: OECD average

Most countries share the same downward sloping trend in avoidable mortality. A distinct increase appears in 2019, with a peak in 2020 and a decrease to previous trend values from 2021 in all countries. The country with the lowest avoidable death rate is Switzerland (CHE) and the highest is Latvia at the beginning of the 2000s, then surpassed by Lithuania (LTU) around 2007. Additionally, there is a strong increase in avoidable deaths from around 400 to almost 650 per 100,000 inhabitants in Mexico in 2019.

Visualising *avoidable_mortality* against independent variables

We can start the data analysis by simply plotting some independent variables against *avoidable_mortality* and observe the general trend.
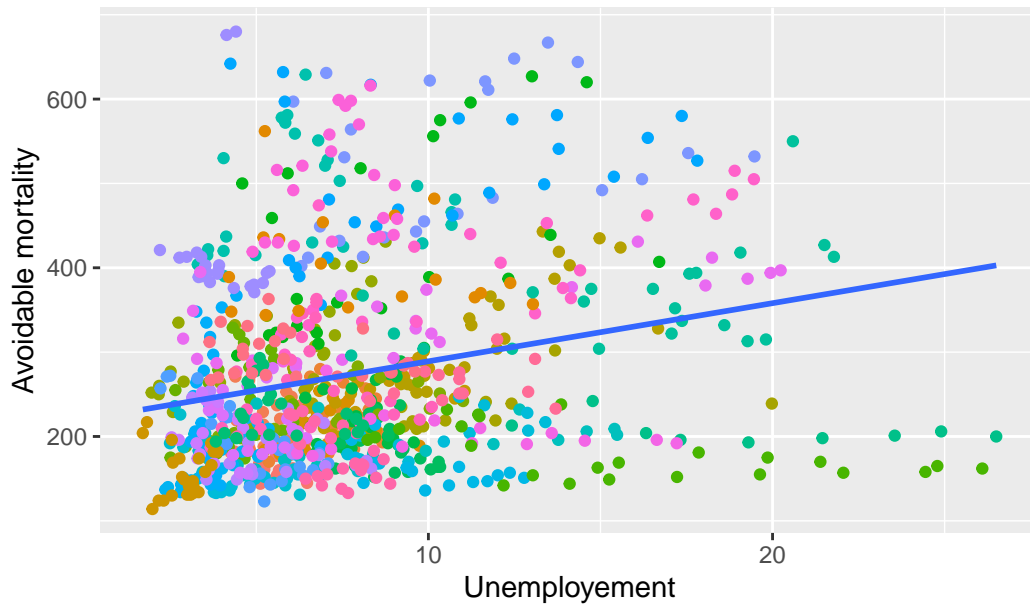

As expected, as the level of health expenditure increases, avoidable mortality tends to decrease. Australia (AUS) detaches itself from other countries by having a particularly high level of health expenditure, but still a moderately high level of avoidable mortality. Most countries are situated on the left side of the graph (with low health expenditure) showcasing countries with high and low rates of avoidable mortality for the same level of health expenditure.

Avoidable mortality against Social Expenditure

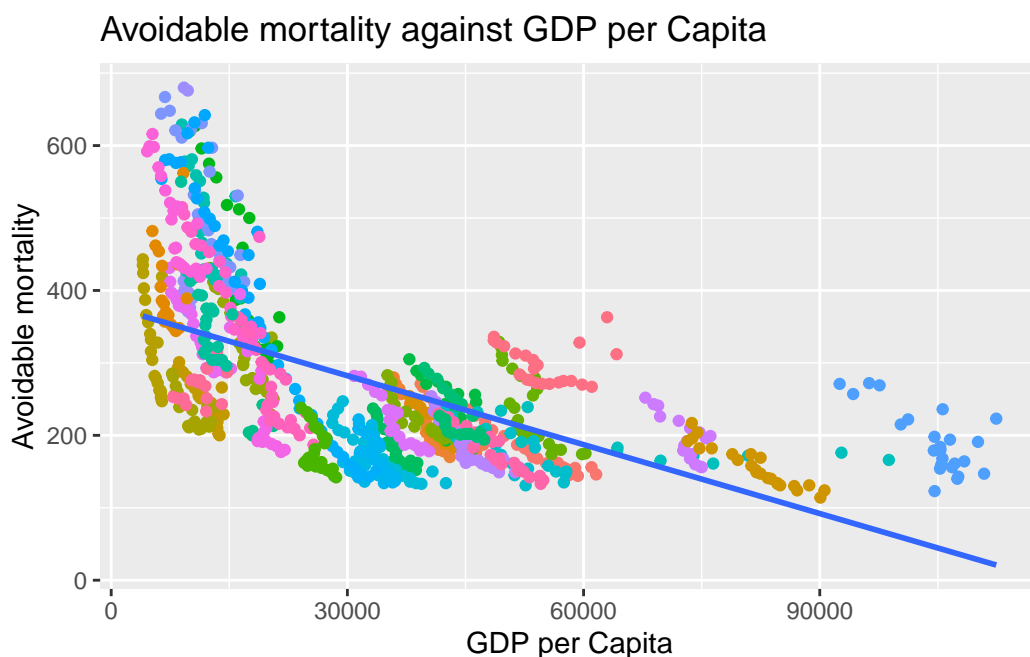As social expenditure increases, the avoidable mortality rate decreases. Most countries' social expenditure is within 10% and 25% of their GDP. However, for these countries similar in social expenditure levels, avoidable mortality goes from 100 to almost 700 per 100,000 deaths, suggesting an important variability. For similar levels of health or social expenditure, avoidable mortality varies significantly between countries.

Avoidable mortality against Unemployment

Here, the trend indicates that as unemployment increases, avoidable mortality does too. Most countries have low levels of unemployment and low rates of avoidable mortality. In the bottom right, Spain (ESP) and Greece (GRC) display very high rates of unemployment while keeping their avoidable mortality rates among the lowest of all countries. On the contrary, countries such as Mexico (MEX) in purple showcase high levels of avoidable mortality with very low rates of unemployment.

Avoidable mortality against GDP per Capita

This graph suggests that as GDP per capita increases, the rate of avoidable mortality decreases. The avoidable mortality rate doesn't go below 100 deaths per 100,000 inhabitants. Thus, as GDP per capita increases, avoidable mortality remains between 100 and 300 deaths per 100,000 inhabitants suggesting marginal change in mortality as a country's population becomes richer.

4. Econometric Analysis : Linear Mixed Model (LMM)

Correlation matrix for control variables

	OBS_VALUE_mortality	OBS_VALUE_gdp	OBS_VALUE_hdi	OBS_VALUE_life_exp	OBS_VALUE_unemp
OBS_VALUE_mortality	1.0000000	-0.6047638	-0.6593684	-0.9559693	0.2393598
OBS_VALUE_gdp	-0.6047638	1.0000000	0.7179984	0.6348101	-0.4273404
OBS_VALUE_hdi	-0.6593684	0.7179984	1.0000000	0.7670182	-0.3755789
OBS_VALUE_life_exp	-0.9559693	0.6348101	0.7670182	1.0000000	-0.2437086
OBS_VALUE_unemp	0.2393598	-0.4273404	-0.3755789	-0.2437086	1.0000000

The correlation matrix reveals strong linear relationships among several control variables.

- Avoidable mortality is highly negatively correlated with life expectancy (-0.96), reflecting the fact that both variables capture very similar aspects of general population health.

- Life expectancy is also strongly correlated with HDI (0.76) and GDP per capita (0.65), meaning there is an overlap between these variables.

To avoid multicollinearity issues in regression models, we choose to exclude HDI and life expectancy from our control variables. We retain GDP per capita and unemployment, which can adjust for differences in economic development between countries.

Creating lag operators/ lagged variables

Before further modelling, we create lagged version of our variables. We believe that past investments in the health sector and social protection have an impact on mortality. We use 1-year lads to study short-term effects. 3-year and 5-year moving averages will be used to capture long-term investments. Indeed, we assume increased hospital capacities or training staff takes a few years before benefiting the health system.

- `mortality_lag1` : 1-year lag of avoidable mortality
- `healthExp_lag1` : 1-year lag of health expenditure (short-run)
- `healthExp_ma3` : 3-year average lag of health expenditure (medium-run)
- `healthExp_ma5` : 5-year average lag of health expenditure (long-run)
- `hospBeds_lag1` : 1-year lag of hospital beds
- `hospBeds_ma3` : 3-year average lag of hospital beds
- `hospBeds_ma5` : 5-year average lag of hospital beds
- `socialExp_lag1` : 1-year lag of social expenditure
- `socialExp_ma3` : 3-year average lag of social expenditure
- `socialExp_ma5` : 5-year average lag of social expenditure

Model testing

First test using multiple lags

We do a first test to see how our variables interact with each other, especially lags

$$Mortality_{i,t} = \gamma_{0,i} + u_i + \beta_1\, mortality_{i,t-1} + \beta_2\, health.expenditure_{i,t-1} + \beta_3\, health.expenditure_{i,t-[1,5]}$$
$$+ \beta_4\, hospital.beds_{i,t-1} + \beta_5\, hospital.beds_{i,t-[1,5]} + \beta_6\, social.expenditure_{i,t-1}$$
$$+ \beta_7\, social.expenditure_{i,t-[1,5]} + \beta_8\, GDP.per.capita_{i,t} + \beta_9\, Unemployment_{i,t}$$
$$+ \epsilon_{i,t}$$

With

- i, index for country i

- t, index for period t (here our periods are years)

- $\epsilon_{i,t}$, the idiosyncratic error term

- u_i, a country-specific random intercept

- γ_0 is the overall mean intercept

All explanatory variables enter the model as fixed effects, while unobserved country-level heterogeneity is captured through a random intercept `\u_{i}`. GDP per capita and unemployment are included as control variables accounting for differences in economic conditions across countries and over time.

Summary of the first regression :

In the regression summary, `REF_AREA (Intercept)` denotes the country-level random intercept.

After this first test, we notice 3 problems :

1. The correlation between `mortality_lag1` and the intercept is very high (-0.729), nullifying the effect of the random intercept (variance and standard deviation equal to 0). This shows that past mortality already explains most differences between countries. We decide to remove `mortality_lag1` from the next regressions.
2. Using both the first lag and the long-run moving average of variables in the same model is not possible. The correlation between the two makes it hard to distinguish which has an effect on mortality and the long-run moving average of explanatory variables end up having incoherent coefficients.
3. There is an issue with the large difference in scale between the variables and `OBS_VALUE_gdp`. Indeed, it is expressed in constant 2015 US$ while other variables are expressed in % of GDP or unit per 1,000 or 100,000. We choose to rescale it in further regressions as `OBS_VALUE_gdp = OBS_VALUE_gdp / 1000`

In addition, we decide to add 2 variables as random effects : social expenditure and health expenditure. This will allow us to better perceive the differences between countries when it comes to social and health policies.
The last explanatory variable, the number of hospital beds will not be added as a random effect since we noticed through testing that it does not have a causal effect on avoidable mortality. Moreover we fear it would clash with the health expenditure variable as they are heavily linked.

We will compare short-run, medium-run and long-run in 3 different models

Models for the short-run, medium-run and long-run

Short-run model :

$$Mortality_{i,t} = \gamma_{0,i} + u_i + \gamma_1\, health.expenditure_{i,t-1} + \beta_2\, hospital.beds_{i,t-1} + \gamma_3\, social.expenditure_{i,t-1}$$
$$+ \beta_4\, GDP.per.capita_{i,t} + \beta_5\, Unemployment_{i,t} \qquad\qquad +\epsilon_{i,t}$$

Summary of the short-run regression :

```
Linear mixed model fit by maximum likelihood  ['lmerMod']
Formula:
OBS_VALUE_mortality ~ healthExp_lag1 + hospBeds_lag1 + socialExp_lag1 +
    OBS_VALUE_gdp + OBS_VALUE_unemp + (1 | REF_AREA) + (0 + healthExp_lag1 |
    REF_AREA) + (0 + socialExp_lag1 | REF_AREA)
   Data: data_lagged
Control: lmerControl(autoscale = TRUE)

     AIC       BIC    logLik -2*log(L)  df.resid
   7858.5    7905.0   -3919.2    7838.5       768


Scaled residuals:
    Min      1Q  Median      3Q      Max
-2.6754 -0.3801 -0.0435  0.2992  8.1851


Random effects:
 Groups      Name             Variance Std.Dev.
 REF_AREA    (Intercept)      35369.29 188.067
 REF_AREA.1  healthExp_lag1     913.88  30.230
 REF_AREA.2  socialExp_lag1      71.02   8.427
 Residual                       808.14  28.428
Number of obs: 778, groups:  REF_AREA, 41


Fixed effects:
               Estimate Std. Error t value
(Intercept)    322.0327    43.5324   7.398
healthExp_lag1 -17.4183     5.6869  -3.063
hospBeds_lag1   29.6867     2.8653  10.361
socialExp_lag1   4.3586     1.8000   2.421
OBS_VALUE_gdp   -4.3308     0.5520  -7.845
OBS_VALUE_unemp -0.5340     0.5736  -0.931
```

```
Correlation of Fixed Effects:
           (Intr) hltE_1 hspB_1 sclE_1 OBS_VALUE_g
hlthExp_lg1 -0.065
hospBds_lg1 -0.536  0.020
soclExp_lg1 -0.105 -0.282  0.095
OBS_VALUE_g -0.452 -0.111  0.255  0.035
OBS_VALUE_n -0.172  0.057 -0.129 -0.149  0.448
```

Interpretation of the results

Firstly, the variance of the random intercept `REF_AREA (Intercept)` is 35369.29, which is much higher than the residual variance (`Residual`) at 808.14. The large random intercept variance indicates substantial between-country differences in baseline avoidable mortality. The standard deviation is 188.07, suggesting a deviation of about 188 deaths per 100,000 between countries.

Second, we will interpret each coefficient : (if $|t| > 1.96$ (for $= 5\%$), the coefficient is statistically significant)

All coefficients are all statistically significant except `OBS_VALUE_unemp` ($t = 0.93$)

- `Intercept` : The baseline number of avoidable deaths per 100,000 when all other predictors are 0 is 322.03.
- `healthExp_lag1` : An increase in health expenditure by 1% of GDP last year is associated with 17 fewer deaths per 100,000. There is a deviation of 30 deaths between countries (`Std.Dev.`).
- `hospBeds_lag1` : The addition of 1 bed per 1,000 inhabitants last year is associated with 29 more deaths per 100,000. This could be a sign that countries with higher mortality happen to invest more in hospital beds.
- `socialExp_lag1`: The increase in social expenditure by 1% of GDP last year is associated with around 4 fewer deaths per 100,000 this year. There is a deviation of 8 deaths between countries (`Std.Dev.`).
- `OBS_VALUE_gdp` : An increase in GDP by 1,000 USD reduces avoidable mortality by 4.33 deaths per 100,000.
- `OBS_VALUE_unemp` : increase in unemployment by 1% of GDP is associated with 0.53 less deaths per 100,000. It is a surprising result but this variable is not statistically significant here.

In conclusion, increases in social and particularly health expenditure reduce the avoidable mortality rate in the short-term and as expected, GDP reduces mortality. Moreover there are huge differences in health and social investments between countries, although it is less the case for social investments.

Medium-run model :

$$Mortality_{i,t} = \gamma_{0,i} + u_i + \gamma_1\,health.expenditure_{i,t-[1,3]} + \beta_2\,hospital.beds_{i,t-[1,3]} + \gamma_3\,social.expenditure_{i,t-[1,3]}$$

$$+ \beta_4\,GDP.per.capita_{i,t} + \beta_5\,Unemployment_{i,t}$$

Summary of the medium-run regression :

```
Linear mixed model fit by maximum likelihood  ['lmerMod']
Formula: OBS_VALUE_mortality ~ healthExp_ma3 + hospBeds_ma3 + socialExp_ma3 +
    OBS_VALUE_gdp + OBS_VALUE_unemp + (1 | REF_AREA) + (0 + healthExp_ma3 |
    REF_AREA) + (0 + socialExp_ma3 | REF_AREA)
   Data: data_lagged
Control: lmerControl(autoscale = TRUE)

     AIC       BIC    logLik -2*log(L)  df.resid
  6970.0    7015.4   -3475.0    6950.0       681

Scaled residuals:
    Min      1Q  Median      3Q     Max
-2.6565 -0.3581 -0.0475  0.2400  8.6197

Random effects:
 Groups      Name            Variance Std.Dev.
 REF_AREA    (Intercept)     34118.23 184.711
 REF_AREA.1  healthExp_ma3     788.73  28.084
 REF_AREA.2  socialExp_ma3      55.75   7.466
 Residual                      785.21  28.022
Number of obs: 691, groups:  REF_AREA, 41

Fixed effects:
                Estimate Std. Error t value
(Intercept)     341.9273    47.4546   7.205
healthExp_ma3   -17.6590     5.8282  -3.030
hospBeds_ma3     27.5718     3.2805   8.405
socialExp_ma3     1.7511     1.9688   0.889
OBS_VALUE_gdp    -3.3749     0.6173  -5.467
OBS_VALUE_unemp  -0.4943     0.6104  -0.810
```

Interpretation of the results First, the variance of the random intercept is still much higher than the residual variance (34118.23 against 785.21). The standard deviation is 184.71, suggesting a deviation of about 184 deaths per 100,000 between countries. This is similar to the short-run results.

Second, all coefficients are all statistically significant except for `OBS_VALUE_unemp` (t = 0.810) and `socialExp_ma3` (t = 0.889).

- Intercept : The baseline number of avoidable deaths per 100,000 when all other predictors are 0 is 341.9.
- `healthExp_ma3` : An increase in health expenditure by 1% of GDP over the past 3 years is associated with 17.66 fewer deaths per 100,000. There is a deviation of 28 deaths between countries (`Std.Dev.`).
- `hospBeds_ma3` : The addition of 1 bed per 1,000 inhabitants over the past 3 years is associated with an increase of deaths per 100,000 by 27.57.
- `socialExp_ma3`: The increase in social expenditure by 1% of GDP over the past 3 years is associated with 1.75 more deaths per 100,000 this year. Given the fact this variable is no longer significant, we can assume the model fails to capture the actual effect of social expenditure on mortality. There is a deviation of 7 deaths between countries (`Std.Dev.`).
- `OBS_VALUE_gdp` : An increase in GDP by 1,000 USD reduces avoidable mortality by 3.37 deaths per 100,000.
- `OBS_VALUE_unemp` : increase in unemployment by 1% of GDP is associated with 0.49 less deaths per 100,000.

Third, the AIC is 6970.0 (lower than for the short-run), indicating better fit for medium-run effects.

In conclusion, health expenditure reduces mortality over 3 years, suggesting that the effects accumulate. Social expenditure is no longer significant. Unexpectedly, there is a positive association between the number of hospital beds and the avoidable mortality rate.
GDP continues to reduce mortality (medium-run effects are slightly smaller than the short-run ones). Moreover, the unemployment effect is much bigger in the medium-run than in the short run with an increase to 9 deaths compared to 1 previously. We still see huge deviations from one country to another in health and social expenditure (although slightly lower).

Long-run model :

$$Mortality_{i,t} = \gamma_{0,i} + u_i + \gamma_1\, health.expenditure_{i,t-[1,5]} + \beta_2\, hospital.beds_{i,t-[1,5]} + \gamma_3\, social.expenditure_{i,t-[1,5]}$$
$$+ \beta_4\, GDP.per.capita_{i,t} + \beta_5\, Unemployment_{i,t}$$

Summary of the long-run regression :

```
Linear mixed model fit by maximum likelihood  ['lmerMod']
Formula: OBS_VALUE_mortality ~ healthExp_ma5 + hospBeds_ma5 + socialExp_ma5 +
    OBS_VALUE_gdp + OBS_VALUE_unemp + (1 | REF_AREA) + (0 + healthExp_ma5 |
    REF_AREA) + (0 + socialExp_ma5 | REF_AREA)
   Data: data_lagged
```

```
Control: lmerControl(autoscale = TRUE)

      AIC        BIC     logLik -2*log(L)   df.resid
     6120       6164      -3050       6100         595

Scaled residuals:
     Min       1Q  Median       3Q      Max
 -3.0926  -0.3223 -0.0530   0.2065   8.6525

Random effects:
 Groups      Name            Variance Std.Dev.
 REF_AREA    (Intercept)     33616.43 183.348
 REF_AREA.1  healthExp_ma5     619.08  24.881
 REF_AREA.2  socialExp_ma5      66.28   8.142
 Residual                     776.99  27.874
Number of obs: 605, groups:  REF_AREA, 41

Fixed effects:
                Estimate Std. Error t value
(Intercept)     332.5441    53.1955   6.251
healthExp_ma5   -13.4040     5.9711  -2.245
hospBeds_ma5     26.6320     3.7660   7.072
socialExp_ma5    -0.4242     2.3199  -0.183
OBS_VALUE_gdp    -2.8131     0.7062  -3.983
OBS_VALUE_unemp  -0.5052     0.6436  -0.785
```

Interpretation of the results First, the variance of the random intercept is still higher than the residual variance (33616.43 against 776.99). There is deviation of about 183 deaths per 100,000 between countries. It is smaller than it was in the medium-run model even though the standard deviation remains more or less the same in the 3 models.

Second, all coefficients are all statistically significant except for `OBS_VALUE_unemp` (t = 0.408) and `socialExp_ma5` (t = 0.183)

- Intercept : The baseline number of avoidable deaths per 100,000 when all other predictors are 0 is 332.
- `healthExp_ma5` : An increase in health expenditure by 1% of GDP over the past 5 years is associated with 13.40 fewer deaths per 100,000. There is a deviation of 24 deaths between countries (`Std.Dev.`).
- `hospBeds_ma5` : The addition of 1 bed per 1,000 inhabitants over the past 5 years is associated with an increase of deaths per 100,000 by 26.63.
- `socialExp_ma5`: The increase in social expenditure by 1% of GDP over the past 5 years is associated with 0.42 fewer deaths per 100,000 this year. It is a small number but once

again this can be explained by the fact that this variable is no longer significant. There is a deviation of 8 deaths between countries (`Std.Dev.`).

- `OBS_VALUE_gdp` : An increase in GDP by 1,000 USD reduces avoidable mortality by 2.81 deaths per 100,000.
- `OBS_VALUE_unemp` : increase in unemployment by 1% of GDP is associated with 0.50 less deaths per 100,000.

Third, the AIC is 6120 (lower than for the short-run and medium-run models), indicating better fit for long-run effects. We see more or less the same deviations between countries in health and social expenditure as before.

Conclusion comparing short-medium-long run models

In the short-run, we can interpret the impact of health and social expenditure on avoidable mortality as an immediate budgetary effect. Having a new monetary flow each year in the health sector is important to face urgent needs or crisis. Health expenditure reduces mortality most strongly in the short run (-17.42 deaths per 100,000). However, the number of hospital beds and social expenditure showcase a positive impact on mortality.

Over a 3-year horizon, cumulative government spending begins to show efficiency: health expenditure reduces mortality by -17.66, hospital beds still show a positive association with mortality (+27.57). This could reflect important investments in the number of beds per 1,000 inhabitants rather than a causal effect on mortality. This highlights that medium-run effects are partly shaped by policy planning and investment patterns, rather than short-run decisions.

In the long-run, the impact of health expenditure and the number of hospital beds can be interpreted as a structural effect. Health expenditure reduces deaths by -13.40 over five years, showing a persistent effect, although slightly smaller than medium-run averages for health expenditure. Hospital beds remain positively associated with mortality (+26.63). GDP consistently reduces mortality across all horizons.

Finally in all 3 models we notice huge deviations between countries when it comes to health and social expenditure (25-30 deaths for health expenditure and ~8 deaths for social expenditure). There is a lot of heterogeneity between countries which we could interpret as differences in institutional quality or past health and social protection history.

Comparing short-run, medium run and long-run

Looking at the AIC :

- For the short-run model : 7858.5
- For the medium-run model : 6970.0

- For the long-run model : 6120

The AIC - Akaike information criterion - is a measure for how well fitted a model is. The AIC verifies both if explanatory variables explain the observed variable correctly.

The lower the AIC is, the better the model is. Here, the long-run model has the lowest AIC. Considering all 3 models have the same number of parameters, a lower AIC means avoidable mortality is better explained by the long-run model.

Testing for the Covid-19 period

We perform regressions using lags and the Covid-19 dummy variable. Moreover, in order to clearly see the impact of previous investment on avoidable mortality during the Covid-19 crisis, we create interaction variables using our lags.

We note that using interaction terms for time t could lead to reverse causality problems. We might be unable to identify whether health expenditure, social expenditure and hospital beds impacted avoidable mortality or if because mortality increased suddenly during the pandemic, health expenditure, social expenditure and hospital beds were increased afterwards.

- `healthexp_pandemic_lag1` = interaction term between the `pandemic` dummy variable and `healthExp_lag1`
- `healthexp_pandemic_ma3` = interaction term between the `pandemic` dummy variable and `healthExp_ma3`
- `healthexp_pandemic_ma5` = interaction term between the `pandemic` dummy variable and `healthExp_ma5`
- `socialexp_pandemic_lag1` = interaction term between the `pandemic` dummy variable and `socialExp_lag1`
- `socialexp_pandemic_ma3` = interaction term between the `pandemic` dummy variable and `socialExp_ma3`
- `socialexp_pandemic_ma5` = interaction term between the `pandemic` dummy variable and `socialExp_ma5`

The following variables will be interpreted as :

- `(Intercept)`, baseline avoidable mortality per 100,000 when all other variables are 0, outside of the pandemic
- `pandemic`, average change in mortality due to COVID (years 2020–2021), all else equal
- `healthexp_pandemic_lag1` and `socialexp_pandemic_lag1`, how the effect of pre-pandemic health/social spending changes during the pandemic

In the short-run :

17

$$Mortality_{i,t} = \gamma_{0,i} + u_i + \beta_1\,health.expenditure_{i,t-1} + \beta_2\,hospital.beds_{i,t-1} + \beta_3\,social.expenditure_{i,t-1}$$

$$+ \beta_4\,pandemic + \beta_5\,hospital.beds_{i,t-1}*pandemic + \beta_6\,social.expenditure_{i,t-1}*pandemic$$

$$+ \beta_7\,GDP.per.capita_{i,t} + \beta_8\,Unemployment_{i,t} \qquad\qquad +\epsilon$$

Summary of the short-run regression :

```
Linear mixed model fit by maximum likelihood  ['lmerMod']
Formula:
OBS_VALUE_mortality ~ hospBeds_lag1 + healthExp_lag1 + socialExp_lag1 +
    pandemic + healthexp_pandemic_lag1 + socialexp_pandemic_lag1 +
    OBS_VALUE_gdp + OBS_VALUE_unemp + (1 | REF_AREA) + (0 + healthExp_lag1 |
    REF_AREA) + (0 + socialExp_lag1 | REF_AREA)
   Data: data_lagged
Control: lmerControl(autoscale = TRUE)

     AIC       BIC    logLik -2*log(L)  df.resid
  7729.8    7790.3   -3851.9    7703.8       765


Scaled residuals:
    Min      1Q  Median      3Q     Max
-3.9325 -0.3929 -0.0177  0.3583  6.2973


Random effects:
 Groups     Name            Variance Std.Dev.
 REF_AREA   (Intercept)     36840.87 191.940
 REF_AREA.1 healthExp_lag1    986.66  31.411
 REF_AREA.2 socialExp_lag1     57.39   7.575
 Residual                    665.79  25.803
Number of obs: 778, groups:  REF_AREA, 41


Fixed effects:
                         Estimate Std. Error t value
(Intercept)              342.9014    42.5044   8.067
hospBeds_lag1             32.6153     2.6836  12.154
healthExp_lag1           -17.2917     5.7270  -3.019
socialExp_lag1             2.8382     1.6614   1.708
pandemic                 111.2662    14.7331   7.552
healthexp_pandemic_lag1    0.2854     1.6989   0.168
socialexp_pandemic_lag1   -3.5745     0.6964  -5.133
```

```
OBS_VALUE_gdp              -4.8443     0.5203  -9.311
OBS_VALUE_unemp            -0.7097     0.5372  -1.321
```

In order to interpret our results we need to sum the coefficient obtained for the health expenditure variable (`healthExp_lag1`) and the interaction term `healthexp_pandemic_lag1`. Same thing for the social expenditure variable. This will tell us how the pandemic impacts the effect of health expenditure and social expenditure on avoidable mortality.

```
healthExp_lag1
     -17.00638
```

```
socialExp_lag1
    -0.7362972
```

Interpretation of the results First, we analyse the random effects.

- `REF_AREA (Intercept)` has a variance of 36840.87, meaning there are large differences between countries in baseline avoidable mortality (around 191 deaths per 100,000)
- `REF_AREA.1 healthExp_lag1` has a variance of 986.66. The effect of health expenditure varies between countries by approximately 31 deaths (given by the standard deviation)
- `REF_AREA.2 socialExp_lag1`has a variance of 57.39. The effect of social expenditure is less important. It varies by around 8 deaths per 100,000.

This suggests that different countries respond differently to health and social expenditures.

Second, we analyse the fixed effects. All variables are significant except `socialExp_lag1`, `healthexp_pandemic_lag1` and `OBS_VALUE_unemp`.

- `(Intercept)` : The baseline number of avoidable deaths per 100,000 if all other variables are null, outside of the pandemic, is 342.90
- `hospBeds_lag1` : Here, we find similar results as in the previous section : adding 1 hospital bed per 1,000 increases short-run mortality by 32.6 deaths.
- `healthExp_lag1` : A 1% of GDP increase in health expenditure last year reduces mortality by 17.3 deaths per 100,000, outside of the pandemic
- `socialExp_lag1`: A 1% of GDP increase in health expenditure last year increases mortality by 2.84 deaths per 100,000, outside of the pandemic. However, this variable is not statistically significant.
- `pandemic` : Covid increases baseline mortality by 111 deaths per 100,000, regardless of spending.
- `healthexp_pandemic_lag1` : We combine `healthexp_pandemic_lag1` with `healthExp_lag1` to appreciate the total effect of health expenditure. Its effect is reduced by 0.29 during Covid, -17.006 deaths in total. The effect is still strong.

- `socialexp_pandemic_lag1` : Outside of the pandemic, social expenditure increases deaths by 2.8382 and decreases it by -0.736 during. In total the effect is slightly negative, thereby having a much smaller impact on mortality compared to health expenditure.

In conclusion, the pandemic increases the number of avoidable deaths sharply. Health expenditure in the short term remains strongly efficient, but this is not the case for social expenditure which has a much smaller effect. Moreover, random intercepts and slopes show important differences between countries.

In the medium-run :

$$Mortality_{i,t} = \gamma_{0,i} + u_i + \beta_1\, health.expenditure_{i,t-[1,3]} + \beta_2\, hospital.beds_{i,t-[1,3]} + \beta_3\, social.expenditure_{i,t-[1,3]}$$

$$+ \beta_4\, pandemic + \beta_5\, hospital.beds_{i,t-[1,3]} * pandemic + \beta_6\, social.expenditure_{i,t-[1,3]} * pandemic$$

$$+ \beta_7\, GDP.per.capita_{i,t} + \beta_8\, Unemployment_{i,t}$$

Summary of the medium-run regression :

```
Linear mixed model fit by maximum likelihood  ['lmerMod']
Formula: OBS_VALUE_mortality ~ hospBeds_ma3 + healthExp_ma3 + socialExp_ma3 +
    pandemic + healthexp_pandemic_ma3 + socialexp_pandemic_ma3 +
    OBS_VALUE_gdp + OBS_VALUE_unemp + (1 | REF_AREA) + (0 + healthExp_ma3 |
    REF_AREA) + (0 + socialExp_ma5 | REF_AREA)
   Data: data_lagged
Control: lmerControl(autoscale = TRUE)

     AIC       BIC    logLik -2*log(L)  df.resid
   6130.4    6188.0   -3052.2    6104.4       608


Scaled residuals:
    Min      1Q  Median      3Q     Max
-3.7127 -0.3373 -0.0207  0.3040  6.5096


Random effects:
 Groups     Name          Variance Std.Dev.
 REF_AREA   (Intercept)   44455.95 210.846
 REF_AREA.1 healthExp_ma3   753.23  27.445
 REF_AREA.2 socialExp_ma5    88.41   9.403
 Residual                  564.17  23.752
Number of obs: 621, groups:  REF_AREA, 41


Fixed effects:
```

```
                        Estimate Std. Error t value
(Intercept)             312.6913    50.7127   6.166
hospBeds_ma3             31.3911     3.4182   9.183
healthExp_ma3           -20.3007     5.9096  -3.435
socialExp_ma3             5.6009     1.8254   3.068
pandemic                127.3569    13.4750   9.451
healthexp_pandemic_ma3    0.1386     1.5438   0.090
socialexp_pandemic_ma3   -4.5283     0.6484  -6.984
OBS_VALUE_gdp            -3.3105     0.6235  -5.309
OBS_VALUE_unemp          -0.8961     0.5723  -1.566
```

Here, if we sum coefficients and interaction terms we have :

```
healthExp_ma3
    -20.16214


socialExp_ma3
     1.072542
```

Interpretation of the results We notice that this model seems to have the same trends as the short-run model.

First, we analyse the random effects.

- `REF_AREA (Intercept)` has a variance of 44455.95, meaning there are large differences between countries in baseline avoidable mortality, more than in the short-run (around 210 deaths per 100,000)
- `REF_AREA.1 healthExp_lag1` has a variance of 753.23. The effect of health expenditure varies between countries by approximately 27 deaths (given by the standard deviation) which is slightly less than in the short-run (about 31 deaths).
- `REF_AREA.2 socialExp_lag1`has a variance of 88.41. The effect of social expenditure is less important. It varies by around 9 deaths per 100,000.

Second, we analyse the fixed effects. All variables are significant except `healthexp_pandemic_ma3` and `OBS_VALUE_unemp`.

- `(Intercept)` : The baseline number of avoidable deaths per 100,000 if all other variables are null, outside of the pandemic, is 312.6.
- `hospBeds_ma3` : Here, we find similar results as in the previous section and the short-run regression : adding 1 hospital bed per 1,000 increases short-run mortality by 20.3 deaths, slightly more than in the short-run (17.3)
- `healthExp_ma3` : A 1% of GDP increase in health expenditure last year reduces mortality by 17.3 deaths per 100,000, outside of the pandemic

- `socialExp_ma3`: A 1% of GDP increase in health expenditure last year increases mortality by 5.6 deaths per 100,000, outside of the pandemic.
- `pandemic` : Covid increases baseline mortality by 127 deaths per 100,000, regardless of spending.
- `healthexp_pandemic_ma3` : The effect of health expenditure is reduced by 0.13 during Covid, -20.2 deaths in total. The effect is barely impacted by the Covid period.
- `socialexp_pandemic_ma3` : The effect of social expenditure is reduced by 4.53 during Covid, increasing deaths by 1.07 in total. In sum, the effect is still positive, suggesting a positive correlation between avoidable mortality and social expenditure.

In conclusion, health expenditure in the medium-term remains strong. This highlights the importance of sustained health spending, which shows positive effects on avoiding deaths druing crises. However, this is not the case for social expenditure which has a positive impact on mortality. This may suggest that countries with higher mortality simultaneously increased social spending in response to Covid. the interaction terms show that the pandemic slightly moderates the impact of health and social expenditures, but the main trends remain similar. Moreover, random intercepts and slopes show even more important differences between countries than in the short-term.

In the long-run :

$$Mortality_{i,t} = \gamma_{0,i} + u_i + \beta_1\,health.expenditure_{i,t-[1,5]} + \beta_2\,hospital.beds_{i,t-[1,5]} + \beta_3\,social.expenditure_{i,t-[1,5]}$$

$$+ \beta_4\,pandemic + \beta_5\,hospital.beds_{i,t-[1,5]} * pandemic + \beta_6\,social.expenditure_{i,t-[1,5]} * pandemic$$

$$+ \beta_7\,GDP.per.capita_{i,t} + \beta_8\,Unemployment_{i,t}$$

Summary of the long-run regression :

```
Linear mixed model fit by maximum likelihood  ['lmerMod']
Formula: OBS_VALUE_mortality ~ hospBeds_ma5 + healthExp_ma5 + socialExp_ma5 +
    pandemic + healthexp_pandemic_ma5 + socialexp_pandemic_ma5 +
    OBS_VALUE_gdp + OBS_VALUE_unemp + (1 | REF_AREA) + (0 + healthExp_ma3 |
    REF_AREA) + (0 + socialExp_ma5 | REF_AREA)
   Data: data_lagged
Control: lmerControl(autoscale = TRUE)

     AIC       BIC    logLik  -2*log(L)  df.resid
  5974.4    6031.7   -2974.2    5948.4        592


Scaled residuals:
    Min      1Q  Median      3Q     Max
-3.6262 -0.3400 -0.0304  0.3362  6.6608
```

```
Random effects:
 Groups      Name            Variance Std.Dev.
 REF_AREA    (Intercept)     36499.9  191.049
 REF_AREA.1  healthExp_ma3     618.3   24.865
 REF_AREA.2  socialExp_ma5      62.5    7.906
 Residual                      578.2   24.047
Number of obs: 605, groups:  REF_AREA, 41

Fixed effects:
                         Estimate Std. Error t value
(Intercept)             347.2838    50.4435    6.885
hospBeds_ma5             30.0932     3.4177    8.805
healthExp_ma5           -17.6259     5.0750   -3.473
socialExp_ma5           -0.2125      2.1396   -0.099
pandemic                123.9684    13.5310    9.162
healthexp_pandemic_ma5   0.2426      1.5749    0.154
socialexp_pandemic_ma5  -4.3975      0.6556   -6.707
OBS_VALUE_gdp           -3.4738      0.6382   -5.443
OBS_VALUE_unemp         -0.8032      0.5641   -1.424
```

Then, if we sum coefficients and interaction terms we have :

```
healthExp_ma5
    -17.38338


socialExp_ma5
    -4.610063
```

Interpretation of the results

First, we analyse the random effects.

- `REF_AREA (Intercept)` has a variance of 36499.9, meaning there are large differences between countries in baseline avoidable mortality, less than in the short- and medium-run (around 191 deaths per 100,000)
- `REF_AREA.1 healthExp_lag1` has a variance of 618.3 The effect of health expenditure varies between countries by approximately 25 deaths (given by the standard deviation) which is slightly less than in the short- and medium-run.
- `REF_AREA.2 socialExp_lag1`has a variance of 62.5 The effect of social expenditure is less important. It varies by around 8 deaths per 100,000.

Second, we analyse the fixed effects. Some variables are not significant : `socialExp_ma5`, `healthexp_pandemic_ma5` and `OBS_VALUE_unemp`.

- `(Intercept)` : The baseline number of avoidable deaths per 100,000 if all other variables are null, outside of the pandemic, is 347.3.
- `hospBeds_ma3` : Here, we find similar results as in the previous section and the short-run regression : adding 1 hospital bed per 1,000 increases short-run mortality by 30.1 deaths, more than in the short- and medium-run.
- `healthExp_ma3` : A 1% of GDP increase in health expenditure last year reduces mortality by 17.6 deaths per 100,000, outside of the pandemic
- `socialExp_ma3`: A 1% of GDP increase in health expenditure last year increases mortality by 0.21 deaths per 100,000, outside of the pandemic. But this variable is not significant.
- `pandemic` : Covid increases baseline mortality by about 124 deaths per 100,000, regardless of spending.
- `healthexp_pandemic_ma3` : The effect of health expenditure is reduced by 0.24 during Covid, -17.4 deaths in total. The effect is, again, barely impacted by the Covid period.
- `socialexp_pandemic_ma3` : The effect of social expenditure is reduced by 4.4 during Covid, decreasing deaths by 4.6 in total. This time, social expenditure is negative, suggesting that its effects are more latent and decrease the avoidable mortality rate in the long-run.

In conclusion, in the long-run, health expenditure remains the most important factor in reducing avoidable mortality. Social expenditure, which showed positive correlations with avoidable mortality in the short- and medium-run now shows a slight negative effect. This might suggest, social expenditure might take longer to have a concrete impact. Random effects indicate that country-level differences remain important, though slightly smaller than in shorter-term models. Overall, one key finding of our analysis is that sustained health investment is important for reducing avoidable deaths, while social spending contributes more gradually, with long-term benefits appearing later.

Moreover, we compare the 3 model's AIC to determine which fits best the data.

- short-run AIC : 7729.8
- medium-run AIC: 6130.4
- long-run AIC: 5974.4

As in the previous section, the long-run model best fits the data.

5. Conclusion

Our research stemmed from wondering whether having a strong health and social policy proved useful in order to have a healthy population. The results seem to confirm our intuition.

Using a linear mixed model allowed us to capture the heterogeneity between countries. It was also key when taking into account the nature of avoidable mortality, which has a very slow evolution, meaning that for most countries the current value of avoidable mortality would be really close to its previous values. In addition a mixed model was also essential since we wanted to account for differences in social and health policies. We have shown using regression analysis that there is significant variability in avoidable mortality both during and outside of crises between countries. However this aspect would require further modelling to determine which countries/health systems are more impacted by crises and tailor policy responses accordingly.

Our main conclusions are that increases in health spending are by far the most important element of reducing avoidable mortality while social expenditure is not as important as we could have thought at the beginning of the analysis. Furthermore according to our model, during the pandemic - which we consider as the years 2020 and 2021 - health expenditure cushioned the increase in avoidable deaths, while social protection was particularly efficient as the population faces difficulties.

It is important to mention that our analysis is limited in many ways. First of all the number of countries studied is still relatively small (41 OECD countries). Then, we could not explore further the structural differences between countries or other potential variables. In spite of these limitations this report can be considered as a first clue in finding how to reduce avoidable mortality.