WEATHER, MOBILITY AND THE EVOLUTION OF THE COVID-19 PANDEMIC

2021

BANCO DE ESPAÑA
Eurosistema

Documentos de Trabajo
N.º 2109

Corinna Ghirelli, Andrea González, José Luis Herrera
and Samuel Hurtado

WEATHER, MOBILITY AND THE EVOLUTION OF THE COVID-19 PANDEMIC (*)

Corinna Ghirelli, Andrea González, José Luis Herrera
and Samuel Hurtado

BANCO DE ESPAÑA

(*) We thank Ángel Gavilán, José Gonzalez, Enrique Moral, Matías Pacce, Alberto Urtasun, Ernesto Villanueva and our anonymous referee for their comments and suggestions. We also thank all participants at the internal seminar of the Bank of Spain for their helpful comments. The views expressed in this paper are solely ours and should not be interpreted as reflecting the views of the Bank of Spain, nor those of the Eurosystem.

Documentos de Trabajo. N.º 2109

2021

The Working Paper Series seeks to disseminate original research in economics and finance. All papers have been anonymously refereed. By publishing these papers, the Banco de España aims to contribute to economic analysis and, in particular, to knowledge of the Spanish economy and its international environment.

The opinions and analyses in the Working Paper Series are the responsibility of the authors and, therefore, do not necessarily coincide with those of the Banco de España or the Eurosystem.

The Banco de España disseminates its main reports and most of its publications via the Internet at the following website: http://www.bde.es.

ISSN: 1579-8666 (on line)

Abstract

We estimate the effective reproduction number (Rt) of the current Covid-19 pandemic, with US daily infections data between February and September of 2020, at the county level. This is then used to estimate the effect of weather and mobility on the spread of the pandemic. We find a strong and significant effect of the weather: lower temperatures are associated with a higher Rt, and this effect is bigger at temperatures below 0ºC. At low temperatures, precipitations are also associated with a higher Rt. We also find that mobility reductions related to certain types of locations (retail and recreation, transit stations, and workplaces) are effective at reducing Rt, but it is an increase of the time spent in parks that helps reduce the spread of the pandemic. The negative effect of increased general mobility is bigger in counties with higher population density, worse numeracy and literacy PIAAC scores, or a lower share of employment in the services sector. Quantitatively, our estimates imply that a 20ºC fall in temperatures from summer to winter would increase Rt by +0.35, which can be the difference between a well-controlled evolution and explosive behavior; and, if this can't be neutralized through general improvements in the fight to stop the pandemic, the additional reduction in mobility that would be needed to compensate for this would be equivalent to returning, from the more relaxed levels observed in the summer, back to the strictest mobility reductions recorded in the US in April.

Keywords: pandemic, Covid-19, coronavirus, temperature, weather, mobility, panel data.

JEL classification: I18, C23, I12.

Resumen

Estimamos el coeficiente de reproducción efectivo (Rt) de la pandemia de Covid-19, con datos de casos detectados en Estados Unidos, a nivel de condado, entre febrero y septiembre de 2020. Con esta información estimamos el efecto de la climatología y de la movilidad sobre el ritmo de expansión de la pandemia. Encontramos un efecto fuerte y significativo de la meteorología: menores temperaturas están asociadas con mayores ritmos de expansión, y este efecto es mayor por debajo de 0ºC. Además, cuando las temperaturas son bajas, las precipitaciones están asociadas con un mayor Rt. En cuanto a la movilidad, encontramos que, en los casos de locales de comercio y recreo, estaciones de tránsito y lugares de trabajo, las reducciones de movilidad son efectivas en términos de frenar el ritmo de expansión de la pandemia, mientras que en el caso de los parques es un aumento de las visitas lo que ayuda a reducir el Rt. El efecto negativo de los aumentos de la movilidad general es mayor en condados con alta densidad de población, peores resultados en las pruebas PIAAC sobre comprensión lectora y de cálculo numérico, o una menor proporción de empleo en el sector servicios. Cuantitativamente, nuestras estimaciones implican que una reducción de 20ºC en las temperaturas entre verano e invierno aumentaría el Rt en +0,35, que puede ser la diferencia entre una evolución bien controlada y un comportamiento explosivo; y si esto no se consigue neutralizar mediante mejoras generales en la estrategia de lucha contra la pandemia, la reducción adicional en la movilidad que sería necesaria para compensar ese efecto es equivalente a retornar, desde los niveles más relajados apreciados en verano, a los niveles más estrictos de reducción de la movilidad observados en Estados Unidos en abril.

Palabras clave: pandemia, Covid-19, coronavirus, temperatura, climatología, meteorología, movilidad, restricciones, datos de panel.

Códigos JEL: I18, C23, I12.

1 Introduction

The Covid-19 pandemic is exacting a very heavy cost in human lives despite the extraordinary measures adopted to contain its spread. It has also caused far-reaching disruption to society and to the global economy throughout 2020. With the pandemic still ongoing, assessing the determinants of its evolution and evaluating the effects of containment measures is of major importance.

In this paper we estimate the effective reproduction number (R_t) of the current Covid-19 pandemic, with US data about the daily number of infections between February and September of 2020, at the county level. This is then used to estimate the effect on the spread of the pandemic of weather and mobility variables. The panel estimation uses county dummies and dummies for state-month interactions, so that the estimated effects capture the differential evolution of, say, warmer or cooler counties within a state and in the same month. Focusing on R_t instead of the cumulative number of cases or deaths allows us to assess the dynamic effects of time-varying determinants, and provides a more useful quantification of the effects of changes in the covariates.

We find a strong and significant effect of the weather, which is troubling for the coming months in the Northern hemisphere: lower temperatures are associated with a higher effective reproduction number; this effect is bigger at temperatures below ºC, and, at low temperatures, precipitations are also associated with a higher R_t. These effects are sizeable: according to our preferred linear specification, a summer-to-winter decrease of 20ºC would imply an increase in the effective reproduction number by approximately 0.35, which can be the difference between a well-controlled evolution and explosive behavior.

We also find that mobility reductions related to certain types of locations (retail and recreation, transit stations, and workplaces) are effective at reducing R_t. On the other hand, and in line with previous literature about the probability of infection in open and closed spaces, it is an increase of the time spent in parks that helps reduce the spread of the virus. The negative effect of increased mobility is bigger in counties with higher population density, worse numeracy and literacy PIAAC scores, or a lower share of employment in the services sector.

2 Previous studies

A few papers have already studied the effect of weather and mobility on the spread of the Covid-19 pandemic. In the areas where they intersect, our results are in line with what these previous articles found, but, compared with them, we put together, in one estimation, the following important factors: (i) use R_t as the variable of interest, instead of cumulative number of cases or deaths, (ii) use data for an area and time period with high heterogeneity in all observed dimensions, (iii) use a big data panel with very granular time and spatial information, and (iv) include both weather and mobility information.

Smith et al (2020) shares three of those factors (and also provides a detailed discussion of why they are important). The one where we differ is the granularity of the data: they run regressions with data for US states before and after lockdown, we use daily information at county level. In this sense, these studies complement each other nicely: they estimate the effect of temperature and mobility using big differences across states and comparing situations before and after stay-at-home orders, whereas in our estimation those big differences go to the fixed effects (together with all unobserved differences in characteristics that differ across counties but not over time, or across months but not day-to-day) and the estimates capture the impact of temperature and mobility by looking at subtle differences within a state in a given month. With a bigger database, our regressions can simultaneously include a relatively high number of covariates and non-linear interactions between them; on the other hand, they also run estimates within a mechanistic epidemiological model, simultaneously identifying both R_t and the effects of temperature and mobility, instead of our two-step approach (first estimate R_t, then see how it relates to the covariates). And they focus more on R0 and the initial wave of the pandemic, whereas we use a longer time period extending until the end of august. Apart from the methodological differences, in terms of the estimation results, both studies are in line: mobility is a major factor in determining the speed of the spread of the pandemic, but temperature also matters, and further reductions in mobility may be needed to face the coming winter in the Northern hemisphere[1].

A second paper that is very close to ours is Wilson (2020): he uses basically the same data (R_t, weather and mobility in the US at county level) and a very similar methodology, but with some important differences[2]. Coming from central bankers, both papers use a methodology that's common in economics, but with emphasis in different directions. We skew the methodology in the direction of what is more common in epidemiology, using R_t, without further alterations, as our variable of interest. This is in line with the request of Smith et al (2020) who conclude that "the role of environment in transmission has become controversial, in part because of the application of models to case prevalence, rather than fundamental epidemiological parameters such as R". One downside of our approach is that it leaves more space for statistical problems in the estimation, which is the reason Wilson (2020) skews his methodology towards the best practices in economics, producing sound estimates that should be more robust to possible issues of identification or endogeneity, at the cost of deviating slightly from the use of "fundamental epidemiological parameters". His derivation of an empirical dynamic panel data model is clearly guided by the standard SIR model of disease spread, but in the end his object of interest is not strictly R_t, but an average of the rate of growth of cases over longer periods, which is a closely related concept but in a form that is less deeply rooted in epidemiology. Nevertheless, the fact that both papers reach similar conclusions is reassuring, because it provides robustness to the common results: he also finds that "holding mobility fixed, temperature reduces Covid-19 infections" and "holding weather fixed, mobility increases infections".

Glaeser et al (2020) looks at the effect of mobility on the spread of Covid-19 using data from New York and four other US cities, and Kapoor et al (2020) looks at the effect of social distancing in the initial spread of the pandemic, with differences in rainfall with respect to the local average as their instrument for social distancing. In terms of the previous discussion about best practices in epidemiology and economics, these two papers fall very close to the latter,

[1] There is one point where the results of these paper may seem to be in disagreement, but they really are not: they find that states with higher population density have a higher R_0, whereas in our non-linear estimation specifications we find that mobility has a bigger effect on R_t in counties with high population density. These estimates are identifying different effects of population density: they are capturing a worse starting point for states with higher population density, we are finding, for high-density areas, a higher effectiveness of reductions in mobility at fighting that disadvantage. These counteracting forces mean that further analysis would be needed in order to identify whether high-density areas require bigger reductions in mobility to bring their local epidemic under control, but in both cases we conclude that a relaxation of control measures would be particularly problematic in high-density areas.

[2] We focus the discussion on the methodological differences. Wilson (2020) also looks at the impact on mobility (and not only on the spread of the virus) of both weather and non-pharmaceutical interventions.

using instrumental variables that ensure exogeneity of the regressors but also using cumulative numbers of cases and deaths as their variables of interest.

Hamada et al (2020) uses data from 25 counties that had big Covid-19 outbreaks, and analyzes how mobility affected the rate of increase in the number of cases. They find a strong relationship between these variables, with Pearson correlation coefficients above 0·7 for 20 of the 25 counties evaluated. Chang et al (2020) provide more nuanced results in terms of mobility than we can assess, but don't investigate the effect of environmental variables such as temperature. They estimate a SEIR model with location data from millions of mobile phones in 10 US cities, and find that a small minority of "superspreader" points of interest account for a large majority of infections and that restricting maximum occupancy at each point of interest is more effective than uniformly reducing mobility. Their model also correctly predicts higher infection rates among disadvantaged racial and socioeconomic groups solely from differences in mobility: they find that disadvantaged groups have not been able to reduce mobility as sharply, and that they visit places that are more crowded and therefore higher-risk. Other recent papers that use very granular or even individual mobility data to study the spread of Covid-19 are Almagro et al (2020) and Couture et al (2020), both of which find that mobile phone data is suitable for quantifying movement and social contact, and is a useful tool for assessing and analyzing the evolution of the pandemic: out-of-home activity is strongly associated with Covid-19 infections.

The OECD's "Walking the tightrope: Avoiding a lockdown while containing the virus" looks at how containment measures (such as workplace or school closures) affect mobility, and how this in turn affects R_t. With a sample at the country level, they find that the biggest effect is for school closures, stay-at-home requirements and workplace closures, but it's not easy to distinguish between these because they have often been imposed at the same time.

Desmet and Wacziarg (2020) is also closely related to our research: they analyze the correlates of COVID-19 cases and deaths across US counties. They consider a wide range of correlates - population density, public transportation, age structure, nursing home residents, connectedness to source countries, etc. - finding that these variables are important predictors of variation in disease severity. Their results are in relation to the cumulative number of cases and deaths, which means they have to put a lot of effort into controlling for the different starting time of the epidemic on each region. By using R_t as our main variable of interest, we can avoid this problem, we can focus on the dynamic effect of time-varying variables such as temperature and mobility, and we also get a more useful quantification of the effects of our covariates, since keeping $R_t<1$ is the main target for a policymaker trying to keep the spread of the pandemic under control.

Baker et al (2020) use a climate-dependent epidemic model to simulate the Covid-19 pandemic, testing different scenarios of climate dependence based on known coronavirus biology. They find that weather matters but is not enough to substantially limit pandemic growth, as levels of susceptibility among the population remain the driving factor for the pandemic.

Poirier et al (2020) examine the spatial variability of (a proxy to) the basic reproductive numbers of Covid-19 across provinces and cities in China, and assess the effect of mobility, and of environmental factors such as temperature and humidity. Because of the limited size of their database (just a few dozen observations), they evaluate the effect of each factor in an independent regression. By using a much more granular database, with daily data for the US at the county level, we can run much more detailed estimations, with many covariates under the same specification, including non-linear interactions, plus time and space fixed effects that allow for a very robust assessment of the effects of our covariates.

Rosario et al (2020) evaluate the relationship between weather factors (temperature, humidity, solar radiation, wind speed, and rainfall) and Covid-19 infection, using data from six cities in the State of Rio de Janeiro, Brazil. Their database only spans 47 days, and doesn't offer much geographical heterogeneity, but they still, they find significant effects from several weather factors. But Sharma et al (2020) use a similar methodology with data for India and find opposite results for important variables such as temperature.

Hamidi, Sabouri and Ewing (2020) find that denser areas are hotspots to spread infections and increase mortality rates.

3 Data

We have constructed a big database of USA data at the county level, from February through September of 2020, combining information from several sources.

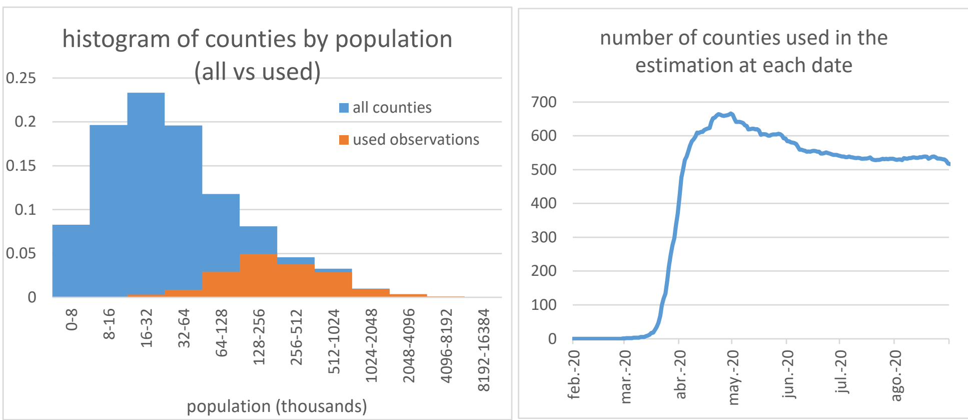
- **Epidemiological data**: we use the daily number of new confirmed Covid-19 cases in the United States at the county level. This data can be accessed through the American Center for Disease Control and Prevention (CDC) website, but is actually elaborated by the usafacts initiative, which compiles information published by different official local sources.
- **Effective reproduction number**: we estimate it following the approach of Cori et al. (2013), as implemented in the EpiEstim software package. As discussed by Gostic et al (2020), this method avoids biases induced by other alternatives, doesn't rely on assumptions that may not always apply, and is also particularly suited for real-time estimation of the evolution of an epidemic. Following Soucy et al (2020), we use rolling windows of 14 days of the number of confirmed Covid-19 cases (longer than the common windows of 7 days, in order to smooth out noise from the county data). The serial interval is specified as a full non-parametric distribution, following the results of Tindale et al (2020) with data from Singapore, which implies an approximately-symmetric distribution with an average serial interval of 4.5 days. When an outbreak is discovered, estimates for R_t can be artificially high; to prevent this from affecting the results, the estimated values are capped at 4 (any values higher than 4, which are typically found in the first days with data on some counties, are substituted by 4, which is a conservative estimate of the basic reproduction number R_0).
- **Weather data**: the variables of interest are the average temperature (degrees Celsius) and precipitations (mm). This is in line with several papers that include weather variables to explain COVID-19 propagation (see Xu et al, 2020; Carson et al, 2020). We use daily observations by weather station and geographical counties' information, which can be accessed through the NOAA's National Centers for Environmental Information (NCEI) and the National Weather Service website. The data at weather station level is assigned to every county by combining the measurements from the ten stations that are closest to the coordinates of the center of the county, with a weighted average where the weight is the inverse of that distance (1/d). This ensures that counties where there is no weather station get a measurement based on several nearby stations, and that stations close to the center of the county receive a disproportionately high weight in the average.
- **Mobility data**: we have collected these data from Google Global Mobility Reports which is a database to analyze how people move in their everyday life during the pandemic. Google elaborates this data using information from mobile phones, collected through

their applications such as Google Maps, which allows the company to record changes in the locations of phones anonymously. It is defined as the change with respect to baseline days before the pandemic outbreak (the median between January 3rd and February 6th). The information is sorted into different location categories: grocery and pharmacy, retail and recreation, public stations, parks, workspaces, and homes; in our estimations, we use both the indicator for parks, and, as an indicator of general mobility, an average of the indicators for retail and recreation, transit stations, and workplaces.

- **PIAAC**: the literacy and numeracy indicator comes from the Program for the International Assessment of Adult Competencies (PIAAC), which can be accessed through the National Center for Education Statistics. The PIAAC is a survey which collects data about the adult's performance in literacy, numeracy and problem solving. In general terms, it is able to assess the cognitive and workplace skills and allows to sort population into three main categories: people with low (P1), intermediate (P2) and high (P3) information-processing level. We use the share of the population that only attains low scores (P1); we use the average of the shares for literacy and numeracy.
- **Density**: inhabitants per squared kilometer, as found in the US census.
- **Services**: share the services sector in employment in 2014-2018, as found in US census.

The timing assumptions are as follows: (i) temperature and mobility today affect contagions today (they have an effect on how many contacts there are and how many of them end up in contagion), and (ii) contagions at t become identified cases at t+7 (the lag would probably be slightly shorter for cases found through contact tracing, and slightly longer for the rest). Following these assumptions, since R_t is estimated with a 14-day window of cases data, we apply the analogous transformation to temperature and mobility data, calculating a 14-day moving average, that then enters the estimations with a lag of one week. Therefore, average temperatures and mobility between t and t+13 affect the R_t estimated with cases data from t+7 to t+20.

Since the data is only available at county level, some of the observations refer to relatively big population groups (the Los Angeles county is home to almost 10 million people), whereas many counties with low population have to be left out of the estimation because they have too few cases and R_t can't be estimated, or because Google didn't publish mobility data because of privacy concerns; this is very common in counties with low population (almost no counties with less than 32,000 people have all the necessary data). Graph 1 represents, first, the histogram of counties by population, for all counties and for the ones that could be used in the estimation

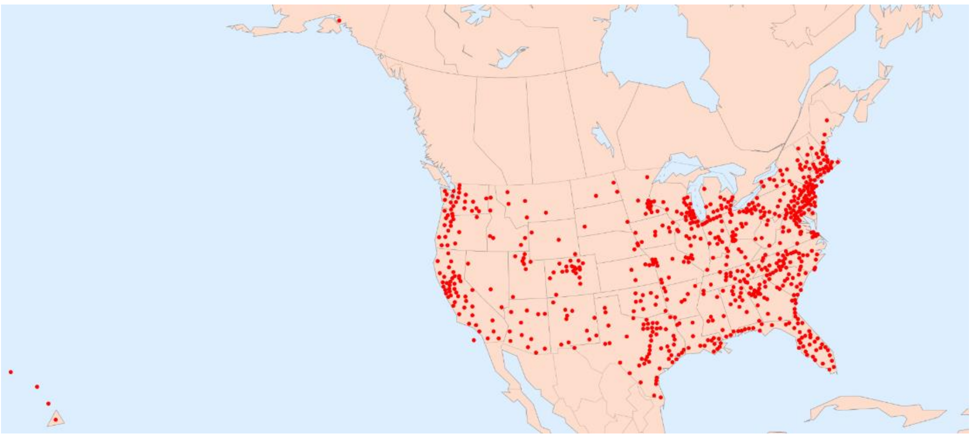


Graph 1: counties used in the estimation

because of data availability, and, second, the evolution of the number of counties used in the estimation across time (some counties disappear from the sample, typically because Google didn't publish mobility data for them during the summer because of increased privacy concerns, or because the number of cases became too small to provide an estimate of R_t).

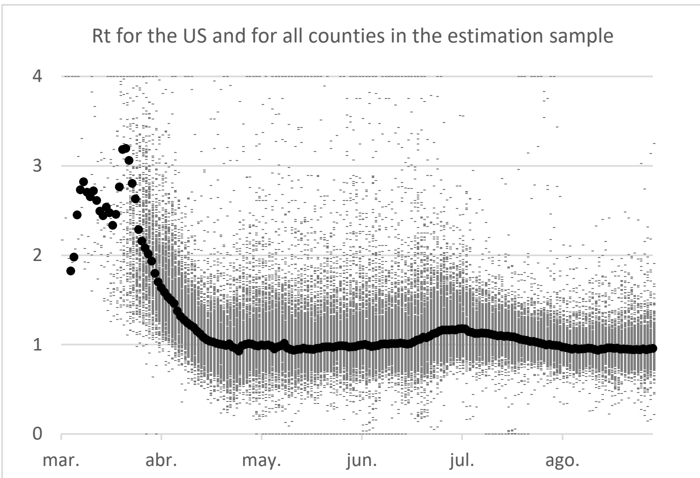
The estimation sample ends on the 31st of August of 2020 because Google drastically reduced the coverage of the mobility indicators for the second half of August, providing data only for a much smaller number of counties. Since this indicator enters the estimation as a two-week average lagged one week, the estimation could still be extended until the end of the month.

The map in Graph 2 represents the counties that appear at least once in the estimation. They are mainly populous areas that have had significant Covid-19 outbreaks. They are approximately 18% of the counties but they represent 49% of the population of the United States.



Graph 2: counties used in the estimation

As implied by this map, there is a lot of variability in the weather data used in the estimation. That is also the case for the epidemiological data; graph 3 presents all the values of R_t that were used in the estimation, compared with the national average. This diversity of observations helps identify the effect of the weather on the evolution of the pandemic without having to rely on broad North-vs-South or spring-vs-summer comparisons, where comparability of the situations would be more questionable.



Graph 3: Rt for the US and for all counties used in the estimation

4 Econometric framework

We rely on a typical fixed effects model, specified as follows:

$$R_t^c = a^c + b_1 Tem_{t-7}^c + +b_2 Pre_{t-7}^c + b_3 Mob_{t-7}^c + b_4 Par_{t-7}^c + \sum_i m_i + u_t^c \qquad (1)$$

Where:

- Superscript "c" refers to counties and subscript "t" refers to days, so that the units of observations varies at the county and day level.
- The dependent variable is the effective reproduction number.
- The main regressors are temperature (Tem), precipitations (Pre), average mobility towards locations in retail and recreation, train stations, and work places (Mob), and mobility in parks (Par). These are all time-varying. All regressors are lagged by one week to reduce endogeneity and correctly incorporate the lag with which contacts become infections and then cases.
- m_i are monthly fixed effects that we add in the equation to control for all unobservable common factors affecting all counties over time.
- a^c are county fixed effects that account for all county-specific characteristics that are constant. To deal with it we implement the fixed effect estimator (within estimator).
- u_t^c is the error term.

In all cases, standard errors are clustered at the county level, to account for the fact that the observations are correlated over time. A standard OLS estimation would assume that residuals are uncorrelated across both space and time; clustering by county relaxes this assumption, allowing for correlation in the residuals across time in each country (but not across counties). This results in bigger adjusted standard errors (which we present in the estimation tables below) and stricter requirements for inference in terms of coefficient significance.

Controlling for county fixed effects allows to control for all county-specific characteristics that are constant over time. To deal with it we implement the fixed effect estimator (within estimator). In addition, in our preferred specification we add to Eq. (1) interactions between time and state fixed effects ($\sum_i \sum_j m_i * s_j$),[3] which take into account all unobservable factors that change over time at the state level, such as the regional economic situation and state-level policies implemented to cope with the spread of the virus.

The inclusion of these detailed fixed effects ensures that the estimated coefficients reflect subtle changes that are not correlated with major changes in circumstances or behavior. For example, the estimated effect of temperature is not the result of comparing the summer with the spring (when schools were open and people didn't have a lot of information about the pandemic), or comparing Arizona with Alaska; instead, it captures the much subtler fact that counties with higher temperatures than other counties in the same state and in the same month observed a more negative deviation of their R_t with respect to the usual level for that county, and to the evolution that was common to the whole state in that month.

The issue of possible reverse causality or variable endogeneity also needs to be discussed: if mobility responds to R_t, our estimates could be biased and the conclusions could be wrong. It

[3] Letter "m" and subscript "i" refer to month fixed effects. Letter "s" and subscript "j" refer to state fixed effects.

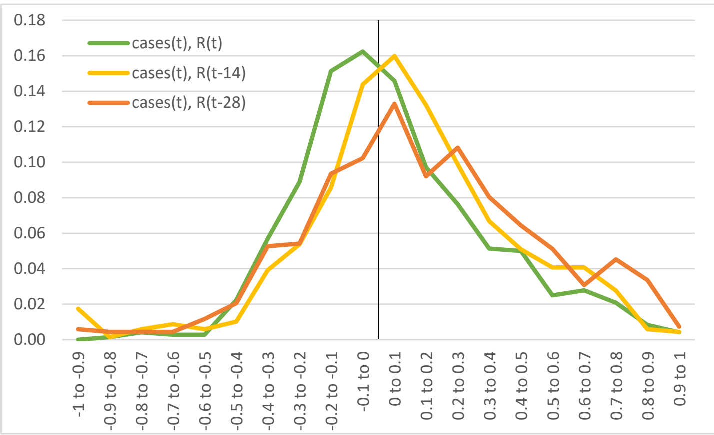
is certain that when the situation worsens and the news are full of reports about the carnage of the pandemic, people reduce their visits to restaurants and shops, and try to work from home if possible. But this is expected in relation to the number of cases and, even more so, of deaths caused by the virus, in levels instead of rate of growth. In terms of our estimation, there would be an endogeneity problem if mobility today responded to the effective reproductive coefficient calculated with the following week's data on identified Covid-19 cases. This is not completely implausible, because R_t can have strong autocorrelation, so it is a good idea to run a formal test.

We run a panel version of the Granger causality test, using the daily data about R_t and mobility at the county level, with a specification that allows up to 28 lags of both variables with respect to our estimation specification. The test is ran for the unbalanced panel of used observations and, following Holtz-Eakin et al (1988), includes county fixed effects:

$$Mob_{t-7}^{c} = a^{c} + \sum_{k=1}^{28} b_1^k Mob_{t-7-k}^{c} + \sum_{k=t}^{28} b_2^k R_{t-k}^{c} + u_t^c$$

We find that no b_2^k is statistically different from zero, for any k between 1 and 28, meaning that the effective reproduction coefficient does not Granger-cause mobility in our county-level database. Based on this result, we can dismiss reverse causality or covariate endogeneity issues in our estimation.

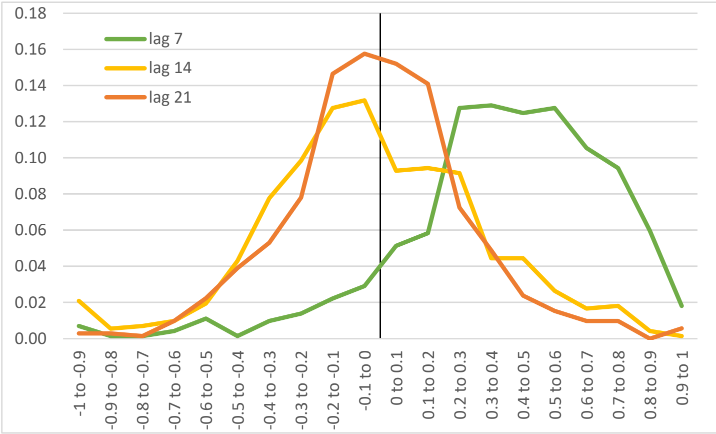
With a higher number of lags in the test (e.g. 35), some remote lags of R_t can be significant in the Granger-causality test. We don't think this is indicative of an endogeneity problem, though: as stated above, it is expected that mobility today will respond to cases and deaths today, and this does not depend on the effective reproduction number today, but it does depend on long-past values of the effective reproduction number. In order to illustrate this claim, Graph 4 shows the distribution across counties of the correlation coefficients between the number of cases in the week leading to t, and the R_t calculated with two weeks of data until t, or t-14, or t-28. When R_t is lagged four weeks, there's a positive correlation in most cases, but when it is contemporaneous, the correlation is negative in 49.3% of the cases (even in this case where R_t is calculated with two weeks of data and the number of cases used is the average of the second of those two weeks).



Graph 4: distribution of correlation coefficients of cases and past R_t In the counties used in the estimation

Part of the explanation for this is that the number of cases depends on a very long sequence of R_t values, and at the county level these don't show strong autocorrelation. Graph 5 shows the distribution of autocorrelation values at the county level, for a lag of one, two or four weeks.

Autocorrelation is clearly positive when looking at a one-week lag (where one week of data is actually common to both estimation windows in the comparison), but it's centered on zero for both two-week and four-week lags.



Graph 5: distribution of autocorrelation coefficients of R_t In the counties used in the estimation

5 Estimation results

5.1 Linear specifications

Table 1 reports the estimation results of our baseline model. We move from a less demanding specification (1) in which we use state fixed effects instead of county fixed effects, to the more demanding specification (4), in which we control for county fixed effects (our units of observations) and also for interactions between time and state fixed effects, as explained in the previous section.

Overall, the estimated coefficients show the expected sign: temperature and mobility in parks are negatively associated with the R_t, while higher precipitations and higher mobility towards recreational, train stations and work related locations increase the spread of the virus.

Results are robust across specifications, although there are some interesting differences. In particular, the effect of temperature is larger whenever the estimation controls for state-level policies against the spread of the Covid-19 virus (column 2 and 4), which suggests that part of the effect of the temperature would be biased by the omission of state policy responses to Covid-19, if the latter happen to have been implemented in periods when the temperature was high. The opposite holds for rainfalls, i.e. the effect of precipitation is reduced when policies are taken into account. The effects of mobility variables are stable across all specifications.

In the rest of this subsection we discuss the estimated coefficients of the specification reported in Column 4, the most demanding one, which is our preferred linear specification. The estimated effect for temperature (when it increases by one degree Celsius, the effective reproduction number decreases by 0.0173 units) is sizeable. According to this, and since our estimation assumes linearity, a summer-to-winter decrease of 20ºC (the national average difference between July and January) would imply an increase in the effective reproduction number by approximately 0.35, which can be the difference between a well-controlled evolution and explosive behavior[4]. The estimation of this effect is also strongly statistically significant. By contrast, the estimate for precipitations shows the expected sign but the coefficient is small and not significant once we control for interactions between state and month fixed effects.

The first mobility variable measures number of visits to three specific categories of locations: retail and recreation, train stations, and workplaces. According to the estimation results, if mobility increases by one percentage point (pp) the effective reproduction number increases by 0.017 units. To contextualize this figure: for one week in April, mobility decreased to around -49% in the US as a whole - the lowest point in the period we consider (from February to August 2020); our estimated coefficient implies that this led to a decrease in the effective reproduction number of 0.83 units. A comparison with the estimated effect of temperature is interesting: compared with the situation in the summer, an additional fall in mobility of 21 points would be needed in order to compensate the increase in R_t associated with a 20ºC fall in temperature; since the average mobility indicator was already at around -25% during the summer, this would bring it close to the maximum recorded in April.

Table 1: Main specifications.

VARIABLES	(1) R_t	(2) R_t	(3) R_t	(4) R_t
Temperature	-0.0047***	-0.0154***	-0.0052***	-0.0173***
	(0.0016)	(0.0016)	(0.0017)	(0.0017)
Precipitations	0.0004***	0.0001	0.0004***	0.0001
	(0.0001)	(0.0001)	(0.0001)	(0.0001)
Mobility (recreation, stations, work)	0.0145***	0.0138***	0.0167***	0.0169***
	(0.0006)	(0.0006)	(0.0007)	(0.0008)
Mobility parks	-0.0012***	-0.0009***	-0.0014***	-0.0012***
	(0.0002)	(0.0001)	(0.0002)	(0.0002)
Fixed effects	state dummies & month dummies	state dummies interacted with month dummies	county dummies & month dummies	county dummies & state dummies interacted with month dummies
Observations	89,669	89,669	89,669	89,669
R-squared	0.302	0.365	0.3028	0.3664
Clusters	county	county	County	county

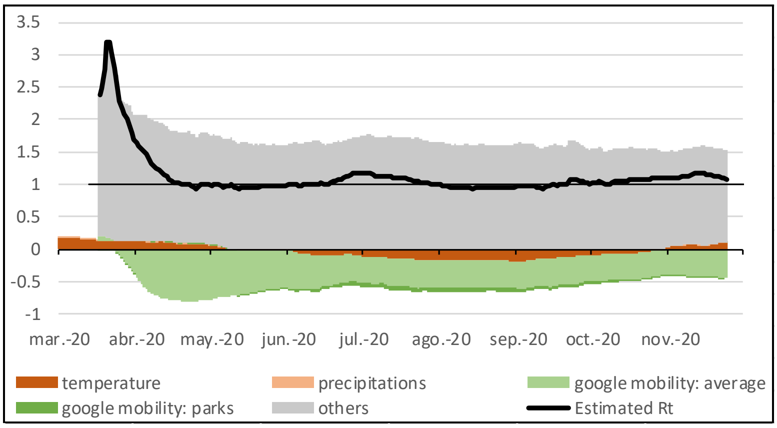
Cluster-robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1

On the other hand, mobility in parks is negatively associated with the rate of contagiousness. The magnitude of the effect is much smaller than the aforementioned measure of mobility (for a one pp increase in park mobility, the spread of the virus decreases by 0.0012 units), but the effect is still highly statistically significant, and quantitatively sizeable: a reduction of the mobility indicator about parks from its level over the summer (around +60%) to what was already observed in November (approximately zero) is associated with an increase of R_t of approximately 0.07, which means it would explain almost a third of the increase over this period.

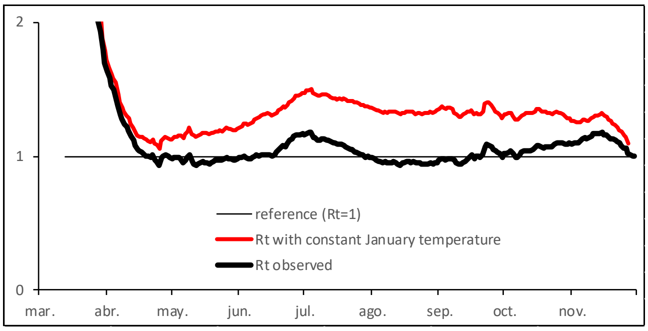
[4] The very slow reduction in daily cases in the US as a whole during May produced estimates of R_t of approximately 0.97, and the increase of the so-called second wave implied a maximum R_t of 1.18 at the beginning of July.

Overall, these results are in line with what previous studies have found about the virus and its transmission. It has been known since early spring that the virus survives much longer at low temperatures: Chin et al (2020) found that the SARS-CoV-2 virus that causes Covid-19 is highly stable at 4ºC, but is sensitive to heat, as it deteriorates approximately ten times faster at 22ºC than at 4ºC, and six times faster at 37ºC than at 22ºC. On the other hand, many studies, and most notably Shen et al (2020), have found that the transmission of the virus is much more likely indoors than outdoors. Our results show that these results from laboratory experiments and contact-tracing studies also translate to the evolution of the epidemic at the macro level.

In order to illustrate the quantitative effect associated to each explaining variable, Graph 6 presents a decomposition of the nation-wide evolution of R_t. For most of the variables, the aggregate national data is used, but temperature and rainfall are population-weighted averages of the county data. The size of each contribution is calculated using the coefficients estimated in column 4 of Table 1.



Graph 6: contributions to the evolution of R_t in the United States



Graph 7: counterfactual scenario for R_t in the USA with constant January temperature

As graph 6 shows, according to these estimates, the increase of R_t above one in the second wave of infections, at the end of June, was mainly caused by an increase of mobility. The third that started in October, on the other hand, happened without an increase in visits to locations of retail and recreation, transit stations, and workplaces, and was associated, instead, to a reduction in park visits and, above all, to the fall in temperatures.

Graph 7 presents a counterfactual scenario where we assess, according to the estimates from specification (4), the evolution of R_t by state that would have been observed if temperatures had remained constant on their levels of January. The difference implied by this exercise is very big, both in terms of levels (with an average of 1.43 in July instead of the observed 1.09, the second wave of Covid-19 in USA would have implied doubling the number of deaths approximately every 9 days instead of every 37 days) and in terms of evolution (with constant temperature, R_t would have been flat or falling from August to November, instead of increasing).

By applying the coefficients of specification (4) to the state-wide data, the results presented in detail in Appendix 1 show how much of the regional variability in R_t can be explained by differences in weather and mobility. According to these results, regional differences in mobility explain approximately half of the state-level variability in R_t, whereas temperature and precipitations explain another 20%, and the unexplained part is just 29%. Further research could try to relate these residuals at the state level –or, even better, the estimated county and state-by-month fixed effects– to observable variables such as population density, income, educational levels, political views, etc.[5]; but, because it is a pressing matter to disclose our current results, we leave this for future work.

5.2 Nonlinear specifications

Adding to specification (4) some interactions between variables allows us to assess nonlinear dynamics in their effects. Table 2 summarizes these results. All specifications retain the general characteristics of (4): county fixed effects, state-month interacted fixed effects, and clustering by county.

Specification (5) adds an interaction of temperature and precipitations, which allows us to decompose the non-statistically-significant coefficient of estimation (4) into two statistically significant ones: the effect of precipitations on R_t is approximately null at high temperatures (when the average is around 25ºC), but it is negative at lower temperatures. Additionally, a non-linearity is allowed at freezing point, and this shows that cold weather is much more negative in terms of the spread of the virus when average temperatures falls below 0ºC.

Specification (6) introduces an interaction of temperature and mobility, and finds that the effect of mobility is reduced in warm weather, both for parks and for general mobility. The implications of these nonlinearities for the coming winter in the Northern hemisphere are complex: a given reduction of mobility generates a bigger fall in R_t, but a mobility relaxation also has bigger negative effects. If we move from a summer situation with an average temperature of 25ºC, general mobility of -25% and parks mobility of +60%, to a winter situation with average temperature of 5ºC and the same -25% and +60% mobility, this estimation indicates an increase of R_t of +0.33, which is very similar to the +0.35 that the linear estimation would imply. And with this change in temperature, the effective coefficient for general mobility goes from 0.013 to 0.023. This makes it easier to compensate, via additional reductions of mobility, for the effect of colder weather, but also increases the repercussions of any relaxation in mobility. Therefore, it is now more important to keep restricting visits to locations related to retail and recreation, transit stations and workplaces; and also to spend more time in parks. And all of this, not only because they have an effect on the effective reproduction number, but because this effect is bigger when it's cold.

Specification (7) adds interactions with characteristics of the counties, and finds that the effect of mobility is enhanced in counties with high population density, where a larger share of the

[5] Note that the effects we analyze for some of these variables in the section about non-linear specifications refer to how they alter the impact of mobility or temperature. It is not their direct impact on R_t, which can't be captured in our specifications because of the use of fixed effects to control for these and all other (possibly unobserved) characteristics of the counties that are constant over time.

Table 2: non-linear specifications

VARIABLES	(5) R_t	(6) R_t	(7) R_t	(8) R_t
Temperature	-0.0158*** (0.0018)	-0.0327*** (0.0023)	-0.0167*** (0.0017)	-0.0336*** (0.0024)
Temperature-negative (<0ºC)	-0.0210** (0.0103)			-0.0127 (0.0100)
Precipitations	0.0010*** (0.0004)	0.0007* (0.0004)	0.0010*** (0.0004)	0.0008** (0.0004)
Temperature x Precipitations	-0.00004** (0.0000)	-0.00003* (0.0000)	-0.00004*** (0.0000)	-0.00004** (0.0000)
Mobility (recreation, stations, work)	0.0168*** (0.0008)	0.0251*** (0.0009)	0.0210*** (0.0017)	0.0260*** (0.0018)
Mobility x Temperature		-0.0005*** (0.0000)		-0.0005*** (0.0000)
MobilityParks	-0.0011*** (0.0002)	-0.0028*** (0.0004)	-0.0010*** (0.0001)	-0.0028*** (0.0004)
MobilityParks x Temperature		0.0001*** (0.0000)		0.0001*** (0.0000)
Mobility x Density			0.1820* (0.0000)	0.0619 (0.0000)
Mobility x PIAAC			0.0153*** (0.0052)	0.0249*** (0.0048)
Mobility x Services			-0.0004*** (0.0001)	-0.0003*** (0.0001)
Fixed effects	county dummies & state dummies interacted with month dummies	county dummies & state dummies interacted with month dummies	county dummies & state dummies interacted with month dummies	county dummies & state dummies interacted with month dummies
Observations	89,669	89,669	89,669	89,669
R-squared	0.3668	0.3729	0.3681	0.3750
Cluster	county	county	county	county

Cluster-robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1

population attained the lowest scores (P1) in the adult numeracy and literacy tests of OECD's Programme for the International Assessment of Adult Competencies (PIAAC), and where the services sector represents a lower percentage of employment (possibly because firms in the industry sector have more workers in bigger spaces)[6].

All the coefficients presented in specifications (5) through (7) are statistically significant, including the ones that identify non-linear effects.

Specification (8) puts all of these elements together and checks that the sign of the estimated coefficients is robust to the simultaneous inclusion of all covariates; the quantification does change in some cases, though, and a few become statistically non-significant.

6 Robustness

We provide two sets of robustness checks. First, we re-estimate our preferred specification (column 4 in Table 1, which is reported in column 1 of table 3 to ease comparisons) by replacing monthly dummies to dummies of two-week periods. That is, we specify time fixed effects at a higher frequency. This allows us to control in a finer way for state-level policy changes, in particular accounting for policies that may have been implemented within the month. Results

[6] We also tried specifications including interactions with median income. This turns out to be significant only when mobility data is not included in the estimation, which is in line with the results from Chang et al (2020), who find that higher infection rates among disadvantaged racial and socioeconomic groups are due to differences in mobility.

for this specification (9) are reported in column 2 of Table 3. Overall, results are consistent with our main specification, and the effect of mobility towards the other locations is quite stable. On the other hand, the size of the effect of temperature and mobility in parks is reduced; however, this is expected, since we are now estimating twice as many time fixed effects, i.e. our specification is more demanding.

In a second robustness check, we check that our results are not affected by the serial correlation induced by the specific construction of our variables. Note that, since R_t, has to be estimated by construction using a window[7] and refers to a lengthy period of time; for consistency, we apply an analogous procedure to all time-varying regressors on the right-hand side of the equation, calculating 14-day moving averages. This use of moving averages generates serial correlation in our sample, since the values of two consecutive days are based on a lot of common information. To control for this, all estimations in Table 1 allow for serial correlation of the residuals by means of clustering at the county level. But still, one may worry that our results are affected by our construction procedure. Therefore, as a robustness check, in specification (10) we estimate a version of our preferred specification that now uses only one set of observations every two weeks. This way, each observation corresponds to a two-weeks average, without overlapping with the previous or later observations. Note, some serial correlation may still be in place because of the nature of the variables, not not, this time, because of the data-compilation procedure. Results are reported in the last two columns of Table 3, defining time fixed effects each month in specification (10), and each two-weeks in specification (11). Overall, results are robust, although the coefficients are slightly smaller in

Table 3. Robustness: frequency and time fixed effects.

	(4)	(9)	(10)	(11)
	sample frequency: daily (each obs is 14-days moving average)		sample frequency: two-weeks (simple average)	
VARIABLES	R_t	R_t	R_t	R_t
Temperature	-0.0173***	-0.00528***	-0.00644**	-0.00662*
	(0.0017)	(0.00191)	(0.00295)	(0.00383)
Precipitations	0.0001	0.00005	-0.00012	0.00010
	(0.0001)	(0.00009)	(0.00014)	(0.00018)
Mobility	0.0169***	0.01216***	0.00649***	0.00335***
	(0.0008)	(0.00076)	(0.00138)	(0.00119)
Mobility in parks	-0.0012***	-0.00078***	-0.00050***	-0.00033*
	(0.0002)	(0.00015)	(0.00017)	(0.00017)
Fixed effects	county dummies & state dummies interacted with month dummies	county dummies & state dummies interacted with two-week dummies	county dummies & state dummies interacted with month dummies	county dummies & state dummies interacted with two-week dummies
Observations	89,669	89,669	6,612	6,612
R-squared	0.36644	0.43024	0.43232	0.51951
Clusters	county	county	county	county

Note. Cluster-robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1

[7] In our case the rolling window is of 14-days: instead of the 7-days window that is common in the literature, we extended it to 14 in order to avoid noise when using county data about Covid-19 cases.

magnitude and less significant. This is reasonable, since the number of observations drops considerably and the specification is very demanding (again, we are estimating a very large number of fixed effects).

A third exercise, presented in Table 4, checks that extending the sample period beyond the period of low data availability in September doesn't alter the results. The first column, once again, restates for convenience the results from the main specification (4). The second column, identified as (12), uses the same specification but extends the sample until the end of October, finding estimates that are qualitatively similar, but obviously with some numerical differences. The coefficient for temperature is slightly lower and the one for mobility is slightly higher, but these differences are not statistically significant (and they wouldn't be either if the exercise was done with two non-overlapping samples, one until July and one starting in August).

The last two columns in this table repeat the estimations of (4) and (12) but limiting the sample to the few counties for which Google provides mobility data at the beginning of September. As this is a non-random subsample of counties, biased towards those with higher population, the estimates change visibly (to some extent, in the direction of the results from the non-linear estimations, that already told us that mobility had a bigger impact in counties with higher population density). As in the comparison of (4) and (12), the difference between the estimation (13) that ends in August and the one (14) that ends in October is relatively small.

We interpret the results from specifications (12) through (14) as indicative that ending the main estimation in August is not biasing the results.

Table 4. Robustness: sample extended until the end of October.

	(4) March-August	(12) March-October	(13) March-August	(14) March-October
	Using all counties for which there is data at each point in time		Limiting the sample to the few counties that have mobility data at the beginning of September	
VARIABLES	R_t	R_t	R_t	R_t
Temperature	-0.0173***	-0.0145***	-0.0146***	-0.0128***
	(0.0017)	(0.0013)	(0.0025)	(0.0018)
Precipitations	0.0001	0.0001*	0.0001	0.0001
	(0.0001)	(0.0001)	(0.0001)	(0.0001)
Mobility	0.0169***	0.0147***	0.0246***	0.0237***
	(0.0008)	(0.0007)	(0.0016)	(0.0014)
Mobility in parks	-0.0012***	-0.000913***	-0.0014***	-0.0014***
	(0.0002)	(0.0001)	(0.0004)	(0.0003)
Fixed effects	county dummies & state dummies interacted with month dummies	county dummies & state dummies interacted with month dummies	county dummies & state dummies interacted with month dummies	county dummies & state dummies interacted with month dummies
Observations	89,669	106,239	18,061	22,549
R-squared	0.3663	0.3525	0.6671	0.6599
Clusters	county	county	county	county

Note. Cluster-robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1

7 Conclusions

In this paper we estimate the effective reproduction number (R_t) of the current Covid-19 pandemic, with US daily infections data between February and September of 2020, at the county level. This is then used to estimate the effect on the spread of the pandemic of weather and mobility. Focusing on R_t instead of the cumulative number of cases or deaths allows us to assess the dynamic effects of time-varying determinants, and provides a more useful quantification of the effects of changes in the covariates. The panel estimation uses county dummies, plus dummies for state-month interactions, so that the estimated effects capture the differential evolution of, say, warmer or cooler counties within a state and in the same month.

We find a strong and significant effect of the weather, which is troubling for the coming months in the Northern hemisphere: lower temperatures are associated with a higher effective reproduction number; this effect is bigger at temperatures below ºC, and, at low temperatures, precipitations are also associated with a higher R_t. These effects are sizeable: according to our preferred linear specification, a summer-to-winter average decrease of 20ºC would imply an increase in the effective reproduction number by approximately 0.35, which can be the difference between a well-controlled evolution and explosive behavior.

We also find that mobility reductions related to certain locations (retail and recreation, transit stations, and workplaces) are effective at reducing R_t, an effect that also becomes bigger at low temperatures. On the other hand, and in line with previous literature about the probability of infection in open and closed spaces, it is an increase of the time spent in parks that helps reduce the spread of the epidemic. The negative effect of increased mobility is bigger in counties with higher population density, worse numeracy and literacy PIAAC scores, or a lower share of the services sector on employment.

BIBLIOGRAPHY

Almagro, M., J. Coven, Gupta and A. Orane-Hutchinson (2020) *Racial Disparities in Frontline Workers and Housing Crowding during COVID-19: Evidence from Geolocation Data*, Federal Reserve Bank of Minneapolis Opportunity and Inclusive Growth Institute Working Paper n.º 37.

Baker, R. E., W. Yang, G.A. Vecchi, C. J. E. Metcalf and B. T. Grenfell (2020). "Susceptible supply limits the role of climate in the early SARS-CoV-2 pandemic", *Science*.

Carson, R. T., S. L. Carson, T. K. Dye, S. L. Mayfield, D. C. Moyer and A. Y. Chu (2020) *COVID-19's US Temperature Response Profile*, MedRxiv.

Chang, S., E. Pierson, P.W. Koh, J. Gerardin, B. Redbird, D. Grusky and J. Leskovec (2020). "Mobility network models of COVID-19 explain inequities and inform reopening", *Nature*.

Chin, A. W. H., J. T. S. Chu, M. R. A. Perera, K. P. Y. Hui, H.-L.Yen, M. C. W. Chan, M. Peiris and L.L.M. Poon (2020). "Stability of SARS-CoV-2 in different environmental conditions", *The Lancet Microbe.*

Cori, A., N. M. Ferguson, C. Fraser and S. Cauchemez (2013). "A New Framework and Software to Estimate Time-Varying Reproduction Numbers During Epidemics", *American Journal of Epidemiology.*

Couture, V., J. Dingel, A. Green, J. Handbury and K. R. Williams (2020) *Measuring Movement and Social Contact with Smartphone Data: A Real-Time Application to COVID-19*, NBER Working Paper No. 27560.

Desmet, K., and R. Wacziarg (2020) *Understanding Spatial Variation in Covid-19 across the United States*", NBER Working Paper nº. 27329.

Gostic K. M., et al. (2020) *Practical considerations for measuring the effective reproductive number, R_t*, preprint on MedRxiv.

Hamada, S. B., H. Du, M. Marshall, E. Dong, M. M. Squire and L. M. Gardner (2020). "Association between mobility patterns and COVID-19 transmission in the USA: a mathematical modelling study", *The Lancet Infectious Diseases.*

Hamidi, S., S. Sabouri and R. Ewing (2020). "Does Density Aggravate the COVID-19 Pandemic?", *Journal of the American Planning Association.*

Holtz-Eakin, D., W. Newey and H. S. Rosen (1988). "Estimating vector autoregressions with panel data", *Econometrica: Journal of the Econometric Society.*

OECD (2020). "Issue Note 4. Walking the tightrope: Avoiding a lockdown while containing the virus", *OECD's Economic Outlook 2020,* issue 2, preliminary version.

Poirier, C., W. Luo, M. S. Majumder, D. Liu, K. D. Mandil, T. A. Mooring and M. Santillana (2020). "The role of environmental factors on transmission rates of the COVID-19 outbreak: an initial assessment in two spatial scales", *Nature.*

Rosario, D. K. A., Y. S. Mutz, P. S. Bernardes and C. A. Conte-Junior (2020). "Relationship between COVID-19 and weather: Case study in a tropical country", *International Journal of Hygiene and Environmental Health.*

Sharma, P., A. K. Singh, B. Agrawal and A. Sharma (2020). "Correlation between weather and COVID-19 pandemic in India: An empirical investigation", *Journal of Public Affairs.*

Shen, Y., et al. (2020). "Community Outbreak Investigation of SARS-CoV-2 Transmission among Bus Riders in Eastern China", *JAMA Internal Medicine.*

Smith, T. P., et al. (2020) *Environment influences SARS-CoV-2 transmission in the absence of non-pharmaceutical interventions*, preprint on MedRxiv.

Soucy, J. -P, R., S. L. Sturrock, I. Berry, N. Daneman, D. R. MacFadden and K. A. Brown (2020) *Estimating the effect of physical distancing on the COVID-19 pandemic using an urban mobility index*, MedRxiv.

TindaleL. C., et al. (2020). "Evidence for transmission of COVID-19 prior to symptom onset", *eLife.*

Wilson D. J. (2020) *Weather, social distancing, and the spread of COVID-19*, Federal Reserve Bank of San Francisco Working Paper nº. 2020-23.

Xu, R., H. Rahmandad, M. Gupta, C. DiGennaro, N. Ghaffarzadegan, H. Amini and M. S. Jalali (2020) *The Modest Impact of Weather and Air Pollution on COVID-19 Transmission*, MedRxiv.

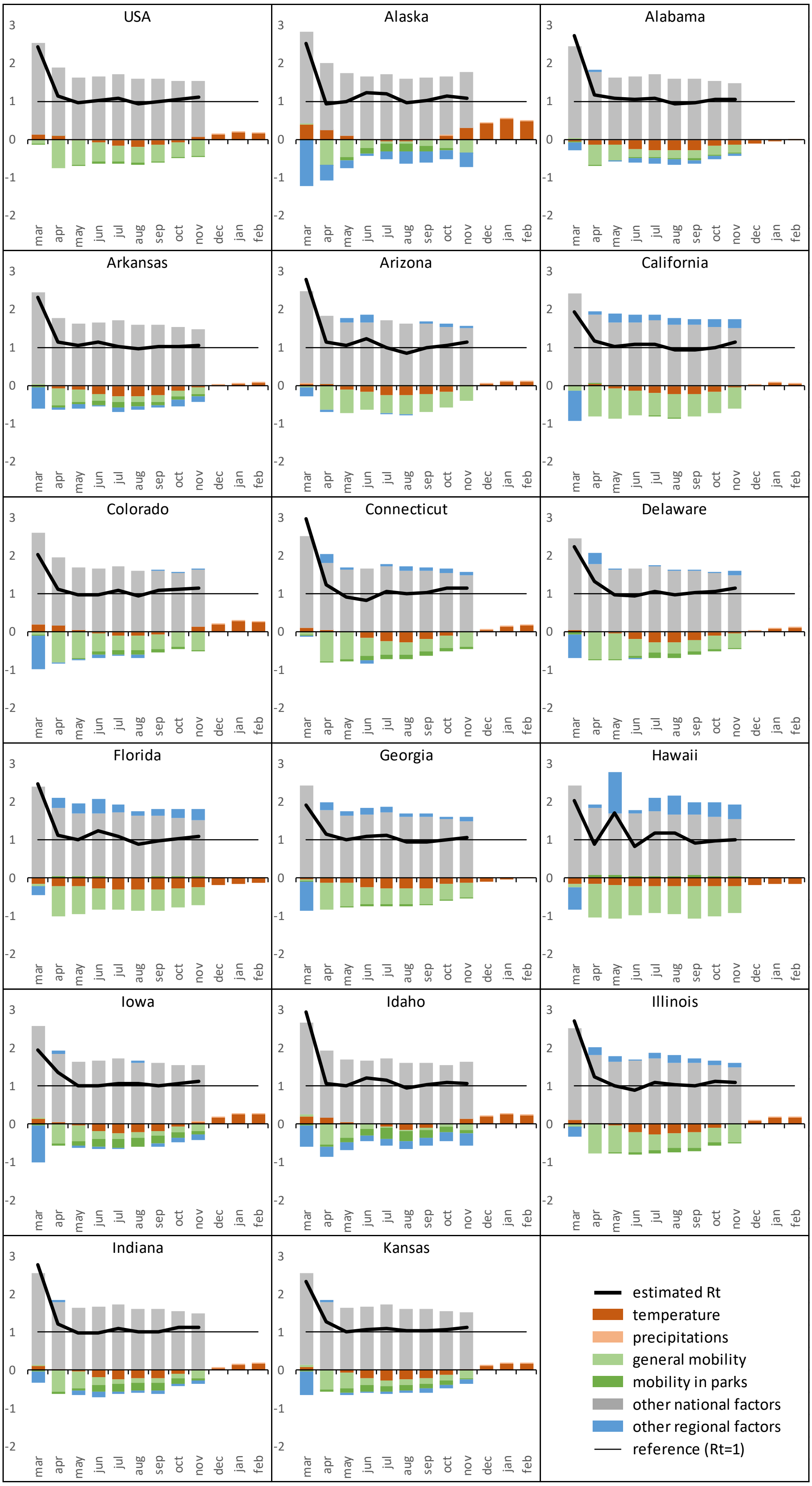
Appendix 1. Decomposition of R_t at the state level

In this appendix we apply the estimated coefficients from specification (4) to the state-level data, and analyze the contributions of the different factors to the regional evolution and to the variability at this level.

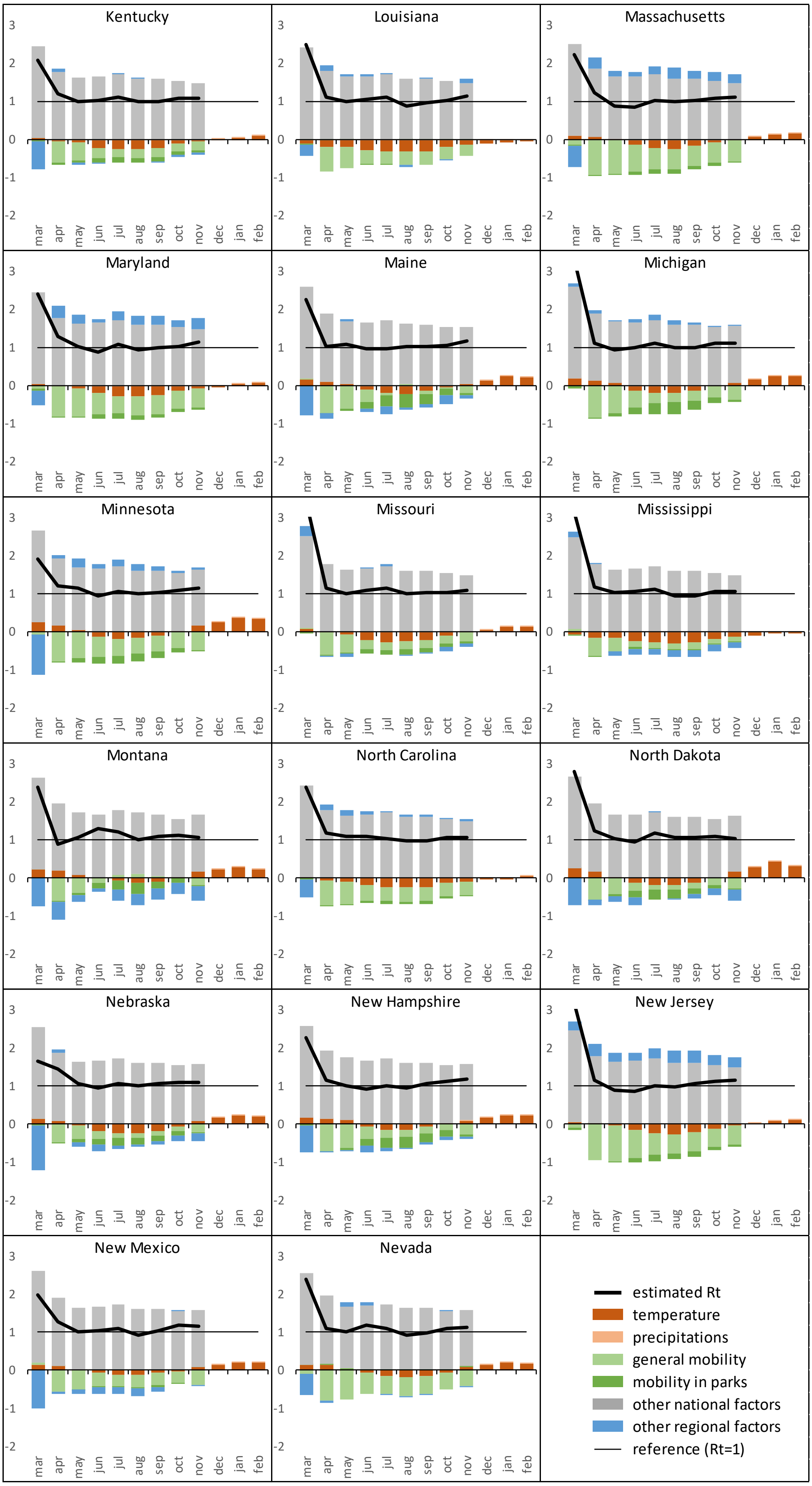
Graph 8 presents contribution charts, similar to those in Graph 6 but now for each state and only with averaged monthly information. The unexplained residual is divided here into two different parts: a common one calculated at the national level (in grey, common to all the panels in the graph) and an idiosyncratic one (in blue) that reflects differences in the regional evolution that are not related to mobility or weather. These blue bars are, in general terms, small.

This is also reflected in table 4, which calculates the contribution of each group of factors to the regional variability in R_t: on average, mobility explains approximately half, weather explains 20%, and just 29% of the (state-level) variability is not explained by the model.

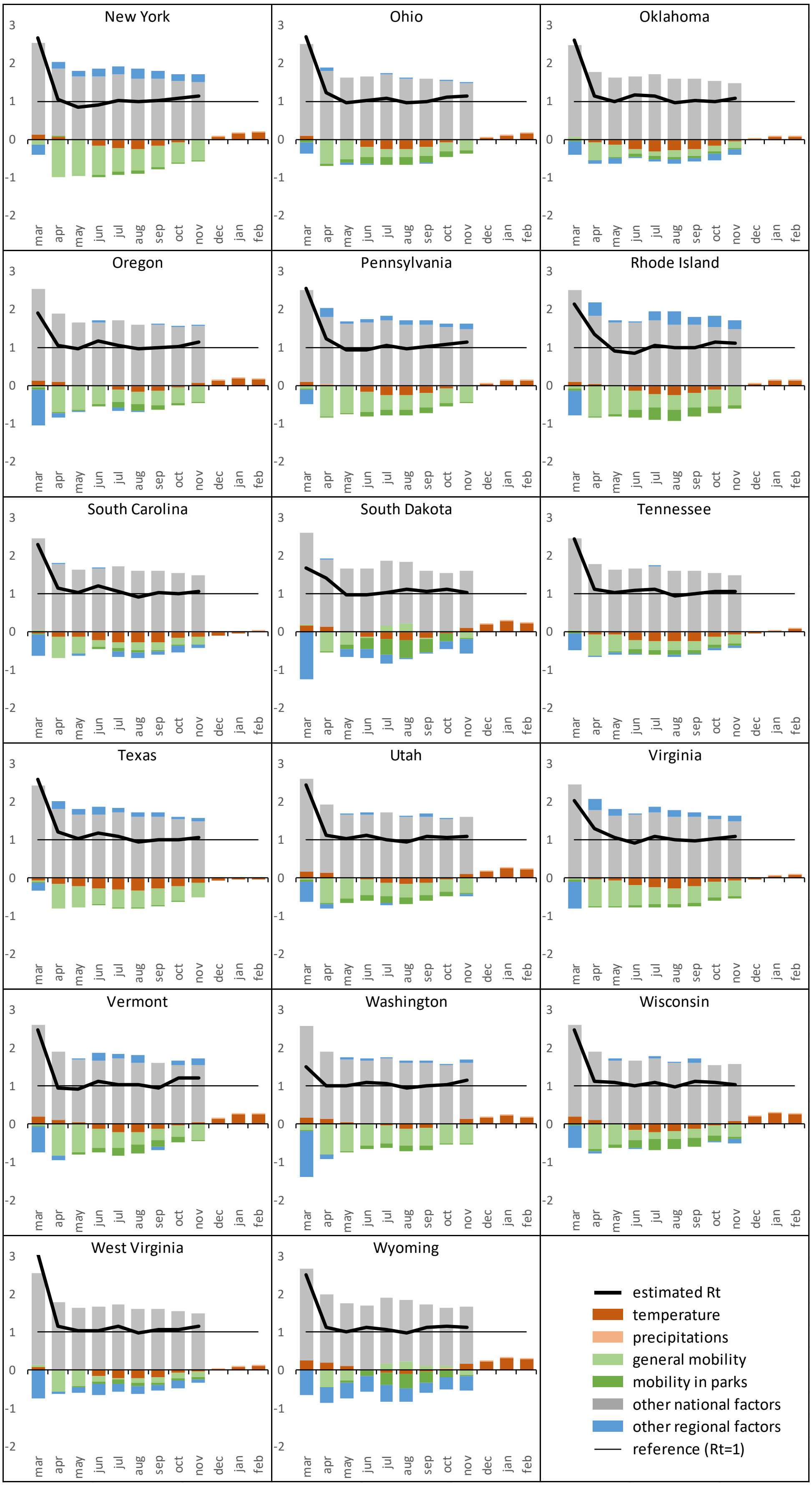
Finally, Graph 9 presents counterfactual scenarios similar to the one in Graph 7 but with regional detail: we assess, according to our estimates, the evolution of R_t that would have been observed in each state if temperatures had remained constant on their regional levels of January. The difference is small in Florida and almost non-existent in Hawaii, but very big in most states and in the national average.



Graph 8: contributions to the evolution of R_t on each state



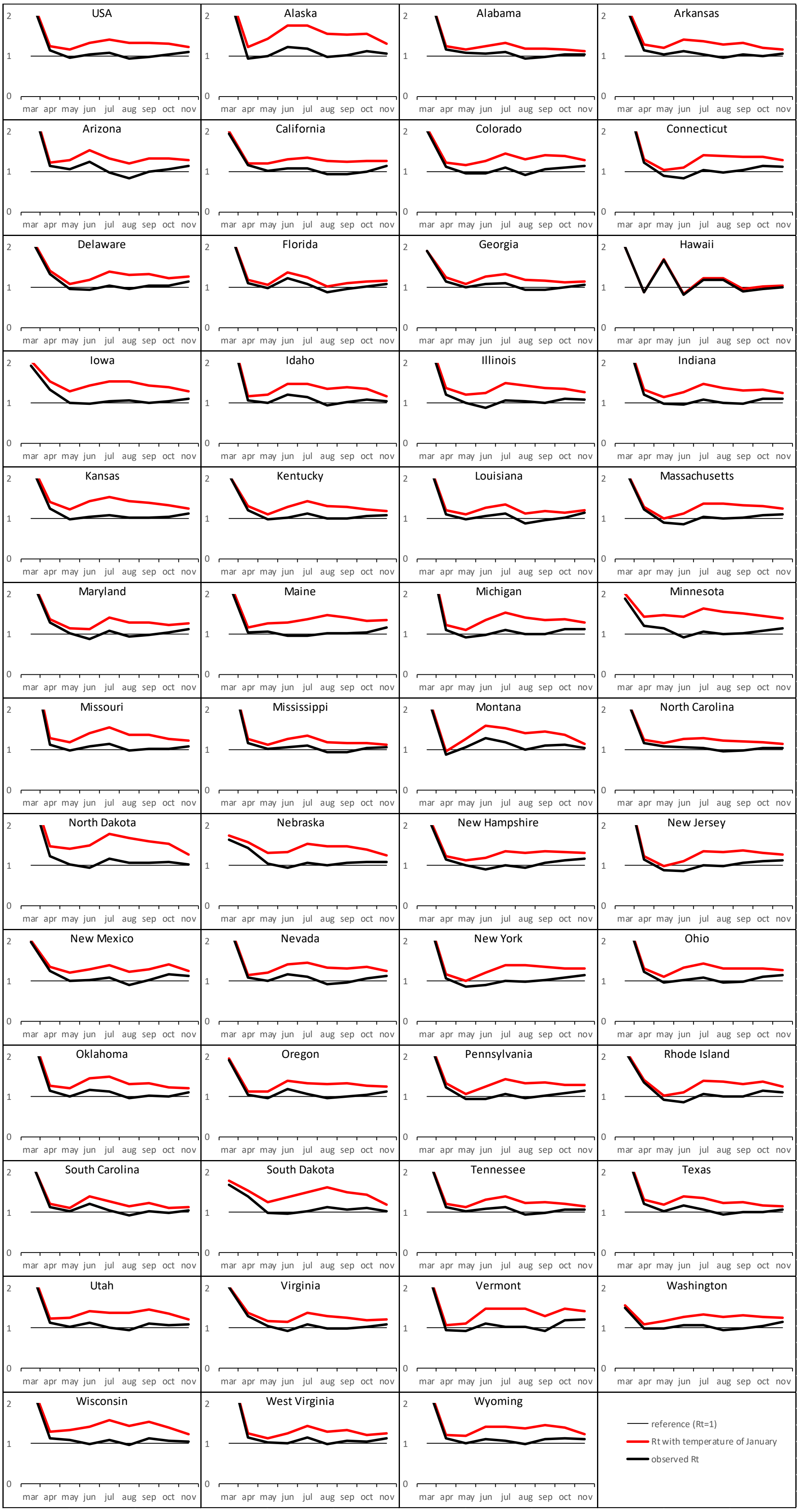
Graph 8 (cont.): contributions to the evolution of R_t on each state



Graph 8 (cont.): contributions to the evolution of R_t on each state

Table 4: contribution of the different factors to the regional variability in Rt

	Weight of the different factors in explaining the differences from the national average (%)			Mean difference on the effect on Rt (with sign)		
	weather	mobility	others (unexplained)	weather	mobility	others (unexplained)
AK Alaska	20	44	36	0.09	-0.38	-0.26
AL Alabama	31	51	18	-0.20	-0.30	-0.07
AR Arkansas	22	61	18	-0.17	-0.31	-0.09
AZ Arizona	31	40	28	-0.14	-0.48	0.05
CA California	16	42	41	-0.13	-0.63	0.18
CO Colorado	41	38	21	0.01	-0.56	-0.02
CT Connecticut	23	41	36	-0.12	-0.54	0.08
DE Delaware	34	41	25	-0.14	-0.49	0.06
FL Florida	35	25	40	-0.27	-0.56	0.24
GA Georgia	48	12	40	-0.20	-0.51	0.12
HI Hawaii	16	41	42	-0.20	-0.73	0.39
IA Iowa	9	75	16	-0.10	-0.39	-0.05
ID Idaho	7	63	30	0.00	-0.34	-0.22
IL Illinois	28	22	50	-0.13	-0.58	0.14
IN Indiana	14	71	16	-0.12	-0.37	-0.07
KS Kansas	18	68	14	-0.14	-0.35	-0.06
KY Kentucky	24	65	11	-0.15	-0.40	-0.01
LA Louisiana	48	34	17	-0.23	-0.40	0.04
MA Massachusetts	10	42	48	-0.09	-0.72	0.22
MD Maryland	23	26	51	-0.16	-0.64	0.23
ME Maine	5	72	23	-0.06	-0.42	-0.11
MI Michigan	10	71	20	-0.04	-0.60	0.07
MN Minnesota	17	41	42	-0.03	-0.65	0.13
MO Missouri	23	64	13	-0.14	-0.38	-0.04
MS Mississippi	27	50	23	-0.21	-0.23	-0.13
MT Montana	9	56	35	0.02	-0.26	-0.29
NC North Carolina	41	34	25	-0.16	-0.48	0.07
ND North Dakota	10	63	27	-0.04	-0.37	-0.14
NE Nebraska	9	67	24	-0.10	-0.34	-0.10
NH New Hampshire	11	70	19	-0.01	-0.49	-0.08
NJ New Jersey	16	36	48	-0.14	-0.73	0.27
NM New Mexico	9	47	43	-0.03	-0.40	-0.12
NV Nevada	11	64	26	-0.06	-0.53	0.01
NY New York	10	45	45	-0.09	-0.72	0.20
OH Ohio	16	77	7	-0.12	-0.46	0.01
OK Oklahoma	23	57	20	-0.18	-0.27	-0.11
OR Oregon	18	53	29	-0.03	-0.53	-0.01
PA Pennsylvania	22	32	45	-0.11	-0.58	0.12
RI Rhode Island	11	42	48	-0.10	-0.69	0.22
SC South Carolina	31	50	19	-0.20	-0.29	-0.08
SD South Dakota	4	75	20	-0.06	-0.32	-0.16
TN Tennessee	28	61	11	-0.16	-0.38	-0.04
TX Texas	46	20	34	-0.23	-0.46	0.13
UT Utah	12	66	22	-0.03	-0.54	0.00
VA Virginia	35	15	50	-0.15	-0.56	0.14
VT Vermont	11	49	40	-0.07	-0.60	0.09
WA Washington	27	47	26	0.00	-0.63	0.04
WI Wisconsin	8	76	16	-0.06	-0.50	0.00
WV West Virginia	11	58	30	-0.11	-0.27	-0.17
WY Wyoming	9	56	35	0.04	-0.21	-0.38
unweighted mean	20	50	29			



Graph 9: counterfactual scenario for R_t, with temperature constant at regional January levels

BANCO DE ESPAÑA PUBLICATIONS

WORKING PAPERS

2001 JAVIER ANDRÉS, PABLO BURRIEL and WENYI SHEN: Debt sustainability and fiscal space in a heterogeneous Monetary Union: normal times vs the zero lower bound.

2002 JUAN S. MORA-SANGUINETTI and RICARDO PÉREZ-VALLS: ¿Cómo afecta la complejidad de la regulación a la demografía empresarial? Evidencia para España.

2003 ALEJANDRO BUESA, FRANCISCO JAVIER POBLACIÓN GARCÍA and JAVIER TARANCÓN: Measuring the procyclicality of impairment accounting regimes: a comparison between IFRS 9 and US GAAP.

2004 HENRIQUE S. BASSO and JUAN F. JIMENO: From secular stagnation to robocalypse? Implications of demographic and technological changes.

2005 LEONARDO GAMBACORTA, SERGIO MAYORDOMO and JOSÉ MARÍA SERENA: Dollar borrowing, firm-characteristics, and FX-hedged funding opportunities.

2006 IRMA ALONSO ÁLVAREZ, VIRGINIA DI NINO and FABRIZIO VENDITTI: Strategic interactions and price dynamics in the global oil market.

2007 JORGE E. GALÁN: The benefits are at the tail: uncovering the impact of macroprudential policy on growth-at-risk.

2008 SVEN BLANK, MATHIAS HOFFMANN and MORITZ A. ROTH: Foreign direct investment and the equity home bias puzzle.

2009 AYMAN EL DAHRAWY SÁNCHEZ-ALBORNOZ and JACOPO TIMINI: Trade agreements and Latin American trade (creation and diversion) and welfare.

2010 ALFREDO GARCÍA-HIERNAUX, MARÍA T. GONZÁLEZ-PÉREZ and DAVID E. GUERRERO: Eurozone prices: a tale of convergence and divergence.

2011 ÁNGEL IVÁN MORENO BERNAL and CARLOS GONZÁLEZ PEDRAZ: Sentiment analysis of the Spanish Financial Stability Report. (There is a Spanish version of this edition with the same number).

2012 MARIAM CAMARERO, MARÍA DOLORES GADEA-RIVAS, ANA GÓMEZ-LOSCOS and CECILIO TAMARIT: External imbalances and recoveries.

2013 JESÚS FERNÁNDEZ-VILLAVERDE, SAMUEL HURTADO and GALO NUÑO: Financial frictions and the wealth distribution.

2014 RODRIGO BARBONE GONZALEZ, DMITRY KHAMETSHIN, JOSÉ-LUIS PEYDRÓ and ANDREA POLO: Hedger of last resort: evidence from Brazilian FX interventions, local credit, and global financial cycles.

2015 DANILO LEIVA-LEON, GABRIEL PEREZ-QUIROS and EYNO ROTS: Real-time weakness of the global economy: a first assessment of the coronavirus crisis.

2016 JAVIER ANDRÉS, ÓSCAR ARCE, JESÚS FERNÁNDEZ-VILLAVERDE and SAMUEL HURTADO: Deciphering the macroeconomic effects of internal devaluations in a monetary union.

2017 JACOPO TIMINI, NICOLA CORTINOVIS and FERNANDO LÓPEZ VICENTE: The heterogeneous effects of trade agreements with labor provisions.

2018 EDDIE GERBA and DANILO LEIVA-LEON: Macro-financial interactions in a changing world.

2019 JAIME MARTÍNEZ-MARTÍN and ELENA RUSTICELLI: Keeping track of global trade in real time.

2020 VICTORIA IVASHINA, LUC LAEVEN and ENRIQUE MORAL-BENITO: Loan types and the bank lending channel.

2021 SERGIO MAYORDOMO, NICOLA PAVANINI and EMANUELE TARANTINO: The impact of alternative forms of bank consolidation on credit supply and financial stability.

2022 ALEX ARMAND, PEDRO CARNEIRO, FEDERICO TAGLIATI and YIMING XIA: Can subsidized employment tackle long-term unemployment? Experimental evidence from North Macedonia.

2023 JACOPO TIMINI and FRANCESCA VIANI: A highway across the Atlantic? Trade and welfare effects of the EU-Mercosur agreement.

2024 CORINNA GHIRELLI, JAVIER J. PÉREZ and ALBERTO URTASUN: Economic policy uncertainty in Latin America: measurement using Spanish newspapers and economic spillovers.

2025 MAR DELGADO-TÉLLEZ, ESTHER GORDO, IVÁN KATARYNIUK and JAVIER J. PÉREZ: The decline in public investment: “social dominance” or too-rigid fiscal rules?

2026 ELVIRA PRADES-ILLANES and PATROCINIO TELLO-CASAS: Spanish regions in Global Value Chains: How important? How different?

2027 PABLO AGUILAR, CORINNA GHIRELLI, MATÍAS PACCE and ALBERTO URTASUN: Can news help measure economic sentiment? An application in COVID-19 times.

2028 EDUARDO GUTIÉRREZ, ENRIQUE MORAL-BENITO, DANIEL OTO-PERALÍAS and ROBERTO RAMOS: The spatial distribution of population in Spain: an anomaly in European perspective.

2029 PABLO BURRIEL, CRISTINA CHECHERITA-WESTPHAL, PASCAL JACQUINOT, MATTHIAS SCHÖN and NIKOLAI STÄHLER: Economic consequences of high public debt: evidence from three large scale DSGE models.

2030 BEATRIZ GONZÁLEZ: Macroeconomics, Firm Dynamics and IPOs.

2031 BRINDUSA ANGHEL, NÚRIA RODRÍGUEZ-PLANAS and ANNA SANZ-DE-GALDEANO: Gender Equality and the Math Gender Gap.

2032 ANDRÉS ALONSO and JOSÉ MANUEL CARBÓ: Machine learning in credit risk: measuring the dilemma between prediction and supervisory cost.

2033 PILAR GARCÍA-PEREA, AITOR LACUESTA and PAU ROLDAN-BLANCO: Raising Markups to Survive: Small Spanish Firms during the Great Recession.

2034 MÁXIMO CAMACHO, MATÍAS PACCE and GABRIEL PÉREZ-QUIRÓS: Spillover Effects in International Business Cycles.

2035 ÁNGEL IVÁN MORENO and TERESA CAMINERO: Application of text mining to the analysis of climate-related disclosures.

2036 EFFROSYNI ADAMOPOULOU and ERNESTO VILLANUEVA: Wage determination and the bite of collective contracts in Italy and Spain: evidence from the metal working industry.

2037 MIKEL BEDAYO, GABRIEL JIMÉNEZ, JOSÉ-LUIS PEYDRÓ and RAQUEL VEGAS: Screening and Loan Origination Time: Lending Standards, Loan Defaults and Bank Failures.

2038 BRINDUSA ANGHEL, PILAR CUADRADO and FEDERICO TAGLIATI: Why cognitive test scores of Spanish adults are so low? The role of schooling and socioeconomic background

2039 CHRISTOPH ALBERT, ANDREA CAGGESE and BEATRIZ GONZÁLEZ: The Short- and Long-run Employment Impact of COVID-19 through the Effects of Real and Financial Shocks on New Firms.

2040 GABRIEL JIMÉNEZ, DAVID MARTÍNEZ-MIERA and JOSÉ-LUIS PEYDRÓ: Who Truly Bears (Bank) Taxes? Evidence from Only Shifting Statutory Incidence.

2041 FELIX HOLUB, LAURA HOSPIDO and ULRICH J. WAGNER: Urban air pollution and sick leaves: evidence from social security data.

2042 NÉLIDA DÍAZ SOBRINO, CORINNA GHIRELLI, SAMUEL HURTADO, JAVIER J. PÉREZ and ALBERTO URTASUN: The narrative about the economy as a shadow forecast: an analysis using Banco de España quarterly reports.

2043 NEZIH GUNER, JAVIER LÓPEZ-SEGOVIA and ROBERTO RAMOS: Reforming the individual income tax in Spain.

2101 DARÍO SERRANO-PUENTE: Optimal progressivity of personal income tax: a general equilibrium evaluation for Spain.

2102 SANDRA GARCÍA-URIBE, HANNES MUELLER and CARLOS SANZ: Economic uncertainty and divisive politics: evidence from the *Dos Españas*.

2103 IVÁN KATARYNIUK, VÍCTOR MORA-BAJÉN and JAVIER J. PÉREZ: EMU deepening and sovereign debt spreads: using political space to achieve policy space.

2104 DARÍO SERRANO-PUENTE: Are we moving towards an energy-efficient low-carbon economy? An input-output LMDI decomposition of CO_2 emissions for Spain and the EU28.

2105 ANDRÉS ALONSO and JOSÉ MANUEL CARBÓ: Understanding the performance of machine learning models to predict credit default: a novel approach for supervisory evaluation.

2106 JAVIER ANDRÉS, ÓSCAR ARCE and PABLO BURRIEL: Market polarization and the Phillips curve.

2107 JUAN de LUCIO and JUAN S. MORA-SANGUINETTI: New dimensions of regulatory complexity and their economic cost. An analysis using text mining.

2108 DANILO LEIVA-LEON and LUIS UZEDA: Endogenous time variation in vector autoregressions.

2109 CORINNA GHIRELLI, ANDREA GONZÁLEZ, JOSÉ LUIS HERRERA and SAMUEL HURTADO: Weather, mobility and the evolution of the Covid-19 pandemic.