Public Health Interventions and Economic Growth: Revisiting The Spanish Flu Evidence

Andrew Lilley, Matthew Lilley, Gianluca Rinaldi*

May 2, 2020

Abstract

Using data from 43 US cities, Correia, Luck, and Verner [2020] find that the 1918 Flu pandemic had strong negative effects on economic growth, but that Non Pharmaceutical Interventions (NPIs) mitigated these adverse economic effects. Their starting point is a striking positive correlation between 1914-1919 economic growth and the extent of NPIs adopted at the city level. We collect additional data which shows that those results are driven by population growth between 1910 to 1917, before the pandemic. We also extend their difference in differences analysis to earlier periods, and find that once we account for pre-existing differential trends, the estimated effect of NPIs on economic growth are a noisy zero; we can neither rule out substantial positive nor negative effects of NPIs on employment growth.

1 Introduction

The 2019 novel coronavirus disease (COVID-19) pandemic has generated a vigorous debate on the appropriate policy responses, and on the resultant economic effects of various public health policies. A central question is whether a policy trade-off exists between reducing the spread of the virus and reducing economic activity. The difficulty in assessing this trade off lies in the paucity of experience with lockdowns or other strict public health measures. Researchers may hope to speak to this trade-off by analyzing the experience of the United States during the Spanish Flu pandemic of 1918, as outlined in Correia, Luck, and Verner [2020].

Motivated by the evidence presented by Correia, Luck, and Verner [2020], we investigate the effects of non-pharmaceutical interventions (NPIs) using data from 43 US cities around the Spanish Flu pandemic

*Harvard University Economics Department and Harvard Business School. We thank Robert Barro for helpful discussions. This comment refers to the April 13th draft of the paper by Correia, Luck, and Verner.

of 1918. We measure the impact of NPIs on manufacturing employment and output in the 1899-1927 period and find that the large apparent positive effects from NPIs are driven by trends across cities which precede the pandemic. In fact, NPIs are strongly related to preexisting patterns in manufacturing employment and output growth from 1899 to 1914, before the onset of the pandemic. As such, measurements of medium run growth in city-level economic activity after 1914, such as those used by Correia, Luck, and Verner [2020], are likely to be picking up long-run trends which were correlated with the NPI treatment variables by chance.

In this brief note we caution against drawing conclusions from this experience, due to the confounding factors that cause divergent growth between cities in this period. After accounting for city-specific pre-trends, we find that both implementing NPIs earlier, and maintaining them for longer, had statistically insignificant effects on economic growth. Specifically, we find that the 95% confidence interval of the effect of implementing NPIs ten days earlier on manufacturing employment ranges from -8% to +12%. The effect on manufacturing employment of implementing three NPIs for an additional ten days each has a confidence interval from -5% to +5%. We conclude this episode does not provide clear evidence of the effect of non-pharmaceutical interventions on economic growth.

2 Data

Our data collection exercise expands on that of Correia, Luck, and Verner [2020].[1] We extend the sample of city-level manufacturing data used by Correia, Luck, and Verner [2020] by adding 1899, 1904, 1925 and 1927 to their 1909-1923 sample. We use the same data as Correia, Luck, and Verner [2020] on influenza mortality at the city-level for 1917 and 1918, which are collected from the Center for Disease Control's Mortality Statistics tables. Additionally, we use city-level population estimates from 1910 and 1917, from *Estimates of Population of the United States, 1910 to 1917*.

We also use the same measures of NPI brought to bear in Correia, Luck, and Verner [2020] - specifically the two NPI measures collected by Markel et al. [2007]. The first NPI measure is the total cumulative number of days that non-pharmaceutical interventions from the three major categories were activated during the 24-week study period from September 1918 through February 1919. The three categories are school closures, public gathering bans, and isolation/quarantine policies. If a city implemented one NPI for a week, the NPI intervention is 7 days, and if it implemented three NPIs for one week, the NPI intervention is 21 days. All 43 cities in the sample implemented at least one of these policies, for between one and ten weeks. The total number of intervention days ranges from 28 to 170, with an interquartile range of 49.5 to 136. Interventions such as business closures are not included in this measure. We refer

[1] We provide a detailed description of the sources in the Online Appendix Table A1. All data is available at https://almlgr.github.io/.

to this measure as Days of NPI or NPI intensity, interchangeably.

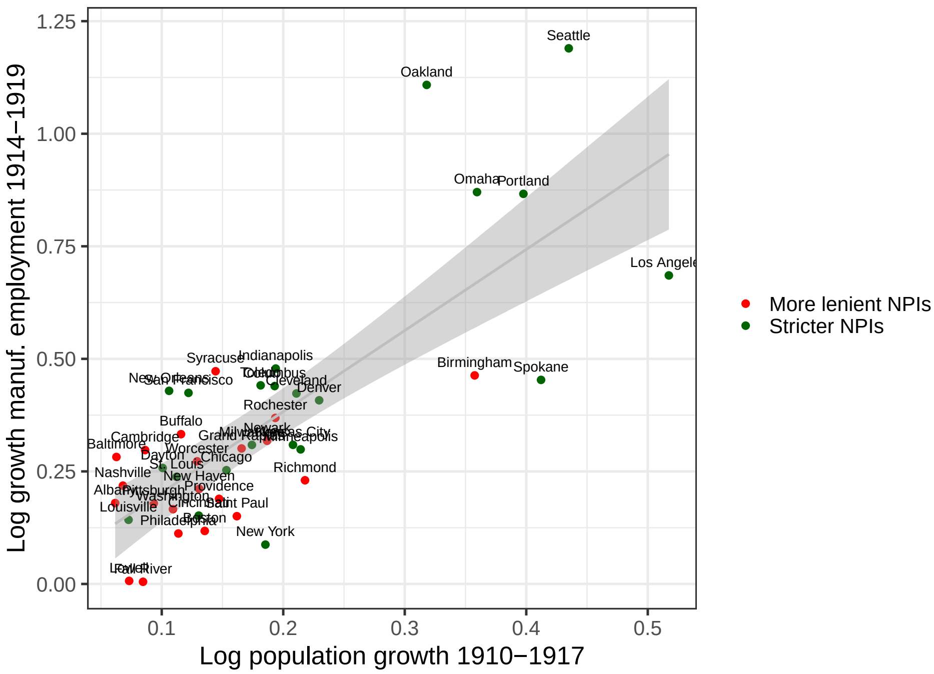
The second NPI measure, Speed of NPI, is the number of days each city took to implement its first NPI, after the excess death rate exceeded two times the baseline influenza and pneumonia death rate. The speed of NPI ranges from 11 days before (denoted +11) to 35 days afterward (denoted -35), with an interquartile range of 2.5 days to 10.5 days afterward. The first NPI for each city was implemented between the 18th of September and the 6th of November 1918.

3 Motivating evidence

Following Correia, Luck, and Verner [2020], we begin by correlating manufacturing employment growth from before the pandemic (in 1914) through to its end (1919) with NPIs. A limitation of this analysis is the long time interval in the measurement of output and employment. While the ideal outcome variable would be employment changes from 1918, the shortest time period overlapping 1918 for which employment data is available is 1914-1919. Therefore, since the first NPI is enacted in September 1918, those can only affect the outcome variable for at most 15 months of the five year period over which the outcome variable is measured.

City level manufacturing employment growth from 1914 to 1919 ranges from 0% to 230% in our sample, and some of this variation may be driven by differences in long-run labor force growth. This can confound the correlation of 1914-1919 employment growth with NPIs.

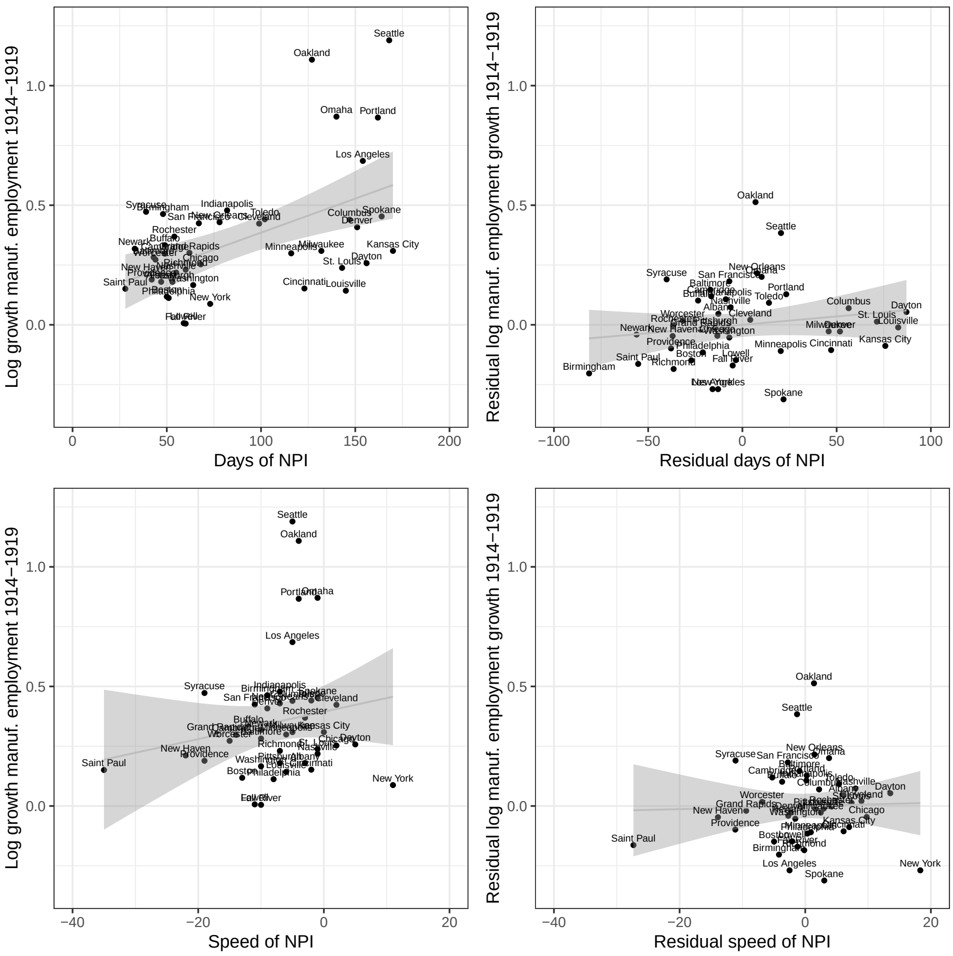
Figure 1: City-level manufacturing employment growth against population growth



This Figure shows the log growth of manufacturing employment from 1914 to 1919 against population growth from 1910 to 1917 at the city level. Red and green points denote cities for which days of NPI were below and above median, respectively.

To determine how much pre-existing trends may explain of total variation in employment, we use an estimate of city-level population growth between 1910 and 1917. The 1910 population data comes from the decennial census, while the 1917 data is estimated by the Census Bureau and published in a 1917 Bulletin. As this growth occurred (and was estimated) before the pandemic, it cannot be affected by it. In Figure 1 we show that much of city-level employment growth over 1914-1919 is explained by this measure of population growth. Moreover, the five most influential observations on the top right all have longer than median NPIs, and are located west of the Mississippi River, suggesting that spatial correlation of errors may pose an additional issue for inference.

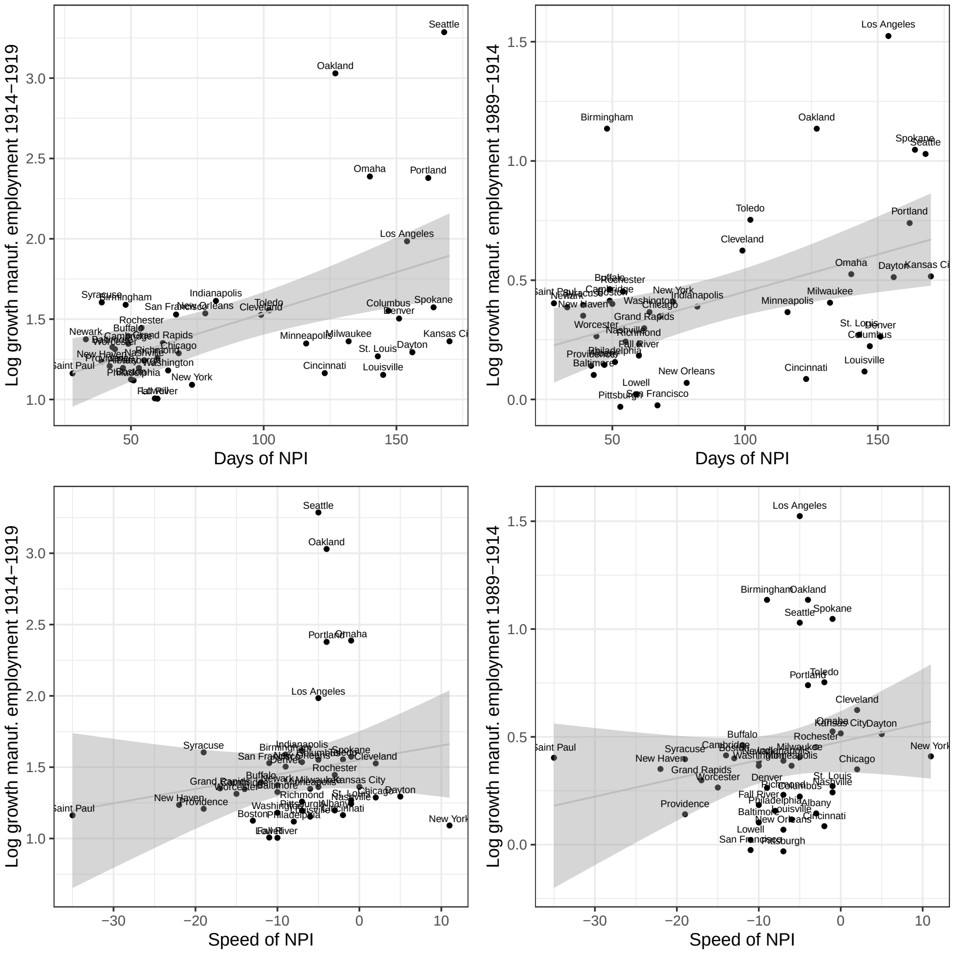
Figure 2: Manufacturing employment growth against Days and Speed of NPIs



The top left Panel of this Figure replicates Panel (a) of Figure 6 in Correia, Luck, and Verner [2020] in showing the relationship of log growth in manufacturing employment from 1914 to 1919 against Days of NPI. The bottom left Panel replicates Panel (b) of Figure 6 in Correia, Luck, and Verner [2020] in showing the relationship of log growth in manufacturing from 1914 to 1919 against the Speed of NPI adoption by city. The top and bottom right Panels report the same relationships after residualizing for log population growth from 1910 to 1917.

We reproduce the motivating evidence of Correia, Luck, and Verner [2020], but also residualize all variables controlling for population growth at the city level from 1910 through 1917. In Figure 2 we show city-level log manufacturing employment growth from 1914-1919 against the intensity (top panel) and speed of NPIs (bottom panel). The left panels of Figure 2 replicate Figure 6 Panel (a) and (b) of Correia, Luck, and Verner [2020] without any discernible differences. We then residualize both the outcome and explanatory variables by city-level population growth from 1910-1917, and show the relationship in the right panels. We find that neither relationship is robust to this population growth control. The absence of a relationship of NPIs and economic growth is consistent with recent work by Barro [2020] and Velde [2020].

Figure 3: Spurious correlation of Speed of NPI and Prior Employment Growth (1899-1914)



The top left Panel of this Figure replicates Panel (a) of Figure 6 in Correia, Luck, and Verner [2020] in showing the relationship of log growth in manufacturing employment from **1914 to 1919** against Days of NPI. The bottom left Panel replicates Panel (b) of Figure 6 in Correia, Luck, and Verner [2020] in showing the relationship of log growth in manufacturing from 1914 to 1919 against the Speed of NPI adoption by city. The top and bottom right panels report the same relationships but instead use the log growth of manufacturing employment from **1899 to 1914**, i.e. prior to the interventions. The heteroskedasticity-adjusted t-statistic of the slope of the coefficient of 1914-19 manufacturing employment growth on Days of NPI is 2.67, while for the period 1899-1914 we find a coefficient t-statistic of 2.60. For Speed of NPI, we find a t-statistic of 1.49 for 1914-1919, and 1.76 for 1899-1914.

Given the prevailing differences in population pre-trends, we investigate whether each of these variables of interest (Days of NPI, and Speed of NPI) are also correlated with employment manufacturing growth from 1899 to 1914, *before* the pandemic begins. In Figure 3, we plot both of these relationships.

City-level employment growth prior to the pandemic is spuriously correlated with future NPIs. This raises concerns for drawing inference from the following difference in differences analysis.

4 Difference in Differences Analysis

To formally investigate the impact of local policy interventions during the Spanish Flu, we follow Correia, Luck, and Verner [2020] in estimating difference-in-differences regressions. These capture the extent to which changes in economic outcomes following the pandemic relative to before covaried with NPIs, after accounting for dynamics explained by a set of control variables. To investigate how these effects evolved over time, we estimate a dynamic difference-in-differences specification of

$$\log(Y_{c,t}) = \alpha_c + \tau_t + \sum_{j \neq 1914} \beta_j NPI_{c,1918} \mathbb{1}_{j=t} + \sum_{j \neq 1914} \gamma_j X_c \mathbb{1}_{j=t} + \varepsilon_{c,t} \tag{1}$$

where $Y_{c,t}$ is manufacturing output or employment in a city c in year t, α_c and τ_t are city and year fixed effects, and $T_{c,1918}$ is the NPI treatment variable of interest. X_c are the city-level time-invariant controls of Correia, Luck, and Verner [2020], which we describe in the notes of the corresponding exhibits.

For each analysis, we first replicate the results of Correia, Luck, and Verner [2020] in their sample period from 1909 to 1923, and then extend the results to the 1899-1927 sample to analyze potential trends that could confound the analysis. Our replication closely, though not perfectly, matches the results in their paper, likely due to minor data discrepancies resulting from the manual digitization process.

4.1 Dynamic Difference in Differences

Starting with city-level NPI intensity, we estimate Equation 1 using the total cumulative days of NPIs per city, denoted $NPI_{c,1918}$. Figure 4 shows the point estimate and 95% confidence interval of the dynamic effect coefficients β_j. We again extend the sample earlier in time through to 1899. Consistent with Correia, Luck, and Verner [2020], we find that cities which implemented NPIs for longer grew faster after the pandemic than those which did not. Yet, the negative coefficient on NPIs in the pre-treatment years reveals that locations which implemented NPIs more aggressively grew faster than those which did not *both before the policy implementation, and afterward.*

To additionally illustrate the importance of these pretrends in driving the results, Figure 4 also shows the estimates from a model which adds a linear interaction between year and the treatment variable (here, Days of NPIs). The post-pandemic treatment coefficients, shown in green, are thus estimated relative to a counterfactual that the observed pre-trend prior to 1918 had continued in a smooth fashion. Here, these estimates are now of the opposite sign - relative to the exhibited pre-trend, the estimated effect of

Figure 4: Effect of NPI intensity on manufacturing employment and output

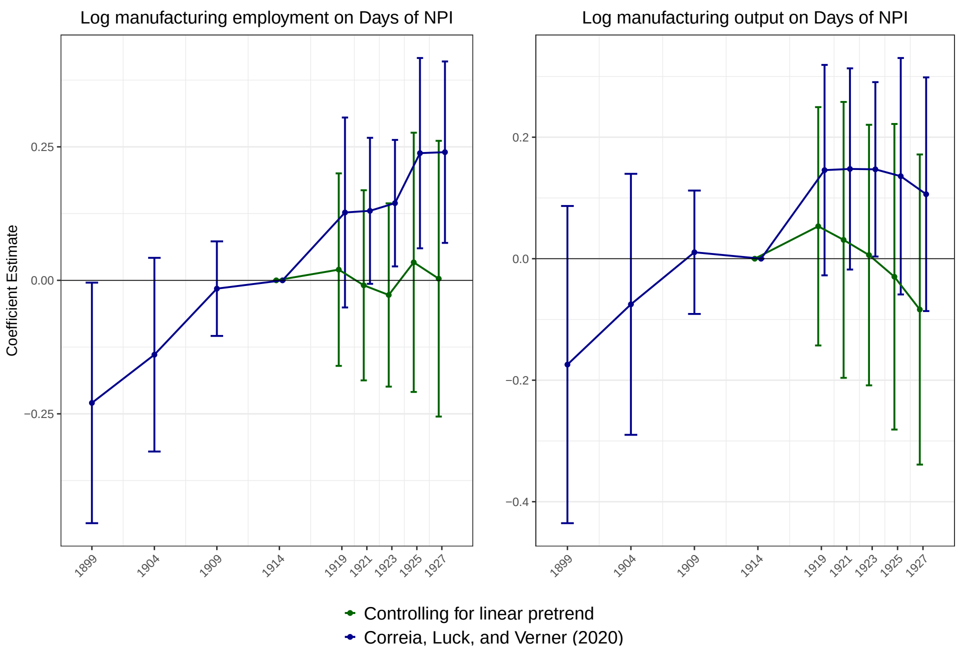


Figure 4 shows β_j for each year in our sample. These coefficients capture the effect of Days of NPI on log manufacturing employment compared to the 1914 baseline. All coefficient estimates are reported in percentage units. The estimates in blue replicate the specification of Correia, Luck, and Verner [2020] extending their 1909-1923 sample to include 1899, 1904, 1925 and 1927. They are obtained estimating a dynamic DiD model of the following form

$$\log(Y_{c,t}) = \alpha_c + \tau_t + \sum_{j \neq 1914} \beta_j NPI_{c,1918} \mathbb{1}_{j=t} + \sum_{j \neq 1914} \gamma_j X_c \mathbb{1}_{j=t} + \varepsilon_{c,t}.$$

The estimates in green attempt to naively correct for pre trends by estimating

$$\log(Y_{c,t}) = \alpha_c + \tau_t + \sum_{j > 1914} \beta_j NPI_{c,1918} \mathbb{1}_{j=t} + \lambda \cdot t \cdot NPI_{c,1918} + \sum_{j \neq 1914} \gamma_j X_c \mathbb{1}_{j=t} + \varepsilon_{c,t}$$

i.e. only keeping year by treatment fixed effects for post years while estimating a linear trend in time interacted with treatment off the variation in the pre period. For the left panel, $Y_{c,t}$ is manufacturing employment in city c and year t while for the right panel $Y_{c,t}$ is manufacturing output. X_c is a vector of city level controls, consisting of log 1910 population, percentage of the city population employed in manufacturing in 1914 and influenza mortality in 1917, as well as state-level measures of the 1910 urban share, agricultural employment share and per capital income. Error bands show 95% confidence intervals obtained from robust standard errors clustered at the city level.

pandemic mortality on manufacturing employment is positive but insignificant. These latter estimates should be interpreted cautiously, and as illustrative, rather than as causal estimates. In the context of strong prevailing trends, there is no guarantee of what precise functional form trends driven by other phenomena take, nor to what extent they will continue. However, the existence of pre-trends in the headline results makes these headline estimates reported in blue unreliable.

We repeat this exercise for the Speed of NPI measure in Figure 5. As with Correia et al. [2020], we find marginally significant results for output, but not employment. Once again, we note a weak trend of better employment outcomes from 1919 onwards for cities which implemented NPIs sooner, yet any inference from this relationship is undercut by the presence of equally strong pre-trends. The post-pandemic treatment coefficients shown in green are estimated relative to a counterfactual that the observed pre-trend prior to 1918 had continued in a linear fashion. We show them with the same aforementioned caveat - that these latter estimates should be interpreted as illustrative of the strength of pre-trends relative to the headline post-treatment estimates.

Figure 5: Effect of NPI speed on manufacturing employment and output

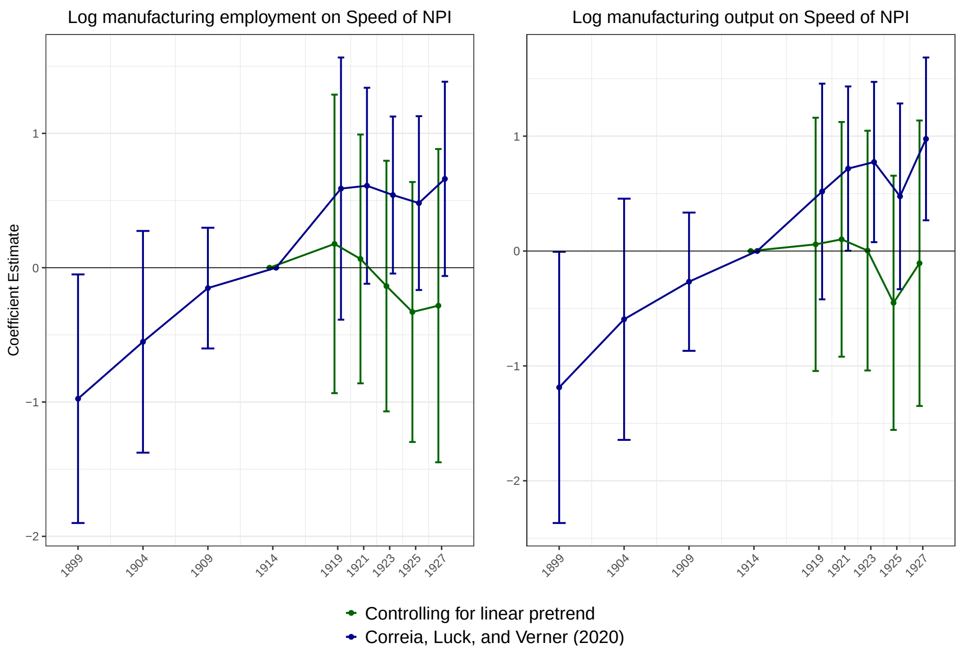


Figure 5 shows β_j for each year in our sample. These coefficients capture the effect of the Speed of NPI on log manufacturing employment compared to the 1914 baseline. All coefficient estimates are reported in percentage units. The estimates in blue replicate the specification of Correia, Luck, and Verner [2020] extending their 1909-1923 sample to include 1899, 1904, 1925 and 1927. They are obtained estimating a dynamic DiD model of the following form

$$\log(Y_{c,t}) = \alpha_c + \tau_t + \sum_{j\neq 1914} \beta_j NPI_{c,1918}\mathbb{1}_{j=t} + \sum_{j\neq 1914} \gamma_j X_c \mathbb{1}_{j=t} + \varepsilon_{c,t}.$$

The estimates in green attempt to naively correct for pre trends by estimating

$$\log(Y_{c,t}) = \alpha_c + \tau_t + \sum_{j>1914} \beta_j NPI_{c,1918}\mathbb{1}_{j=t} + \lambda \cdot t \cdot NPI_{c,1918} + \sum_{j\neq 1914} \gamma_j X_c \mathbb{1}_{j=t} + \varepsilon_{c,t}$$

i.e. only keeping year by treatment fixed effects for post years while estimating a linear trend in time interacted with treatment off the variation in the pre period. For the left panel, $Y_{c,t}$ is manufacturing employment in city c and year t while for the right panel $Y_{c,t}$ is manufacturing output. X_c is a vector of city level controls, consisting of log 1910 population, percentage of the city population employed in manufacturing in 1914 and influenza mortality in 1917, as well as state-level measures of the 1910 urban share, agricultural employment share and per capital income. Error bands show 95% confidence intervals obtained from robust standard errors clustered at the city level.

4.2 Pooled Difference in Differences

We also estimate pooled difference-in-differences models, where estimates capture the additional change in outcomes after the occurrence of the pandemic relative to prior, for each additional unit higher value of $NPI_{c,1918}$. This replaces the year-specific β_j treatment effects in Equation 1 with a single interaction between the post-period and $NPI_{c,1918}$ but leaves the specification otherwise unchanged.

Table 1 displays the results of the pooled difference-in-differences regressions. Column (2) shows the estimates from estimating the specification proposed by Correia, Luck, and Verner [2020], with our results closely matching the estimates they provide, which we provide for comparison in Column (1). Since this base specification is known to violate the parallel pre-trends assumption, in Column (3) we allow for city-specific time trends,[2] and in (4) we include a linear interaction between time and and the NPI intervention.

We find that NPI intensity and speed do not have statistically significant effects on employment growth after including these controls. Moreover, the confidence intervals for the effect of NPIs on manufacturing employment growth resulting from our specifications in columns (3) and (4) imply substantial uncertainty on the sign and magnitude of the effects. Using our estimates in column (3), we estimate that the 95% confidence interval of the effect of implementing NPIs ten days earlier on manufacturing employment ranges from -8% to +12%. The effect on manufacturing employment of implementing three NPIs for an additional ten days each has a confidence interval from -5% to +5%. We conclude this episode does not provide clear evidence of the effect of non-pharmaceutical interventions on economic growth.

We showed that the presence of differential pre-trends make drawing inference from the Spanish Flu episode challenging. Simple difference in differences models appear unreliable in this setting.

[2]If adding such unit-specific time trends drives to zero the coefficient of interest in a difference-in differences regression, the result was likely driven by pre existing differential trends across units Angrist and Pischke [2008].

Table 1: Manufacturing Employment and NPIs

	Manufacturing Employment			
	(1)	(2)	(3)	(4)
Panel A: NPI Days				
	Correia et.al.	Replication	City Trends	Pretrend
Post × NPI Days	0.133***	0.134**	0.002	0.002
	(0.058)	(0.060)	(0.082)	(0.078)
R-squared	0.39	0.75	0.91	0.75
Pretrend P-value				0.016
Cities	43	43	43	43
Observations	172	172	387	387
Panel B: NPI Speed				
	Correia et.al.	Replication	City Trends	Pretrend
Post × NPI Speed	0.565*	0.580	0.162	0.162
	(0.325)	(0.346)	(0.503)	(0.475)
R-squared	0.39	0.74	0.91	0.74
Pretrend P-value				0.076
Cities	43	43	43	43
Observations	172	172	387	387
Specification:				
City Time Trends	No	No	Yes	No
Treatment by Year Trend	No	No	No	Yes
Time FE	Yes	Yes	Yes	Yes
City FE	Yes	Yes	Yes	Yes

All coefficients are reported in percentage points (i.e multiplied by 100). Column (1) reports results from Correia et al. [2020] for comparison. The difference in R-squared compared to the other columns is due to a difference in the level of time interaction. Changing the year fixed effects interacted with controls in our specification (2) with a single post-1918 fixed effect recovers an R-squared very close to that of Correia et al. [2020]. Note that the coefficient estimate is not affected by this choice of time controls since it is identified off the covariance of average changes from pre to post 1918 with the treatment variable, and extracting year-specific (rather than period specific) constants leaves this unaffected. Column (2) includes 4 waves of data between 1914 and 1923 and estimates the following specification

$$\log(Y_{c,t}) = \alpha_c + \tau_t + \beta T_{c,1918}\mathbb{1}_{t\geq 1918} + \sum_{j\neq 1914} \gamma_j X_c \mathbb{1}_{j=t} + \varepsilon_{c,t}.$$

where $T_{c,1918}$ is Days of NPI for Panel A and Speed of NPI for Panel B. Columns (3) and (4) include 9 waves of data between 1899 and 1927 and each add additional controls to account for potential pre-trends. Column (3) adds linear city specific time trends $\alpha_c \times t$ while Column (4) adds a linear trend interacted with treatment $T_{c,1918} \cdot t$, the p-value for the test that the coefficient on this control equals zero is reported below the R-squared. X_c is a vector of city level controls, consisting of log 1910 population, percentage of the city population employed in manufacturing in 1914, influenza mortality in 1917, as well as state-level measures of the 1910 urban share, agricultural employment share and per capital income. Robust standard errors clustered at city level reported in parentheses. All regressions include a vector of controls interacted with year. * p<0.10, ** p<0.05, *** p<0.01.

References

Joshua D Angrist and Jörn-Steffen Pischke. *Mostly harmless econometrics: An empiricist's companion*. Princeton university press, 2008.

Robert J Barro. Non-pharmaceutical interventions and mortality in u.s. cities during the great influenza pandemic, 1918-1919. Working Paper 27049, National Bureau of Economic Research, April 2020. URL http://www.nber.org/papers/w27049.

Sergio Correia, Stephan Luck, and Emil Verner. Pandemics depress the economy, public health interventions do not: Evidence from the 1918 flu. *SSRN*, 2020. URL https://ssrn.com/abstract=3561560.

Howard Markel, Harvey B. Lipman, J. Alexander Navarro, Alexandra Sloan, Joseph R. Michalsen, Alexandra Minna Stern, and Martin S. Cetron. Nonpharmaceutical Interventions Implemented by US Cities During the 1918-1919 Influenza Pandemic. *JAMA*, 298(6):644–654, 08 2007. ISSN 0098-7484. doi: 10.1001/jama.298.6.644. URL https://doi.org/10.1001/jama.298.6.644.

Francois Velde. What happened to the us economy during the 1918 influenza pandemic? a view through high-frequency data. Working Paper 2020-11, Federal Reserve Bank of Chicago, April 2020. URL https://www.chicagofed.org/publications/working-papers/2020/2020-11.