¿Lo que no mata, engorda? The Impact of the 1918 Spanish Flu Epidemic on Economic Performance in Sweden[‡]

Martin Karlsson
Technische Universität Darmstadt[*]

Therese Nilsson
Lund University[†]

Stefan Pichler
Technische Universität Darmstadt[**]
Goethe University Frankfurt[††]

October 14, 2011

[‡]We would like to thank seminar participants of Bocconi University for excellent comments, and Ingrid Magnusson and Amanda Stefansdotter for excellent research assistance. We take responsibility for all remaining errors in and shortcomings of the paper.

[*]Corresponding author: Technische Universität Darmstadt(TU Darmstadt), Marktplatz 15 - Residenzschloss, 64283 Darmstadt, e-mail: karlsson@vwl.tu-darmstadt.de

[†]Department of Economics, Lund University, and Research Institute of Industrial Economics (IFN) e-mail: therese.nilsson@nek.lu.se

[**]Technische Universität Darmstadt(TU Darmstadt), Marktplatz 15 - Residenzschloss, 64283 Darmstadt, e-mail: pichler@vwl.tu-darmstadt.de

[††]Goethe University Frankfurt, Grüneburgplatz 1, 60323 Frankfurt

Abstract

We study the impact of the 1918 influenza pandemic on economic performance in Sweden. The pandemic was one of the severest and deadliest pandemics in human history, but is has hitherto received only scant attention in the literature – despite important implications for modern-day pandemics. In this paper, we exploit seemingly exogenous variation in incidence rates between Swedish regions to estimate the impact of the pandemic. Using difference-in-differences and high-quality administrative data from Sweden, we estimate the effects on earnings, capital returns, poverty and manufacturing production. We find that the pandemic lead to a significant increase in poverty rates. There is also relatively strong evidence that capital returns were negatively affected by the pandemic. On the other hand, we find robust evidence that the influenza had no discernible effect on earnings. This finding is surprising since it goes against most previous empirical studies as well as theoretical predictions.

Keywords: Spanish Flu; Difference-in-Differences.

JEL classification: I18

1 Introduction

In 1918 the world is severely hit by a first wave of the very contagious Spanish flu. Estimates indicate that 500 million individuals worldwide were infected by the deadly virus, and that 50-100 million people died in the aftermaths of an infection between 1918 and 1920 (Johnson and Mueller, 2002). Unlike when customary strains of influenza periodically circulate the world, the majority of the victims of the Spanish flue were healthy young people in the age of 15-40 – not weakened patients, nor children or elderly.

While much has been written and argued about the medical causes of the Spanish flu, the origins of the deadly virus, and its connection to the 2006 bird flu (see e.g. Tumpey et al., 2005; Bos et al., 2011), limited attention has been given to the societal effects of the epidemic. What are the short term economic consequences following from a health shock where a large share of the population of working age falls dead within a very limited time period?

Gaining knowledge on this issue is relevant for several reasons. Along the recent urgings from the WHO (2007), studying the effects of the Spanish flu can give insights intothe effects that future pandemics may have on various economic outcomes. An understanding of the link between such a health shock and subsequent economic performance may also be helpful in establishing appropriate policy responses. Moreover, mainly affecting the young and healthy, the 1918 influenza shares striking similarities with the HIV/AIDS pandemic in developing countries. To analyse the economic consequences using HIV/AIDS prevalence or related mortality rates however is deemed problematic as the disease is a much slower process. Given the heightened awareness of economic issues associated with pandemics, it seems timely and relevant to acquire knowledge of any consequences of an event such as the Spanish flu. Not the least since the influenza, although severe, was appeared during a very short time which facilitates the identification of the effects of interest and serves as a severe test of a health shock on economic outcomes.

Using high-quality administrative data from Sweden and an extension of the standard Difference-in-Differences estimator we examine the impact of the 1918 influenza pandemic on economic performance. Close to one percent of the Swedish population died from the Spanish flue, but across counties influenza mortality varied widely and idiosyncratically, with some areas experiencing almost three times the number of deaths per capita of other areas where the pandemic was less severe (Åman, 1990). Besides high variation across analysed units, using data on

Sweden to study the subject matter is advantageous since the population is very homogenous. More importantly the country did not take part in the World War that occurred before and during the spread of the flu pandemic. Previous studies on the economic effects of the Spanish flu use data for countries participating in the war which clearly challenges identification, in turn aggravating interpretation of results. Finally, the richness and the high quality of the administrative data allows for estimating the effects of the influenza on a number of outcomes, and carefully testing key methodological assumptions.

Our results suggest that the influenza significantly increased regional poverty. Moreover, the pandemic appears to have affected capital returns negatively – and besides, the estimated reduction is surprisingly large. However, contrary to theoretical predictions and previous empirical evidence, we find robust evidence that the Spanish flu had no evident effect on earnings.

The next section outlines theory and empirical evidence on the Spanish flu, along with the details of the worldwide pandemic and its development within Sweden. Section 3 describes the methodology used in this study, while section 4 presents the data. Section 5 visualizes support of the common time trend hypothesis, and discusses the empirical results, namely the difference-in-difference estimates of the effect of regional exposure of the influenza on earnings, capital returns, poverty, and manufacturing production.

2 The Spanish Flu Pandemic: Facts, Theory and Empirical Evidence

In this section we present some stylized facts about the Spanish Flu. We start with an overview of the worldwide consequences. After this we summarize the progression of the epidemic in Sweden and review the literature theoretical and empirical literature on the economic consequences.

There is widespread agreement that the global consequences in terms of infections and death toll are unparalleled. Taubenberger and Morens (2006) for instance calls it "the mother of all diseases" and estimate that 500 million individuals were infected. However, due to lack of reliable data, there is considerable variation in the number of deaths estimated to have been caused by the 1918-1919 influenza virus. Estimates range from 20-40 million (Patterson and Pyle, 1991) to 50 million, but the highest estimates are as high as 100 million (Johnson and Mueller, 2002). According to Morens and Fauci (2007) most deaths resulted from "respiratory complications" (e.g. pneumonia).

It should be clear from the above that a prominent characteristic of the Spanish flue were the exceptional mortality rates. In the normal case approximately 0.1 per cent of all individuals infected by a flue perish. The second wave of the epidemic in the fall of 1918 was 5 to 20 times more deadly. The main reason for why the Spanish flu was so extraordinarily aggressive is that the virus did not only attack the bronchus, but also the lungs, leading to many people dying from severe pneumonia. Recent research indicates that it was the combination of three genes that made it possible for the virus to attack human lungs.[1] What furthermore characterizes the disease is the heavy toll on young adults. It is estimated that around half of the death toll was "paid" by individuals between 20 and 40 (Simonsen et al., 1998). This is very unusual. Unlike other (influenza) diseases which exhibit a U shape in terms of mortality, the Spanish Flu is known to have had a W shape.

Researchers of medical history agree that the disease accrued in three to four waves (even though not all regions of the world experienced these three waves, most actually did). The first appearance was in the spring of 1918, then the disease returned in the fall of the same year and finally in 1919. The last wave accrued mainly in Scandinavia and some isolated islands in the South Atlantic. Due to the relatively low spread researchers do not agree on the terminology of the "fourth wave". The second wave is estimated to have been the most severe one, and only areas that were exceptionally isolated escaped the influenza altogether. Historically, the spread of the pandemic was accelerated by increased troop movement due to World War I (Patterson and Pyle, 1991).

2.1 The Spanish Flu Pandemic in Sweden

With respect to the number of deaths caused, the Spanish flu is one of the most severe catastrophes ever affecting Sweden. Unequalled by any epidemic in modern times, the flu killed close to 38,000 individuals in total, representing almost one percent of the population. As in other parts of the world, morbidity rates were even higher. According to official records more than ten per cent of the Swedish population was infected in 1918 (Medicinalstyrelsen), but case studies suggest that incidence was as high as 75 per cent in some areas. There is consequently large variation in the estimates of the number of infected individuals, but generally it is believed that mortality rates amongst infected approached 2 per cent.

[1]The combination of these three genes pioneered the creation of the protein RNA-polymerase. Due to the existence of this protein the virus could mass-produce in human lungs.

The first case of the Spanish flu in Sweden was reported in the southern part of the country in late June 1918. A month later, not all regions had been affected yet, but in early August an increasing number of cases are reported to have died from the flu in the northern provinces. In particular the regions *Jämtland* and *Västernorrland*, and villages along the main inland railway line *Inlandsbanan*, were severely hit. The high mortality rates in the remote northern areas partially have a demographic explanation as these regions had a rather young population at the time and the Spanish flu mainly affected people between 15-35 years. However, following the same line of reasoning as to why the young population were affected to a larger extent than the older, it has also been hypothesised that the high regional variation in mortality rates may be explained by remoteness, and that people living in these areas had less immunological protection against the virus as they generally had been less exposed to earlier flu waves.

Normal flu waves affecting Sweden typically have their outbreak and peaks in February and March, but the Spanish flu reached its peak in Sweden in October and November. During these two months only, the number of victims of the epidemic reached 20,000 individuals. Another, less severe, wave hit the country in March 1919 and new waves appear until early 1920. Due to the fast spread of the disease in the North the national government tried to mobilize medical resources to these areas. Moreover, local authorities took actions to limit the spread of the disease and implemented various public health measures, such as the banning of public gatherings, closure of theatres, and closure of schools. These extraordinary actions however had limited effectiveness as the virus was transmitted through the air and since inhaling the virus coming from another person sneezing or coughing could infect you.

2.2 Theoretical Perspectives

The Spanish Flu had severe effects. Even though only a small proportion died of the disease many were infected. Moreover, individuals of prime age (between 20 and 40 years old) where disproportionally hit. Therefore the most immediate consequences for the economy were a smaller labour force due to absenteeism or deceased workers. This reduction becomes even more important if one considers the fact that manual labour was more important in those days than it is nowadays. Due to this decrease in labour supply, an increase in wages is the most likely consequence.

In terms of economic consequences, Boucekkine et al. (2008) evaluate existing growth models to predict the impact on economic growth. They distinguish between the effect of short term

diseases like the influenza and long term ones, like HIV-AIDS. According to the authors the effects, which are driven by human capital and population dynamics, depend mainly on whether a single- or a two-sector framework is used. In the single sector model the restriction of the fixed ratio between physical and human capital will lead to reaccumulation of human capital that was destroyed by the disease. This together with an increase in population growth (with population only playing a marginal role) will lead to further human capital accumulation and thus higher economic growth.

In the two-sector model, with education representing a separate sector, the effects of a pandemic are different. During the spread of the disease, the labour force will be reduced, and wages will increase – leading to a growing labour force (i.e. more people leave school earlier to take advantage of the higher wages). Due to the reduction in human capital, Boucekkine et al. (2008) predict that the period of disease is followed by low growth and high wages in the short and also in the medium term. The effects of the pandemic on population growth remains an open issue. Population growth is an exogenous parameter in the model. If population growth increases to balance the increased death rate due to the pandemic human capital will increase rapidly and boost the economy in the long run. However in a different paper Boucekkine and Laffargue (2010) derive population growth endogenously and predict that population decreases both in the short and long run.

In a more recent paper, Boucekkine and Laffargue (2010) consider distributional effects of epidemics. In an overlapping generations model with heterogeneity in worker skills, it is shown that income inequality may increase in the short and medium term – even though the epidemic leaves no permanent traces in the income distribution. Since we consider poverty as one outcome variable in this paper, Boucekkine and Laffargue (2010) provide a very useful theoretical backdrop. However, in their case the mechanism is a reduced investment in human capital due to orphanhood in low-skilled families. Since our analysis excludes orphans, our results may be driven by a different mechanism.

Finally, Boucekkine and Laffargue (2010) also briefly analyse the effects on output and productivity, which we also consider as outcome variables. In their model output per worker is not affected over the short-term horizon, while the medium and long term effects are not clear.

2.3 Empirical Evidence

Up to now there are not many empirical studies estimating the effects of the pandemic.[2] Most empirical studies will face two serious problems. Firstly, there is often a lack of reliable data from the time period. The second problem is that identification is difficult, inter alia due to the fact that the flu occurred during and shortly after the First World War.

Garrett (2008) analyses the determinants of influenza incidence and finds that rural areas are less affected by the pandemic. Moreover he analyses differences in incidence by race and finds that non-whites and people on lower incomes were more heavily affected. In terms of economic effects, Garrett (2008) relies on newspaper articles which suggest losses for many businesses, especially in the service sector – while businesses specialized in health care had an increase in revenues.

Brainerd and Siegler (2003) is one of the few papers that consider the effects on growth. They study changes in real personal incomes between 1919/21 and 1930. Due to data restrictions they are only able to analyse long term effects and are not able to distinguish whether the effect was due to recovery or growth. In any case, the authors find a significant positive effect, in the sense that those states that where hit harder by the flu experienced a higher income growth rate from 1919/1921 to 1930. From a theoretical point of view, this result might have driven by increased investment in human capital and higher population growth after the occurrence of the Spanish Flu.

In a more recent paper, Garrett (2009) analyses the effects of the pandemic on manufacturing wages. Garrett (2009) uses the same mortality data as Brainerd and Siegler (2003), but he has access to wage growth between 1914 and 1919; thus, the study is able to compare before and after the pandemic, but is only able to estimate effects in the very short term. The paper concludes that the epidemic appears to have had a positive impact on manufacturing wages. However, it is not always clear to what extent the results are not attributable to the First World War.

In summary, there have been some attempts to estimate the effects of the pandemic in the US, but there is still no study which rigorously applies the methods typically used to conduct causal inference. The main reason appears to be a severe lack of reliable data. As we will show

[2] A growing literature however analyses the long-term consequences environmental shocks of survivors, and in utero exposure of the influenza, on later health and labor market outcomes (cf. Almond and Mazumder, 2005; Maccini and Yang, 2009).

in the following sections, Swedish data appear to offer a significant improvement compared with previous studies.

3 Econometric Approach

In this section, we first present our main treatment indicator, then discuss the assumptions required for identification. Next, we present our regression equations and then we discuss how to estimate standard errors in the presence of non-spherical disturbances. Finally, we present the methods we use to support our identification strategy.

3.1 Defining the Treatment Indicator

Our analysis is conducted at the level of Swedish counties (Swedish: län) of which there were 25 at the time of the Spanish flu. As already mentioned, the incidence of and mortality related to the pandemic exhibits considerable variation across regions. The main assumption underlying our analysis is that the regional exposure to the Spanish influenza represents an exogenous shock and that regions that were particularly hardly affected would have followed the same time trend as other regions in the absence of the pandemic. Thus, we define treatment as the degree of exposure to the pandemic, measured by total excess regional influenza mortality through the years 1918-20. In our baseline specifications, we furthemore assume that the effects of Spanish flu mortality is constant over time and a linear function of the excess mortality.

Thus, the main treatment variable used in our analysis is defined as

$$w_{it} = \sum_{s=1918}^{\min\{t,1920\}} m_{is}^{a} - m_{i,1917}^{a} \tag{1}$$

where i indexes counties, t years, and m_{is}^{a} is the mortality rate due to influenza, expressed as the number of deaths per 100,000 inhabitants.

3.2 Identifying Assumptions

The method we use is an extension of the standard Difference-in-Differences estimator; our extension is simply that we need to allow for varying treatment intensity. Thus, the functional form imposed adds a further assumption to the standard assumptions, and it should clearly be formally tested. In addition, it is important to spell out the standard DID assumptions since it

is possible to assess their credibility.

In order to gauge the plausibility of the assumptions, it is useful to introduce some additional notation. As is standard in the programme evaluation literature, we refer to the concept of potential outcomes, although in this setting the potential outcomes have a support along the entire range of the treatment variable. Also, we denote the vector of treatment intensities for all time periods for a single individual i by $\mathbf{w}_i$. Moreover, we denote possible conditioning variables, that may possibly be included in the regression equations below, by X_{it}. Denoting the potential outcome function by $Y_{it}(\cdot)$, we may thus state the first assumption.

Assumption 1 (SUTVA)

$$y_{it} = Y_{it}(\mathbf{w}_i) \tag{2}$$

The SUTVA assumption states that a) only one amongst the potential outcomes is observed, and b) there are no relevant interactions between units (i.e. only $\mathbf{w}_i$ matters in the determination of unit i:s outcome). For most variables that we consider in this study, the assumption seems plausible. It is unlikely that the ifnluenza exposure in one county had a strong impact on economic outcomes in other regions, espcially if they are far away from each other. It should furthermore be noted that the assumption is reconcilable with *spatial autocorrelation*, i.e., that the residual variation in the outcomes is spatially correlated.

We now introduce the *Common Trend* assumption, which is absolutely essential for the DID method:

Assumption 2 (CT)

$$\mathbb{E}\left[Y_{it}(\mathbf{w}_i) - Y_{i,t-1}(\mathbf{w}_i)\right] = \mathbb{E}\left[Y_{jt}(\mathbf{w}_j) - Y_{j,t-1}(\mathbf{w}_j)\right] \;\; \textit{if } \mathbf{w}_i = \mathbf{w}_j \tag{3}$$

This assumption, which is made somewhat stronger in comparison to the standard DID assumption, states that differences in time trends between counties (here denoted i and j) are only attributable to differences in treatment exposure. The assumption does not require that counties are at similar levels, but only that the dynamics of the outcomes would have been similar if the exposure to treatment had also been similar. This assumption may be sensitive to functional form assumptions, and for this reason, we always consider outcomes in levels as well as in natural logarithms as a robustness check.

In our view, the CT assumption is the most problematic of them all, since it is possible

that diverging time trends among regions confound the estimated effects of the pandemic. The assumption is not directly testable, but we can assess the extent to which it is plausible by looking at periods during which there were no significant differences in influenza mortality between counties. We will provide this analysis in two steps below: firstly, we will provide visual evidence of how the outcomes evolved in different regions, sorted by degree of affectedness, before the pandemic. Secondly, we will run regressions using a 'placebo' treatment, in which we counterfactually assume that Sweden was hit by the pandemic in a year prior to 1918 (and only in that year).

Next, we consider the *No Effect Prior to Treatment* assumption:

Assumption 3 (NEPT)

$$Y_{it}\left(\mathbf{w}_i\right) = Y_{it}\left(w_{it}\right) \tag{4}$$

This assumption rules out the possibility that the excess influenza mortality in one year has an impact on (potential and observed) outcomes in previous years. It should be noted, however, that due to the way w_{it} has been constructed, the converse does not hold: the excess mortality in one year may very well have an impact after that year – indeed, this is one of the main issues analysed in this paper. The NEPT assumption rules out economic effects based on an *anticipation* of the pandemic. This assumption seems very reasonable for the years preceding 1918, since the first report worldwide on the pandemic appeared only in January of that year. Once the influenza had started spreading, an anticipation of further deaths in subsequent years may of course in theory have had an impact on the behaviours of economic agents; however, this aspect seems very unlikely to matter in practice.

Since a rich set of possible conditioning variables are available, we could also have included covariates, denoted X_{it}. However, these variables would need to satisfy an exogeneity assumption which might not be warranted:

Assumption 4 (EXOG)

$$X_{it}\left(w_{it}\right) = X_{it} \tag{5}$$

In short, the exogeneity assumption states that conditioning variables are unaffected by the treatment. This seriously limits the number of conditioning variables that we may consider. The assumption is obviously going to hold for variables which do not vary over time, but these are washed out by the fixed county effects anyway. As far as time-varying covariates are concerned,

it is difficult to think of any economically relevant variable which is not also potentially affected by the pandemic. Thus, we only condition on time and county fixed effects in what follows.

3.3 A Linear Specification

For all the various outcome variables we consider, our main baseline specification is

$$y_{it} = \alpha_i + \beta w_{it} + \lambda_t + \epsilon_{it} \tag{6}$$

where y_{it} is the outcome variable (e.g. wages, production or prices), α_i is a county fixed effect, λ_t is a year fixed effect, and ϵ_{it} is a residual disturbance. Imposing the assumptions above, it is possible to show that $\mathbb{E}(\epsilon_{it} \mid w_{it}, i, t) = 0$ so that an OLS estimate of β using equation (6) is unbiased.

In an alternative set of specifications, we also allow the impact of the influenza pandemic to vary over time:

$$y_{it} = \alpha_i + \beta w_{it} + \gamma w_{it} \mathbf{1}(t > \tau) + \lambda_t + \epsilon_{it} \tag{7}$$

where γ captures treatment effect heterogeneity over time, and $\mathbf{1}(t > \tau)$ is a dummy variable indicating that the year is after some cutoff year denoted τ.

The placebo regression will take a very similar form:

$$y_{it} = \alpha_i + \delta w_{i,t+3} + \lambda_t + \epsilon_{it} \text{ if } t < 1918 \tag{8}$$

In words, we estimate the 'effect' of a counterfactual placebo epidemic, which is assumed to have occurred in the years 1915-17 with the incidence rates of 1918-20. If the placebo parameter δ is precisely estimated and close to zero it can be seen as evidence for the common time trend. Moreover, it will give us an indication of whether spatial autocorrelation is a problem in the dataset.

3.4 Estimating Standard Errors

Inference in DID models has attracted considerable attention in the literature over the past decade. Particular attention is devoted to two issues related to the estimation of standard errors:

autocorrelation and common group effects. Since we use data aggregated at the regional level throughout, the common group error problem discussed by Donald and Lang (2007) is unlikely to be a problem. On the other hand, our estimates are based on relatively long panels – we have as much as 20 time periods in some specifications below – so the possibility of autocorrelation needs to be taken into account.

In a seminal paper, Bertrand et al. (2004) discuss the problems associated with autocorrelation in difference-in-differences studies and also compare different solutions. Whenever variables are (positively) serially correlated, DID estimates not correcting for autocorrelation are likely to lead to over-rejection, i.e., the null hypothesis of no treatment effect is rejected too often. Several solutions are available to handle this problem, each one of them with their specific drawbacks. Given the low number of counties available, only two alternatives seem appealing for this study: to rely on the GLS estimator originally suggested by Kiefer (1980), or to reduce the time dimension by collapsing the data. The GLS approach is based on estimating the full covariance matrix, allowing for any pattern of correlation between time periods. In a recent paper, Hausman and Kuersteiner (2008) analyse the properties of this GLS estimator. Their main conclusions are that a FGLS procedure generally outperforms procedures where the time dimension is reduced by aggregating observations. They do, however, find that there might still be size distortions when the number of observations are small – and they suggest a size correction that is shown to have better properties. Even though their size correction is promising, we decided not to follow that route here. The estimated correlation matrix exhibited positive autocorrelation in the short term but negative autocorrelation in the long term – and thus, standard errors often turned out smaller than in the original OLS specification.

Thus, we decided to reduce the number of time periods instead, by collapsing the data. Hence, in a set of robustness checks, we shrank the time dimension into five or less time periods, and checked whether results were robust to this modification. The estimating equations remain the same as those above, but we now used a collapsed version of the outcome variable, defined as follows:

$$\tilde{y}_{it} = \begin{cases} \frac{1}{T_0} \sum_{s=t_0}^{1917} y_{is} & \text{if } t = 1917 \\ y_{it} & \text{if } t \in (1918, 1920) \\ \frac{1}{T_1} \sum_{s=1921}^{t_1} y_{is} & \text{if } t = 1921 \end{cases} \tag{9}$$

where T_0 is the number of time periods before 1918; t_0 is the first year covered by the panel; T_1 is the number of time periods after 1920, and t_1 is the last year covered by the panel. Obviously, the treatment variable $\tilde{w}_{it}$ is defined analogously:

$$\tilde{w}_{it} = \begin{cases} 0 & \text{if } t = 1917 \\ w_{it} & \text{if } t \in (1918, 1921) \end{cases} \tag{10}$$

Thus, we require estimated effects to be robust to this change in specification.

4 Data and Variables

There are two sources of influenza statistics available for Sweden and they differ to some extent. It is generally believed that the figures from *Medicinalstyrelsen* tend to underestimate the number of cases. The figures from Statistics Sweden are however viewed to be more correct, in particular since the bureau implemented more detailed and stricter reporting procedures in 1911 which improved the reporting from rural areas.[3]

In Figure 1, we show the 1918-20 excess influenza mortality by region. The front row of bars show the influenza mortality in the overall population, and the back row shows the influenza mortality amongst people aged 0-40. Since 90 per cent of influenza mortality occurred in this age group, it makes little difference to use this alternative definition. However, since the age structures of regions differed significantly, it is crucial to take age into account – as otherwise differences in trends between regions with different age structures might be a confounding factor.

[Insert Figure 1 about here]

Figure 1 shows that there is a distinct north/south divide in terms of influenza mortality, even if the pattern is much weaker when the younger group is considered. One notable exception, however, is the southernmost *Malmöhus län*, which belongs to the most heavily affected counties. Moreover, there is no corresponding urban/rural divide, since the densely populated counties tend to be in the middle, whereas the more thinly populated regions tend to the extremes.

Further variables that we use are **earnings** and **capital incomes** per capita, which both are taken from official tax records. We also include **poverty rates**, defined as the number of

[3]Priests had to make monthly reports to the Statistics Sweden on the likely cause of death of persons in cases where no doctor already had been involved. These notes and reports were then reviewed and confirmed by a GP who reported the final cause of death to the bureau.

inhabitants in public poorhouses as a proportion of the total population.[4] Finally, we include **industrial production** per capita, measured as the total value added in production over the county population. In order to avoid spurious effects of the influenza working through the denominators of the per capita variables, we use the average population over the year throughout. Descriptive statistics for all variables are provided in Table 1.

[Insert Table 1 about here]

5 Results

In this section, we present the main results from our analysis. First, we present visual evidence supporting the common time trend hypothesis for the main outcome variables. We then turn to estimates of the effects of the pandemic on these variables.

5.1 Common Time Trend: Visual Evidence

Before we turn to regression analysis, we provide some preliminary evidence concerning the common time trend assumption. As already mentioned, our case differs from the standard DID setting in the sense that we have more than two degrees of treatment intensity, and hence the counties included do not form two distinct groups. However, in terms of the total excess influenza mortality experienced over the entire 1918-20 period, we may distinguish two different strata of exposure. Most of the counties fall within the range of 750-950 additional deaths per 100,000 population. Above that, there is a smaller group of seven counties which experienced between 984 and 1481 additional deaths. Thus, in what follows, we present visual evidence contrasting these two groups.[5] In the graphs that follow, counties are weighted by their 1917 population size, and all monetary variable are expressed in 1917 crowns (adjusted according to the CPI).

In Figure 2, we present evidence for capital income per capita. The brown curve pictures the growth in capital incomes (compared with the 1912 level) for counties which were particularly hardly hit by the epidemic. The blue curve plots the corresponding series for the less severely affected counties. The dotted curves show the 95 % confidence intervals. The figure seems to indicate that the common time trend is a reasonable assumption before the pandemic hit: the curves are very close and their confidence intervals have a significant overlap. During and after

[4] It should be noted that although the variable used includes dependants of adult poorhouse residents, it does not include orphans as these resided in special orphanages.

[5] Allowing for more groups did not make much difference, but made the figures more difficult to read.

the pandemic, the two groups diverge. The group more severely affected by the pandemic appears to suffer a quite significant relative loss in capital incomes due to the pandemic. However, there is at the same time a rapid increase in capital income, so it is unclear whether this observation holds for closer scrutiny.

[Insert Figure 2 about here]

In Figure 3, the corresponding information for earnings per capita is presented. For this variable, the common time trend assumption seems slightly harder to maintain: the two curves are somewhat further apart, but in fact, after 1913 they are fairly parallel. Also in this case, the figure suggests that more hardly hit counties also experienced slower earnings growth during and after the pandemic.

[Insert Figure 3 about here]

Next, we provide visual evidence for the poverty variable, which captures the proportion of the population that was living in public poorhouses. For this variable, the year 1919 is missing. The common time trend assumption appears to be satisfied, even though the more severely affect group is outside the other group's confidence interval in 1917. Nevertheless, the poverty rates appear to be following a reasonably parallel trend before the pandemic, and then start diverging from 1921 onwards. The more severely affected counties experienced faster growth in poverty rates: on average, they experienced an increase from 4.6 per cent in 1921 to 7.1 per cent in 1930, whereas the corresponding increase for the rest of the counties was from 4.2 per cent to 5.5 per cent.

[Insert Figure 4 about here]

Thus, Figure 3 suggests that more strongly affected counties experienced slower earnings growth, and Figure 4 suggests that poverty rates were positively affected by the pandemic. Obviously, both these observations might be driven by a change in poverty rates. Thus, we defined an alternative earnings variable, where total annual earnings at county level are divided by the number of inhabitants who are not poor. Visual evidence for this new variable is provided in Figure 5.[6] For this variable, the common time trend assumption appears to be even more plausible than for the original earnings variable, but otherwise, no important changes are discernible.

[6]Since the poverty variable is missing for 1919, we used linear intrapolation for the denominator in 1919.

[Insert Figure 5 about here]

Finally, we consider industrial production per capita. For this variable, it appears that the high-influenza counties might have been somewhat more strongly affected by the slump starting in 1917; however, the common time trend assumption nevertheless seems to be a reasonable approximation. Also for this variable it is difficult to draw any conclusion about the influenza effect based on this visual evidence only.

[Insert Figure 6 about here]

In conclusion, we have seen that there are no blatant violations of the common time trend in our data, and that the pandemic appears to have had an impact on all three outcome variables. However, the evidence is too crude and summaric to provide a reliable estimate of the effects. This goes in particular for capital income and earnings, which seem to have increased rapidly throughout the country during the epidemic. Hence, we now turn to more rigorous regression-based evidence.

5.2 Regression Analysis

We now present regression results for the three main variables of interest. Each variable is analysed in five different specifications: first, we estimate the overall impact of the pandemic using the entire sample. Then we allow for the effect to be different *during* and *after* the pandemic. After that, we turn to placebo estimates: restricting the sample to the years before 1918, we assume counterfactually that the pandemic struck Sweden between 1915-17 and estimate the effects of this 'placebo' intervention. Finally, we address the issue with autocorrelation by presenting estimates where the dataset has been collapsed into five time periods: before, 1918, 1919, 1920 and after.

It is well known that the DID estimator is sensitive to functional form assumptions. In our case, the natural alternatives would be to use euither levels or logarithms of the outcome variables. Since the different counties are at very different levels at the outset, we think that a logarithmic specification is to be preferred. However, for the sake of completeness, we also provide estimates for the outcome variables in levels. If both estimates deliver the same result, it of course adds credibility to the evidence. If not, it suggests that the functional form deserves particular attention.

In Table 2, we present results for capital income per capita. In the left part of the panel, the dependent variable is log capital income, and in the right part, it is capital income in levels. The first column presents the overall effect of the pandemic, averaging over all years from 1918 onwards. According to our estimate, each additional death on 100,000 population was associated with a reduction in capital income per capita by 0.046 per cent. Since the median total mortality rate between 1918-20 was 900, this coefficient indicates that the median county suffered a reduction in capital incomes by 41 per cent due to the pandemic. However, in order to avoid extrapolation, one could compare the 25th and the 75th percentile, with an incidence of 842 and 984 respectively. The difference between these two counties would correspond to a reduction in capital incomes by 6.5 per cent, which is still a fairly large amount.

In the second column, we contrast the effects during (1918-20) and after (1921-30) the pandemic. Clearly, the bulk of the effect materialised afterwards, whereas only a smaller decline was noticeable during the epidemic. Finally, the third column presents estimates for a 'placebo epidemic' which is assumed to have occurred between 1915-17 with the county-specific incidence rates of 1918-20. This estimate is nowhere near statistical significance and it is very precisely estimated. Thus, our observation from Figure 2 is confirmed: the common time trend assumption can be maintained.

In columns 4-5, we collapse the time period into five in order to reduce problems related to autocorrelation. The estimates in these column clearly indicate that autocorrelation is an issue. Nevertheless, the effects observed after the epidemic are still significant at the ten per cent level.

In columns 6-10, we repeat the same exercise for capital income in levels. Clearly, this specification fares much worse in terms of explanatory power, which is discernible from the reported R^2 and the statistical significance is lost for most estimates. Nevertheless, findings appear to be generally reconcilable with our estimates for the logarithmic specifications.

[Insert Table 2 about here]

In Table 3, we turn to the earnings variable. Again, the first five columns deal with log earnings. For this variable, there is much less evidence of an effect of the pandemic. The point estimate of the overall effect is -0.00008 which, according to our previous comparison would imply a reduction in wages by 7.2 per cent in the median county; or a relative decline of 1.1 per cent in the 75th percentile county compared with the 25th percentile. On the other hand, it should be noted that the placebo estimate is even smaller and with a similar degree of precision.

Hence, the common time trend assumption cannot be rejected, and we may thus conclude that the epidemic appears to have had no effect at all on earnings.

[Insert Table 3 about here]

We report results for poverty rates in Table 4. For this variable, the results are generally very strong: the pandemic appears to have had a strong and lasting positive effect on poverty. The overall effect is estimated at 0.00036, which would imply that the median county had its poverty rate increased by 32 per cent due to the pandemic. Comparing again the 25th and th 75th percentile, the difference in flu mortality would give rise to an in crease in poverty by 5.1 per cent. Again, this effect is quite substantial. And if we focus on the post-pandemic period, it is actually almost twice as high. Again, the placebo estimate is insignificant, small and relatively precisely estimated. Thus, the common time trend assumption seems to be confirmed also in this case.

[Insert Table 4 about here]

Next, we turn to the earnings of the non-poor population. Results are presented in Table 5. However, normalising earnings using the non-poor instead of the total population does not change much: the estimated effect is still insignificant, albeit somewhat higher than before.

[Insert Table 5 about here]

Finally, in Table 6 we present estimates for industrial production per capita. Also for this variable, the placebo estimates are reassuring, but we don't observe a treatment effect either. Thus, it appears that there was no significant effect of the pandemic on industrial production in general.

[Insert Table 6 about here]

In conclusion, we have found very strong evidence for the pandemic having a long-term positive impact on poverty. We have found relatively strong evidence that capital incomes were negatively affected by the pandemic. However, there is no evidence whatsoever that earnings or productivity have been affected by the pandemic. Moreover, using placebo regressions we found that the common time trend assumption can be retained in all cases – placebo estimates are generally insignificant and very close to zero.

6 Conclusion

In this paper we estimate the economic consequences of the Spanish Flu in Sweden. In order to exploit the varying levels of intensity of the Spanish flu, measured by death rates, across different regions and time periods we use a modified Difference-in-Differences approach. Using placebo estimates we first demonstrate that the main identifying assumption of a common time trend can not be rejected for the economic outcome variables we analyse.

Our results reveal that, in line with economic theory (Boucekkine and Laffargue, 2010), the Spanish flu increased poverty in Sweden, while we find no effects on earnings and production. Capital returns were also negatively affected. In a further step we separate immediate and long term effects and find that the long term effects are more severe for both lower capital returns and poverty.

These results are in parts surprising, so it is useful to consider whether they are possibly driven by selective mortality. However, we find it difficult to reconcile our findings with a selection-on-health story, irrespective of whether one considers positive or negative selection (and the literature is inconclusive on this point). *Positive* selective mortality – in the sense that healthier people are less likely to survive – would be consistent with increased poverty and reduced capital returns (even if the effects seem to be large) but it is difficult to imagine a health-based selection that is so strong that it cancels out the effects on earnings of the double withdrawal of poor and deceased people from the labour force. Conversely, *negative* selection (i.e., unhealthy individuals more likely to die) would certainly be inconsistent with observing no effect on earnings, and also possible preclude an increase in the relative number of poor people. In conclusion, neither of these two types of selection appear to sit well with our results.

Our results would need to be corroborated further using data from other countries. If they are confirmed, they suggest that theoretical models might need to be refined to accomodate with the inconsistencies identified in this study.

References

Almond, D., Mazumder, B., 2005. The 1918 influenza pandemic and subsequent health outcomes: an analysis of sipp data. American Economic Review, 258–262.

Åman, M., 1990. Spanska sjukan: den svenska epidemin 1918-1920 och dess internationella bakgrund. Ubsaliensis Academiae.

Bertrand, M., Duflo, E., Mullainathan, S., 2004. How much should we trust differences-in-differences estimates?*. Quarterly Journal of Economics 119 (1), 249–275.

Bos, K. I., Schuenemann, V. J., Golding, G. B., Burbano, H. A., Waglechner, N., Coombes, B. K., McPhee, J. B., DeWitte, S. N., Meyer, M., Schmedes, S., Wood, J., Earn, D. J. D., Herring, D. A., Bauer, P., Poinar, H. N., Krause, J., 2011. A draft genome of yersinia pestis from victims of the black death. Nature.

Boucekkine, R., Diene, B., Azomahou, T., 2008. Growth economics of epidemics: A review of the theory. Mathematical Population Studies 15 (1), 1–26.
URL http://www.tandfonline.com/doi/abs/10.1080/08898480701792410

Boucekkine, R., Laffargue, J., 2010. On the distributional consequences of epidemics. Journal of Economic Dynamics and Control 34 (2), 231–245.

Brainerd, E., Siegler, M., 2003. The economic effects of the 1918 influenza epidemic. Centre for Economic Policy Research.

Donald, S., Lang, K., 2007. Inference with difference-in-differences and other panel data. The Review of Economics and Statistics 89 (2), 221–233.

Garrett, T., 2008. Pandemic economics: the 1918 influenza and its modern-day implications. Federal Reserve Bank of St. Louis Review 90 (2), 75–93.

Garrett, T., 2009. War and pestilence as labor market shocks: Us manufacturing wage growth 1914–1919. Economic Inquiry 47 (4), 711–725.

Hausman, J., Kuersteiner, G., 2008. Difference in difference meets generalized least squares: Higher order properties of hypotheses tests. Journal of Econometrics 144 (2), 371–391.

Johnson, N., Mueller, J., 2002. Updating the accounts: global mortality of the 1918-1920 "spanish" influenza pandemic. Bulletin of the History of Medicine 76 (1), 105–115.

Kiefer, N., 1980. Estimation of fixed effect models for time series of cross-sections with arbitrary intertemporal covariance. Journal of econometrics 14 (2), 195–202.

Maccini, S., Yang, D., 2009. Under the weather: Health, schooling, and economic consequences of early-life rainfall. American Economic Review 99 (3), 1006–1026.

Morens, D., Fauci, A., 2007. The 1918 influenza pandemic: insights for the 21st century. Journal of Infectious Diseases 195 (7), 1018.

Patterson, K., Pyle, G., 1991. The geography and mortality of the 1918 influenza pandemic. Bulletin of the History of Medicine 65 (1), 4.

Simonsen, L., Clarke, M., Schonberger, L., Arden, N., Cox, N., Fukuda, K., 1998. Pandemic versus epidemic influenza mortality: a pattern of changing age distribution. Journal of Infectious Diseases 178 (1), 53.

Taubenberger, J., Morens, D., 2006. 1918 influenza: the mother of all pandemics. Rev Biomed 17, 69–79.

Tumpey, T., Basler, C., Aguilar, P., Zeng, H., Solórzano, A., Swayne, D., Cox, N., Katz, J., Taubenberger, J., Palese, P., et al., 2005. Characterization of the reconstructed 1918 spanish influenza pandemic virus. Science 310 (5745), 77.

Figure 1: 1918-20 Influenza mortality in different counties

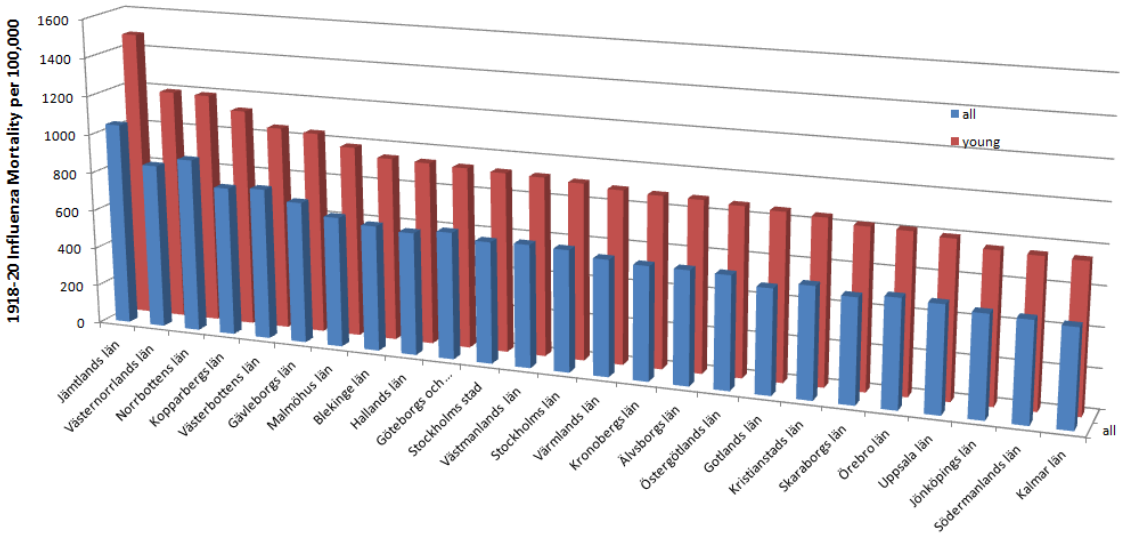

Figure 2: Common Time Trend for Capital Income

Figure 3: Common Time Trend for Earnings

Figure 4: Common Time Trend for Poverty

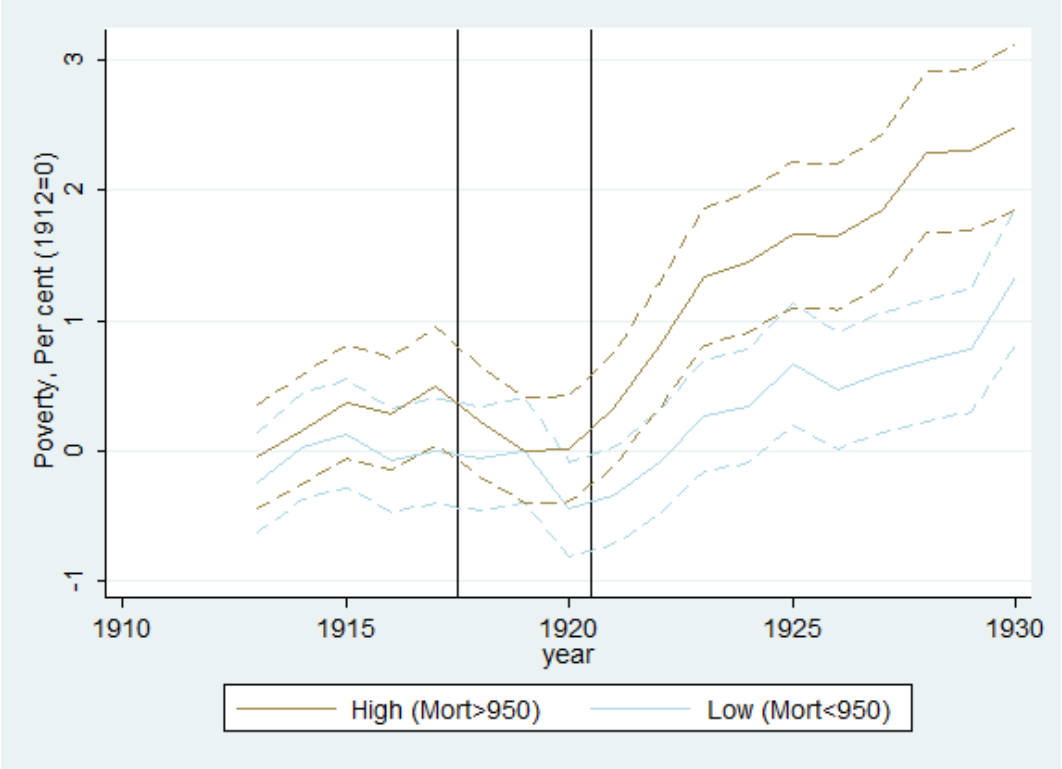

Figure 5: Common Time Trend for Non-Poor Earnings

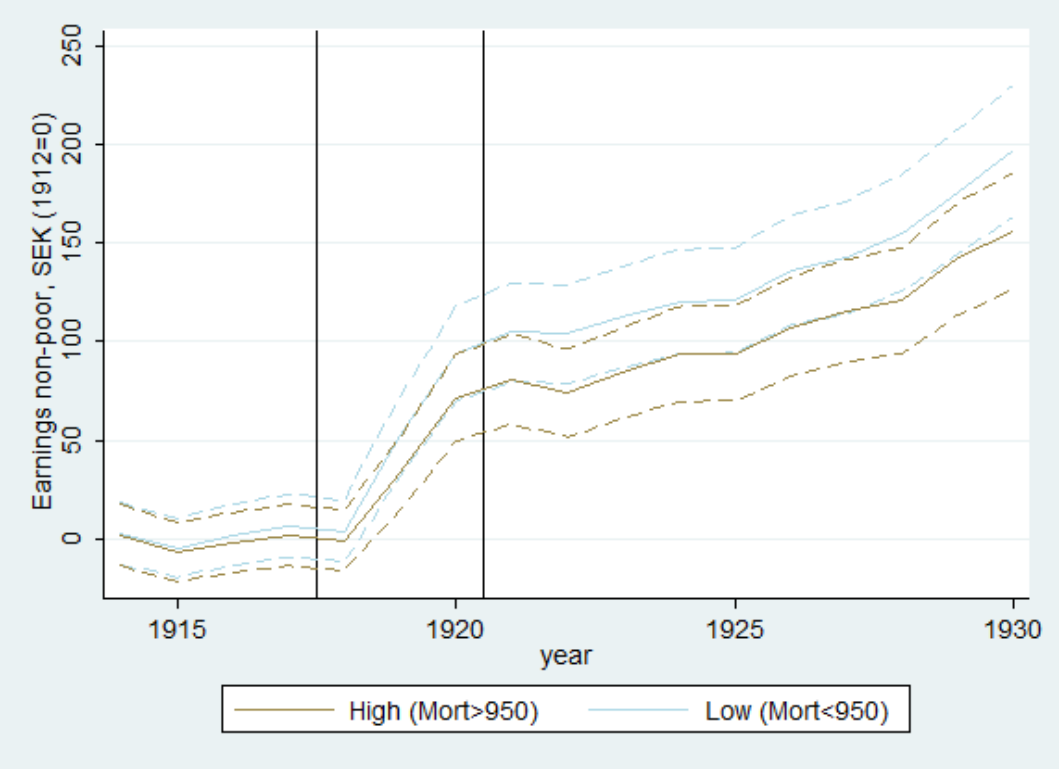

Figure 6: Industrial Production per Capita

Table 1: Descriptive statistics

Variable	Mean	Std. Dev.	Min.	Max.	N
Capital Income (MSEK)	11.7397	19.5075	0.4149	136.3964	475
Earnings (MSEK)	108.5246	117.1777	7.0840	894.8387	475
Poverty, Per cent	4.6445	1.473	0.55	12.1	450
Capital Income	216.6215	127.4703	65.1708	967.0460	450
Industrial Production (MSEK)	146.1584	121.5026	6.9886	602.6663	375
Cumulative Flu Mortality	0.3211	0.4531	0	1.4817	475

Table 2: Estimation results, capital income per capita

	Dependent Variable: *ln(Capinc)*					**Dependent Variable: *Capinc***				
	All (1)	**All (2)**	**Placebo (3)**	**Collapsed (4)**	**Collapsed (5)**	**All (6)**	**All (7)**	**Placebo (8)**	**Collapsed (9)**	**Collapsed (10)**
D_{it}	-0.45628***	-0.17652		-0.15490	-0.07685	-15.29	9.62		27.01	34.29
	0.13462	0.10404		0.11511	0.10902	20.26	24.48		41.41	43.69
$D_{it} \times \mathbf{1}(t > 1920)$		-0.37210***			-0.32481**		-33.13*			-30.29*
		0.12018			0.12761		18.02			15.96
Placebo			0.03795					1.35		
			0.06686					1.77		
Constant	3.03484***	3.03484***	3.03484***	2.96821***	2.91499***	27.92***	27.92***	27.92***	4.25	-0.71
	0.02447	0.02440	0.01057	0.08116	0.07676	4.59	4.60	0.48	33.28	34.84
Observations	475	475	150	125	125	475	475	150	125	125
R^2	0.94421	0.94621	0.56701	0.94782	0.94940	0.54	0.54	0.23	0.54	0.54

* p<0.1, ** p<0.05, *** p<0.01. Robust standard errors reported. Time and county fixed effects suppressed.

Table 3: Estimation results, earnings per capita

	Dependent Variable: *ln(Earnings)*					**Dependent Variable: *Earnings***				
	All	**All**	**Placebo**	**Collapsed**	**Collapsed**	**All**	**All**	**Placebo**	**Collapsed**	**Collapsed**
	(1)	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**	**(7)**	**(8)**	**(9)**	**(10)**
D_{it}	-0.07712	-0.01048		-0.00264	0.01504	-5.54	42.93		71.92	84.14
	0.14490	0.08687		0.08829	0.07530	69.27	60.45		87.56	90.61
$D_{it} \times \mathbf{1}(t > 1920)$		-0.08864			-0.07356		-64.47			-50.83
		0.09959			0.09358		79.02			71.14
Placebo			0.01769					12.39		
			0.07078					19.93		
Constant	5.48130***	5.48130***	5.48130***	5.66849***	5.65643***	281.67***	281.67***	281.67***	271.51***	263.18***
	0.08881	0.08888	0.07694	0.06011	0.05115	28.40	28.43	18.71	70.64	73.35
Observations	475	475	150	125	125	475	475	150	125	125
R^2	0.89228	0.89248	0.16408	0.94490	0.94518	0.77	0.77	0.15	0.78	0.78

* p<0.1, ** p<0.05, *** p<0.01. Robust standard errors. Time and county fixed effects suppressed.

Table 4: Estimation results, poverty

	Dependent Variable: *ln(Poverty Rate)*					**Dependent Variable: *Poverty Rate***				
	All (1)	**All (2)**	**Placebo (3)**	**Collapsed (4)**	**Collapsed (5)**	**All (6)**	**All (7)**	**Placebo (8)**	**Collapsed (9)**	**Collapsed (10)**
D_{it}	0.35846*	-0.13874		0.04277	-0.13911	2.36***	-0.45		0.53	-0.51
	0.17749	0.09025		0.11589	0.08384	0.81	0.38		0.60	0.40
$D_{it} \times \mathbf{1}(t > 1920)$		0.61212***			0.59345***		3.45***			3.39***
		0.16576			0.16260		0.70			0.69
Placebo			-0.05153					0.14		
			0.15611					0.52		
Constant	1.44180***	1.44180***	1.44180***	1.57166***	1.18608***	4.40***	4.40***	4.40***	4.72***	2.51***
	0.02401	0.02414	0.01599	0.10405	0.18098	0.12	0.13	0.05	0.52	0.83
Observations	450	450	150	100	100	450	450	150	100	100
R^2	0.40476	0.42586	0.06142	0.52459	0.60934	0.44	0.47	0.09	0.49	0.61

* p<0.1, ** p<0.05, *** p<0.01. Robust standard errors. Time and county fixed effects suppressed.

Table 5: Estimation results, non-poor earnings

	Dependent Variable: *ln(Earnings)*					**Dependent Variable: *Earnings***				
	All (1)	**All (2)**	**Placebo (3)**	**Collapsed (4)**	**Collapsed (5)**	**All (6)**	**All (7)**	**Placebo (8)**	**Collapsed (9)**	**Collapsed (10)**
D_{it}	-0.05829	-0.02193		-0.01318	-0.00432	4.26	23.58		37.37	41.87
	0.14215	0.08605		0.08854	0.07496	34.92	32.69		43.99	45.83
$D_{it} \times \mathbf{1}(t > 1920)$		-0.04857			-0.03686		-25.80			-18.72
		0.09476			0.08957		41.19			36.78
Placebo			-0.03068					-0.00		
			0.05408					7.43		
Constant	5.58369***	5.46125***	5.05308***	5.02848***	5.02244***	274.56***	231.31***	171.54***	143.34***	140.27***
	0.13347	0.08421	0.05267	0.05999	0.05074	32.40	33.00	7.16	35.66	37.28
Observations	450	450	125	125	125	450	450	125	125	125
R^2	0.94473	0.94480	0.12790	0.95249	0.95256	0.78	0.78	0.15	0.79	0.79

* p<0.1, ** p<0.05, *** p<0.01. Robust standard errors. Time and county fixed effects suppressed.

Table 6: Estimation results, industrial production per capita

	Dependent Variable: *ln(Production)*					**Dependent Variable: *Production***				
	All	**All**	**Placebo**	**Collapsed**	**Collapsed**	**All**	**All**	**Placebo**	**Collapsed**	**Collapsed**
	(1)	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**	**(7)**	**(8)**	**(9)**	**(10)**
D_{it}	0.02097	0.07261		0.10232	0.11654	-24.73	-13.18		19.43	19.43
	0.13767	0.20494		0.15273	0.20605	75.35	109.27		75.00	107.40
$D_{it} \times \mathbf{1}(t > 1920)$		-0.07001			-0.04277		-15.67			0.01
		0.27075			0.27120		154.19			153.23
Placebo			-0.13287					-76.53		
			0.19130					149.03		
Constant	6.11547***	6.11547***	6.11547***	6.21908***	6.24583***	510.99***	510.99***	510.99***	589.64***	589.64***
	0.03937	0.03945	0.03781	0.14627	0.17401	18.42	18.45	19.24	75.99	90.39
Observations	375	375	175	100	100	375	375	175	100	100
R^2	0.51953	0.51978	0.61673	0.09994	0.10038	0.52	0.52	0.61	0.14	0.14

* p<0.1, ** p<0.05, *** p<0.01. Robust standard errors. Time and county fixed effects suppressed.