Economics and Human Biology 52 (2024) 101318

Contents lists available at ScienceDirect

Economics and Human Biology

journal homepage: www.elsevier.com/locate/ehb

Socioeconomic mortality differences during the Great Influenza in Spain

Sergi Basco[a], Jordi Domènech[b], Joan R. Rosés[c,*]

[a] Departament d'Economia, Universitat de Barcelona, Spain
[b] Department of Social Sciences, Universidad Carlos III de Madrid, Spain
[c] Historical Economic Demography Group, Department of Economic History, London School of Economics and CEPR, United Kingdom

ARTICLE INFO

JEL Classification:
N34
J1
I14

Keywords:
Pandemics
Health inequality
Mortality inequities
Urban penalty

ABSTRACT

Despite being one of the deadliest viruses in history, there is limited information on the socioeconomic factors that affected mortality rates during the Great Influenza Pandemic. In this study, we use occupation-province level data to investigate the relationship between influenza excess mortality rates and occupation-related status in Spain. We obtain three main results. Firstly, individuals in low-income occupations experienced the highest excess mortality, pointing to a notable income gradient. Secondly, professions that involved more social interaction were associated with a higher excess of mortality, regardless of income. Finally, we observe a substantial rural mortality penalty, even after controlling for income-related occupational groups. Based on this evidence, it seems that the high number of deaths was caused by not self-isolating. Some individuals did not quarantine themselves because they could not afford to miss work. In rural areas, home confinement was likely more limited because their inhabitants did not have immediate access to information about the pandemic or fully understand its impact due to their limited experience handling influenza outbreaks.

1. Introduction

Mortality rates during modern pandemics are unequal (Bambra et al., 2020; Chen and Krieger, 2021; Feigenbaum et al., 2019; Turner-Musa et al., 2020). The pandemics impact different countries, regions, genders, age groups, and social classes with heterogenous levels of severity. A straightforward explanation of these differences is challenging. The timing of the pandemic's arrival and the precautionary measures adopted can explain a significant geographic variation in mortality rates (Markel et al., 2007; Simonsen et al., 2018). Some intrinsic characteristics of the affected locations like population density, climate, and pollution, can also account for these geographical patterns (Bloom-Feshbach et al., 2013; Tamerius et al., 2013; Mamelund, 2011; Clay et al., 2018, 2019; Franke, 2022). Genetic differences or previous immunization to the pandemic also shape sex and age mortality differentials (Acosta et al., 2019; Noymer and Garenne, 2000). Explaining the differences in mortality rates among social groups is an even more difficult task (Mamelund, 2017). Social-related illnesses and living conditions, such as poor housing, nutrition, and sanitation, can cause high pandemic mortality (Brown and Ravallion, 2020). During pandemic outbreaks, poor individuals are at higher risk of infection due to their inability to avoid social contact (Jay et al., 2020). Finally, some jobs, particularly those in which workers are crowded (mining, factories), or have many social interactions (medical professions, retail trade, police, and army), have higher infection rates and mortality risks than others (Creighton et al., 2022).

The main contribution of this paper is to uncover the substantial unequal mortality differentials by occupational and income groups during the Great Influenza Pandemic in Spain. Furthermore, we also show that mortality was higher in the countryside than in cities.[1] Spain is an excellent country to study the dramatic consequences of the 1918 pandemic for different reasons. First, Spanish authorities collected highly detailed data, which we use to compute excess mortality rates by occupation and province. Second, flu-driven mortality rates were very high in Spain, among the highest in Europe and North America, and displayed significant spatial differences (Basco et al., 2022). Third, the population was representative of all age groups since the country remained neutral during World War I, which implies that the young-adults mortality was not affected by war-related deaths. Finally, Spain promptly informed its population of the pandemic and

* Corresponding author.
E-mail addresses: sergi.basco@ub.edu (S. Basco), jdomenec@clio.uc3m.es (J. Domènech), J.R.Roses@lse.ac.uk (J.R. Rosés).

[1] Despite it is beyond the scope of this paper, a substantial literature has underlined the long-lasting impact of pandemics on health and socioeconomic attainment (e.g., Boberg-Fazlic et al. 2021; Cook et al. 2019; Helgertz and Bengtsson, 2019; and Arthi and Parman, 2021).

https://doi.org/10.1016/j.ehb.2023.101318
Received 15 January 2023; Received in revised form 21 September 2023; Accepted 18 November 2023
Available online 25 November 2023
1570-677X/© 2023 The Author(s). Published by Elsevier B.V. This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/).

implemented some (unsuccessful) measures to fight it, unlike belligerent countries where governments censored information (Basco et al., 2022). Furthermore, Spaniards, particularly urbanites, had substantial experience dealing with pandemic outbreaks, and the (urban) population knew that self-isolation was an effective measure against the spread of pandemics (Rodríguez-Ocaña, 1994; Pérez Moreda et al., 2015).

Early research on the Great Influenza Pandemic downplayed the significance of socioeconomic factors in determining mortality rates (Pearl, 1921; Bengtsson et al., 2018; Mamelund et al., 2021). Instead, most of the recent literature points to the existence of a link between some locations' socioeconomic characteristics and influenza-related mortality. The most common approach to assess the importance of the socioeconomic factors on influenza mortality is to collect influenza-related mortality data at low levels of disaggregation and analysing their correlation with a series of indicators collected at equally granular levels (Chowell, Erkoreka et al., 2014; Herring and Korol, 2012; Mamelund, 2006; Mamelund, 2018; Økland and Mamelund, 2019; Tuckel et al., 2006; Vaughan, 1920; and Wilson et al., 2014). The standard explanatory variables used to explain these differences are population density, illiteracy rates, homeownership rates, number of rooms per household and unemployment (Grantz et al., 2016; Mamelund, 2006, 2018; Vaughan, 1920). Clay et al. (2018, 2019) show that population health, poverty, air pollution, and the timing of onset and proximity to military bases were correlated with excess pandemic mortality in the US. Similarly, Franke (2022) finds that influenza-related mortality was correlated in Germany with poverty and air pollution. Less developed places with less affluent families tend to have higher mortality rates. Nonetheless, the causal link and the specific channel between these spatial socioeconomic indicators and mortality differentials across occupations/social classes are hard to establish (Chowell and Viboud, 2016).

To our knowledge, the recent studies of Bengtsson et al. (2018) about Sweden and Rijpma et al. (2022) about the Netherlands are the only research papers with direct evidence of occupations of the deceased. The Swedish study discovered sizable class variations in excess mortality but not a clear-cut class gradient. During the pandemic, the death rate among Swedish skilled workers was statistically much lower than that of low-skilled and unskilled workers. However, there was no statistically significant difference between low-skilled and unskilled workers' mortality, and farmers had the lowest mortality levels. The results of the Netherlands' study are more unambiguous regarding socioeconomic mortality differentials since they found a sizeable socioeconomic gradient to excess mortality among men during the 1918–19 pandemic. This study also found that mortality was higher in occupations with high levels of social contact and among farmers.

There are several previous studies on the 1918 pandemic in Spain, but none has considered the occupation mortality differentials and their causes.[2] The previous research has typically identified three waves like in other developed countries (Taubenberger and Morens, 2006): a mild one in the spring and summer of 1918, the deadly second wave in the fall of 1918, and a final wave in the winter of 1919 (Echeverri Dávila, 1985; Chowell, Erkoreka et al., 2014). The Spanish experience also shares other international traits of the 1918 influenza: women had higher mortality than men, and mortality differentials by age-group were very large, with mortality peaking for the "younger than 5" age group, the "15–34" age group and the "older than 60" age group. Finally, this literature has also identified the extraordinary spatial heterogeneity of the influenza-driven mortality within Spain (Chowell, Erkoreka et al., 2014). Moreover, each wave impacted each region with different intensity.[3] When analysing the factors driving the spatial variation in mortality differentials, authors have considered several indicators like urbanization, proxies for poverty such as child mortality or army recruits' heights, latitude to proxy for climatic effects, population density, and demographic factors, especially the share of the population in the age group hardest hit by the pandemic (the 20–40 year olds) or the share of children (Echeverri Dávila, 1985; Chowell, Erkoreka et al., 2014).[4]

Our research makes significant contributions to the ongoing discussion surrounding the 1918 Influenza Pandemic. First, we uncover exceptional excess mortality differences across occupations (excess mortality ranged from 102% for miners to 19% for landlords). Second, these differences are also substantial when we aggregate occupations for broader socioeconomic groups. The high-income group (liberal professions and landlords) had an average excess mortality rate of 29% compared to 69% in the low-income group (agriculture and mining) and 62% in the mid-income group (industry, trade, and transport).[5] Third, we employ multiple regressions to further investigate excess mortality differentials. By regressing influenza-driven mortality by occupational group at the province level with occupation dummies and province controls, we document a robust income-related ranking. Indeed, the coefficient on the low-income occupation dummy is larger than the mid-income, which, in turn, is larger than the coefficient on the high-income dummy. More importantly, the high-income dummy is smaller than the other two and the differences are statistically significant. Then, we regress influenza-driven mortality for each occupational group with province controls. We do not find that neither provincial GDPs per capita or literacy rates are correlated with mid-income or high-income occupations' excess mortality. For the low-income groups, we find that the only spatial variable with explanatory power is atmospheric pressure, which is negatively correlated with influenza-driven mortality. Fourth, our research shows that rural excess mortality rates exceeded urban ones in each occupation during the pandemic thereby leading to a substantial urban premium. We attribute the higher mortality rate in rural areas during the pandemic to a lack of home quarantine measures. Delays in receiving information about the severity of the pandemic or a lack of experience in dealing with influenza-type pandemics may exacerbate rural excess mortality rates. Finally, the previous findings along with the higher death rates of the police and the military forces, a mid-income occupation with substantial social interactions, support the hypothesis that the capacity of each social group to reduce social contact was a key factor in the mortality differences observed between groups.

The rest of the paper is organized as follows. Section 2 discusses the different sources and the various excess mortality measures. Section 3 reviews the main characteristics of the influenza-driven mortality in Spain: the pandemic's chronology, the diffusion of information during the pandemic, and the different nonpharmacological measures taken by the authorities. Section 4 discusses the heterogeneity of excess mortality across occupations and the urban premium. We also employ several regression specifications to further analyse the determinants of pandemic mortality differentials and examine several alternative hypotheses discussed in the literature. The last section concludes.

[2] An exception is the study of Bernabeu (1991) on the city of Alicante which looks at mortality differentials by occupations. However, from that local study is extremely difficult extrapolate conclusions to the rest of Spain given the substantial spatial differences in the flu impact and economic development across the country (see below).

[3] The province of the capital Madrid, for example, experienced an intense first wave and less intense second and third waves in comparison with other provinces hard hit by the second wave (Echeverri Dávila, 1985: 83–88).

[4] The contributions studying variation across provinces in Spain were complemented by a series of papers using local information (Erkoreka, 2010; Cilek et al., 2018; Bernabeu, 1991, which do not contradict previous findings.

[5] Furthermore, the differences in the age structure of the different socioeconomic groups only accounted partly for these excess mortality differentials.

2. Measuring mortality differentials

The main objective of this paper is to study the socioeconomic characteristics of pandemic mortality. This research objective shapes this study's sources and methodologies. Spanish authorities have issued death certificates since the 19th century. Municipal registers have collected and preserved these certificates, which contain relevant information on the deceased (including the cause of death and occupation). However, this source is difficult to use for nationwide research, given the dispersed nature of the registers and their different levels of preservation and public access. Furthermore, given the substantial heterogeneity in economic development across Spanish provinces, this approach would require the compilation of an extensive database covering many different and dispersed locations.

An alternative to the use of individual death certificates is to employ the official mortality statistics. Spanish government compiled national aggregate statistics (*Movimiento de la población de España*) employing direct information gathered by the provincial registers. Specifically, Spain's vital aggregate statistics give consistently comparable information since 1900. We will utilize the province-level mortality rates classified by occupations, the province capital occupational mortality rates and the province-level mortality rates by age and cause of death.

There are two main methodologies to measure pandemic mortality. The first one uses the official causes of mortality and the second infers pandemic-related mortality indirectly with the excess mortality methodology. Previous studies on Spain adopted the first methodology (Chowell, Simonsen et al., 2014; Echeverri Dávila, 1985). As mentioned above, the data on causes of death is available in Spain's Vital Statistics. However, using the official causes of death is fraught with reporting problems. During the pandemic, flu-testing technologies did not exist, therefore the cause of death was established by the external symptoms of the sick patient. To address this reporting problem, the literature typically identifies all deaths related to pulmonary illnesses as influenza related.[6] However, we cannot use official mortality causes data to study the socio-economic mortality differentials because the Spanish authorities did not produce a table that cross-classifies mortality causes and occupation. As an alternative, we have chosen to follow the excess mortality methodology. To ensure the validity of our findings, we have compared them to the official figures on mortality causes. Our measure of flu-driven excess mortality is computed using the following equation,

$$FED_i = 100 * \left[\frac{Death \quad in 1918_i}{Average \quad Death \quad Pre - 1918_i} - 1\right] \tag{1}$$

where FEDi is flu excess mortality for the interest group *i*. The average number of deaths pre-1918 is calculated as average deaths between 1911 and 1917. In our main specification *i* refers to both province and occupational group. That is, for each occupation and province, we compute the number of deaths in 1918 and the average number of deaths before 1918. Note that this measure allows us, when controlling for province characteristics, to compare mortality across occupational groups. To provide suggestive evidence on the importance of occupations, we aggregate all country deaths by occupation. That is, in this case, *i* will be just occupational group. Note that our measure of excess mortality rate accounts for potential seasonal mortality effects, previous mortality related to the same causes and controls for age differentials of mortality (within socioeconomic groups and occupations). In fact, it is akin to differences in differences regression approach.

As a comparative assessment, we compute overall excess mortality in Spain in 1918 and relate it to the alternative method using the reported causes of death. The Spanish registrars identified in 1918 a total of 264,892 lung-related deaths (147,114 influenza deaths and 117,778 deaths caused by other respiratory problems). Instead, with our method, the total figure of excess deaths is 245,406. The year before the pandemic (1917), the number of lung-related deaths was 52,212. Therefore, our estimation lies between the registrar figures of influenza and all pulmonary-related deaths (264,892 deaths) and this number of deaths minus the latter year's recorded deaths by seasonal influenza and other respiratory illnesses (264,892 - 52,212 = 212,680 deaths). In sum. the excess mortality method delivers plausible estimates of flu-related deaths.

We next show that that the main characteristics of our mortality estimation (sex and age) closely resembles those obtained with the alternative sources and other methods.

Fig. 1 reports the excess death rates by age group and gender (see Table 1A at the appendix for the actual numbers). Excess death rates follow an inverse U-shape: they were lower for children and older adults than for young people. The age-group with the higher mortality rates was the age-group from 25 to 34 years old. This age-group has an excess death rate above 200 per cent. In contrast, people above 55 years old had an excess death rate below 40 per cent. The lower excess mortality for children and older adults is consistent with previous studies on the Great Influenza Pandemic (Schoenbaum, 1996; Luk et al., 2001) and other related literature on this pandemic in Spain (Echeverri Dávila, 1985; Chowell, Erkoreka et al., 2014). Additionally, this mortality age pattern is specific to the Great Influenza Pandemic and has been utilized to trace it (Taubenberger and Morens, 2006). Existing research has argued that it was likely that older adults had already gone through previous influenza episodes and had acquired immunity to the influenza strain of the 1918 pandemic (Schoenbaum, 1996; Luk et al., 2001).

3. The development of the Great Influenza Pandemic in Spain

In this section, we will examine the timeline of the pandemic, how information on the pandemic spread, and the authorities' response to its progression. Our goal is to demonstrate that information about the pandemic extended at the same rate as the illness itself and that public measures to prevent the pandemic impact did not seem successful.

A substantial 1918 influenza literature argues the existence of three waves in this pandemic: a first wave during the summer of 1918; a second one in the fall of 1918 and a third, milder one, during the winter 1918/19 (see, for example, Chowell, Erkoreka et al., 2014, for Spain; (Pearce et al., 2011), for England and Wales; and Taubenberger and Morens, 2006, for a summary review). The following Fig. 2 shows the diffusion of the pandemic employing a variation of the Eq. (1) (in this case, *i* refers to the month).

Fig. 2 represents the overall series of monthly excess mortality. Our data confirms the presence of three waves albeit the second is several times bigger than the other two.[7] For the first wave, we can observe a small peak in June of 1918 (excess mortality was around 30 per cent). The second wave started in September, reaching a peak in October, and bottoming up in December. Quantitatively, the peak in excess mortality in October is shocking. The number of deaths more than quadrupled those commonly observed in October (to be precise, the increase was

[6] It is important to note that this approach does not fully solve the problem of identifying the mortality impact of the Great Influenza Pandemic. The world suffered deaths related to seasonal flu strain and other respiratory illnesses before the 1918 wave. Therefore, it is likely that these pulmonary illnesses caused some mortality during the Great Influenza Pandemic. Chowell, Erkoreka et al. (2014) addressed this reporting issue by computing excess mortality of the mortality causes associated with influenza and, hence, eliminated the typical annual deaths related to these causes from the final estimations.

[7] Echeverri Dávila, 1985 and Chowell, Erkoreka et al. (2014) also claim the presence of three waves in the influenza pandemic in Spain. Cilek et al. (2014) detect three waves, but their results are circumscribed to the city of Madrid. However, the Spanish evolution of the pandemic was clearly different to the English and Wales one (Pierce et al. 2010: Fig. 1), where the three waves were noticeably at first sight, with the peak in excess mortality in the third wave of around half the second.

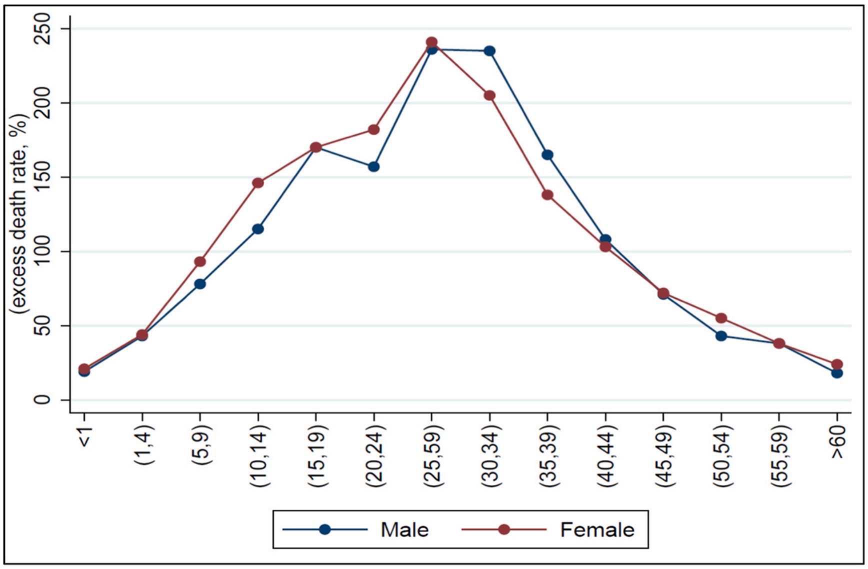

Fig. 1. Excess Mortality Rate by Age Group. Notes: Excess mortality computed with Eq. (1). Sources: Movimiento de la población de España.

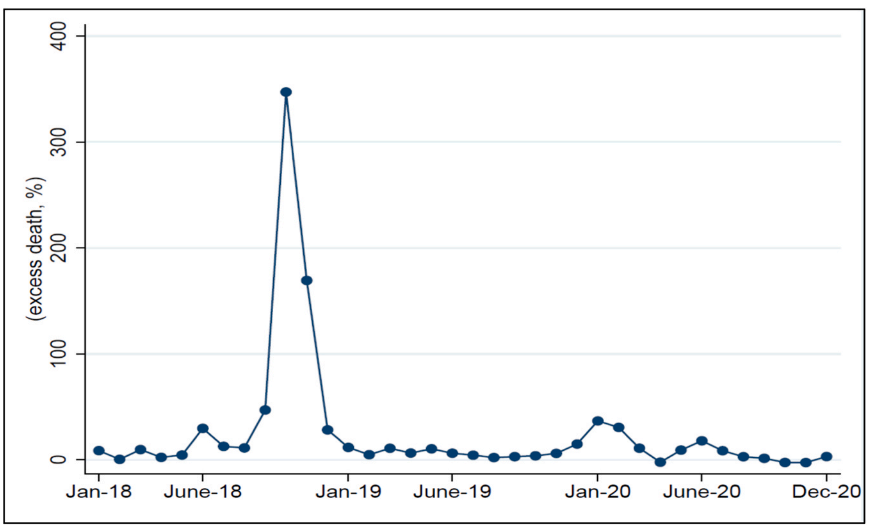

Fig. 2. Evolution of Monthly Excess Mortality. Notes and Sources: see Fig. 1.

347 per cent). Even though excess mortality in November was significantly lower than in October, it still implies that the number of deaths more than double the typical number in November (169 per cent increase). The third wave can be observed in January of 1920, but the mortality peak is much weaker smaller than the previous two. A plausible explanation for the small significant third wave in Spain is that the second fall wave was so intense and the exposure so widespread that inhabitants gained immunity protection (Barry et al., 2008; and Pierce et al., 2010).

Contrary to what happened in World War I belligerent countries, the Spanish press regularly informed on the pandemic. To illustrate the evolution of the presence of the influenza in the media, Fig. 4 displays the timeline of the number of mentions to "epidemic" and "grippe" (the French word for influenza typically used in Spain) in the leading Barcelona newspaper, *La Vanguardia*, (hereafter *LV*). This newspaper had a daily section that published news on every Spanish province.

A straightforward comparison of Figs. 2 and 3 shows a significant temporal correlation between our measure of excess mortality and news coverage on influenza. There was a mild increase in newspaper hits for these two words in April, which peaked in June 1918 at around seven times the value in January 1918. However, the press ignored the modest wave of January 1920.[8]

A question difficult to answer with the available data is what population share got access to the news published in the newspapers, given the high levels of illiteracy among the Spanish population (Núñez, 1992), and if they really understand the perils associated with the pandemic. A related issue is the low quality of the news and if they downplayed sometimes the pandemic impact (Porras Gallo, 1995). Recent research on Spanish media during the period has rebuffed the idea that the press did not influence the population's beliefs and culture

[8] We have replicated the same exercise for the Madrid-based newspaper ABC. Analogously, we find that the interest on the influenza was high between September 1918 and March 1919, peaking in October 1918. The same happens in searches for "grippe" or "gripe" in digitized Spanish newspapers for all Spain's provinces located at the on-line newspaper library of Spain's Ministry of Culture and Sports. From the first of April of 1918 to the end of March 1920, the searches yield 17,492 hits in over 150 newspapers. The number of hits is particularly high from September 1918 to March 1919, peaking in October 1918. All this evidence is available from the authors upon request.

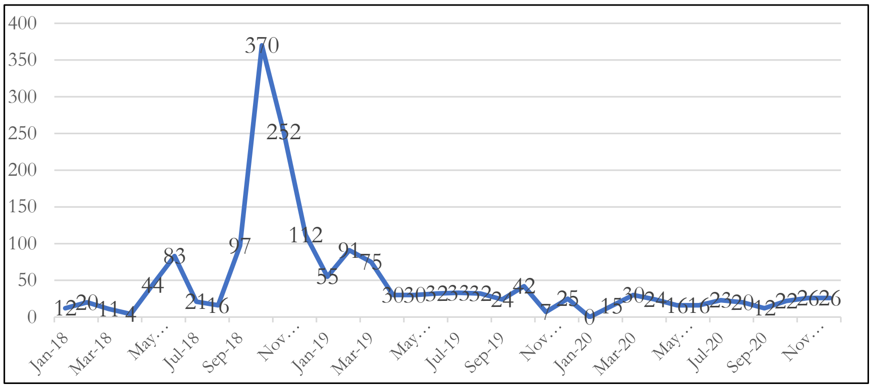

Fig. 3. Evolution of Pandemic Mentions in the Spanish News. Notes: Times the word "epidemic" and "grippe" are mentioned in LV, from January 1918 to December 1920. Source: Underlining data has been obtained from LV's online (https://www.lavanguardia.com/hemeroteca).

(Archilés and Carrión, 2012). Furthermore, the pandemic epicenter – North Castile- was the Spanish region with the highest literacy levels (see Fig. 4) and pandemic mortality was not negative correlated with literacy (see Table 4). However, previous research shows that the urban population had more experience dealing with pandemics than the rural population (Rodríguez-Ocaña, 1994; Pérez Moreda et al., 2015).

Given the general awareness on the advance of the influenza epidemic, Spanish authorities implemented some measures to limit the spread of the disease. In Spain, the adoption of measures against pandemics was decentralized: in each province, the prefect (the highest government political authority) and the health commission ("Junta Provincial Sanitaria") could officially declare the existence of a pandemic and implement different actions. The contemporary scientific understanding of the contagion channel was reasonably accurate. Moreover, the set of measures taken by the authorities reveal the existence of some prior, accumulated knowledge about how to deal with epidemics, especially in urban contexts (Rodríguez-Ocaña, 1994; Pérez Moreda et al., 2015). However, some of the implemented hygienic measures —as in other Mediterranean countries— were ineffective because they were like the ones used during Cholera outbreaks. Spanish authorities combined the correct advice to stay away from crowded, poorly ventilated locations and large crowds with Cholera-type prevention measures like the extensive application of disinfectants, the recommendation of washing of fruits and vegetables, and the closure of public restrooms (Tognotti, 2003). Councils heavily publicized these cleaning efforts to calm the population (Basco et al., 2022).

The Spanish health authorities did not introduce quarantine measures that could directly impact basic economic activities (Basco et al., 2022). There were no strict lockdowns or closures of businesses, workshops, and shops unless all workers were sick and unable to operate. Agricultural production did not appear to suffer as farmers collected harvests (Basco et al., 2021). While the authorities eventually implemented some restrictive policies, they were not in place for as long as necessary. Businesses and even the Church sometimes contested these measures (Trilla et al., 2008). Instead, authorities mainly focused on implementing social preventive measures for non-essential economic activities, such as canceling festivals and local fairs.[9] Schools and universities did not immediately reopen after the summer break, and entertainment venues like dance halls and theaters were closed.[10] The government even canceled the replacements of military conscripts to stop the spread of contagion in military quarters. In prisons, sick inmates were set apart and quarantined.[11]

The Spanish healthcare system's response was wholly insufficient. When doctors passed away or were severely ill, it was difficult to find replacements. In many remote areas, little medical care was available and health support practically disappeared during the main influenza outbreaks. Additionally, various therapies that were tried were ineffective, and some experimental vaccines also failed (Trilla et al., 2008).

The evidence collected in this section underlined two main characteristics of the pandemic in Spain: the population promptly received information on the pandemic, but the authorities did not take sufficient and consistent measures to stop its diffusion. Consequently, it fell to individuals to make the decision to self-isolate themselves or not. In the Spanish society of the early 20th Century, where a large share of the population did not have savings,[12] social isolation was an expensive choice for many.

4. Results

4.1. Descriptive evidence

We examine the potential heterogeneity of excess mortality (defined in the Eq. (1)) during the Great Pandemic across occupations and gender separately by urban and rural locations.[13] Furthermore, we also aggregate mortality in three different income groups. Table 1 reports the excess occupational mortality rates by gender and urban and rural settings.

Looking at Panel A (column 9), we observe that the occupations with the largest flu-driven excess mortality rates were workers in mining, armed forces and police, and transportation (with increased excess death rates above 80 per cent). These results seem consistent with the view that influenza contagion was higher in occupations in which people were in close contact. Mining workers had the additional

[9] *LV*, 29th September 1918, p. 17.

[10] *LV*, 29th September 1918, p. 17; *LV*, 1st October 1918, p. 12.

[11] *LV*, 19th September 1918, p. 15; *LV*, 1st October 1918, p. 12.

[12] Unfortunately, there are no available historical reconstructions of family savings in Spain. However, the available evidence on savings accounts (Martínez Soto and Cuevas Casaña, 2004) and family expenditure (Borderías et al. 2022) show that family savings were low in Spain during the period.

[13] Given the paucity of the Spanish mortality data, we use province capitals as identical to urban and the rest of the province as rural. Spanish urban penalty literature (Reher, 2001) typically uses this assumption. This assumption could introduce some bias in our results. However, we should not exaggerate our skepticism. In all Spanish provinces except one (Pontevedra), the most populated location was the province capital. Capitals had a different socioeconomic structure since they had a larger labor share at the secondary and tertiary sectors than the remaining provincial locations. In 1920, only four Spanish provinces had more than 50 per cent of their inhabitants in cities (that is, in locations of more than 25,000 people). Two of these provinces (Cádiz and Murcia) had two relatively large cities including the capital, Madrid had only one big city, but Barcelona had several industrial cities. Interestingly, none of these provinces experienced higher mortality rates during the flu pandemic.

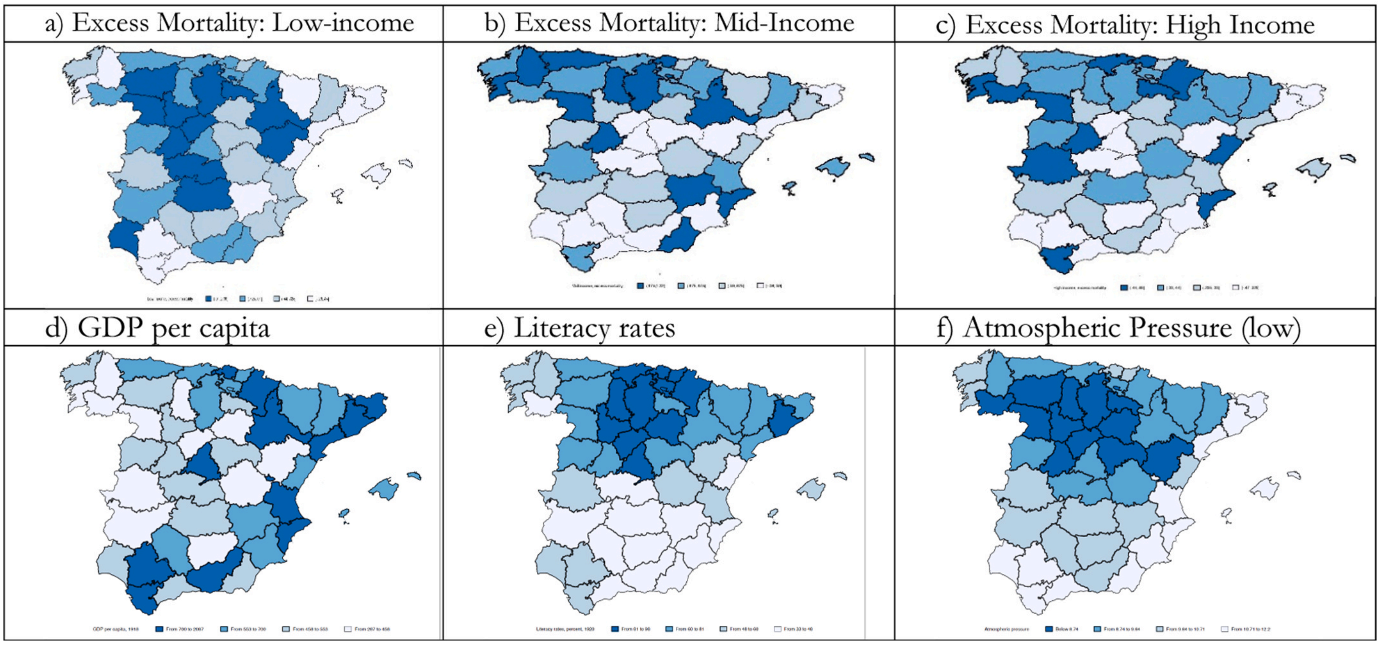

Fig. 4. Distribution of Excess Mortality by socioeconomic groups and its determinants. Sources: See Table 3.

Table 1
Excess pandemic mortality by occupation, income, location, and sex in 1918 (per cent).

	Urban			*Rural*			*Total*		
	Male (1)	*Female* (2)	*Total* (3)	*Male* (4)	*Female* (5)	*Total* (6)	*Male* (7)	*Female* (8)	*Overall* (9)
Panel A (occupations)									
(i) Agriculture	51.98	43.81	51.49	73.15	50.86	69.13	72.41	50.79	68.59
(ii) Mining	58.67		58.67	105.82		105.82	101.80		101.80
(iii) Industry	39.05	31.04	38.48	68.44	52.48	67.32	60.94	46.93	59.95
(iv) Transport	53.42		53.21	90.43		90.48	81.44		81.43
(v) Trade	6.70		3.62	57.75		50.55	42.52		37.23
(vi) Armed forces and Police	95.06		95.06	104.69		104.69	99.72		99.72
(vii) Administration	12.02		12.45	50.74		50.55	36.61		36.65
(viii) Liberal Professions	34.54	26.75	32.77	59.56	41.83	55.94	51.14	36.29	48.00
(ix) Landlords	16.95	-5.86	13.60	17.42	29.62	19.47	17.35	24.71	18.56
(x) Domestic Workers	25.98	54.00	53.87	65.05	78.87	78.84	49.65	73.51	73.45
(xi) No Occupation	47.37	62.99	47.48	62.95	104.73	65.64	57.47	102.42	59.49
(xii) Unknown/Non-productive	32.83	26.04	29.95	42.01	43.11	42.51	40.04	39.81	39.94
Overall	38.57	41.55	40.01	54.69	61.51	58.05	51.77	57.63	54.49
Panel B (income groups)									
(a) Low-income	52.17	43.81	51.68	73.51	50.86	69.47	72.76	50.79	68.92
(b) Mid-income	38.74	24.95	37.93	71.59	45.54	69.97	63.18	40.48	61.79
(c) High-income	26.41	15.81	24.37	30.13	33.95	30.81	29.31	29.66	29.37
Overall	38.30	21.67	36.55	68.12	48.41	64.87	65.24	46.81	62.30

Notes: Urban are deaths in the provincial capital while rural are deaths in the rest of the province. The Spanish literature uses the same definition of the urban and rural population (Reher, 2001). The official statistics classified children as non-productive. There are very few women working in the transport, trade sector and administration; so, their data is not presented but considered for overall calculations. Agricultural and mining workers make the low-income group. Workers in industry, trade and transport make the mid-income. Liberal professions and landlords make the high-income one. We do not consider in Panel B domestic workers, armed forces & police, no occupation, and unknown, and non-productive given uncertainty on their income levels. The distribution of occupations among income-groups is based on the underlying data of Prados de la Escosura (2008) and Rosés and Sánchez-Alonso (2004).
Sources: See text.

disadvantage of working in places with poor ventilation, where it has been shown that there is faster diffusion of the virus (Brundage and Shanks, 2008). Similarly, people employed in the armed forces and police, and transportation had to move across places and, arguably, had high exposure to the influenza. Also, the military personnel lived in barracks and slept in communal dorms, which facilitated the spread of the pandemic. In contrast, the occupation with the lowest excess mortality rates was landlords (only 18 per cent).[14] It is plausible that landlords who were aware of the dangers of the influenza and had considerable savings did not need to leave their homes to work and

[14] We have translated the Spanish word in the sources “rentistas” by landlords. Spanish sources define “rentistas” as people who mainly live on their income (rents) and excludes managers and agrarian owners-cultivators.

could maintain a high level of social distancing.

To summarize, the data presented in panel A provide strong evidence that occupations affected excess mortality. A plausible interpretation of panel A is that workers in high-income occupations had the economic means (savings) to shield themselves more effectively against the pandemic. To give further support to this interpretation, we classify occupations in three income groups: (a) low-income, (b) mid-income and (c) high-income. Panel B reports excess mortality in 1918 for each of these groups. It becomes apparent from the table that high-income occupations have substantially lower excess mortality. Quantitatively, excess mortality in 1918 was 29 per cent higher than their historical average among high-income workers. Excess mortality was higher among low- and mid-income occupations, with an extra mortality of 69 per cent and 62 per cent relative to the 1911–1917 trend, respectively.

We now discuss the gender differences in the number of excess deaths by occupation. We report these computations in Table 1, Panel A, columns 7 and 8. The goal of this exercise is twofold. Firstly, we want to examine whether females experienced different flu-driven mortality performing the same job as males. Secondly, we use the excess mortality differentials by gender as a robustness check of the result that high-income occupations had low excess mortality. It is important to note that a relevant characteristic of Spanish labor markets in the 1910s was the substantial segregation by gender: men were not typically employed in the domestic sector, while there were few women employed in mining, transport, trade, armed forces and police, and administration (Nicolau, 2005). In the case of agriculture, many active women employed in the sector were classified as not being gainfully employed or domestic workers (Gil Ibáñez, 1978).

We conclude from this exercise that higher female mortality rates were mainly due to a compositional effect. On the one hand, mortality was higher in the two "female" occupations (domestic workers and the group with no reported occupation). On the other hand, male mortality was higher in occupations with the two sexes (agricultural, industrial and liberal professions). The only exception is landlords, for which female mortality was higher than male mortality. Similarly, in Panel B, where we group occupations in three large income groups, we observe a very similar mortality pattern for male and female, which is consistent with the aggregate numbers discussed above. Females and males working in high-income occupations had substantially smaller excess mortality than those working in low- and mid-income occupations. The numbers for high-income occupations are almost the same for males and females. Excess mortality is high for males in low- and mid-income groups, which is partly due to the overrepresentation of males in some occupations.

We now turn to the descriptive analysis of the heterogeneity between urban and rural locations (Table 1, columns 1–6). When looking at the urban-rural mortality differentials, our results are very consistent: excess mortality was higher in rural areas in all different occupations. For example, among workers in the industry, rural locations had a 75 per cent higher excess mortality than cities. For the remaining occupations, the rural excess mortality was between 40 and 70 per cent higher. Given that this urban mortality advantage was not due to structural health factors (life expectancy was historically lower in the cities), it is likely that social distancing played a substantial role in these excess deaths' differentials. Furthermore, these results also cast doubt about the hypothesis that previous sanitary or health conditions were decisive for differentials in mortality rates (this result is discussed further in Section 4.3).

Next, we turn our income-group classification in Panel B. As expected, urban-rural differences in this table are lower than their counterparts in Panel A. However, we observe substantial heterogeneity: high-income occupations had a minor urban mortality advantage. One potential explanation is that workers in high-income occupations in rural and urban areas had enough resources to maintain social distancing, which protected them against influenza. Interestingly, the most substantial urban advantage is in mid-income jobs (industry, trade, and transport).

Although excess mortality varies dramatically among occupations, the observed excess mortality values can be biased due to age differences (Colvin and McLaughlin, 2021). The concern is that high-income occupations may have had lower excess mortality because their members were generally older. As previously discussed, this is because older populations had lower excess mortality compared to the younger groups during the Great Pandemic (see Fig. 1). Unfortunately, the Spanish data do not allow us to test this issue directly and we have resorted to an indirect method to infer the importance of this bias.

We proceed as follows. First, we compute, from the 1930 census (the earliest census with detailed information on occupation and age-brackets), the average number of years of workers employed in each occupation. For example, the average age of workers in liberal professions was 37.2 years. Then, we use the excess mortality that each occupation should have according to the average number of years of the workers in that occupation. Following with the same example, according to Table 1A, a person with 37.2 years had an excess flu-driven mortality rate of 151 per cent. Thus, we define the age-implied flu mortality of liberal professions as 151 per cent. Finally, we define age-related penalty as the relative change between the age-related mortality of two different occupations. For example, the age-implied flu-mortality of agricultural workers is 151 per cent (the average age of agricultural workers is 36.9 years). Thus, the age-implied penalty of agricultural workers vs liberal professions is zero.[15] The actual penalty is computed using the excess mortality reported in Table 1.

The results from Table 2 clearly show a significant difference in mortality penalties between the low-income and high-income groups after discounting age-implied mortality. The low-income male workers had an actual penalty of 135 per cent (column 6), while the age penalty for the low-income group compared to the high-income group was only 44 per cent (column 3). In sum, age differences explain about one-third of the excess mortality among low-income workers. It is also important to note that the numbers for age-related excess mortality may be biased upward as older workers tend to work in jobs with less social interaction.

To continue with this analysis in the last two rows of the Table 2, we compared the age-implied and actual mortality penalties for agricultural

Table 2
Estimation of age-related influenza mortality.

	Age-Implied penalty			*Actual Penalty*		
	Male (1)	Female (2)	Total (3)	Male (4)	Female (5)	Total (6)
Low vs High income	53%	92%	44%	146%	72%	135%
Agricultural vs Liberal	0%	0%	0%	42%	40%	43%
Mining vs Liberal	42%		45%	99%		112%

Notes: Age implied penalty is computed as: (excess mortality of an individual with average age in group i / excess mortality of an individual with average age in group $n-1$) * 100. Actual penalty is computed as: (excess mortality of group i / excess mortality of group $n-1$) * 100. Sources: Data on group average age is drawn from the 1930 census (Unfortunately, the population censuses of 1910 and 1920 do not contain information on age and occupation. However, this data is for the first time available in the 1930 population census.). Table 1A (excess death by age) at the appendix contains the data on age-related mortality.

[15] According to the census, the average years of workers in low- and high-income groups is 36.8 and 45.0 years, respectively. For mining, the average number is 33.8.

workers and miners versus liberal professions.[16] Age did not account for any difference in the mortality penalties between agricultural workers and liberal professions. In contrast, the age-implied penalty explained about 40 per cent of the difference between miners (a group working in crowed locations with higher contagion risk) and liberal professions. After discounting age-implied mortality (actual minus estimated age penalty), the difference in mortality penalties between miners and liberal professions was still higher substantial (more than 50 per cent).

4.2. Regression analysis

In this sub-section, we demonstrate with a straightforward regression that there were statistically significant differences on excess mortality rates between the occupational groups. We then analyse the determinants of excess mortality for each occupational group. To perform this exercise, we consider the following equation,

$$FED_{op} = \delta_o + \beta X_p + \varepsilon_{op} \quad (2)$$

where FED_{op} is flu-driven excess mortality in province p and occupational group o, δ_o is a set of occupational-group dummies and X_p are province-level variables. Specifically, for each province, we include GDP per capita, literacy rates and atmospheric pressure.[17] These variables should control for the income of the province, the education level and the climatic conditions which have been emphasized in the literature as having explanatory power (see Section 1).

Table 3 reports the main results. Columns 1–3 consider the overall excess mortality in the province. In the first column, we impose the restriction that there are no mortality differences by occupational group. In this case, we find that the average excess mortality across occupations is 86 per cent (the coefficient of the constant term). Notice that only atmospheric pressure is mildly significant (10 per cent). The negative sign of this coefficient implies that provinces with lower atmospheric pressure (linked with rainy and cold weather) had higher excess mortality. In column 2, we exclude the constant and include the three income-level dummies. The coefficients on these three dummies are significant and have the expected ranking. The coefficient of the low-income occupation dummy is 1.03. Therefore, it is larger than the mid-income (0.94) and the high-income (0.61) ones. When we test for the statistical differences of the coefficients (lower part of the table), we find that the coefficient on the high-income dummy is smaller than the low-income and mid-income dummy. Differences among coefficients are also statistically significant. A potential concern with this estimation is that there may still be some province-level differences which we are not controlling for. To address this concern, in column 3, we include province-fixed effects. That is, we absorb all aggregate shocks in the province which are common across occupations (for example, public awareness of the great influenza or efficacy of the local institutions in implementing health protections). The three dummy coefficients remain largely unchanged while maintaining their statistical significance and rankings: the excess mortality of high-income groups is lower than that of the other two groups. Specifically, mortality among low-income occupations is 96 per cent higher than among high-income occupations.

In columns 4–6, we reproduce the results for capitals, our measure of urban mortality. In our preferred regression, presented in column 6, we also observe the same ranking as in the overall province. In this case, only the coefficient of the low-income dummy is significant. However, outliers are driving this result: when we eliminate the outliers, the coefficients of each of the income-level dummies are all significant (see Appendix 2). Lastly, in columns 7–9, we report the results for the rural component. In our preferred regression, displayed in column 9, the coefficients of the three income-level dummies are significant and have the expected ranking. When we test if the differences in the coefficients of the income-level dummies are statistically significant, we also find that the excess mortality among high-income occupations is significant and lower than in the other groups. We perform several robustness checks in Appendix 2 to corroborate these results.

We now turn to explaining the determinants of influenza-related mortality for each occupational group. Fig. 4 shows the geographic variation of excess mortality in each occupational group and its likely determinants. The three maps at the top of the figure present evidence, respectively, on the geographical variation of low-, middle- and high-occupational group excess mortality and the three maps at the bottom show, respectively, the geographical dispersion of GDP per capita, literacy rates and atmospheric pressure as a measure of climatic differences. The maps suggest a spatial correlation between atmospheric pressure and excess mortality in low-income occupations. However, the geographical distributions of literacy and GDP per capita fit poorly the spatial variation in excess mortality. Furthermore, the spatial correlation of deaths in the different income groups is weak. Our next step is confirming these graphical interpretations with a regression analysis. To do this, we perform the following regression for each income-level group:

$$FED_p = \alpha + \beta X_p + \varepsilon_p \quad (3)$$

where FED_p is the winsorized version of flu-driven mortality in the province. In the Appendix 3, we show that the results are robust to using the non winsorized version.

Table 4 displays the coefficients for each income-level group employing the above equation. Columns 1–3 consider the total population in the province. Only the coefficient of atmospheric pressure in column 1 (low-income) is negative and statistically significant. It implies that low atmospheric pressure was correlated positively with excess mortality for workers in low-income occupations. The specification explains about 30 per cent of the variation in excess mortality in low-income workers across provinces. This result contrasts sharply with the other occupational groups where the R-squared of a similar regression is below 10 per cent. In other words, the evidence strongly suggest that climatic conditions were an important determinant of mortality only among workers in low-income occupations. In contrast, GDP per capita is not significant for any of the three income-level groups. Perhaps surprisingly, the literacy rate is positive and mildly significant (10 per cent), but with an unexpected sign (the more educated population had higher mortality) for mid-income occupations. However, the latter result is not robust when we exclude outliers from the regression.

Columns 4–6 report the same coefficients for the urban population, but we do not find any significant coefficient. Columns 7–9 report the results for the rural population. The only coefficient statistically significant is atmospheric pressure in column 7 (low-income occupations). The negative coefficient implies that rural areas with lower atmospheric pressure had higher excess mortality in low-income occupations. The coefficient (and its explanatory power) are very similar to the coefficient in column 1 because most low-income workers lived in rural areas.

4.3. The urban mortality premium

The consensus in the literature is that there was an urban mortality penalty in industrial countries until the discovery of several new treatments and massive and widespread investments in urban sanitation (e.g., Cain and Hong, 2009; Evans, 2006, and Haines, 2001). This urban penalty persisted in Spain until the Civil War, but it was not as severe as in other European countries (Reher, 2001; Ramiro and Sanz, 1999;

[16] In the other columns, we collect the same information for males and females. Note that for males, we have the same pattern. The age-implied penalty is lower than the actual penalty. For females, the age-implied penalty is higher. However, we should take these data with a grain of salt since females in low-income occupations were underrepresented (according to the 1930 Census, the share of women in low-income jobs was only 7 per cent).

[17] We have tested alternative measures of climate with similar results. This evidence is available upon request from the authors.

Table 3
The determinants of pandemic excess mortality: regressions.

Dependent var.: FED_{op}	*Total* (1)	*Total* (2)	*Total* (3)	*Urban* (4)	*Urban* (5)	*Urban* (6)	*Rural* (7)	*Rural* (8)	*Rural* (9)
Explanatory Variables									
Low-income		1.028***	0.854***		2.34	0.945**		0.995***	1.149***
		(0.250)	(0.059)		(2.094)	(0.396)		(0.264)	(0.061)
Mid-income		0.944***	0.770***		1.666	0.269		0.981***	1.135***
		(0.224)	(0.040)		(1.628)	(0.204)		(0.243)	(0.042)
High-income		0.608***	0.435***		1.571	0.175		0.583**	0.736***
		(0.227)	(0.035)		(1.626)	(0.199)		(0.243)	(0.039)
GDP per capita	-1.41	-1.41		-3.72	-3.723		-0.926	-0.926	
	(0.89)	(0.899)		(2.41)	(2.427)		(0.724)	(0.729)	
Literacy	0.245	0.245		0.269	0.269		0.245	0.245	
	(0.157)	(0.158)		(0.537)	(0.541)		(0.176)	(0.177)	
Atmospheric Pressure	-0.029*	-0.029*		-0.106	-0.106		-0.028	-0.028	
	(0.017)	(0.017)		(0.126)	(0.124)		(0.017)		
Constant	0.860***			1.86			0.853***		
	(0.230)			(1.768)			(0.2441)		
Province FE	No	No	Yes	No	No	Yes	No	No	Yes
R-squared	0.064	0.799	0.86	0.017	0.131	0.426	0.05	0.796	0.86
No Observations	144	144	144	144	144	144	144	144	144
Stat. Diff. Coefficients									
Low-Mid		0.084	0.084		0.676	0.676		0.014	0.014
		(0.078)	(0.095)		(0.494)	(0.598)		(0.081)	(0.098)
Low-High		0.420***	0.420***		0.771	0.771		0.412***	0.412***
		(0.074)	(0.089)		(0.489)	(0.593)		(0.773)	(0.093)
Mid-High		0.335***	0.335***		0.095	0.095		0.398***	0.398***
		(0.039)	(0.047)		(0.063)	(0.076)		(0.044)	(0.053)

Notes: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Standard errors in brackets. Sources: see Table 1; Atmospheric pressure is a daily average during September and October of 1918. Data obtained from Goerlich (2012). Gross Domestic Product (GDP) per capita, in real terms, in 1910 from Rosés et al. (2010). Literacy rates are defined as the share of people aged over 10 who could write and read. Data from Nuñez (1992).

Table 4
The determinants of excess mortality by income group: regression analysis.

Dependent var.: FED_p		*Total*			*Urban*			*Rural*	
	Low (1)	*Mid* (2)	*High* (3)	*Low* (4)	*Mid* (5)	*High* (6)	*Low* (7)	*Mid* (8)	*High* (9)
Explanatory Variables									
GDP per capita	-1.39	-1.66	-1.257	-5.69	0.65	-1.33	-0.91	-1.09	-0.639
	(1.12)	(1.52)	(0.814)	(2.08)	(3.48)	(2.281)	(1.15)	(1.26)	(0.984)
Literacy	0.215	0.460*	0.242	0.473	0.527	0.337	0.169	0.454	0.258
	(0.207)	(0.268)	(0.168)	(0.527)	(0.432)	(0.344)	(0.216)	(0.268)	(0.199)
Atmospheric Pressure	-0.068**	0.013	0.009	0.014	-0.016	0.015	-0.074***	0.022	0.01
	(0.027)	(0.030)	(0.027)	(0.460)	(0.419)	(0.041)	(0.027)	(0.029)	(0.032)
Constant	1.414***	0.383	0.206	0.562	0.276	0.05	1.492***	0.315	0.188
	(0.347)	(0.385)	(0.351)	(0.679)	(0.543)	(0.550)	(0.360)	(0.380)	(0.408)
R-squared	0.33	0.07	0.07	0.11	0.04	0.024	0.32	0.06	0.05
No Observations	48	48	48	48	48	48	48	48	48

Notes: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Standard errors in brackets. Sources: see Table 3.

García Gómez, 2016). Unskilled and semi-skilled urban workers were particularly affected, as seen in the heights of conscripts (Martínez-Carrión and Moreno-Lázaro, 2007). This urban penalty was mainly due to poor sanitation, inadequate housing, food quality, and harsh working conditions (Escudero and Nicolau, 2014). The infectious pandemics did not typically provoke the demise of the urban penalty (Hardt, 2016). During pandemics, urban areas could be more contagion-prone due to overcrowding, but city dwellers also had access to better information and took more precautions to limit the spread of disease (Haines, 2001). In addition, Spanish urban workers generally had higher incomes and savings than rural workers, which could allow them to maintain social distancing (Rosés and Sánchez-Alonso, 2004).

To further investigate whether the Spanish Flu reduced the urban penalty, as it seemed from on the aggregate results in Table 1, we implement the following equation:

$$\textit{Urban Excess Mortality Premium}_{ip} = \beta_1 \mathrm{Low}_i + \beta_2 \mathrm{Middle}_i + \beta_3 \mathrm{High}_i + \Delta_p + \varepsilon_{ip} \quad (4)$$

where $\textit{Urban Excess Mortality Premium}_{ip}$ is the urban excess mortality premium in income group *i* and province p. We computed this value as excess rural mortality_{ip} minus excess urban mortality_{ip}. The coefficient of each income group shows the average urban excess mortality premium in each group.

Table 5 shows that excess mortality was lower in urban areas for all occupational groups. Indeed, note that all income-group coefficients are significant and positive. The value of the coefficients does not change when weighting by population or excluding the provinces with the two largest metropolises (Madrid and Barcelona). It is worth noting that the middle-income group has the largest coefficient in our findings. We run a further regression with the three occupations forming the middle-income group (see Appendix 4). Industry and trade coefficients are similar, but the transport one is lower. However, the difference among them is not statistically significant.

What could explain this urban mortality premium during the Great Influenza? Economic factors like income and savings do not seem to be

Table 5
The urban excess mortality premium: regression analysis.

Dependent var.: *Urban Mortality Premium$_{ip}$*	*Baseline*	*Population Weighted*	*No Madrid and Barcelona*
	(1)	(2)	(3)
Explanatory Variables			
Low-income	0.296***	0.296***	0.289***
	(0.674)	(0.672)	(0.069)
Mid-income	0.555***	0.560***	0.559***
	(0.050)	(0.043)	(0.051)
High-income	0.318***	0.313***	0.321***
	(0.044)	(0.044)	(0.046)
Province FE	Y	Y	Y
R-squared	0.538	0.549	0.535
No Observations	144	144	138

Notes: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Standard errors in brackets. Province clustered standard errors. This is the winsorized regression (non winsorized version is available upon request from the authors). Sources: see Table 4.

the cause, as the coefficients for low and high-income groups were similar. Delays in pandemic information and lack of pandemic experience in rural areas may have led to a higher death rate. The higher premium in the mid-income group was also the result of these information factors and some intrinsic characteristics of rural mid-income jobs. For example, in the trade sector, street markets and itinerant traders were common in the small towns and villages, while shops and closed markets were ubiquitous in the Spanish cities (Mirás Araujo, 2008; Nielfa, 1985). Therefore, personal contact was harder to avoid in rural than urban trade.

4.4. *Discussing alternative hypotheses*

The literature on the urban penalty has emphasized the correlation between urbanization rates and mortality, including those linked to pandemics, and that the most populated cities had higher mortality rates (Cain and Hong, 2009). Moreover, given that aerosol transmission is predominant in influenza, it is likely that, *ceteris paribus*, the denser locations (most populated cities) had higher mortality rates. According to our previous results, the correlation between urbanization and higher pandemic mortality does not hold for the Great Influenza in Spain. We will now test the second part of the hypothesis to see if cities with larger populations have higher mortality rates during the Great Influenza Pandemic.

Fig. 5 reports the correlation between province capitals' excess mortality and its population in 1910 (the latest census available before the Great Influenza). Colors blue, red, and green represent occupational groups with low, middle, and high incomes, respectively. The correlation between excess mortality and urban population is weak but, if anything, negative for each occupational group. Thus, we can rule out the hypothesis that more crowded cities experienced higher mortality rates.

Another popular theory links increased mortality rates during the influenza outbreak with industrialization levels. A version of this argument associated industrial air pollution with exceptional pandemic mortality. However, our previous evidence on urban excess mortality during the Great Influenza contradicts this hypothesis since a substantial part of the Spanish industry was in cities. Fig. 6 provides additional suggestive evidence reporting the correlation between urban excess mortality and the levels of industrialization in 1910 (the closest year with a sectoral breakdown of the working population before the pandemic), measured as the per centage of the population working in manufacturing in each province. As can be seen, there is no clear correlation between industrialization rates and excess mortality for any of the occupational groups. Compared to the previous Fig. 5 (urban population), the correlation is even smaller (in absolute terms).

5. Concluding remarks

In this paper, we have documented the unequal effects of pandemics on mortality across socioeconomic dimensions. This question has received renewed interest after the outburst of Covid-19. We have focused on the mortality consequences of the 1918 Great Influenza in Spain (1918). By using a historical episode, we have the advantage of examining a fully completed event. Also, Spain in 1918 resembles today's developing economy in which social distancing is difficult due to low incomes and poverty, which can help us to extrapolate our results to developing countries in the Covid-19 pandemic.

Our main result is that the excess mortality rate substantially varied

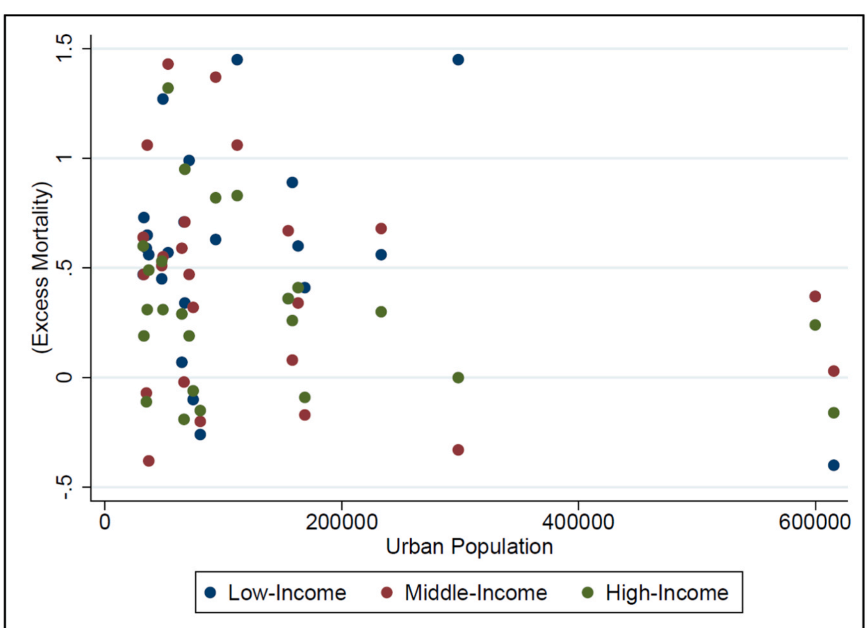

Fig. 5. Correlations between socioeconomic groups excess mortality and urban population. Notes and Sources: Excess mortality is from Table 2 and the province capital's population data from the Population Census of 1910.

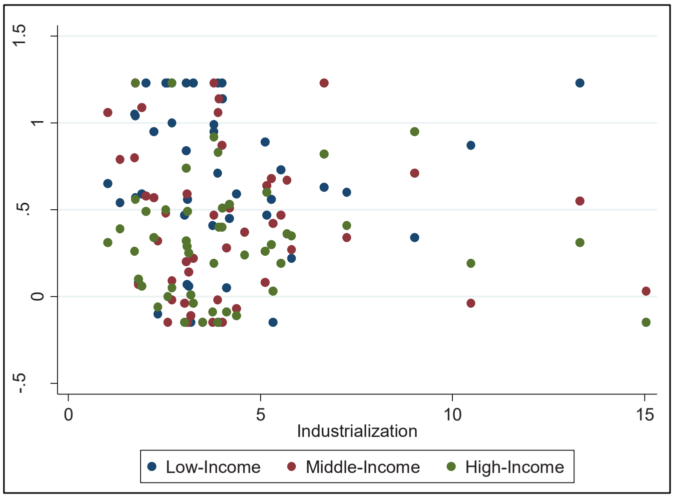

Fig. 6. Correlations between socioeconomic groups excess mortality and industrialization. Notes and Sources: Excess mortality is from Table 2 and Industrial employment from the Population Census of 1910.

across occupations with a sizeable income gradient. One interpretation is that whereas workers with middle and high-income could keep social distancing, less affluent workers could not and continued their normal daily activities. When analysing the determinants of excess mortality, we find that climate is the only explanatory variable, albeit mildly significant, for overall mortality (mainly due to the low-income workers effect). We also document an urban mortality premium: excess mortality was lower in the provincial capitals for all occupations and income-levels. Given that this urban premium was prevalent among all socio-economic groups, the likely cause was delays in information and the absence of experience dealing with pandemics in rural locations.

What is the implication of this study for the Covid-19 crisis? Access to information, health measures and medical knowledge have dramatically improved compared to 1918. Furthermore, the 1918 influenza has a substantially different age-related component since the most affected were the young adults. However, a plausible interpretation of our findings is that pandemic awareness and the ability to keep social distancing were critical to reduce pandemic-driven mortality. Therefore, fully informed individuals who could avoid working outside their home experienced much lower mortality rates. This result strongly resonates with the current socioeconomic differences in Covid-related mortality.

Declaration of Competing Interest

None.

Data availability

Data will be made available on request.

Appendix 1. Excess mortality data

Table 1A
Excess death by age and gender.

Age	Gender	Excess Death Rate	Excess Death Absolute	Contribution to total Excess Death Rate
<1	Male	19%	9878	8%
	Female	21%	8624	7%
(1, 4)	Male	43%	16,717	14%
	Female	44%	16,278	13%
(5, 9)	Male	78%	5983	5%
	Female	93%	7320	6%
(10, 14)	Male	115%	4132	4%
	Female	146%	5868	5%
(15, 19)	Male	170%	8376	7%
	Female	170%	9000	7%
(20, 24)	Male	157%	9122	8%
	Female	182%	11,348	9%
(25, 29)	Male	236%	11,982	10%
	Female	241%	13,665	11%
(30, 34)	Male	235%	11,454	10%
	Female	205%	11,968	9%
(35, 39)	Male	165%	8441	7%
	Female	138%	7724	6%
(40, 44)	Male	108%	6527	6%
	Female	103%	6020	5%
(45, 49)	Male	71%	4777	4%
	Female	72%	3900	3%
(50, 54)	Male	43%	3733	3%
	Female	55%	3857	3%
(55, 59)	Male	38%	3613	3%
	Female	38%	2961	2%
over 60	Male	18%	12,680	11%
	Female	24%	18,267	14%
Unknown	Male	48%	357	0%
	Female	56%	278	0%
Total	**Male**	**51%**	**117,774**	
	Female	**57%**	**127,078**	

Source: See Fig. 1.

Appendix 2. Robustness check of Table 3

Table 2A performs robustness checks to our preferred specification (including province fixed effects). A first concern could be that our results are driven by the relatively low mortality in the provinces of Barcelona and Madrid. Column 1 shows that the occupation dummies for total mortality barely change when excluding these provinces. A second concern is that there may be some small provinces with very extreme values that are driving the results. Column 2 performs a population weighted regression and show that the quantitative results hold. A related concern is the presence of outliers specially when separating capital and rural population. Columns 3–5, report the coefficients of interest when eliminate the outliers (10 per cent). For overall mortality (column 3) we obtain the same ranking and significance but, as it could be expected, the difference between groups diminishes a bit. For example, mortality among workers in low-income occupations is 67 per cent larger than in high-income. The most important difference is in urban mortality. In this case, we find that the three dummies are positive and significant and the difference between the high dummy and the rest is significant. In the capitals of province, workers in low-income occupation had a 93% larger mortality than workers in high-income occupations. Lastly, column 6 reports the results for the rural population. As it can be seen all results hold.

Table 2A
The Determinants of Pandemic Excess Mortality: Regressions

Dependent var.: FED_{op}	*Total No Madrid & BCN* (1)	*Total Popul. Weighted* (2)	*Total 10 W* (3)	*Urban 10 W* (4)	*Rural 10 W* (5)
Explanatory Variables					
Low-income	0.853***	0.829***	0.804***	0.632***	0.956***
	(0.062)	(0.053)	(0.060)	(0.084)	(0.081)
Mid-income	0.773***	0.784***	0.781***	0.429***	0.992***
	(0.042)	(0.038)	(0.060)	(0.082)	(0.081)
High-income	0.433***	0.447***	0.474***	0.328***	0.641***
	(0.037)	(0.035)	(0.059)	(0.082)	(0.077)
Province FE	Yes	Yes	Yes	Yes	Yes
R-squared	0.86	0.87	0.92	0.79	0.92
No Observations	138	144	144	144	144
SDC					
Low-Mid	0.08	0.045	0.023	0.203**	-0.036
	(0.099)	(0.086)	(0.050)	(0.080)	(0.051)
Low-High	0.420***	0.380***	0.331***	0.304***	0.315**
	(0.093)	(0.081)	(0.045)	(0.071)	(0.047)
Mid-High	0.34***	0.336***	0.308***	0.101	0.351***
	(0.049)	(0.050)	(0.041)	(0.065)	(0.043)

Notes: See text.

Appendix 3. Robustness check of winsorized regressions

Fig. 1A reproduces the coefficients of the occupational dummies for the three population samples with the winsorized measure. We want to emphasize three facts. First, there is a clear ranking on the occupation dummies. Second, mortality among workers in high-income occupations is statistically smaller than in the other two. Third, excess mortality in the capital (urban) is lower than in rural locations.

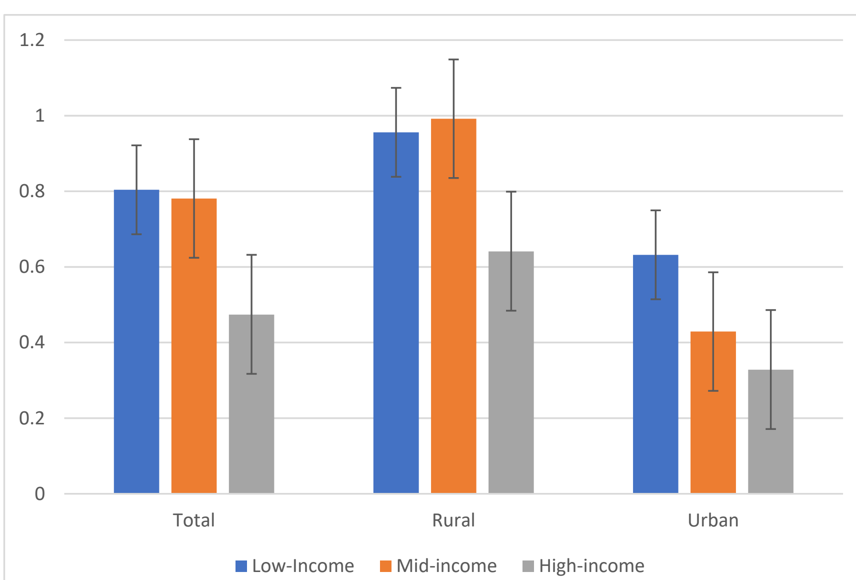

Fig. 1A. Excess mortality across occupations.

.

Appendix 4. Mid socioeconomic group: urban excess mortality

Table 3A
Urban excess mortality premium: mid socioeconomic group.

Dependent var.: *Urban Mortality Premium$_{ip}$*	*Baseline*	*No Madrid and Barcelona*
	(1)	(2)
Explanatory Variables		
Industry	1.135***	1.139***
	(0.109)	(0.114)
Transport	0.704***	0.678***
	(0.188)	(0.194)
Trade	1.041***	1.063***
	(0.191)	(0.199)
Province FE	Y	Y
R-squared	147	141
No Observations	0.388	0.391

Notes: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Standard errors in brackets. Province clustered standard errors. This is the winsorized regression (non winsorized version is available upon request from the authors). Sources: see Table 4.

References

Acosta, E., Hallman, S.A., Dillon, L.Y., Ouellette, N., Bourbeau, R., Herring, D.A., Inwood, K., Earn, D.J., Madrenas, J., Miller, M.S., Gagnon, A., 2019. Determinants of influenza mortality trends: age-period-cohort analysis of influenza mortality in the United States, 1959–2016. Demography *56* (5), 1723–1746.

Archilés, F., Carrión, M.G., 2012. En la sombra del Estado. Esfera pública nacional y homogeneización cultural en la España de la Restauración. Hist. Contemp. 45, 483–518.

Arthi, V., Parman, J., 2021. Disease, downturns, and wellbeing: economic history and the long-run impacts of COVID-19. Explor. Econ. Hist. 79, 101381.

Bambra, C., Riordan, R., Ford, J., Matthews, F., 2020. The COVID-19 pandemic and health inequalities. J. Epidemiol. Community Health *74* (*11*), 964–968.

Barry, J., Viboud, C., Simonsen, L., 2008. Cross-protection between successive waves of the 1918–1919 influenza pandemic: epidemiological evidence from US Army camps and from Britain. J. Infect. Dis. *198*, 1427–1434.

Basco, S., Domènech, J., Rosés, J.R., 2021. The redistributive effects of pandemics: evidence from the Spanish flu. World Dev. *141*.

Basco, S., Domènech, J., Rosés, J.R., 2022. Pandemics, Economics and Inequality: Lessons from the Spanish Flu. Palgrave Macmillan.

Bengtsson, T., Dribe, M., Eriksson, B., 2018. Social class and excess mortality in Sweden during the 1918 influenza pandemic. Am. J. Epidemiol. 187 (12), 2568–2576.

Bernabeu, J., 1991. La Ciutat davant el Contagi. Alacant i la Grip de 1918-19. València: Generalitat Valenciana. Conselleria de Sanitat i Consum.

Bloom-Feshbach, K., Alonso, W.J., Charu, V., Tamerius, J., Simonsen, L., Miller, M.A., Viboud, C., 2013. Latitudinal variations in seasonal activity of influenza and respiratory syncytial virus (RSV). A global comparative review. PLoS One *8* (2), e54445.

Boberg-Fazlic, N., Ivets, M., Karlsson, M., Nilsson, T., 2021. Disease and fertility: evidence from the 1918–19 influenza pandemic in Sweden. Econ. Hum. Biol. 43, 101020.

Borderías, C., Muñoz-Abeledo, L., Cussó Segura, X., 2022. Breadwinners in Spanish cities (1914-1930). Rev. Hist. Ind. Ind. Hist. Rev. 31 (84), 59–98.

Brown, C.S., Ravallion, M., 2020. Inequality and the Coronavirus: Socioeconomic Covariates of Behavioral Responses and Viral Outcomes Across US Counties (NBER Working Papers 27549). Cambridge: National Bureau of Economic Research.

Brundage, J.F., Shanks, G.D., 2008. Deaths from bacterial pneumonia during 1918–19 influenza pandemic. Emerg. Infect. Dis. *14* (*8*), 1193.

Cain, L., Hong, S.C., 2009. Survival in 19th century cities: the larger the city, the smaller your chances. Explor. Econ. Hist. 46 (4), 450–463.

Chen, J.T., Krieger, N., 2021. Revealing the unequal burden of COVID-19 by income, race/ethnicity, and household crowding: US county versus zip code analyses. J. Public Health Manag. Pract. *27* (*1*), S43–S56.

Chowell, G., Viboud, C., 2016. Pandemic influenza and socioeconomic disparities: lessons from 1918 Chicago. Proc. Natl. Acad. Sci. USA *113*, 13557–13559.

Chowell, G., Erkoreka, A., Viboud, C., Echeverri-Dávila, B., 2014. Spatial-temporal excess mortality patterns of the 1918–1919 influenza pandemic in Spain. BMC Infect. Dis. *14*, 371.

Chowell, G., Simonsen, L., Flores, J., Miller, M.A., Viboud, C., 2014. Death patterns during the 1918 influenza pandemic in Chile. Emerg. Infect. Dis. *20* (11), 1803.

Cilek, L., Chowell, G., Ramiro, D., 2018. Age-specific excess mortality patterns during the 1918-1920 influenza pandemic in Madrid, Spain. Am. J. Epidemiol. 187 (12), 2511–2523.

Clay, K., Lewis, J., Severnini, E., 2018. Pollution, infectious disease, and mortality: evidence from the 1918 Spanish influenza pandemic. J. Econ. Hist. 78 (4), 1179–1209.

Clay, K., Lewis, J., Severnini, E., 2019. What explains cross-city variation in mortality during the 1918 influenza pandemic? Evidence from 438 US cities. Econ. Hum. Biol. *35*, 42–50.

Colvin, C.L., McLaughlin, E., 2021. Death, demography and the denominator: age-adjusted Influenza-18 mortality in Ireland. Econ. Hum. Biol. 41, 100984.

Cook, C.J., Fletcher, J.M., Forgues, A., 2019. Multigenerational effects of early-life health shocks. Demography 56 (5), 1855–1874.

Creighton, M., Capistrano, D., Sorokowska, A., et al., 2022. Physical spacing and social interaction before the global pandemic. Spat. Demogr. 10, 107–116.

Echeverri Dávila, B., 1985. La gripe española. La pandemia de 1918-1919. Madrid: Centro de Investigaciones Sociológicas-Siglo XXI.

Erkoreka, A., 2010. The Spanish influenza pandemic in occidental Europe (1918-1920) and victim age. Influenza Other Respir. Virus 4 (2), 81–89.

Escudero, A., Nicolau, R., 2014. Urban penalty: nuevas hipótesis y caso español (1860-1920). Historia social, pp. 9–33.

Evans, R., 2006. Death in Hamburg. Society and politics in the Cholera Years 1830-1910. Penguin, London.

Feigenbaum, J., Muller, C., Wrigley-Field, E., 2019. Regional and racial inequality in infectious disease mortality in U.S. cities, 1900–1948. Demography 56, 1371–1388.

Franke, R., 2022. Poverty, pollution, and mortality: the 1918 influenza pandemic in a developing German economy. Econ. Hist. Rev. 75 (4), 1026–1053.

García Gómez, J.J., 2016. Urban penalty en España: el caso de Alcoy (1857-1930). Rev. Hist. Ind. *25*, 49–78.

Gil Ibáñez, S, 1978. Un intento de homogeneización de las clasificaciones profesionales en España (1860-1930). Rev. Int. Sociol. *25*, 7–40.

Goerlich, F.J., 2012. Datos climáticos históricos para las regiones españolas. CRU TS 2.1, Investigaciones de Historia Económica, 8(1), pp. 29-40.

Grantz, K.H., Rane, M.S., Salje, H., Glass, G.E., Schachterle, S.E., Cummings, D.A., 2016. Disparities in influenza mortality and transmission related to sociodemographic factors within Chicago in the pandemic of 1918. Proc. Natl. Acad. Sci. *113*, 13839–13844.

Haines, M.R., 2001. The urban mortality transition in the United States, 1800-1940. Ann. Démogr. Hist. (1), 33–64.

Hardt, M.D., 2016. History of Infectious Disease Pandemics in Urban Societies. Lanham: Lexington Books.

Helgertz, J., Bengtsson, T., 2019. The long-lasting influenza: the impact of fetal stress during the 1918 influenza pandemic on socioeconomic attainment and health in Sweden, 1968–2012. Demography *56* (*4*), 1389–1425.

Herring, D.A., Korol, E., 2012. The north-south divide: social inequality and mortality from the 1918 influenza pandemic in Hamilton, Ontario. In: Fahrni, M., Jones, E.W., (Eds), Epidemic Encounters: Influenza, Society, and Culture in Canada. Vancouver, BC: UBC Press, pp. 97–112.

Jay, J., Bor, J., Nsoesie, E.O., Lipson, S.K., Jones, D.K., Galea, S., Raifman, J., 2020. Neighbourhood income and physical distancing during the COVID-19 pandemic in the United States. Nat. Hum. Behav. *4* (*12*), 1294–1302.

Luk, L., Gross, P., Thompson, W.W., 2001. Observations on mortality during the 1918 influenza pandemic. Clin. Infect. Dis. 33, 1375–1378.

Mamelund, S.E., 2006. A socially neutral disease? Individual social class, household wealth and mortality from Spanish influenza in two socially contrasting parishes in Kristiania 1918–19. Soc. Sci. Med. 62, 923–940.

Mamelund, S.E., 2011. Geography may explain adult mortality from the 1918–20 influenza pandemic. Epidemics 3, 46–60.

Mamelund, S.E., 2017. Social inequality – a forgotten factor in pandemic influenza preparedness. J. Nor. Med. Assoc. *13*, 911–913.

Mamelund, S.E., 2018. 1918 Pandemic morbidity: the first wave hits the poor, the second wave hits the rich. Influenza Other Respir. Virus 12 (3), 307–313.

Mamelund, S.E., Shelley-Egan, C., Rogeberg, O., 2021. The association between socioeconomic status and pandemic influenza: systematic review and meta-analysis. PLOS One 16 (9), e0244346.

Markel, H., Lipman, H.B., Navarro, J.A., Sloan, A., Michalsen, J.R., Stern, A.M., Cetron, M.S., 2007. Nonpharmaceutical interventions implemented by US cities during the 1918-1919 influenza pandemic. JAMA 298 (6), 644–654.

Martínez Soto, A.P., Cuevas Casaña, J., 2004. La expansión y consolidación de las Cajas de Ahorros en el sistema financiero español, 1880–1936. Rev. Hist. Econ. J. Iber. Lat. Am. Econ. Hist. 22 (1), 65–110.
Martínez-Carrión, J.M., Moreno-Lázaro, J., 2007. Was there an urban height penalty in Spain, 1840–1913? Econ. Hum. Biol. 5, 144–164.
Mirás Araujo, J., 2008. The commercial sector in an early-twentieth century Spanish City, La Coruña 1914-1935. J. Urban Hist. 34 (3), 458–483.
Nicolau, R., 2005. Población, salud y actividad. In Estadísticas históricas de España: siglo XIX-XX. Madrid: Fundación BBVA, pp. 77–154.
Nielfa, G., 1985. Los sectores mercantiles en Madrid en el primer tercio del siglo XX. Tiendas, comerciantes y dependientes de comercio. Madrid: Ministerio de Trabajo y Seguridad Social.
Noymer, A., Garenne, M., 2000. The 1918 influenza epidemic's effect on sex differentials in mortality in the United States. Popul. Dev. Rev. 26 (3), 565–581.
Núñez, C.E., 1992. La fuente de la riqueza. Educación y desarrollo económico en la España contemporánea. Madrid: Alianza Universidad.
Økland, H., Mamelund, S.E., 2019. Race and 1918 influenza pandemic in the United States: a review of the literature. Int. J. Environ. Res. Public Health 16 (14), 2487.
Pearce, D.C., Pallaghy, P.K., McCaw, J.M., McVernon, J., Mathews, J.D., 2011. Understanding mortality in the 1918–1919 influenza pandemic in England and Wales. Influenza Other Respir. Virus *5* (2), 89–98.
Pearl, R., 1921. Influenza studies: further data on the correlation of explosiveness of outbreak of the 1918 epidemic. Public Health Rep. 36, 273–298.
Pérez Moreda, V., Reher, D., Sanz Gimeno, A., 2015. La conquista de la salud. Mortalidad y modernización en la España contemporánea. Madrid: Marcial Pons.
Porras Gallo, M.I., 1995. La prensa madrileña de información general ante la epidemia de gripe de 1918-19. Med. Hist. Rev. Estud. Hist. Hist. Cienc. Méd. 57, 1–16.
Prados de la Escosura, L., 2008. Inequality, poverty and the Kuznets curve in Spain, 1850–2000. Eur. Rev. Econ. Hist. 12 (3), 287–324.
Ramiro, D., Sanz, A., 1999. Cambios estructurales en la mortalidad infantil y juvenil en España, 1860-1930. Bol. ADE 17, 40–87.
Reher, D.S., 2001. In search of the 'urban penalty': exploring urban and rural mortality patterns in Spain during the demographic transition. Int. J. Popul. Geogr. *7*, 105–127.
Rijpma, A., Van Dijk, I.K., Schalk, R., Zijdeman, R.L., Mourits, R.J., 2022. Unequal excess mortality during the Spanish Flu pandemic in the Netherlands. Econ. Hum. Biol. 47, 101179.
Rodríguez-Ocaña, E., 1994. La salud pública en España en el contexto europeo, 1890–1925. Rev. Sanid. Hig. Pública *68*, 11–27.
Rosés, J.R., Sánchez-Alonso, B., 2004. Regional wage convergence in Spain 1850–1930. Explor. Econ. Hist. *41*, 404–425.
Rosés, J.R., Martínez-Galarraga, J., Tirado, D.A., 2010. The upswing of regional income inequality in Spain (1860–1930). Explor. Econ. Hist. *47* (2), 244–257.
Schoenbaum, S.C., 1996. Impact of Influenza in Persons and Populations. Options for the Control of Influenza III. Elsevier, Amsterdam, pp. 17–25.
Simonsen, L., Chowell, G., Andreasen, V., Gaffey, R., Barry, J., Olson, D., Viboud, C., 2018. A review of the 1918 herald pandemic wave: importance for contemporary pandemic response strategies. Ann. Epidemiol. *28* (5), 281–288.
Tamerius, J.D., Shaman, J., Alonso, W.J., Bloom-Feshbach, K., Uejio, C.K., Comrie, A., Viboud, C., 2013. Environmental predictors of seasonal influenza epidemics across temperate and tropical climates. PLoS Pathog. *9* (3), e1003194.
Taubenberger, J., Morens, D., 2006. 1918 Influenza: the mother of all pandemics. Emerg. Infect. Dis. *12*, 15–22.
Tognotti, E., 2003. Scientific triumphalism and learning from facts: bacteriology and the 'Spanish flu' challenge of 1918. Soc. Hist. Med. 16 (1), 97–110.
Trilla, A., Trilla, G., and Daer, C. (2008). The 1918 "Spanish flu" in Spain. *Clinical infectious diseases*, 47 (5): 668-673.
Tuckel, P., Sassler, S., Maisel, R., Leykam, A., 2006. The diffusion of the influenza pandemic of 1918 in Hartford, Connecticut. Soc. Sci. Hist. 30 (2), 167–196.
Turner-Musa, J., Ajayi, O., Kemp, L., 2020. Examining social determinants of health, stigma, and COVID-19 disparities. Healthcare 8 (2), 168.
Vaughan, W.T., 1920. Influenza. An epidemiological study. Am. J. Hyg.
Wilson, N., Oliver, J., Rice, G., Summers, J.A., Baker, M.G., Waller, M., Shanks, G.D., 2014. Age-specific mortality during the 1918–19 influenza pandemic and possible relationship to the 1889–92 influenza pandemic. J. Infect. Dis. 210 (6), 993–995.

Sources

ABC. *Archivo ABC*. ⟨https://www.abc.es/archivo/periodicos/⟩.
Dirección General del Instituto Geográfico y Estadístico. Censo de población de 1910, Madrid.
Dirección General del Instituto Geográfico y Estadístico. Censo de población de 1920, Madrid.
Dirección General del Instituto Geográfico y Estadístico. Censo de población de 1930, Madrid.
Dirección General del Instituto Geográfico y Estadístico. Movimiento natural de la población española, several years, Madrid.
La Vanguardia. *Hemeroteca Histórica*: ⟨https://www.lavanguardia.com/hemeroteca⟩.
Ministerio de Cultura y Deportes. *Biblioteca virtual de Prensa Histórica*: ⟨https://prensahistorica.mcu.es/es/consulta/busqueda.do⟩.