Since January 2020 Elsevier has created a COVID-19 resource centre with free information in English and Mandarin on the novel coronavirus COVID-19. The COVID-19 resource centre is hosted on Elsevier Connect, the company's public news and information website.

Elsevier hereby grants permission to make all its COVID-19-related research that is available on the COVID-19 resource centre - including this research content - immediately available in PubMed Central and other publicly funded repositories, such as the WHO COVID database with rights for unrestricted research re-use and analyses in any form or by any means with acknowledgement of the original source. These permissions are granted for free by Elsevier for as long as the COVID-19 resource centre remains active.
Long-lasting effects of pandemics: The case of the 1918 influenza pandemic in Argentina

Fernando Antonio Ignacio González a,*, Juan Antonio Dip b, Silvia London a

a Instituto de Investigaciones Económicas y Sociales del Sur, Universidad Nacional del Sur/CONICET (IBESS, UNS-CONICET), San Andrés 800, Bahía Blanca-Buenos Aires
b Facultad de Ciencias Económicas, Universidad Nacional de Misiones, Ruta 12 Km 12.5, Posadas-Misiones

Abstract

The 2019 novel coronavirus disease pandemic poses a serious threat. While its short-term effects are evident, its long-term consequences are a matter of analysis. In this work, the existence of long-lasting negative effects derived from exposure in utero to a great pandemic -1918 influenza pandemic- is analysed for the Argentine case.

Outcomes of interest include educational achievement and unemployment status in adulthood ~50 years after the pandemic. Based on a regression analysis, temporal differences in the spread of the pandemic and between close birth cohorts are exploited.

The results indicate a significant reduction in educational achievement for people exposed in utero to the pandemic. In the region with the highest incidence of cases (Noroeste), this reduction is 0.5 years of education. There are no significant changes in the chances of being unemployed. In the context of climate change, these results constitute a call of attention for the implementation of child protection policies from gestation.

© 2021 Elsevier Ltd. All rights reserved.

1. Introduction

The 2019 novel coronavirus disease pandemic (COVID-19) constitutes a profound disruption in the functioning of affected communities. The immediate negative socio-economic consequences of this biological natural disaster are evident. Globally, there is a wide increase in the unemployment rate (Béland et al., 2020), falls in the gross domestic product (GDP) (Boissay & Rungcharoenkitkul, 2020) and increases in the incidence of poverty (World Bank, 2020). The latter is particularly worrying given that since 1998 there has been no increase in the incidence of income poverty globally. For Argentina the contraction of the GDP, in the second quarter of 2020, reached 19.1% (National Institute of Statistics and Censuses, 2020).

In addition, for this country an increase of 4 percentage points is forecasted for the incidence of income poverty. This is equivalent to 1.1 million additional poor considering a poverty line of US$ 5.5 dollars a day (World Bank, 2020).

The persistence, in the long term, of the adverse consequences of the pandemic is a matter of debate. At the moment a fundamental restriction operates: the long term has not yet arrived. That is, generating estimates that allow causal inference regarding the persistence of the negative effects derived from exposure to the pandemic requires having information sufficiently distant -in time- from the pandemic itself. The realization of causal inference exploits the sudden, unexpected and brief nature of a pandemic. This sets up a natural experiment.

Despite the above, it is possible to examine the persistence of negative effects for previous pandemics. In particular, the 1918 influenza pandemic -frequently called the Spanish flu- (1918-1919) is the one that has caused the highest mortality, to date, in the case of Argentina. Between October 1918 and August 1919, 14,9972 deaths were recorded due to this pandemic (Carbonetti & Álvarez, 2017). While in 1917 the flu -as a cause of death- was responsible for 0.7% of all deaths, in 1920 it represented 20.7% (Carbonetti, 2010). As in the present, the health authorities...
implemented measures aimed at reducing social interactions. This included schools closing, reduced business hours, prohibition of massive shows, changes in the use of public transport (Carbonetti & Álvarez, 2017), although without reaching generalized traffic bans.

In the decades prior to the occurrence of the pandemic, Argentina experienced the period of greatest economic growth in its history (Maddison Project, 2018) based on the export of agricultural products. In this context, its population was growing at an accelerated rate due to increasing immigration (Arceo et al., 2019), a longer life expectancy -annual average increase of almost one year per calendar year- and a lower mortality rate (Grushka, 2018). In this sense, living conditions greatly improved in this period. At no time did Argentina participate in the First World War.3

The persistence of negative effects derived from an environmental shock (pandemic) has been studied from different perspectives. In particular, exposure to shocks during gestation in utero has gained wide interest in the literature (Bennet & Gunn, 2006; Parman, 2015; Ogasawara, 2018). The hypothesis of fetal origin (Barker, 1990) suggests that those growth restrictions experienced in utero constitute a risk factor for diseases in adulthood, including hypertension, diabetes and obesity (Giraldo et al., 2014), or mental illness (Liu et al., 2016). It has also been observed that malnutrition in utero -measured as birth weight- is associated with lower cognitive skills, reduced educational achievement and lower income in adulthood (Lin et al., 2007; Lin & Liu, 2009; Almond & Currie, 2011). This does not ignore the fact that there are other periods in which experiencing growth restrictions can have persistent negative effects -especially during early childhood- (Alderman et al., 2006; Martins et al., 2011), but emphasis is placed on a particularly critical one (gestation in utero).

Determining the channels through which the in utero environment affects long-term outcomes is a complex task, since there could be unobservable characteristics of the mother or family that affect both (Lin & Liu, 2014). Typically, the literature has highlighted channels such as maternal stress and malnutrition and infectious diseases (Schlotz & Phillips, 2009; Beijers et al., 2014; Duque et al., 2018). For the particular case of influenza, Kawana et al. (2007) find that those exposed to the 1918 influenza pandemic in Japan experienced loss of appetite (45% of cases) or vomiting and diarrhoea (22%). That is, influenza can reduce the nutritional intake of the pregnant woman and the fetus. Maternal stress -measured from cortisol levels- generated by a shock such as the influenza virus is negatively associated with the child’s motor development and IQ (Huizink et al., 2003). Also, increases in interleukin-6 (IL-6) during pregnancy due to exposure to influenza are associated with higher blood pressure in adulthood (Mazumder et al., 2010). Additionally, and most consistent with the hypothesis of fetal origin, exposure to the 1918 influenza have implied an increase in the chances of having a premature delivery in England (Reid, 2005). This intraterine growth restriction may experience multiple negative consequences in later stages of development (Barker, 1990) and constitutes another potential channel to explain the persistence of these harmful effects.

Mixed results have been reported in relation to the previous hypothesis and the consequences of the 1918 influenza pandemic. In the case of Sweden, when using microdata from 1970, it is observed that those born during the peak of the pandemic (cohort 1919) present higher chances of being hospitalized in a specific year and experiencing longer hospitalizations (+3.5%), in relation to those born before or after (1914–1925). However, there are no significant differences in terms of income or the chances of being unemployed (Helgertz & Bengtsson, 2019). For the United States, and using census microdata from 1960, 1970 and 1980, Almond (2006) finds that individuals in the 1919 cohort have a lower educational achievement (0.17 and 0.12 years of education less in men and women, respectively) in relation to close cohorts (1912–1922). For the 1919 cohort, the same author also reports greater chances of being income poor or having some type of disability to work.

For the case of six cities in Brazil and using repeated waves of microdata from 1986 to 1998, Nelson (2009) observes that people in the 1919 cohort have, on average, 0.046 fewer years of education, fewer chances of being employed (5%) or graduating from university (13%), in relation to those of close cohorts (1912–1922). The author finds no negative effects on the literacy rate. Concordant results, in terms of a reduction in the educational achievement of the cohorts exposed in utero, have been reported for Switzerland (Neelsen & Stratmann, 2012) or Taiwan (Lin & Liu, 2014).

When considering a group of 53 countries, for which there are census microdata available, Vollmer and Wójcik (2017) report that the existence of long-lasting negative effects derived from the 1918 influenza pandemic, in terms of educational achievement, employment and disabilities, is not robust and argue that the apparent consensus in the evidence may be due to publication bias. In their estimates, they included Argentina and the 1970 census microdata. They found no significant effect on the chances of completing primary or secondary education, or on the chances of being employed. The education estimates arose from considering a dummy variable as dependent (1 if the educational level was completed, 0 in the other cases). For both the education and employment dimensions, they considered a regression by ordinary least squares, with a polynomial specification of degree two, and a dummy indicating the 1919 cohort. The sample comprised the 1910-1928 cohorts. Similar conclusions, lack of robust evidence, were reported in Cohen et al. (2010) when evaluating mortality indicators for a group of 24 countries.

In this context, this paper examines the existence of long-lasting negative effects derived from the exposure in utero to the 1918 influenza pandemic for the Argentine case. Outcomes of interest include educational achievement and unemployment status in adulthood -50 years after the exposure to the pandemic. To do this, we follow the regression discontinuity framework, exploiting the temporal differences in the propagation of the pandemic and between birth cohorts close to each other. The robustness of the results is also analysed by considering different bandwidths, placebo tests and an additional estimation method (differences in differences). To the best of our knowledge, this work adds value to the literature on the long-term effects of natural disasters on human well-being in two aspects. In the first place, it is the first work to examine the persistence of the effects of the 1918 influenza pandemic considering an identification strategy that compares between people born just months apart, and a continuous specification of educational achievement, for the case of Argentina. Vollmer and Wójcik (2017) provide estimates for 53 countries -including Argentina- but use an identification strategy in which they consider cohorts distant up to 10 years from that affected by the pandemic and a binary definition of educational achievement. Second, unlike most of the background, this paper explores socioeconomic regional differences -within the country- in the consequences of the pandemic and how these can contribute to widening territorial inequality.

The structure of the papers is as follow: Section 2 describes the sources of information. Section 3 presents the estimation methodologies used. Section 4 analyses the results and, finally, Section 5 discusses the main conclusions of the work.

---

3 This avoids possible biases due to differences in the characteristics of the parents, as highlighted in Brown (2011).
2. Sources of information

For this work, we combine two sources of information. On the one hand, the microdata from the 1970 National Census of Population, Households and Housing provided by the Integrated Public Use Microdata Series platform ([IPUMS] IPUMS, 2017) are used. This is the oldest census whose microdata is publicly available. Using the oldest census allows to capture a greater number of people whose gestation took place during the pandemic (1919). In case of using the most recent census (2010), people born during the pandemic would have an average age of 91 years -which, logically, would cause inconvenience due to loss of observations.-

Census microdata allow knowing the district of residence, composition of the family group, educational achievement, age, among others, for a random sample of 466,892 people (2% of the census population). As in Almond (2006), it is not possible to identify the exact day of birth -given the census microdata. For this, we infer the membership cohort from the reported age. Age-heaping constitutes another limitation from our source of information (those born in 1919 could report age 50 instead of age 51, leaving the birth year to be incorrectly identified) and, as remarked by Lin and Liu (2014), this could bias the results downward. We assume that the district of birth (N=342) is the same as that of residence at the time of the census. This is reasonable given that the decision to migrate is not expected to correlate with the infrequent occurrence of a major pandemic.

On the other hand, it is necessary to have records of the incidence of the pandemic and its temporal and geographical differences. In the Argentine case, the National Hygiene Department (the predecessor of today's Ministry of Health) was the body responsible for recording the evolution of the pandemic (National Hygiene Department, 1921 cited in Carbonetti & Álvarez, 2017). However, these records -which were published in 1921- are not available online. Carbonetti (2010) and Carbonetti and Álvarez (2017) manage to access, and process, this information. From this source, it is possible to know the mortality rate caused by the 1918 influenza pandemic (deaths per 10,000 inhabitants) and its temporal distribution.

The National Hygiene Department prepared records by province. At that time (1918–1919) there were 15 provinces, while at present there are 24 (including the City of Buenos Aires). The difference (9) is due to the fact that these jurisdictions constituted national territories. This group included 3 of the 4 provinces of the Noreste region (Misiones, Chaco and Formosa) -since Corrientes already constituted a province. The provinces of the Patagonia region were also national territories (Rio Negro, Neuquén, La Pampa, Chubut, Santa Cruz and Tierra del Fuego). The remaining regions that exist in Argentina are: Noroeste ([NOA] Salta, Jujuy, Tucumán, La Rioja, Catamarca and Santiago del Estero), Cuyo (Mendoza, San Juan and Sal Luis) and Centro (Córdoba, Santa Fe, Entre Ríos, Province of Buenos Aires and City of Buenos Aires). From now on, the term Argentina refers to the 24 jurisdictions that exist today -including the provinces of the early twentieth century (15), but also the national territories (9).

The records of the National Hygiene Department suggest that the pandemic generated the highest number of deaths, in the Argentine case, between the last quarter of 1918 and the first half of 1919 (Carbonetti & Álvarez, 2017). In 1918, the City and Province of Buenos Aires presented the first cases in the country. However, the NOA and Cuyo regions were more affected in relative terms. Graph 1 reflects these temporal and geographic differences.

3. Methodology

For the empirical identification strategy, we follow the regression discontinuity framework. We build on the methodological proposal of Sotomayor (2013), where the occurrence of natural disasters (hurricanes) is analysed for the case of Puerto Rico by using local linear regressions. In particular, Sotomayor (2013) exploits differences between close birth cohorts in exposure (in utero) to two hurricanes, and analyses the health impacts of these natural disasters more than 70 years after their occurrence. Since the proposal of Thistlethwaite and Campbell (1960), where the effect on academic performance of a scholarship program is analysed, the use of regression discontinuous designs has become more frequent in Economics (Lee & Lemieux, 2010). This method allows causal inference to be made about the impact of a non-randomized treatment on a given outcome of interest. It exploits the existence of a threshold that separates the target population into two groups: those who receive the intervention of interest (treatment) given that they exceed the critical threshold, and those who do not receive it (control). Therefore, when comparing those people who are located just below or above the separating threshold, it is reasonable to assume that the pre-treatment differences are negligible and, therefore, any observed difference after treatment can be causally attributed to treatment.

In the context of this work, the separating threshold is given by the gestation period in utero at a time when the 1918 influenza pandemic was spreading throughout Argentina. Given that the pandemic covered, especially, from the last trimester of 1918 to the first semester of 1919, the gestation in utero of this period refers to the births reported throughout 1919. This definition is consistent with previous estimates for other countries (Almond, 2006; Nelson, 2009; Helgertz & Bengtsson, 2019). Thus, people in this cohort constitute the treatment group. Those born just a few months before and after constitute the control group (Table 1).

The choice of bandwidth is essential to ensure adequate comparability between the two groups. The narrower the bandwidth, the greater comparability can be achieved, although at the cost of a loss of observations. In this case, since the assignment variable is discrete, the choice becomes less complex. The minimum interval to consider is one year before (1918) and one after (1920). Other bandwidths -greater- are considered in order to check for the robustness of the results.

Eq. 1 presents the regression specification used. A local linear regression is employed, avoiding the use of higher-order polynomial regressions (Gelman & Imbens, 2014). $y_{id}$ is the outcome of interest of person $i$ of the cohort $t$ born in district $d$; $Exposed_{id}$ is a dummy that takes value 1 in case the person belongs to the treatment group (cohort 1919), and 0 if they belongs to the control group (cohorts 1918 or 1920); $Cohort_{id}$ is the cohort of the person; $d_t$ are district fixed effects; $X_{id}$ is a vector of covariates (it includes sex, number of children, and marital status); and $μ_{id}$ is the model error term.

$$y_{id} = β_0 + β_1 Exposed_{id} + β_2 Cohort_{id} + d_t + π X_{id} + μ_{id} \quad (1)$$

**Table 1**

| Variable | Unit of measure | Control | Treatment | Mean difference |
|----------|-----------------|---------|-----------|----------------|
| Sex      | % of women      | 50.66   | 50.97     | -0.31          |
| Number of children | Average number of children | 3.02 | 3.06 | -0.04 |
| Marital status | % of married people | 72.18 | 72.89 | -0.71 |
| Years of education | Average number of completed years of education | 5.58 | 5.52 | 0.06 |
| Unemployment rate | % of unemployed people | 0.53 | 0.54 | -0.01 |

Source: own elaboration.

Note. The treatment group refers to the 1919 cohort, while the control group refers to the 1918 and 1920 cohorts. * Significant at 10%, ** significant at 5%, *** significant at 1%.
Logically, the pandemic presented geographic differences in the incidence of infections and mortality across Argentine regions. The NOA was the most affected, followed by Cuyo, while the Central region had the fewest cases in relation to its population (Graph 1). Therefore, in order to improve comparability, Eq. 1 is estimated for each region—in addition to the national total.

In order to strengthen the estimates from Eq. 1, we proceed with two complementary analyses. First, placebo tests are carried out in which people who were in utero at times other than those covered by the pandemic are considered as affected—treatment group-. No significant differences are to be found in this case.

Second, a specification of differences in differences is estimated. This allows causal inference to be made from global estimates—that is, they use the entire sample and not just the local environment of the critical threshold—by exploiting differences in the geographic incidence of the pandemic. For this, the treatment group is defined based on a double condition: that the person was born in 1919 (as in the previous case) and, in addition, that they had done so in one of the districts of the provinces most affected by the pandemic. In particular, those districts of provinces whose mortality rate from the pandemic has been at least 20 people per 10,000 inhabitants are considered as affected. This threshold is slightly higher than that used in Helgertz and Bengtsson (2019) of 17.5 people per 10,000 inhabitants.

Eq. 2 describes the difference-in-difference regression specification, \( y_{id} \) is the outcome of interest of person \( i \) of cohort \( t \) born in district \( d \); \( Exposed_{id} \) is a dummy that takes value 1 in case the person belongs to the treatment group (cohort 1919 and born in a district belonging to an affected province), and 0 if they belong to the control group; \( \beta_d \) are district fixed effects; \( X_{id} \) is a vector of covariates (it includes sex, cohort, number of children, and marital status); and \( \mu_{id} \) is the model error term.

\[
y_{id} = \beta_0 + \beta_1 Exposed_{id} + \beta_d + \pi X_{id} + \mu_{id}
\]

The main outcome of interest is the educational achievement, that is, the number of years of education completed by each person. Estimates referring to the unemployment status of each person are also provided (defined as a dummy with a value of 1 if the person is unemployed, and 0 otherwise).

Finally, before proceeding to the results section, descriptive statistics are presented for the control and treatment groups used in our main design. It can be seen that both groups are balanced.

Graphically, it is possible to observe the average educational achievement (years of education) by birth cohorts. Two different polynomials were fitted for the cohorts before and for those after the 1919 cohort. (Graph 2).

4. Results

Table 2 presents the results that arise from estimating Eq. 1. People who were in utero during the 1918 influenza pandemic present, in adulthood, a significantly lower educational achievement than their unexposed peers. This is especially true for those born in districts of the NOA region whose educational attainment is reduced by almost 0.5 years of education. The results are robust when considering different subsamples of the census microdata. In the first place, the results are robust when controlling for internal migration by using the subsample of people who lived—at the time of the census—in the same province in which they were born (Table A.1 in the Annex). Second, a negative effect is also found for the Argentine regions when considering a slightly different control group (cohorts 1917 and 1921). This suggests that the results are not explained by imprecision in identifying the exact day of birth (Table A.2 in Annex).

The reduction in educational achievement is not homogeneously distributed across the different regions of Argentina. Table 2 evidences that the NOA region is the most affected. This situation accounts for the country’s territorial inequalities: the NOA is the region that showed the highest mortality during the pandemic and it is also the one with the highest incidence of poverty in Argentina—in conjunction with the NEA- (Arévalo & Paz, 2015; González & Santos, 2020). The reduction in educational achievement is maintained when other bandwidths are considered. Table 3 provides estimates for three alternatives: 1917-1921, 1918-1921, and 1917-1920.

The results are also robust to the implementation of placebos in which the treatment group is falsely defined based on cohorts that were not exposed to the pandemic (Table 4). As expected, no significant differences are observed.

By expanding the time horizon, including all the cohorts registered in the census, by specifying a differences-in-differences model, a reduction in the educational achievement of people exposed to the pandemic in utero is also observed (Table 5). This result is maintained when considering the group of 15 provinces for which there are records of the incidence of the pandemic (Provinces), but also when estimating for the entire country (Argentina). Surprisingly, for a dimension broadly related to edu-

Graph 1. Mortality Rate From Spanish Influenza in Argentine Provinces.
Note. Mortality rate per 10,000 inhabitants. The most affected region (Noroeste) includes Catamarca, Jujuy, La Rioja, Salta, Santiago del Estero and Tucumán. Corrientes is the only province from the Noreste region. The Cuyo region includes Mendoza, San Juan and San Luis. The Centro region includes Capital Federal (or Buenos Aires City), Buenos Aires, Córdoba, Entre Ríos and Santa Fe.
Source: own elaboration based on Carbonnetti and Álvarez (2017).
Educational attainment, such as unemployment, our estimates do not denote the existence of long-lasting effects derived from exposure to the pandemic. From the census microdata, it is not possible to inquire about effects on the income dimension in which long-term effects seem more likely to exist.

The results reported here are in line with most of the previous evidence (Almond, 2006; Nelson, 2009). The reduction observed in Argentina in terms of years of education for the 1919 cohort, 0.08 (Table 2), is less than that reported in Almond (2006) with microdata from 1970 for the United States of 0.18 and 0.12 for men and women, respectively. In turn, it is slightly higher than what was found in Nelson (2009) for the case of Brazil with a reduction of 0.046 years of education. These findings differ from those reported in Vollmer and Wójcik (2017) considering binary educational indicators (1 if the person finished primary/secondary education, and 0 otherwise), whose results suggest the absence of significant effects. The differences may come from aspects different from the particularities of the case study and the intensity of the pandemic in each country.

Graph 2. Educational Achievement by Birth Cohort.
Note. Each marker represents a cohort. The vertical line refers to the 1919 cohort. Cohorts of at least 18 years of age at the time of the census are included. 95% confidence intervals.
Source: own elaboration.

Table 2
Educational Achievement and the 1918 Influenza Pandemic (Cohorts 1918–1920).

| Dependent: years of education | NOA | Cuyo | Centro | NOA + Cuyo + Centro | NOA + Cuyo + Centro | NOA + Cuyo + Centro | NOA + Cuyo + Centro | NOA + Cuyo + Centro |
|-------------------------------|-----|------|--------|---------------------|---------------------|---------------------|---------------------|---------------------|
| Treatment                    | -.4635*** | -.0809 | -.0642*** | -.0931*** | -.0787*** |
| Control variables            | Yes | Yes | Yes | Yes | Yes |
| Adj R^2                      | 0.18 | 0.18 | 0.18 | 0.18 | 0.21 |

Source: own elaboration.
Note. Robust standard errors clustered by cohort in parentheses. * significant at 10%, ** significant at 5%, *** significant at 1%. Control variables include sex, number of children, and marital status (Table A.3 in Annex).

Table 3
Educational Achievement and the 1918 Influenza Pandemic Under Different Bandwidths.

| Dependent: years of education | NOA | Cuyo | Centro | NOA + Cuyo + Centro | NOA + Cuyo + Centro | NOA + Cuyo + Centro | NOA + Cuyo + Centro | NOA + Cuyo + Centro |
|-------------------------------|-----|------|--------|---------------------|---------------------|---------------------|---------------------|---------------------|
| Cohorts 1917–1921             | -.3408*** | .0055 | -.0501* | -.0672* | -.0549* |
| Cohorts 1918–1921             | -.3742** | -.0387 | -.0398 | -.0632* | -.0540* |
| Cohorts 1917–1920             | -.4364*** | -.0813** | -.0866** | -.1126*** | -.0916*** |

Source: own elaboration.
Note. Each coefficient corresponds to a linear local regression. In all cases, those of the 1919 cohort constitute the treatment group. The 1917–1921 cohorts include 23383 observations, the 1918–1921 include 18237, and the 1917–1920 include 19114 observations. Robust standard errors clustered by cohort in parentheses.
* significant at 10%.
** significant at 5%.
*** significant at 1%.
The estimation strategies present interesting differences. While in this work a reduced bandwidth is considered (just one cohort before and one after that exposed in utero), Almond (2006) and Nelson (2009) consider longer periods to compare (1912–1922). Vollmer and Wójcik (2017) analyse an even longer period (1910–1928). This allowed them to have a greater number of observations, although at the cost of reducing the comparability between the exposed cohort (1919) and the unexposed ones. Linked to this, Almond, Nelson, and Vollmer and Wójcik (op. cit.) also adopt a quadratic specification to try to reflect the trend present in the data. In this work, instead, a linear specification was chosen as recommended in Hahn et al. (2001) and Gelman and Imbens (2014). This was possible since we decided to narrow the control group as much as possible.

In terms of unemployment, no significant effects were observed. This is consistent with Helgertz and Bengts-son (2019) where no significant worsening of occupational status was observed either, except for those exposed to the pandemic during the first trimester of pregnancy. Vollmer and Wój- cik (2017) also found no significant effect on the chances of being employed. Significant differences are reported in Nelson (2009) for those born in the first and second quarter of 1919, but not for those born in the second half of the same year. These levels of disaggregation are not feasible for the Argentine case.

5. Conclusions

Throughout this work, the existence of long-lasting effects derived from exposure to a major pandemic has been examined. In particular, the hypothesis of fetal origin (exposure in utero) has been verified in relation to the 1918 influenza pandemic for the particular case of Argentina. The results suggest a broad reduction in the educational achievement of children born in the most affected districts (Noroeste region) of almost 0.5 years of education less than their unaffected peers.

In the current context of the COVID-19 pandemic, it is relevant to inquire about the potential socio-economic consequences derived from the exposure to a biological natural disaster during.

Table 4
Placebos for Educational Achievement and the 1918 Influenza Pandemic.

|                  | 1921–1923 | 1929–1931 | 1937–1939 |
|------------------|-----------|-----------|-----------|
| Dependent: Years of education in NOA | -0.0274   | 0.0410    | 0.0548    |
| Control variables | Yes       | Yes       | Yes       |
| N                | 1142      | 1489      | 1571      |
| Adjusted R²      | 0.13      | 0.16      | 0.16      |

Source: own elaboration.

Note: The treatment group is the intermediate cohort in each case (1922, 1930, and 1938). Robust standard errors clustered by cohort in parentheses. ∗ significant at 10%, ∗∗ significant at 5%, ∗∗∗ significant at 1%.

Table 5
Educational Achievement, Unemployment and the 1918 Influenza Pandemic.

|                | Education | Unemployment |
|----------------|-----------|--------------|
|                | Provinces | Argentina    | Provinces | Argentina |
| Treatment      | -1.552*** | -1.594***    | -0.0817   | -0.0997   |
|                | (.0411)   | (.0468)      | (.1070)   | (.1040)   |
| Control variables | Yes       | Yes         | Yes       |           |
| N              | 363361    | 333573      | 144025    | 154836    |
| Adjusted R²    | 0.38      | 0.37        | 0.08      | 0.08      |

Source: own elaboration.

Note: The treatment group comprises people from the 1919 cohort born in districts of the affected provinces. Provinces include the 15 provinces for which data existed around 1919. Unemployment is estimated using a logistic regression and excludes the inactive. Robust standard errors clustered by cohort in parentheses. ∗ significant at 10%, ∗∗ significant at 5%, ∗∗∗ significant at 1%.

The results of this work are especially worrying and constitute a clear call for attention in favour of the implementation of public policies aimed at protecting children from the gestation stage. Considering the education context, Dip and Gamboa (2019) have pointed out the potential benefits of early schooling in the short and medium-term. In this sense, public policies should equally extend to this stage of children’s life.

In the first place, persistent negative consequences emerge even in a context of vast improvements in living conditions, such as that experienced by Argentina between 1880 and 1930. This is certainly not the case during the last decade (2011–2020). Second, given that we included, within the control group, those people exposed to the pandemic during their first year of life (cohort 1920), the true long-term effects of the pandemic could be underestimated. For this reason, the estimates in this work are only minimal estimates.

Third, the long-lasting negative effects do not seem to be homogeneously distributed within Argentina. The poorest districts (Noroeste region) show the greatest reduction in the educational achievement of their inhabitants. This without considering that there is no information available for the districts of the Noroeste region, with the exception of those of Corrientes, which have similar levels of poverty to those of the Noroeste. This could indicate the existence of a close inter-relationship between the material deprivations experienced by households and the long-term consequences of a pandemic. This, in turn, implies that a pandemic constitutes an environmental shock that contributes to increasing regional disparities. In this sense, recent evidence has highlighted the limitations that a situation of poverty imposes on coping with a pandemic (Paz, 2020; Levi Yeyati & Sartorio, 2020).

Some caveats should be made regarding the results of this work. For the Argentine case, the estimates are subject to the limitations associated with the main sources of information. First, mortality records are not available online and were carried out only for the provinces existing at that time (excluding national territories) -that is, for 15 of the 24 current jurisdictions-. In turn, the microdata from the 1970 census do not allow knowing some sensitive parameters for our estimates -such as the exact day of birth- or incorporating control variables that are frequently used.

In the future, it is desirable to explore the specific channels through which exposure to biological natural disasters can generate long-lasting negative effects. Two of these channels widely indicated in the literature, malnutrition and maternal stress, take on special relevance in the current context of economic crisis and obligatory social distancing. To advance in this direction of analysis, the availability and easy access to microdata for a wide population group is essential.

Acknowledgements

We are grateful to Emiliano Amaro (Agencia Crédima) and Leticia Giles (IIESS, UNS-CONICET) for technical support. We thank the anonymous reviewers and the permanent economics seminar participants at the Universidad de San Andrés for their valuable comments. This work was supported by SCyT-UNS (24/E143 and 24/E160).
### Annex

#### Table A.1

Educational Achievement and the 1918 Influenza Pandemic (Cohorts 1918–1920) When Controlling for Internal Migration.

| Dependent: years of education | NOA   | Cuyo  | Centro | NOA+Cuyo+Centro | Argentina |
|-------------------------------|-------|-------|--------|-----------------|-----------|
| Treatment                     | -2.919** | .1001 | -.081*** | -.0851*** | -.0848** |
| (0.0568)                      | (0.0556) | (0.0098) | (0.0086) | (0.0031) |
| Control variables             | Yes   | Yes   | Yes    | Yes             | Yes       |
| N                             | 754   | 598   | 6189   | 7541            | 8174      |
| Adjusted R²                   | 0.25  | 0.24  | 0.25   | 0.26            | 0.28      |

Source: own elaboration.

Note: Robust standard errors clustered by cohort in parentheses. * significant at 10%. ** significant at 5%. *** significant at 1%. Only those who lived -at the time of the census- in the same province in which they were born are included.

#### Table A.2

Educational Achievement and the 1918 Influenza Pandemic (Cohorts 1917–1921).

| Dependent: years of education | NOA   | Cuyo  | Centro | NOA+Cuyo+Centro | Argentina |
|-------------------------------|-------|-------|--------|-----------------|-----------|
| Treatment                     | -2.475*** | .0911 | -.0276 | -.0354*** | -.0275    |
| (0.0426)                      | (0.0593) | (0.0075) | (0.0072) | (0.0107) |
| Control variables             | Yes   | Yes   | Yes    | Yes             | Yes       |
| N                             | 958   | 755   | 9461   | 11174           | 12217     |
| Adjusted R²                   | 0.20  | 0.23  | 0.2    | 0.21            | 0.22      |

Source: own elaboration.

Note: Robust standard errors clustered by cohort in parentheses.

* significant at 10%.

** significant at 5%.

*** significant at 1%.

#### Table A.3

Control Variables of Table 2.

| Dependent: years of education | NOA   | Cuyo  | Centro | NOA+Cuyo+Centro | Argentina |
|-------------------------------|-------|-------|--------|-----------------|-----------|
| Sex (Men is the reference category) | -.4276 | 1.026 | .0294* | .0738           | .0639     |
| (1.0178)                      | (.6578) | (.0094) | (.0444) | (.0483) |
| Number of children            | .0016  | .0194*** | .0081** | .0084** | .0084*** |
| (0.0070)                      | (.0005) | (.0008) | (.0008) | (.0009) |
| Cohort                        | .1543*** | -.1215* | .0020 | .0071** | -.0083*** |
| (0.0086)                      | (.0370) | (.0019) | (.0010) | (.0004) |
| Marital status (Single is the reference category): | | | | | |
| Married                       | .7763*  | 1.3088* | .7020* | .7386*** | .8099**   |
| (2.246)                      | (1.3067) | (1.0000) | (0.0585) | (0.0926) |
| Consensual union              | -.1099* | -.13057*** | -.13992*** | -.13718*** | -1.2904*** |
| (3.756)                      | (0.8084) | (1.356) | (1.346) | (1.082) |
| Separated/Divorced            | -.1312 | 1.2528 | -.3496 | -.2334 | -1.083    |
| (4.448)                      | (1.4780) | (.2157) | (.2563) | (.2360) |
| Widowed                       | -.0032* | .7583   | .2360* | .2418* | .3552**   |
| (2.0693)                     | (.6688) | (.0555) | (.0425) | (.0655) |
| District fixed effects        | Available upon request to authors | | | | |

Source: own elaboration.

Note: Robust standard errors clustered by cohort in parentheses.

* significant at 10%.

** significant at 5%.

*** significant at 1%.

District fixed effects are omitted from the table for simplicity (e.g., district fixed effects include 341 dummy variables).
