Response to Lilley, Lilley, and Rinaldi (2020)

Sergio Correia, Stephan Luck, Emil Verner

May 15, 2020

We would like to thank Andrew Lilley, Matthew Lilley, and Gianluca Rinaldi for their comment on our paper "Pandemics Depress the Economy, Public Health Interventions Do Not: Evidence from 1918 Flu" (henceforth CLV). We welcome an open academic debate about the topic, and their comment will help us improve our paper.

Our paper makes two key observations:

1. Using variation across U.S. states and cities, we find that areas more severely affected by the 1918 Flu experience have lower economic activity after the pandemic. This finding is in line with cross-country evidence on the 1918 Flu Pandemic from Barro, Ursúa, and Weng (2020).

2. Using variation in the speed and intensity of the implementation of non-pharmaceutical interventions (NPIs) across U.S. cities during the second wave of the 1918 Flu, we find that cities that responded more aggressively do not experience worse economic outcomes, and, if anything, perform better after 1918.

The note from Lilley et al. (2020), henceforth LLR, raises concerns about the second finding. LLR argue that our main finding on the effect of NPIs is confounded by differential population growth rates prior to 1918 in locations that implement stricter NPIs during fall 1918. If cities with more aggressive NPIs were growing faster before 1918, our estimates may be confounded by these pre-trends.

We raise two main points in response to the findings presented by LLR:

1. LLR's argument that "city-level population growth from 1910-1917" is a confounding factor hinges on the use of population estimates from 1917. However, the Census Bureau emphasizes that these population values "are not to be considered in any sense as a census" and "are not based upon any detailed investigation of local conditions." Instead, the 1917 value is purely based on a linear extrapolation from the 1900 and 1910 census, which leads to large and systematic measurement error.

2. The extended dynamic difference-in-differences analysis suggested by LLR is sensitive to the estimation period and the control variables used. While there are no apparent pre-trends when using data from 9 years prior to the pandemic, concerns arise when using data from

19 years before the pandemic. Adding a control variable for past population growth— thus addressing LLR's original concern regarding differential population growth— alleviates the presence of pre-trends and leads to results that are closer to the original results in CLV, even when extending the estimation period.

In the following we elaborate on both points in more detail.

1 City-level population estimates from 1917 are unreliable

In their abstract, LLR write "[Correia, Luck, and Verner's (2020)] 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." The central element in this argument is the 1917 city-level population numbers, which LLR explain are "estimated by the Census Bureau and published in a 1917 Bulletin" (p. 4).

In the introduction to the 1917 Bulletin (the source used by LLR), the Census Bureau cautions against interpreting these numbers as reflecting any change in local conditions between 1910 and 1917:

> "It must be evident to any person who reads the text of the bulletin that the figures presented are not to be considered in any sense as a census and must not be so regarded. Moreover, they are not based upon any detailed investigation of local conditions which would affect the movement of the population..." *Estimates of the Population of the United States Bulletin 138* (page 5).

Instead of capturing local conditions, the 1917 city-level population estimates reported in the 1917 Bulletin are merely linear extrapolations using population from the 1900 and 1910 census. Specifically, the Bulletin writes that "The method of arithmetical progression was adopted for computing the estimates of population", and that "It is based on the assumption that the increase each year since the enumeration is equal to the annual increase from 1900 to 1910."

Given the importance of local population estimates for the WWI draft apportionment, local population estimates were an important topic of debate at the time. Contemporary observers emphasized that the 1917 population estimates were unreliable. For example, in their discussion of the controversy surrounding the WWI draft, the *New York Tribune* noted that "The [War] department had available in 1917 the estimates of population which the census bureau issues annually. These are not very trustworthy."[1]

To understand the magnitude of the error and potential bias in the 1917 Bulletin estimates, it is useful to compare the 1917 Bulletin figures with the 1910 and 1920 census counts. Let P_{1917}^{Bul} be the 1917 Bulletin value. An imperfect but more standard approach is to estimate 1917 population by

[1] See `https://chroniclingamerica.loc.gov/lccn/sn83030214/1920-10-12/ed-1/seq-10/` (New York Tribune, Tuesday October 12, 1920).

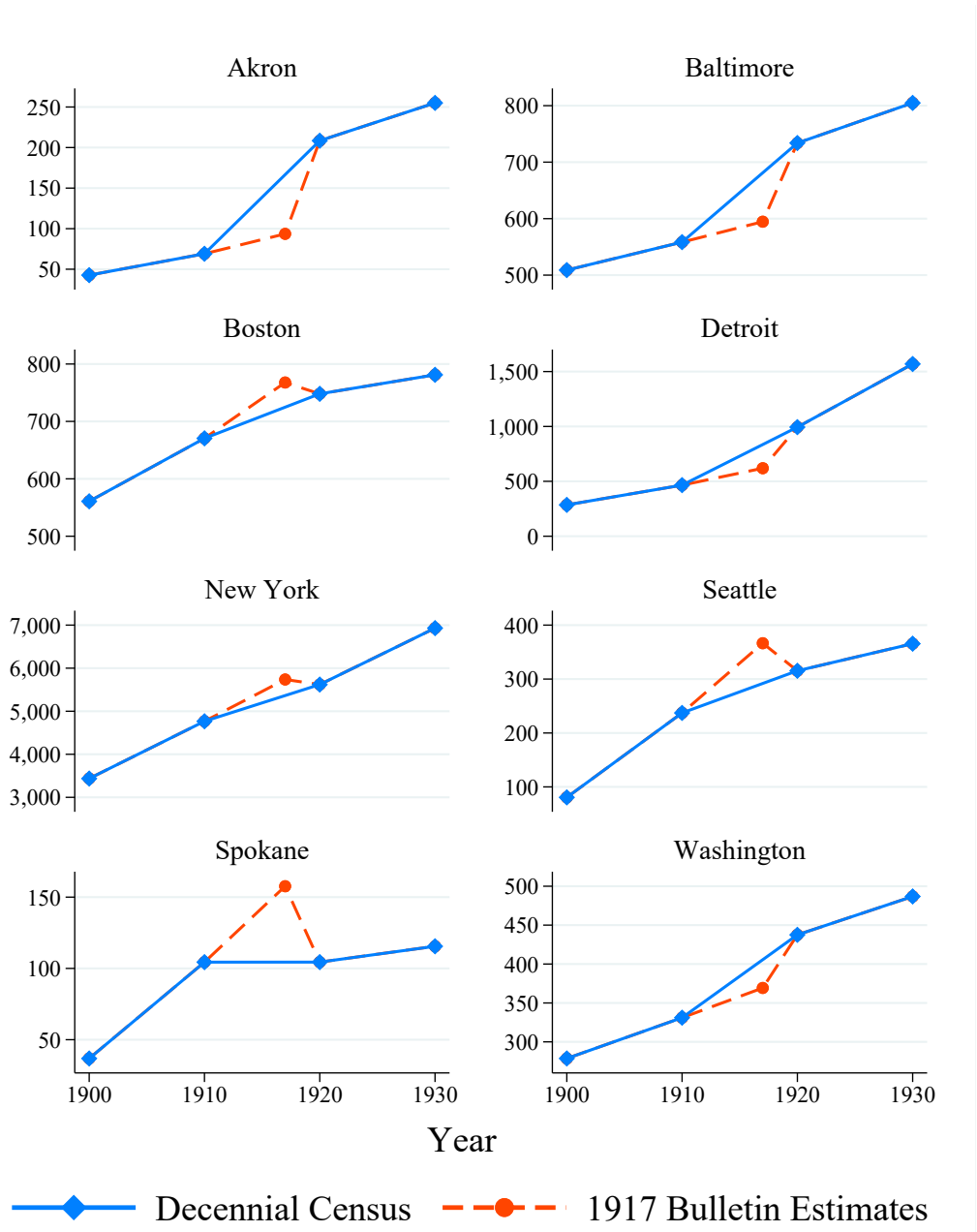


Figure 1: Population figures for eight cities. Numbers from the Decennial Census are in blue diamonds, and the 1917 Bulletin population figures are in red circles.

linearly interpolating between the 1910 and 1920 census estimates, P_{1917}^{Int}. Of course, using the 1920 census is confounded by the 1918 Flu Pandemic, but this exercise nevertheless serves to highlight the large and systematic errors in the 1917 Bulletin estimates.

Figure 1 presents examples for several cities. As an example, the 1917 Bulletin estimate for Seattle is not only above the 1920 Decennial Census; it is also above the 1930 Decennial Census. The 1917 Bulletin value for Spokane, another city in our sample, is over 50% higher than

population in the 1910 and 1920 Censuses. This pattern exists for many cities in the sample.[2] The mean absolute deviation between the linearly interpolated 1917 values and the 1917 Bulletin estimates, $\frac{1}{N}\sum_c |P_c^{Bul} - P_c^{Int}|$, is 6.7% for the sample of 43 cities with information on NPIs.[3] This approximate error is an order of magnitude larger than estimated mortality from the 1918 Flu Pandemic (0.66%).

Figure 2a uses the discrepancy between the 1917 Bulletin and the 1910-1920 linearly interpolated estimates to categorize cities. The figure shows large and regionally concentrated discrepancies in the two estimates. Cities in the west are significantly more likely to have upwardly biased 1917 Bulletin estimates, as these cities grew faster over 1900-1910 than over 1910-1920 (see Figure 2b). As a result, the error in the 1917 Bulletin estimates are likely to be systematically and spuriously correlated with variables of interest. In particular, as argued in CLV, cities further west were quicker to implement NPIs as they learned from cities affected earlier by the pandemic in the east (see CLV for details). This makes the 1917 Bulletin population figures not only mismeasured, but also biased.[4]

In sum, we fully acknowledge and share the concern that other omitted variables may drive the positive correlation between NPIs and manufacturing growth. Table A7 in the April 10, 2020 CLV draft addressed some, but certainly not all, of the potential omitted variables. However, using mismeasured and biased estimates of population growth, as suggested by LLR, is unlikely to address the concern they raise and may introduce significant bias to the estimates.

2 Pre-trends in manufacturing outcomes

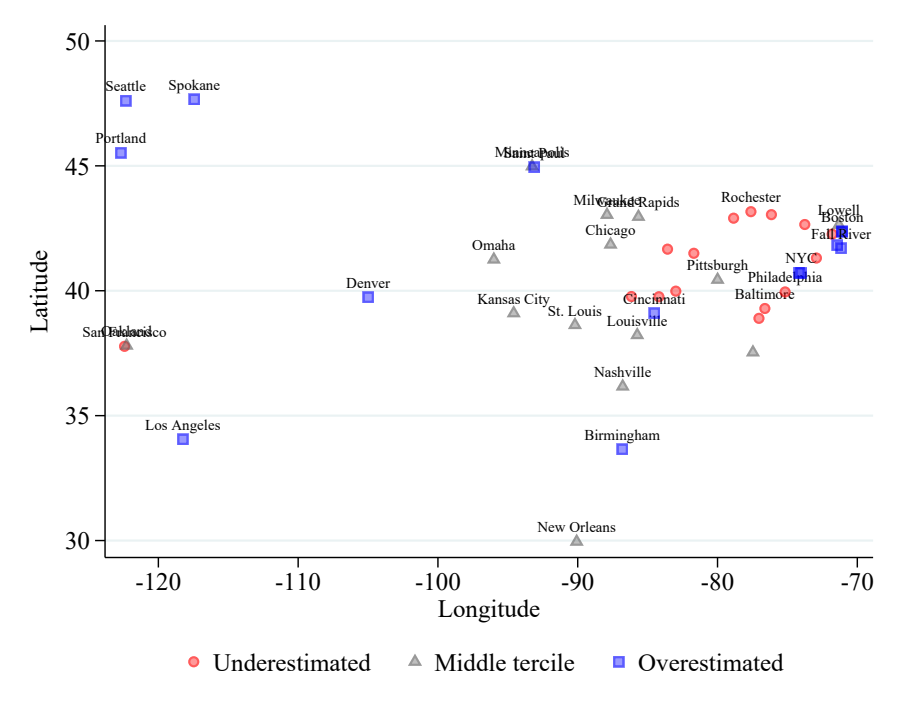
In the original sample in CLV, we examined manufacturing data back to the 1909 Census of Manufactures, nine years before the 1918 Pandemic. This allowed us to examine pre-trends in manufacturing employment from 1909 to 1914. LLR extend the analysis back to 1899, adding data for 1899 and 1904. Using data from up to 19 years before the 1918 Pandemic instead of 9 years, LLR find that there is a pre-trend in manufacturing outcomes in high NPI cities from 1899 to 1914. In particular, LLR write, "We also extend [CLV's] 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..."

To make their argument, LLR start out by correlating manufacturing growth from 1899 to 1914 with the intensity of NPIs and find a positive correlation (Figure 3 in LLR). LLR conclude that

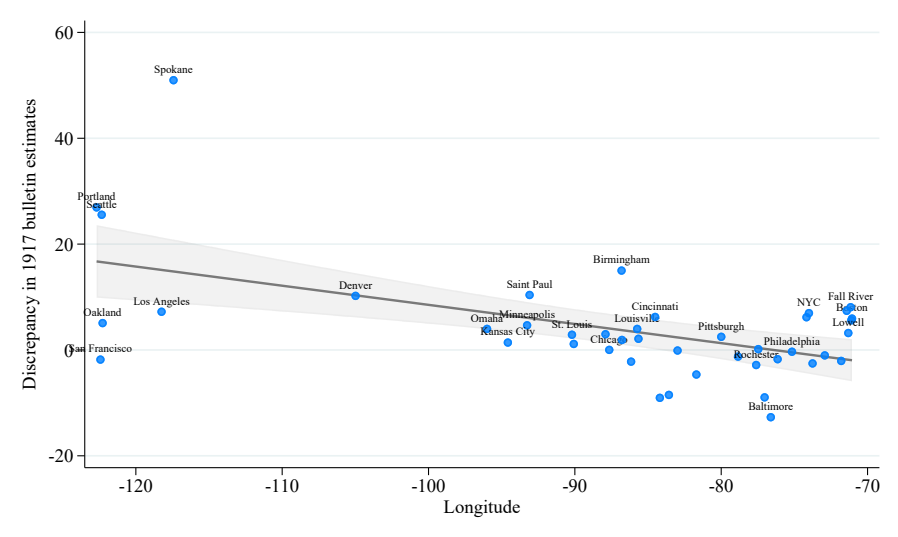
[2] See `http://scorreia.com/research/llr-all-cities.pdf` for plots of every city.

[3] Results are similar using a second estimate for 1917 population, $\tilde{P}_{1917}^{Int}$, which interpolates between the counts of the 1910 and 1920 decennial census, while adjusting for excess influenza and mortality rate from the 1918 pandemic, as compiled by Markel et al. (2007). The corresponding mean absolute error using the mortality adjusted interpolation is 6.72%.

[4] Controlling for population growth from 1910 to 1917 based on the linearly interpolated trend between the 1910 to 1920 Censuses leads to conclusions that are more in line with the original finding in CLV. This control is, of course, problematic, as it uses information from after the 1918 Pandemic, but it nevertheless contains substantially less measurement error than the 1917 Bulletin estimate and provides a better picture of population dynamics in the years around the pandemic.



(a) "Map" of the cities in the sample, colored by discrepancy of the 1917 population estimates. Cities in the lower tercile (where the bulletin estimates underestimated 1917 population) are in red circles; cities in the middle tercile are in the gray triangle; cities in the upper tercile (where the bulletin estimates overestimated 1917 population) are in blue squares. This figure shows mismeasurement is correlated with geography.



(b) Error in 1917 Bulletin estimates are large and correlated with longitude. This figure reports the deviation between 1917 Bulletin population estimates and estimates based on interpolation between the 1910 and 1920 Census (y-axis) against city longitude (x-axis).

Figure 2: Systematic bias of the 1917 population estimates from the Census Bulletin

"city-level employment growth prior to the pandemic is spuriously correlated with future NPIs."

Correlating employment growth in the period prior to the pandemic with measures of NPIs is equivalent to a placebo test. However, in the context of our paper, we believe it is more informative to conduct this type of test not using data with up to 19 years prior to the pandemic, but rather use the data from the censuses most close to the Pandemic (1909 and 1914). After all, the structure of the U.S. economy and U.S. cities changed quickly around the turn of the 20th century. For instance, cities like San Francisco, Los Angeles or Seattle changed considerably between 1899 and 1909, raising the question of how suitable they are for a placebo test and about the appropriate length of pre-trends. Moreover, using data too far in the past, such as the 1899 Census, has also practical limitations: because city boundaries change over time, comparisons across distant census years can be problematic.[5]

Figure 3 shows the correlation between NPIs and manufacturing growth 1909-1914 and 1914 and 1919. There are two important patterns. First, cities with stricter NPIs experienced higher manufacturing employment growth between 1914 and 1919. Second, cities with stricter NPIs did not exhibit faster manufacturing growth between 1909 and 1914. Thus, the pre-trend concern raised by LLR only applies when using data from 9 to 19 years prior to the pandemic, but not when using data from 4 to 9 years before the pandemic.

Nevertheless, we take the concern about pre-trends raised by LLR seriously. We thus estimate the original specification used in CLV that interacts treatment and control variables with year fixed effects using the longer time series series used by LLR.[6] Estimating this model with the baseline controls used in CLV (labelled "Controls" in panel (a) of Figure 4) shows indeed that cities with stricter NPIs grew faster than those with more lenient NPIs over 1899-1904 and 1904-09 (although not over 1909-14).

We then estimate the model again, also controlling for log population in 1900 in addition to log population in 1910, (labeled "Controls + ln(pop1900)" in panel (a) of Figure 4). This arguably helps

[5]To address this, the Census Bureau standardizes the definition of a city within a given publication, recomputing city-level totals whenever boundary changes occur. As an example, note the Boston annexation of Hyde Park in 1912: the Thirteenth Census from 1910 reported 69,637 workers in Boston for 1909, but the Fourteenth Census from 1920 reported 73,957 workers *also for 1909*. This discrepancy is due to Hyde Park workers, who were counted as part of Boston in the 1920 Census to facilitate for comparisons over time. Because LLR obtained the 1909 data from the 1910 Census, they selected the lower, earlier number of 69,637 workers, and thus they overestimate growth rates between 1909 and 1914. This issue is explicitly discussed by the Census Bureau. For instance, the 1909 Census of Manufactuters (Vol. 9, Table 5) states that "the figures for some cities do not agree with those published in 1904, because it was necessary to revise the totals in order to include data only for those establishments located within the corporate limits of the city." Moreover, there are footnotes in every table indicating the largest boundary changes, such as the Boston one. This issue can be avoided for the 1904 and 1909 data by simply using the same source as for the 1914 and 1919 data. However, this is not possible for 1899 as we do not know of any document encompassing both 1899 and 1919.

[6]This is the same specification as used in our original draft:

$$Y_{ct} = \alpha_c + \tau_t + \sum_{j \neq 1914} \beta_j NPI_{c,1918} \mathbf{1}_{j=t} + \sum_{j \neq 1914} X_s \gamma_j \mathbf{1}_{j=t} + \varepsilon_{ct}, \quad (1)$$

. In the analysis in this note, we focus on manufacturing employment as the dependent variable and measure NPIs by the cumulative total of days in which three main NPIs were implemented, $NPIIntensity_{c,1918}$. In our April 10, 2020 draft and in LLR, X_c include 1910 state-level agriculture employment share, 1910 state-level urban population share, 1910 state-level income per capita, 1914 manufacturing employment to 1910 population, and log 1910 population.

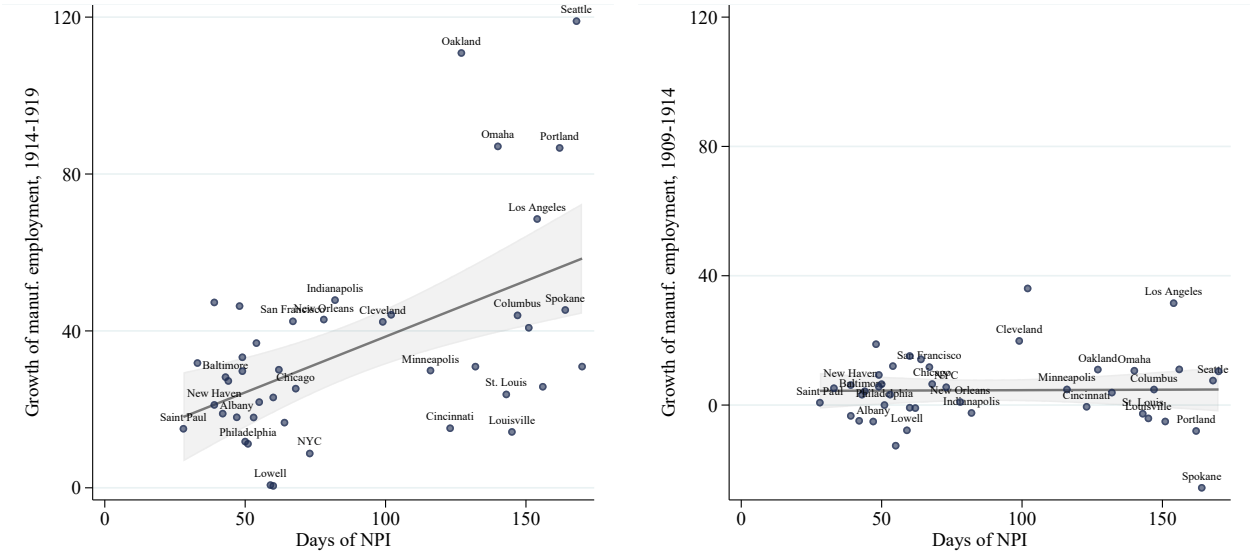


Figure 3: NPI Intensity and employment growth over 1914-1919 and 1909-1914.

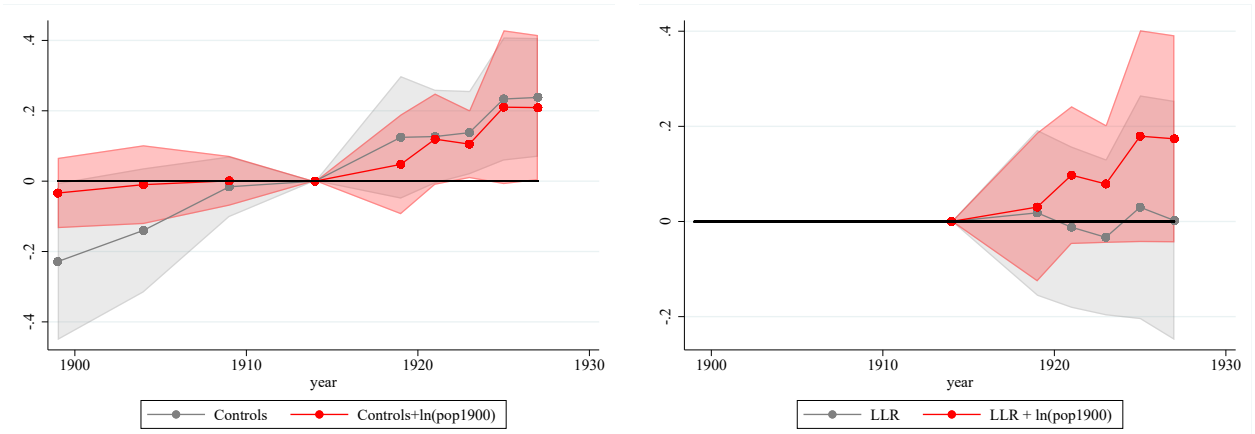
to address the original concern that cities with stricter NPIs were growing faster in the past. Once we control for population growth using reliable population estimates from the census, the presence of pre-trends is significantly mitigated: after controlling also for past population growth, cities with more and less strict NPIs during Fall 1918 have about the same dynamics in manufacturing employment in 1899, 1904, 1909, and 1914. Moreover, the estimates from 1919 onward are closer to the original finding in CLV: cities with stricter NPIs do not seem to perform worse. If anything, the confidence bands suggest that cities with strict NPIs exhibit better economic performance in the years after the 1918 Flu Pandemic.

Further, panel (b) of Figure 4 shows estimates from replicating the detrended difference-in-differences equation estimate by LLR, using data back to 1899 (labeled "LLR").[7] The findings suggest that when controlling for the pre-trend "naively" (as suggested by LLR), there are no effects on manufacturing employment after 1918. We then also estimate their model controlling for past population growth as above. Again, we believe that controlling for city-level population growth is a more direct way to address the original concern than solely a linear interaction term, which also imposes strong parametric assumptions. As in panel (a), panel (b) suggests that once we control for population in 1900 (labelled "LLR + ln(pop1900)"), the estimation reveals a pattern closer to the one found in CLV.

In our view, the main concern in CLV is that we do not have annual manufacturing census data from 1914 to 1919. This is a concern that we emphasized several times (e.g., on page 26 and in the conclusion on page 30). We do not believe that adding data back to 1899 is a substitute for

[7]Specifically, LLR suggest estimating a model of the following form:

$$Y_{ct} = \alpha_c + \tau_t + \sum_{j>1914} \beta_j \times NPI_{c,1918}\mathbf{1}_{j=t} + \lambda \times t \times NPI_{c,1918} + \sum_{j\neq 1914} \gamma_j X_c \mathbf{1}_{j=t} + \epsilon_{ct} \quad (2)$$



(a) (b)

Figure 4: Effects of NPIs (NPI intensity) on manufacturing employment. This figure shows the results from estimating Equation (1) (panel (a)) and Equation (2) (panel (b)) using manufacturing employment outcomes from 1899, 1904, 1909, 1914, 1919, 1921, 1923, 1925, 1927.

having data between 1914 to 1919.[8]

Does this mean that pre-trends are not a concern in our setup? No. As in any difference-in-difference specification, they are a concern that needs to be taken seriously. LLR point out that our finding that NPIs do not hurt the economy is not necessarily robust to removing certain controls and thus subject to some uncertainty. Nonetheless, this response shows that our original conclusion that there is no obvious trade-off between "flattening the curve" and economic activity is largely robust to additional data from 1899 and 1904 and to changing the econometric specification to account for pre-trends.

Summary The broader discussion in CLV centers around the question whether an economy would do better in the absence of NPIs, or whether the pandemic itself already is a source of economic slowdown. Is there something we can learn from the 1918 Flu Pandemic? We believe the answer is yes. We interpret the combined evidence presented in CLV and in this response as an indication that a) the 1918 Flu Pandemic itself had adverse economic consequences and b) it is not a foregone conclusion that NPIs exacerbated the economic disruptions.

We believe that our analysis can be one data point in a wider discussion on the economic effects of pandemics. However, we also believe that our findings have to be taken with a grain of salt and the evidence presented is indicative, not conclusive. As pointed out in CLV, there are several important caveats that give rise to some uncertainty regarding our findings on the effects of NPIs: Our sample that allows to study the effect of NPIs is restricted to only 43 cities. The second wave of the 1918 Flu pandemic coincided with the end of WW1. Moreover, NPIs were often stricter cities located in the west, giving rise to the concern that differences in outcomes can

[8]Our approach was to instead examine annual data on bank assets, which constitute mostly of loans to local businesses and households, as a proxy for local economic conditions.

be confounded by different city growth trajectories that are independent of the 1918 Flu. Carefully controlling for observable differences across cities, as done in this response, can alleviate some, but certainly not all of those concerns.

The 1918 Flu Pandemic provides an exciting historical laboratory, especially in the light of recent events. We welcome the academic debate and comments made by LLR, as they help to strengthen and extend our joint understanding of this important historical episode. We are currently updating our original draft, incorporating comments from LLR and others.

References

Barro, R. J., J. F. Ursúa, and J. Weng (2020). The coronavirus and the great influenza pandemic: Lessons from the "spanish flu" for the coronavirus's potential effects on mortality and economic activity. Working Paper 26866, National Bureau of Economic Research.

Correia, S., S. Luck, and E. Verner (2020). Pandemics depress the economy, public health interventions do not: Evidence from the 1918 flu. *Unpublished. Available online at https://ssrn.com/abstract=3561560*.

Lilley, A., M. Lilley, and G. Rinaldi (2020). Public health interventions and economic growth: Revisiting the spanish flu evidence. *Unpublished. Available online at https://ssrn.com/abstract=3590008*.

Markel, H., H. B. Lipman, J. A. Navarro, A. Sloan, J. R. Michalsen, A. M. Stern, and M. S. Cetron (2007). Nonpharmaceutical Interventions Implemented by US Cities During the 1918-1919 Influenza Pandemic. *JAMA 298*(6), 644–654.