I Z A

Economic Activity and the Spread of Viral Diseases: Evidence from High Frequency Data

Jérôme Adda

September 2015

Forschungsinstitut
zur Zukunft der Arbeit
Institute for the Study
of Labor

Economic Activity and the
Spread of Viral Diseases:
Evidence from High Frequency Data

Jérôme Adda
Bocconi University, IGIER and IZA

ABSTRACT

Economic Activity and the Spread of Viral Diseases: Evidence from High Frequency Data[*]

Viruses are a major threat to human health, and - given that they spread through social interactions - represent a costly externality. This paper addresses three main issues: i) what are the unintended consequences of economic activity on the spread of infections? ii) how efficient are measures that limit interpersonal contacts? iii) how do we allocate our scarce resources to limit their spread? To answer these questions, we use novel high frequency data from France on the incidence of a number of viral diseases across space, for different age groups, over a period of a quarter of a century. We use quasi-experimental variation to evaluate the importance of policies reducing inter-personal contacts such as school closures or the closure of public transportation networks. While these policies significantly reduce disease prevalence, we find that they are not cost-effective. We find that expansions of transportation networks have significant health costs in increasing the spread of viruses and that propagation rates are pro-cyclically sensitive to economic conditions and increase with inter-regional trade.

Corresponding author:

Jérôme Adda
Department of Economics
Università Bocconi
Via Roentgen 1
20136 Milano
Italy
E-mail: jerome.adda@unibocconi.it

I. INTRODUCTION

Viruses are a major threat to human health. Over the last century, viruses were responsible for many more deaths than all armed conflicts that took place during that period. For instance, smallpox killed about 300 millions individuals during the twentieth century. Influenza was responsible for about 100 million deaths, a large part of them during the major outbreak in 1918-1919, known as the Spanish flu, but also through more recent pandemics. Other viruses also contribute to this sombre statistics, which include HIV (about 30 million deaths so far) and a number of viruses leading to gastroenteritis (about one million deaths per year worldwide).[1] In modern societies, with individuals in better health, viruses are also an important cause of morbidity and loss of productivity. For instance, in the US, about 30 million cases of gastro-enteritis are reported each year, leading to 120,000 hospitalizations. Likewise, influenza results in about 200,000 hospitalisations per year. Viruses impose therefore a cost on society, through premature deaths, long-lasting morbidity (Almond (2006), Kelly (2011)), increased health-care utilisation and loss of schooling or hours of work. More broadly, as emphasised by Fogli and Veldkamp (2013), the prevalence of diseases can be linked to social network structures and have long-term effects on growth.

In this paper, we provide a number of contributions to the understanding of the role of social interaction and economic activity in the spread of viruses. First, we exploit data that are unusually detailed, describing the incidence of three major viral diseases: influenza, gastro-enteritis and chickenpox. The data spans a period of up to a quarter of a century, at a weekly frequency, across geographical locations in France. Moreover, the data allow us to distinguish between age groups, an important element, as children and the elderly are particularly vulnerable populations, with potentially different transmission patterns. These data allow us to analyse the spatial and temporal evolution of diseases in a developed economy. This focus is novel in the economics literature, which has mostly studied the incidence of one particular virus, HIV, usually in developing countries (Oster (2005) or Oster (2012)).[2]

---

[1]For an overview of the determinants of mortality in developed and developing countries over time, see Cutler et al. (2006).

[2]The economic literature has been fruitful because the spread of HIV relies on the choice of individuals whether to protect them-selves or not and part of the trade-off is through the beliefs of the prevalence of

Second, we use quasi-experimental variation to assess the effectiveness of various measures such as closing down schools, shutting down public transportation or expanding railways on the transmission of a number of viruses. Closing down schools during influenza epidemics is routinely done in countries such as Japan or Bulgaria. In addition, a number of countries such as China, France, the United Kingdom and the United States have used this measure during the 2009 pandemic (Cauchemez et al. (2014)). Closing down temporarily public transportations or declaring a general curfew is a more drastic measure to curb an epidemic. There is a long history of quarantines to prevent the spread of diseases, which dates back at least to the plague epidemics in Europe and Asia in the Middle Ages. It has been used selectively during the Ebola outbreak in 2014, but not for seasonal epidemics.

The paper improves on the literature in epidemiology, which has developed models of disease diffusions dating back to Kermack and McKendrick (1927). [3] The work in epidemiology provides little direct and data based evidence on the effectiveness of policies aiming at reducing population contact rates. When this is the case, this literature also ignores important issues such as measurement errors, serial correlations and endogeneity, which are likely to bias the results. We assess the role of transportation infrastructures in the propagation of diseases in two ways. First, we exploit public transportation strikes throughout this period. In France, this is a relatively common occurrence, with a variety of general and more local strikes. Second, we exploit the extensive development of high speed railways across regions of France, and we use openings of new lines to assess their effect on the speed of disease transmissions. We further evaluate the role of peer-effects and spill-over effects by using information on school closures due to holidays. School holidays vary within the year and across the country, allowing for variability across time and space. Moreover, these closures are decided well in advance and are therefore unrelated to the prevalence of diseases, which helps to remediate reverse-causality issues.

the disease. Early such contributions focussing on "rational epidemics" include for instance Geoffard and Philipson (1996) or Kremer (1996). Auld (2006), Lakdawalla et al. (2006) and Dupas (2011) investigate the role of beliefs and prevalence on behavior, whereas Greenwood et al. (2013) look at equilibrium models of disease transmission.

[3] Anderson and May (1991) provide an overview of this literature. New developments include Epstein et al. (2007) or Lempel et al. (2009) that develops calibrated agent-based models.

A third contribution is to investigate the long-run economic determinants of the spread of viruses. This is an under-explored area, which is important to better understand how epidemics and economic activity are related. In the context of developed countries, and for common viral diseases, there are many unanswered questions: do viruses propagate faster during economic recessions or booms; does the unemployment rate contribute to epidemics; what is the role of inter-regional trade in the transmission of diseases[4]; do viruses spread symmetrically across space or do they follow a gradient determined by economic factors? To answer these questions, we develop a dynamic and spatial model of diffusion, based on work in epidemiology, which is estimated, paying particular attention to measurement error, serial correlation and endogeneity, using instrumental variable techniques.

We find that school closures have a pronounced effect on the incidence of influenza, reducing its incidence for a period of about three to four weeks. This effects is stronger for children, but also sizable for adults. We find similar effects for gastro-enteritis and chickenpox, although the effect is more short-lived and less important. Moreover, we find that for the elderly, such measures can *increase* the incidence in the short-run. Our results suggest therefore important spill-overs across age groups, within the extended family. We find significant decreases in the transmission rates of diseases during public transportation strikes, which mirrors the increased transmission rates following extensions of railway lines. By exploiting data for three viral diseases, which differ in several aspects, such as incubation time or infectiousness our results also shed light on how efficient those measures can be on emerging diseases that share some of the same characteristics as the ones we consider.

We next calculate the expected monetary benefits of such measures, using cost estimates from the literature. We find that although closing down schools for a period of 2 weeks would reduce the total annual incidence of some of the diseases by up to 12 percent across all age groups, such measures represent an overall cost to society. We find similar results, although with smaller costs, for a policy that would close down public transportations for a week. We show that those measures would become beneficial for epidemics characterised by a slightly more deadly strain, or for a policy that would match the closure of schools

---

[4]Prior work exists on the link between trade and HIV infection as Oster (2012) shows that increased exports in African countries lead to higher infection rates.

during epidemics with a longer term time in the summer. We also evaluate the health costs of expanding transportation infrastructures. Connecting two regions with a high-speed rail link generates a cost of about 80 million euros per year, through higher medical costs and loss of productivity. This cost is of the same order of magnitude as the time saved by faster transportation.

We find that epidemics spread faster during economic booms. During booms more people are traveling which increase inter-personal contacts and the spread of diseases. This is in contrast with the work in epidemiology that has found that diseases are more prevalent in bad times, although this literature relies mostly on aggregate cross-country comparison using annual data (see for instance the meta-analysis by Suhrcke et al. (2011) and the papers reviewed therein). Our results also add to the literature linking health and in particular mortality to business cycles, which was pioneered by Ruhm (2000). This literature finds that mortality declines during recessions and has been attributed to changes in health behaviour (Ruhm (2003), Ruhm (2005) and Evans and Moore (2012)). Our results point to an additional channel, which mitigates the health behaviour channel, at least for mortality due to infectious diseases or respiratory causes.

The rest of the paper is organized as follows. Section II. gives medical information on the viruses we study, which are important to interpret the results in the next sections. It also describes the various data set used in the analysis. Section III. presents an event analysis describing the effect of school closures and transportation strikes. Section IV. presents a dynamic and spatial model of viral spread, discusses econometric issues and presents the results. Section V. investigates the cost efficiency of different policy measures. Section VI. evaluates the health cost of transportation infrastructures. Finally section VII. concludes.

II. Viral Diseases: Data and Descriptive Statistics

In this section, we provide an overview of the characteristics of the viral diseases we study as well as details on public health measures in France between 1984 and 2010, the period we consider in the empirical section. We then present the data we analyse.

II.A. *Disease Characteristics and Prophylactic Measures*

We consider three major viral diseases, influenza, gastroenteritis and chickenpox. All diseases are caused by specific viruses, and as such, cannot be treated by drugs such as antibiotics. They usually result in infections that last for two to ten days. In some cases, infected individuals can develop complications, that may result in death, especially for people with health predispositions or for elderly individuals.

Table I provides a brief comparison of all three diseases, with information on incubation time, approximate duration of symptoms and the period when an infected individual becomes contagious. For influenza, the incubation period ranges from 2 to 5 days. Acute diarrhea caused by viruses incubates in about a day. On the other hand, chickenpox takes much longer to incubate, a period of between two to three weeks. For this particular disease, infected individuals acquire a life-long immunity. This is not the case for influenza and gastroenteritis, as the viruses involved mutate very quickly. While it is rare to get a second bout of influenza in the same year, it is possible to contract gastroenteritis several times in a year.

For each of these diseases, vaccines exist, but they vary in their efficacy and their use by public health authorities. In the case of influenza, in France, as in many developed countries, vaccination is targeted towards the elderly, health care workers and a small minority of people at risk. The coverage was about 70 percent for the elderly, 16 percent for adults and only 6 percent for children in 2010. These number have risen over time, as the practice was to provide immunization for people above 75 in 1985, and then above 65 from year 2000. In the case of gastro-enteritis, vaccination is difficult as many viruses, such as rotaviruses, noroviruses or adenoviruses, can be the cause of the disease. A first vaccine against the rotavirus A was developed in 1999 but then withdrawn because of side-effects. A new vaccine is on the market since 2006, but few countries have a systematic vaccination plan. French public health authorities advised against the use of the vaccine in 2010 and recommended instead a better use of re-hydration therapies in grave cases. For chickenpox, a vaccine was developed in 1988. It is widely used in the US, but not in Europe, where it is considered potentially more dangerous than the disease itself. As a consequence, immunity is acquired

almost entirely through exposure to disease when children are in day care or in primary school.

The average yearly incidence (percentage of population that get infected in a given year) differs widely across diseases. Table I presents typical ranges for the US and European countries. Gastroenteritis is the most prevalent disease, with an incidence rate varying between 20 and 40 percent. The incidence of influenza varies widely as well, depending on the viral strain that is present, ranging from 5 to 20 percent in the case of seasonal influenza. In pandemic years, the figure can be much higher for both gastro-enteritis and influenza. The 1918 pandemic infected about 30% of the population (Taubenberger and Morens (2006)). The incidence of chickenpox is much lower in the total population for two reasons. It is almost exclusively a childhood disease and countries differ in the use of vaccines. In the UK or France where no vaccination occurs, the incidence corresponds roughly to the number of babies born each year, as eventually almost all of them will get contaminated by the chickenpox virus. In the US, due to vaccination, the prevalence is only about 0.1%.

II.B. *High Frequency Data on Disease Prevalence*

We use detailed, high frequency (weekly) data on disease incidence in France, covering a period of up to 25 years. The data is available for three age groups - children (aged 0-18), adults (18 to 64) and the elderly - and at regional level.[5] The data is unique in the sense that no other countries have tried to gather systematic measures of incidence over such a long period and for several viral diseases. For instance, the data on influenza for the US, collected by the CDC, cover only the period from 1997 onwards, and for broad regions. Similarly, Google[6] reports data on flu epidemics, but only from 2003. Although their estimations track public health data quite well on average, it is unclear how to break it down by smaller geographical regions, and even more by age groups. There is also an issue with trends in internet usage, across region or age groups that could lead to measurement error. For other viral diseases, such as gastroenteritis or chickenpox, the data is even more

---

[5]Mainland France is subdivided into 21 regions. The average population size of a region was 2.1 million in 2009, ranging from 0.7 million (Limousin, around Limoge) to over 6 million (Paris).

[6]See http://www.google.org/flutrends/.

scarce.

The data we use on disease incidence come from the "Reseau Sentinelles", a network of about 1,300 general practitioners which was set-up in 1984 as a public health surveillance system by the Institut de Veille Sanitaire (InVS).[7] Those physicians report each week the number of cases of many diseases diagnosed on the basis of the symptoms and medical examination of the patient. However, their assessment is not based on the analysis of biosamples, which would ascertain the viral origin of the disease. Hence, to be precise, the data pertain to influenza-like illnesses rather than influenza and acute diarrhea rather than gastroenteritis and we shall label them as such in the reminder of the paper. The data is aggregated up to regional level, before its release.

Figure I displays the time series patterns of incidence rates at the national level and on a weekly basis between 1984 and 2010. Flu-like illnesses, acute diarrhea as well as chickenpox have recurrent peaks each year during the winter season. Given the seasonal fluctuations, it is difficult to see whether the incidence of these diseases has increased or decreased over this period. Table II displays the average annual percentage increase in the incidence of all three diseases, by age groups.[8] The situation is contrasted. The incidence of flu-like illnesses is stable for children, but has decreased for both adults and the elderly, by about two to five percent per year. One reason for this downward trend is the increased uptake of vaccination, especially for older individuals. In contrast, the incidence of acute diarrhea has increased by one to two percent per year over the period 1990-2010, and for all age groups. The incidence of chickenpox for children has not changed in a significant way during that period.

Figure II displays the average incidence rate within a year, by calendar week, from the first week of January to the last week of December. The graph distinguishes the incidence rate by age groups. The incidence rate is inversely related to age. Some diseases such as chickenpox are almost exclusively a childhood disease, while influenza and acute diarrhea can also affect older individuals. As seen in the previous graphs, flu-like illnesses and acute diarrhea show strong seasonal patterns with a peak in winter and a low incidence between mid-spring to mid-fall. Chickenpox has a different pattern across the year. The incidence is

---

[7]The Reseau Sentinelles represents about 1-2 percent of all general practitioners in France over the period we analyze and cover the whole of France. Appendix A provides further details about the network.

[8]The results are obtained from regressing the log average annual incidence per 100,000 on a linear trend.

roughly constant up to July, then decreases substantially during the summer and increases thereafter. The decrease coincides with the summer vacations of French schools. We return to the influence of school closure on epidemics below in more detail.

One reason for the higher incidence of viral diseases in recent times (at least among the non-vaccinated population) as shown in Table II can be due to the fact that diseases tend to propagate faster than before. Over time, improvement in public transportation and communication infrastructures, increases in trade across regions may help to spread diseases faster. This is important as a fast-evolving epidemic is more difficult to contain. We investigate this point in Table III, where we test whether yearly epidemics reach a peak earlier in the year in more recent times. It is sometimes difficult to identify a unique peak in the incidence rate, as in some regions or years, there could be several of them. We therefore define the seasonal peak as the date when the cumulative infection rate reaches 80% of the total infection rate. This often coincides graphically with the maximum incidence. We experimented with different cut-offs and got very similar results. Given the seasonality of these diseases, we start the year in the first week of July, when infection rates are generally at their lowest. We calculate the time to the peak (expressed in weeks since July) for each region r, each year t and each age group g and denote this variable TTP_{rgt}. This resulted in 1,573 durations (combination of year, region and age groups) for flu-like illnesses and 1,259 durations for acute diarrhea, as the data start only in 1990. We do not attempt the analysis for chickenpox, given that its seasonal pattern is not as marked. We regress the duration to reach this threshold on a constant, a trend (expressed in years), indicator variables for each age group and region indicators:

$$TTP_{rgt} = a_0 + a_1 year_t + a_2 AgeGroup_g + a_3 Region_r + u_{rgt} \tag{1}$$

Table III displays the results. For each additional year, the time to reach the peak decreases by 0.07 weeks for flu-like illnesses and 0.4 weeks for acute diarrhea. In other words, the results indicate a faster spread of these diseases over the period of observation. Note also that the peak is reached first for children and last for the elderly, with a difference of about a week.

II.C. Potential Determinants of the Spread of Diseases

We give details on the data used to explore the determinants of the spread of diseases. We refer the reader to appendix A for additional information on data on weather and population density.

1. School Closures. In France, schools close for vacation five times a year. The summer break lasts for about eight weeks, usually from the beginning of July to the beginning of September. The other vacation breaks last between one and two weeks and take place around October-November (All Saints break), the end of December (Christmas break), February-March (winter break) and April-May (spring break). The calendar is set by the Ministry of Education about two years in advance, and is binding for all public and private schools. Some breaks apply to all regions at the same time, mostly for the autumn, Christmas and summer breaks. For the winter and spring holidays, France is divided into two or three zones, depending on the year, and these zones have a staggered break. Some regions have school holidays earlier than others, with a difference that can reach four weeks. For the epidemics we study, a period of four weeks is large as they evolve quickly. Moreover, some regions have shifted from one zone to another. Hence, with up to 25 years of data, there is variability across years and across regions in school closures, which we exploit to infer the causal effect of school closures on the spread of diseases. This variability allows us to control for region fixed effect and for week fixed effects, as vacation periods have additional variation.

We obtained data from the Ministry of Education on all school holidays for all regions between 1984 and 2010. The school holidays often coincide with vacation taken by parents, especially at Christmas, and often during the winter break. French workers are entitled to 5 weeks of vacations, since 1981. From the year 2000, adults also obtained more vacation as the hours worked per week were capped at 35 hours on average during the year. A large number of workers stayed on a 39 hour week schedule, but were then entitled to 25 additional days of vacation, leading to a total of 8 weeks of vacations. This implies that parents can often take vacations at the same time as their children. Hence, the effect of school holidays on epidemics has to be interpreted in this broader sense.

2. Public Transportation Strikes. We focus on strikes affecting the French national railways (SNCF) which lasted more than three days in a row. In France, railways are an important mean of transportation, as governments have massively invested in railway after the Second World War, when railway lines were targeted by bombings. As a result, French passengers travel on average 1,370km per year by rail, which puts France in the top three countries for rail travel in the world, and the first one in Europe. In the case of a railway strike, travelers can switch to road transportation, but given the limited capacity of roads and highways, these become rapidly blocked, which limits transportation further. Hence a railway strike seriously limits population movements within and across regions.

We searched the popular press through the LexisNexis interface between the year 1984 to 2010 for all strike events that lasted more than three days. Most of the strikes are national, but there are some instances of regional strikes in particular in the south-east of the country. Train strikes often occur when unions and the government negotiate over employment or pay. We recorded between 19 and 28 weeks of strike, depending on the region, between 1984 and 2010, roughly one week of strike per year. Figure IIIa plots the frequency of these strikes during the calendar year and by duration. Strikes occur most often between October and May, with no strikes during the summer. Hence, railway strikes are more likely to occur during the influenza and gastroenteritis epidemic seasons. To eliminate this confounding effect, we control in all the regressions below for week (and year) effects and we assume that the date of the strike is exogenous once we control for such variables. Strikes last on average for about a week and a half, with half of the strikes lasting less than 8 days (see Figure IIIb).

3. High Speed Railway Openings. Since the seventies, the French government has developed a network of high speed trains, which travel at speeds up to 200 miles per hour. We use precise data on the date of openings of high speed rail lines. The first line opened on September 27, 1981 and connected Paris to Lyon, the second largest city in France. Since that date, the network expanded throughout France, linking major cities, with the latest addition in 2007, between Paris and Strasbourg in the east of France. Figure IV plots the main high speed rail-lines in France together with the date they opened. These lines became major transportation routes, as they cut transportation time substantially. For instance a

typical journey from Paris to Marseille on the French riviera was cut from about 7 hours to only 3 hours. The number of passengers travelling on these trains grew from 1.26 million in 1981 to almost 100 million in 2010. Between 1981 and 2013, 2 billion passengers travelled on the high speed train network.[9]

At first, the lines were connecting Paris to major cities. However traveling between two cities other than Paris remained complicated as one had to connect in Paris, where different train stations were handling trains to and from the south, the north or the west. The connection time could add a couple of hours to a journey as well as the hassle of transferring luggage through the Parisian public transportation system. Gradually, major cities became linked without the need of transfer as trains avoided the connection in the centre of Paris. We also use these openings in our empirical model (although the dates do not appear on Figure IV to avoid cluttering). We include a series of indicator variables which take the value of one from the moment two regions become connected by a high-speed train without a connection in Paris. In our empirical model, these networks are only affecting the inter-regional transmission of diseases, as the lines have few, if any, stops within a region.

4. Economic Activity. One aim of this analysis is to evaluate how economic activity influences the rate of transmission of viral diseases. There are several potential indicators of economic activity, which have been used in previous work, which include income (see Ettner (1996), Adams et al. (2003) or Adda et al. (2009)), per capita GDP, unemployment rates (Ruhm (2000) and the literature that enfolded), or trade and exports in a context of HIV infections (Oster (2012)). Although these indicators all capture some aspect of the business cycle, they are also specific to certain aspects of health that the previous literature has focused on. Changes in income may affect use of health care or health behavior such as alcohol consumption or smoking. Unemployment may affect more particularly mental health. We follow the empirical literature on infectious diseases and opt for inter-regional trade as our main measure of economic activity. Higher trade may be associated with higher rates of traveling within and across regions, determining the rate of interpersonal contacts. The choice is also due to lack of data on income at regional level for the earlier part of our

---

[9]The development of high-speed trains reduced the share of travellers by plane and cars along those routes, but also attracted many new travellers.

sample. Nonetheless, trade is highly correlated with aggregate GDP (correlation=0.92) and track expansions and recessions well. We also explore the role of unemployment and how it may affect propagation rates. Further details on the trade data we use can be found in appendix A.

III. School and Public Transportation Closures: Evidence from an Event Analysis

We start by analyzing the dynamics of the incidence of the diseases within regions, following two particular events: school closures due to holidays and strikes that shutdown the public transportation system. Analyzing the global effect of these events within and across regions requires further modeling. This is also the case with other determinants that do not have as sharp temporal variation as the two variables we study. We defer such modeling to Section IV. and start with an event analysis, where an event is defined as the first week of school closure or of a public transportation strike. Denote the incidence of a disease in region r, and week t by I_{rt} and by E_{rt} an indicator variable equal to one if schools are closed or public transportations are on strike in that period. We also define as T_{rt} the average weekly temperature. We estimate the following equation by OLS, for each disease and age group:

$$I_{rt} = \sum_{k=-3}^{K} b_E^k E_{rt-k} + b_T T_{rt} + b_X X_{rt} + v_{rt} \tag{2}$$

The matrix X_{rt} includes a constant, region fixed effects, week of the year fixed effects and year effects. We cluster the standard errors by region.

The coefficients b_E^k are displayed in Figures V and VI. For ease of interpretation, we normalise these coefficients by b_E^{-1} to display the relative incidence of the disease compared to the one in the week prior to the event. In the case of school closures, we find a sharp decline in the incidence of flu-like illnesses, of a magnitude between 20 to 30 percent. This decrease is sustained for at least 4 weeks. The effect is more pronounced for children than for adults or the elderly, but still statistically significant between the weeks 3 to 5 for the latter. For acute diarrhea, the effect of school closure is less pronounced (a decreace in incidence of about 10 percent) and more short-lived. The effect disappears after the third week. Finally,

for chickenpox, we find a statistically significant effect for children after 4 weeks, with a reduction of about 10 percent in the incidence. Figure VI displays similar statistics in the case of a public transportation strike. For flu-like illnesses, the effect is sustained for about 2 weeks for children and adults. We do not find statistically significant effects for the elderly, nor for the other diseases (acute diarrhea and chickenpox).

In summary, the event analysis suggests a marked effect of school closures and public transportation strikes on the dynamics of diseases, and for various age groups. It is interesting to see that school closures have an effect not only on children, but also on adults and even the elderly. This suggests sizable interaction effects between various age groups.

IV. MODEL OF VIRAL SPREAD WITHIN AND ACROSS REGIONS

We now develop a dynamic model of the spread of viral diseases, within and across regions, which is based on models of infections developed in the epidemiology literature. This model is suited to analyse the complex interactions between several determinants of the spread of diseases, both over time and across space. We then discuss its empirical implementation as well as important identification issues.

IV.A. Standard Inflammatory Response Model

The medical literature on infectious diseases has modeled epidemics using a Standard Inflammatory Response (SIR) model. This model is able to describe the dynamics of an epidemic in a concise but accurate way (Kermack and McKendrick (1927), Anderson and May (1991)). Let S denote the fraction of individuals who are susceptible to contract the disease, I the fraction of individuals who are infected and R the fraction of individuals who have recovered but are still immune. At any point in time, the equality $S + I + R = 1$ holds. The model is

usually expressed in continuous time and the flows into each category are expressed as:

$$\frac{dI}{dt} = \alpha SI - \beta I$$
$$\frac{dR}{dt} = \beta I - \lambda R \qquad (3)$$
$$\frac{dS}{dt} = -\alpha SI + \lambda R$$

The first equation in (3) describes how the number of infected individuals evolves over a short interval of time. A fraction β of those infected recover from the disease, while new cases develop, at a rate α. β^{-1} is the average infectious period. New infections occurs from the interaction of susceptible and infected individuals, hence the multiplicative formulation. The stock of recovered individuals is increased by the number of individuals who exit the infectious state and is decreased by the fraction of individuals who lose their immunity, at a rate λ. The stock of susceptible individuals is increased in each period by the flow of recovered individuals who lose their acquired immunity. At the same time the stock is decreased by the fraction of individuals who become infected. The rate λ varies considerably across the various viral diseases we consider. For chickenpox, λ is equal to zero, as individuals acquire a lifetime immunity. Hence, within a birth cohort, the stock of susceptible individuals decreases towards zero as they age. In contrast, individuals who had a spell of gastro-enteritis have almost no immunity, and λ is expected to be high. The case of influenza is an intermediate one. We return to this issue in the empirical section, where we explain how we construct a measure of the stock of susceptible individuals.

The dynamics of the epidemic depends on the ratio α/β, the basic reproduction number. If this ratio is greater than one, an epidemic develops, whereby the fraction of infected individuals sharply increases in a very short time-span. The epidemic eventually dies out as the population of susceptible individuals approaches zero or as the proportion of infectious individuals vanish. The model describes the propagation of viruses within a closed society. In reality, there are also individuals who are sick coming from other regions, so the fraction of infected individuals also depends on the fraction of infected individuals in other regions or countries. The model can be extended such that the number of new cases is equal to $\alpha_{within}SI + \alpha_{between}S\tilde{I}$, where $\tilde{I}$ is the fraction of infected individuals from outside the region of interest who meet susceptible people from within the region.

For influenza, typical parameter values are $\beta = 1/2.6$ days and $\alpha/\beta = 1.8$ (Ferguson

et al. (2006)). Public health interventions are unlikely to change β, which is a biological parameter. Its value depends on the type of virus and the particular strain, but there is usually no treatment that shortens the infectious period. On the contrary, it is likely that the transmission rate, α, varies across space and time. Its value can also be changed by suitable interventions as the transmission rate depends on the frequency of contacts between infectious and susceptible individuals. Hence, the model predicts that keeping sick individuals at home, closing schools or workplaces will lower, at least temporarily, the spread of an epidemic.

The parameter α is also a behavioral parameter as sick individuals decide whether or not to mix with susceptible individuals (and vice-versa). The economic literature on rational epidemics (see Kremer (1996), Geoffard and Philipson (1996), Philipson (2000) or Chan et al. (2015)) has extended the epidemiological models to take into account how individuals react to changes in prevalence. However, in the case of many viral diseases, contaminated individuals become infectious before the symptoms become apparent, which mitigate the scope for avoidance. Moreover, as most viral diseases are benign, the cost of avoidance can outweigh the benefits of avoidance of potentially infected individuals. The literature on rational epidemics has therefore focused mostly on the case of sexually transmitted diseases such as HIV. [10] Taking into account behavioral changes for the diseases we study is difficult as it would require data on whether individuals protect them-selves by interacting less with others, or whether they wash their hands more often during epidemics.[11]

The transmission parameter α will also depend on many factors such as population density or the rate at which individuals travel. Hence, the spread of epidemics depends both on long-run and short-run demographic and economic factors. In the long-run, the spread of viruses depends on how a society is organized and how integrated the economy is, as discussed in Fogli and Veldkamp (2013). Developing roads, highways, rail connections or airports may increase the speed of propagation and the number of cases.

In the short-run, the strength of an epidemic may depend on fluctuations in economic

---

[10]Empirical evidence that individuals alter their behavior in response to the perceived risk they face include Adda (2007), Thornton (2008), Neidell (2009), Moretti and Neidell (2011) and de Paula et al. (2013).

[11]If behavioral adjustment is important, it would imply that the parameter α is decreasing with the perception individuals have of the likelihood of contracting the disease. We test this conjecture in Appendix D.

activity, for at least two reasons. First in good times, more individuals may travel and meet people for business purposes. Second, unemployed individuals may have different patterns of socializing, which may lead to different propagation rates. Ruhm (2000) shows that individuals are in better health during recessions. Part of the effect is due to changes in health behavior, which leads to fewer deaths during recessions. Whether epidemics and viral contagion fluctuates with the business cycle remains an open question we address in this paper.

The literature in epidemiology has produced an important body of research to devise public health strategies to curb viral epidemic, especially for influenza. [12] This work often relies on calibrated parameters, which are chosen partly based on medical information (infectious period) and partly on average durations and scope of an epidemic. With such an approach, it is difficult to empirically evaluate the role of policies as one lacks sharp and exogenous variation for identification. Another strand of the literature has estimated models of diffusion using OLS, Bayesian or simulation techniques, using high frequency data. However, this literature has failed to consider the endogenous nature of important explanatory variables.[13]

IV.B. Econometric Model 1: Within Region Spread

We start by taking to the data a simple form of the model developed in Section IV.A., which we estimate by OLS. This serves as a baseline model.

Let I_{rt} denote the incidence rate, i.e. the number of new cases, of a particular virus in region $r \in \{1, \ldots, R\}$ and S_{rt} the proportion of susceptible individuals in a given region. We write equation (3), in discrete time as:

$$I_{rt} = \alpha_{within} I_{rt-\tau} S_{rt-\tau} + X_{rt}\delta + \eta_{rt} \tag{4}$$

---

[12] Much less work exist on other diseases such as gastroenteritis or childhood diseases such as chickenpox.

[13] For instance, some authors have investigated the effect of public health measures such as travel restrictions and school closures (Ferguson et al. (2006)). However, these estimations rely on assumptions on how transmission rates would decline were a school to close, but are not estimated from data. The same argument applies for travel bans. Recent examples of this literature are Hufnagel et al. (2004) and an important exception is Cauchemez et al. (2008), which estimate a model using a simulated maximum likelihood method.

where the matrix X_{rt} includes region fixed effects, week effects and year effects in levels. The parameter τ represents the incubation time. For acute diarrhea and flu-like illnesses, we set τ to one week, as both diseases incubates in less than a week (see Table I). For chickenpox, we set τ to 3 weeks.

We identify the model without relying solely on time variation or cross-section variation, which could confound many of the effects. The implementation using a difference-in-difference identification strategy is also a departure from the existing literature in epidemiology. The shock η_{rt} is possibly serially correlated, which would lead to bias since equation (4) is essentially a lagged dependent model. We deal with this issue below. We use a panel-corrected standard error procedure allowing for heteroskedasticity, spatial correlation across regions and for serial correlation of the form of an autoregressive process of order one.

The estimation of equation (4) requires the computation of the stock of the susceptible population, S_{rt} for each region and year. We proceed in different ways for each viral disease, taking into account the differences on how immunity is acquired.

For influenza, the chances to get a second bout in the same year is low (although not zero). We assume that immunity lasts for a year, until a new epidemic starts. Given the seasonality patterns of the disease, we define the start of the year to be at the beginning of July, when the infection rate is at its lowest. The stock of susceptible individuals in a given week is therefore the entire population minus those who are vaccinated and those who have been infected previously since the end of the previous July.[14]

For gastro-enteritis, we assume that immunity only lasts for a week after the end of the infection. Hence the stock of susceptible individuals consists of the entire population minus the fraction who were ill the previous week. Given the low efficiency of the vaccine and its rare usage in France (the vaccine was first produced in 2006), we do not consider it in the computation of the susceptible population.

In the case of chickenpox, a contaminated individual acquires a life-time immunity. Hence the susceptible population consists mostly of very young children. We construct the stock

---

[14]Data on vaccination rates per age groups were obtained from GEIG, a French public health institute (see http://www.grippe-geig.com). The data is at the national level, but broken down by age groups and we predict the incidence of vaccination in each region and year based on the age structure of the population.

of susceptible individuals for each period and region by combining data on the incidence of diseases in the previous years, as well as data on the size of birth cohorts across region and time. The exact procedure is detailed in Appendix B.

For ease of interpretation, we have normalised the susceptible population by the total population in the region in a particular age group. Hence, the coefficients presented are to be interpreted as marginal effects of a change in the infection rate on the future infection rate, when the entire population is susceptible to the disease. However, as a disease progresses, the pool of susceptible individuals decreases, which implies smaller marginal effects. The results are displayed in Table IV, Panel A. Each column represents a separate regression for different diseases or age groups. In the case of flu-like illnesses, each infected individual infects 0.28 children, 0.46 adults and 0.1 elderly individual. In the case of acute diarrhea, the estimates indicate that an infected individual transmit the disease to 0.18 children, 0.26 adults and 0.03 elderly individuals. Finally, a child infected with chickenpox spread it to 0.18 other children. These estimates are significant at the one percent level, but are on the low side, as each infected individual is transmitting the disease to fewer than one other individual in total. If that is indeed the case, then these diseases should not evolve into epidemics, which is counterfactual. The reason for this discrepancy is that the estimation is likely to be biased for several reasons, including measurement error and omitted variables.

IV.C. Econometric Model 2: Within and Between Region Spread

We next introduce a spatial dimension to the model, and we let the incidence rate of region r be determined by the incidence rate of other regions as well. Given that the pattern of socialisation across regions is likely to be different from the ones within a region, we allow the transmission parameter α to be different within and between regions:

$$I_{rt} = \alpha_{within} I_{rt-\tau} S_{rt-\tau} + \alpha_{between} \sum_{c \in R \backslash r} I_{ct-\tau} S_{rt-\tau} + X_{rt}\delta + \eta_{rt} \tag{5}$$

The set of all other regions includes all regions R minus region r, denoted $R\backslash r$. In this specification, each region has the same effect, regardless of distance or connection. We introduce differential effects across region below.

1. OLS Results. We estimate model (5) by OLS and display the results in Table IV, Panel B. Each individual infected with a flu-like illness is spreading the disease to 0.24 children, 0.41 adults and 0.1 elderly individuals within the same region. In addition, this individual infects 0.005 children, 0.006 adults and 0.003 elderly individuals in each of the other regions. Although the number of further infected individuals in the same region is lower than the one found in Section IV.B., the total number of infected individuals is now slightly above one.[15] The estimated propagation rates are also higher for acute diarrhea and chickenpox, but still well below one in total. This points out that it is important to take into account spatial diffusion across regions, but also that the estimates are still potentially biased.

2. Instrumental Variable Results. As argued above, there are several reason why OLS estimates of equation (5) may be biased. First, the possible serial correlation in the error term would lead to bias the OLS estimates, given the dynamic structure of the model. Second, measurement error would also lead to bias. The data on the incidence of diseases are estimated from the number of cases seen by a network of general practitioners, which is subject to sampling error. Given that the incidence of diseases appears as an explanatory variable, it is important to take this fact into account. We show in Appendix C that measurement error leads to a complex error term, with serial dependence. Hence, equation (5) cannot be consistently estimated by OLS.

To get consistent estimates of the transmission parameters, we use instrumental variables. We have two variables to instrument, the within-region lagged incidence times susceptible rates, $I_{rt-\tau}S_{rt-\tau}$, and the one across regions, $\sum_{c\in R\backslash r} I_{ct-\tau}S_{rt-\tau}$. We use lagged weather episodes as instruments. There are biological reasons to consider those variables as instruments. Viruses, especially influenza, do not survive well in warmer temperatures, but can cope well with cold and dry conditions (e.g. Lowen et al. (2007)). Viruses responsible for gastroenteritis are also sensitive to warmer temperatures (Moe and Shirley (1982)). Moreover, adverse weather may also influence socialising patterns and the rate at which viruses are

---

[15]The total number of infected individuals is obtained by summing the coefficients across age groups, and, in the case of between region coefficients, by multiplying by the number of regions minus one. The analysis uses 21 regions.

passed on from person to person. In the case of diseases with some degree of acquired immunity (influenza or chickenpox), the number of susceptible individuals, $S_{rt-\tau}$, displays inertia and previous weather episodes are particularly suited instruments. For the within region term, we construct the number of weeks with temperatures below 0, 5, 10, ..., 25 degrees Celcius, from the end of the preceding summer (when incidence rates are at the lowest). Similarly, we construct the amount of rain in the preceding four and eight weeks. We therefore use a total of 8 instruments.[16] For the between region term, we use 11 instruments, consisting of the ones detailed above and three measures of the number of regions with weekly temperatures below 0, 10 and 20 degrees Celsius. This aggregate measure of temperature determines the sum of the incidences across all regions (except r).

As seen in Table I, each disease vary in terms of length of incubation and duration of symptoms. Hence it is likely that the effect of weather is stronger at various lags, depending on the disease we study. This is precisely what we find. We use a one week lag for acute diarrhea and for flu-like illnesses and a three week lag for chickenpox. For the latter disease we also use the size of the birth cohorts, at regional level, in the preceding six years. Indeed, as those who contract the disease acquire near full immunity, the pool of susceptible individuals consists mainly of young children and a larger birth cohort in a given year and a region leads to a larger pool of susceptible individuals a few years laters, when this cohort enters kindergarten or primary school. The regression also includes region, year and week fixed effects. We test for the joint significance of the instruments and present the F-test and its associated p-value in Table V.[17] The instruments work well, with all the p-values well below the 5 percent level. The table also displays R-squares, which vary between 0.3 and 0.8. Note that our first stage is stronger for children and adults, compared to the elderly. It is also stronger for the susceptible population interacted with the incidence rate in other regions, which is equivalent to the national incidence rate minus the one in the region of interest. The reason is that lagged weather is a better predictor of the national incidence rate than the regional ones.

The results are displayed in Table IV, Panel C. Compared with the OLS results presented

---

[16]In the case of chickenpox, we also include the size of the birth cohort within the region in the last 6 years.

[17]Table I in the appendix provides a full list of the first stage coefficients.

in Panel A and B of the same table, the transmission rates are globally larger. One individual with flu-like illnesses spread the disease to about 1.27 individuals, within and across regions and age groups. In the case of acute diarrhea and chickenpox, the number of infected individuals is equal to 1.08 and 1.9 respectively.[18] The Hausman test displayed at the bottom of the table shows that exogeneity of the regressors is rejected in six out of seven cases. The results in the last panel highlight the importance of taking into account the endogenous nature of the dependent variables and of allowing for spatial effects.

3. Econometric Model 3: Full Model. We now consider a complete model, where the transmission rate is time-varying and region specific. Allowing for heterogenous effects in the coefficient of interest, α, in the context of a dynamic model, is rare in the economic literature, especially relying on differential variations across space and time. It is feasible, given that we are relying on a small cross-section of regions and a large time span.

Denote a set of K region-specific variables W_{rt}^k that potentially influence the transmission rate of diseases within a region. They include those already mentioned earlier, such as school closure and transportation strike indicators. We add to the model the effect of temperature, heavy rainfall or snow, population density, a measure of economic activity (volumes of trade within the region), quarterly dummies and a time trend. For the transmission rates across regions, denote a set of $\tilde{K}$ variables $\tilde{W}_{rct}^k$, which measure characteristics specific to regions r and c. For instance, the variable can be a binary indicator equal to one when schools in both regions are closed, when both regions experience a train strike or when they are linked by a high speed train connection. We add to those measures inter-regional trade, distance between regions (defined as the distance between the most populous cities in each region), population ratios, log regional GDP ratios and temperature differences. These latter variables capture potential asymmetries in the spatial transmission of viruses. We also allow for quarterly dummies and a time trend to affect the inter-regional spread. Although the model includes many determinants of viral spread, some of them have of course been left

---

[18]The literature in epidemiology has produced numbers which are of similar magnitude, although slightly higher. The basic reproduction number for influenza is usually around 2. The one for chickenpox is closer to 7, compared to the one found in the analysis of about 4, taking into account the fact that the disease lasts for about 2 weeks.

out, due to data limitation. Any event that lead to changes in social interaction could affect the spread of diseases. Omitted variables include the flow of travellers by road or by air between regions. These data are unavailable over the period of analysis and even more so at a weekly frequency. These variables are obviously correlated with holidays, the expansion of high speed trains and to some extent with public transportation strikes. The interpretation of the effect of these variables on the spread of diseases is contingent on this omission. More individuals may opt for cars during a public transportation strike, which means that the effect we are estimating is the combined effect of strikes and higher usage of alternative transportation. Similarly, trade is proxying for economic activity, but we cannot break it down into various components, including changes in income, travel or social interactions with the data at hand. Note that the model includes trends both for the within and between transmission of diseases, in addition to region, year and week fixed effects, which could proxy for long-term changes in patterns of transportation for instance. The estimated model is written as:

$$I_{rt} = I_{rt-\tau}S_{rt-\tau}\sum_{k=1}^{K}\alpha_{within}^{k}W_{rt-\tau}^{k} \tag{6}$$

$$+ \sum_{c\in R\backslash r}I_{ct-\tau}S_{rt-\tau}\sum_{k=1}^{\tilde{K}}\alpha_{between}^{k}\tilde{W}_{rct-\tau}^{k} + X_{rt}\delta + \eta_{rt}$$

As in the previous specifications, X_{rt} contains region, year and week-in-the-year fixed effects. We use a panel-corrected standard error procedure allowing for heteroskedasticity, spatial correlation across regions and for serial correlation of the form of an autoregressive process of order one. The interaction variables W_{rt}^{k} and $\tilde{W}_{rct}^{k}$ are assumed to be exogenous, conditional on the week, year and regional fixed effects. We instrument the endogenous variables using the same instruments as in Section 2. and interacting them with exogenous variables. The regressions are done separately for each disease and each age group. For a clearer presentation, we group similar parameters across diseases and age groups and present them in Table VI across different panels.

Factors Reducing Interpersonal Contacts. We first turn to factors that potentially reduce interpersonal contacts, presented in Table VI, Panel A. We focus first on school closures due to holidays. School closures have a large and significant effect on the rate

23

of transmission to children, with a decrease of 0.14 individuals infected in the case of flu-like illnesses, 0.07 individuals for acute diarrhea and 0.19 individuals for chickenpox. The effect is relatively strongest for flu-like illnesses and weakest for chickenpox, in line with the results obtained with the event analysis in Section III. The incubation period for this disease is longer than the usual school vacation, which ranges from one to two weeks, except for summer holidays. School holidays also decrease the transmission rate of flu-like illnesses for adults with a reduction of about 0.1 people infected. However, the effect for acute diarrhea in adults is positive but not significant. The elderly are affected in different ways. There is a small but insignificant decrease (at the 5% level) for flu-like illnesses, but a significant *increase* in the propagation rate for acute diarrhea. In principle, the effect of school closure could go in opposite directions. School closures may decrease the general incidence of a disease and lead to an indirect effect on the elderly. On the other hand, school holidays are also a period when children are more likely to interact with their grand-parents, and could therefore be a period of heightened infectiousness. The overall effect is therefore ambiguous, with a detrimental effect for acute diarrhea.

The second row of the panel displays the effect of school holidays on the transmission across regions. As noted in Table IV Panel C, discussed above, the average propagation rates are generally much smaller than the ones within a region. The overall effect is also ambiguous, as school closures would reduce interpersonal contacts, but are a period when families are traveling, which could actually increase the propagation rates. For most age groups and diseases, school closures lead to a decrease in the propagation rate, except for children in the case of acute diarrhea. Each infected individuals transmit the disease to 0.003 more children across regions during a school closure.

We now turn to changes in the availability of public transportation, either through a shutdown due to strikes or through the openings of new high-speed train links. Transportation strikes should reduce the propagation of diseases by limiting travel. However, shutting down public transportation does not necessarily hinder people from traveling, as they can still travel by car or other means. Most often, transportation strikes lead to chaotic situations where people may be more exposed than usual, as they try to cram into crowded vehicles. One could argue that such events are actually relevant for public policy, as health

authorities may not have the power to shut down effectively all means of transportation in case of a major disease outbreak. The analysis provides some evidence that public strikes reduce the propagation rate of diseases. We find significant decreases in the propagation rate for adults and the elderly in the case of flu-like illnesses within a region. For the other diseases and age groups, the estimates are not precise enough to rule out no effect of such events. We also find a decreased rate of propagation across regions, for all age groups in the case of flu-like illnesses and acute diarrhea.

The fifth row shows the effect of opening up a new high speed rail line. The effect is striking as we find a marked increase in the transmission rates for two of the three diseases we analyze, flu-like illnesses and acute diarrhea and for all age groups. The opening of a new high-speed rail line leads to an increase of 0.001 to 0.014 more individuals infected for each sick individual in the regions being connected. The effect is larger for flu-like illnesses than for acute diarrhea. The effects are sizeable as they compare to the cross-region effects in Table IV, Panel C.

To check further the hypothesis of a lower propagation rate of diseases when transportation is hampered, we look at how the transmission rates vary in case of adverse weather. This is defined as the occurrence of either torrential rain (precipitation above the 90th percentile within the region) or a combination of freezing temperatures and above median precipitations, which would lead to heavy snow. Note that we also control for temperature in the model, to capture the fact that diseases are more prevalent in cold weather. We find consistent evidence that such extreme weather episodes reduce the propagation of flu-like illnesses or acute diarrhea within a region. Again, we do not find evidence of an effect in the case of chickenpox. The results are consistent with a reduction in travels and less interpersonal contacts.

Transmission Rates and Economic Activity. We now turn to the effect of economic activity on the transmission of diseases. We present the effect of intra and inter-regional trade in Table VI, Panel B. The trade variable has been standardised to show the effect of a one standard deviation change. Such variation within a region decreases the transmission rates for children by 0.08 cases for flu-like illnesses, but this effect is not statistically significant. We

do not find much evidence of a link between trade and disease propagation within a region. However, increases in trade between two regions significantly increase disease transmission for all diseases and almost all age groups. Our results point to a pro-cyclical effect of economic activity on disease transmission. This is in contrast with the counter-cyclical effects in the case of mortality found by Ruhm (2000). However, for the elderly, which is the population most at risk of dying from viral infections, we find significant increases for flu-like illnesses only.

Panel C of Table VI uses local unemployment rates within a region and national employment rates across region to investigate further the effect of economic activity.[19] As discussed above, unemployment is a marker of economic cycles, but can also capture additional effects such as different socialisation patterns for those out of the labor force. The results are more mixed than the ones using trade flows. We find a reduction in the transmission of acute diarrhea within a region for adults, but contrasted effects across regions and age groups.

Within and Between Region Characteristics. Table VI, Panel D presents the effect of temperature and population density on within region transmission rates. As expected, average weekly temperatures (expressed in Celsius) have a marked effect on propagation rates. We use various lags depending on the disease we study, using one week lag for flu-like illnesses and acute diarrhea, and a three week lag for chickenpox.[20] Higher temperatures reduce the propagation rates for all diseases, a combination of the difficulty of the virus to survive warmer temperatures and behavioural changes in socialisation patterns.

We find evidence that regions with higher population density have higher transmission rates for children and for adults and a negative effect for the elderly, for flu-like illnesses, acute diarrhea and chickenpox.

Table VI, panel E investigates heterogeneous effects across regions. The first variable we

---

[19]As the unemployment rate and trade volumes are correlated, including both in the regression leads to non-significant estimates. The results we present use either one or the other measure of economic activity. The coefficients displayed in Table VI for the other variables are obtained from a regression including trade. The results are not quantitatively different if unemployment is included rather than trade.

[20]Note that the use of lagged temperature does not invalidate the instruments as they are constructed as cumulative days of cold weather over longer periods, and because temperature series, while persistent to some degree, have independent variation condition on past values.

consider is distance. While for flu-like illnesses and chickenpox we do not find much evidence of a role for distance, in the case of acute diarrhea the spread is decreasing with distance. We next look at asymmetric transmission rates across regions. These asymmetric effects can arise because of differential socialisation patterns or different propensities to travel. We find evidence that viral diseases transmit faster from less populated regions to high populated regions. The population ratio is expressed in logs. Transmission of acute diarrhea between a sending region twice as large as the receiving one are increased by 0.005 cases for children. We find similar patterns for other diseases. The next row displays the effect of differential regional GDP, expressed as the log of the ratio of the GDP of the sending to the receiving region. For adults and the elderly, viruses tend to move from poorer to richer areas, ceteris paribus. The third row presents the effect of differential temperatures. We find consistent effects that transmission rates are higher going from colder to warmer regions. Finally, the regressions also control for the incidence in the Paris region, given the centrality of that region and we find that transmission rates from Paris to other regions is higher.

V.   Closing down Schools or Public Transportation?

We now investigate the efficiency of policy measures aimed at reducing the prevalence of diseases, by analysing their costs and their benefits. We rely on the model estimated in the previous section and we simulate two counterfactual policies: school closures or public transportation shut-downs. We explore their effects on flu-like illnesses and on acute diarrhea, but not on chickenpox, for which we do not find large or significant effect of such measures.[21]

We first simulate a baseline scenario without these policies. We track how the disease spreads across space, time and age groups over a year. To this end, we draw incidence shocks (labelled η_{rt} in equation (6)) as well as values for the parameters of the model from their estimated distribution. Each draw leads to a particular incidence path over the year, and we average the incidence paths over 500 replications. Given the seasonal nature of flu-like

---

[21]In the case of chickenpox, it is not clear public health authorities would want to halt an epidemic, especially among children, as the disease is more serious for older individuals.

illnesses and acute diarrhea, we start the simulations in the first week of September when incidence is very low. We allow the temperature to change over the year, given that it is an important determinant of the spread of these diseases. We use the average temperature observed week by week over the period of analysis. We allow for usual school vacations but assume that no public transportation strikes take place during that year. We keep the other explanatory variables used in the estimation fixed at their mean.[22] In the first weeks, the incidence increases slowly, a combination of warm or mild weather in autumn, frictions in the spread of diseases as captured in the estimated propagation coefficients, and regular school holidays. During winter times, a critical mass of individuals are infected and with low temperatures, the disease reaches epidemic proportions. Eventually, as more people become immune - in the case of flu-like illnesses - and warmer temperature during the spring and summer, the epidemic dies out.

For the evaluation of the first policy, we introduce a spell of two weeks of school closure, which is the typical variation observed in the data. This spell is in addition to regular holidays. We investigate the effect of this policy introduced in any week during the year, to find out the optimal timing of such a policy and its largest effect. The second policy is a public transportation shut-down for a week and again, we evaluate its effect as a function of the week when it is introduced. We assess these policies in two ways. We first compute the total number of individuals who contract the disease over the whole year and we compare it to the baseline. We also perform a cost-benefit analysis, where we take into account the short and long-run consequences of the policies we evaluate.

Figure VII displays the effect of the two policies on the annual prevalence, relative to the baseline. Figure VIIa shows the effect of closing schools at various times during the year on the prevalence of flu-like illnesses. The effect of the policy is largest in early December, a period of high prevalence. At this point, closing down schools would decrease the total annual incidence rate by about 12 percent.[23] Figure VIIb shows the effect of a similar policy in the case of acute diarrhea. The largest effect of the policy is a reduction in the total incidence rate of about 4 percent and the optimal timing is in early January.

---

[22]For flu-like illnesses we fix the proportion of vaccinated individuals to the observed levels in 2005.

[23]The effect is discontinuous during the regular school holidays, when the effect of the policy is trivially zero. The effect is smaller one week before holiday periods, as the policy then only lasts for one week.

Figures VIIc and VIId display the effect of a public transportation closure on the annual incidence of these diseases. For flu-like illnesses, the highest reduction is equal to 8 percent, in mid January. As apparent from Figure VIIc, there is a complementarity between school closures and public transportation shutdowns. The reduction in incidence is lower for acute diarrhea, with a maximum reduction of about 2 percent in mid December.

We next perform a cost-benefit analysis of these policies. We draw on the epidemiological literature for the evaluation of the costs of these diseases, in terms of medication, health care use and costs. We also take into account the cost of death by considering mortality risks, which are important for the elderly, and to a smaller extent for newborns. We extend this literature by considering the effect of school closures on the human capital of children.

The empirical literature calculating the value of a statistical life has provided quite wide ranging estimates, depending on the method used. Ashenfelter and Greenstone (2004) found a rather low value of about 1.5 million dollars using mandated speed-limits. Viscusi and Aldy (2003) review the literature and find values between 5.5 and 7.5 million dollars. Murphy and Topel (2006) use a value of 6.3 million dollars. In this study, we use a range of values between 1.3 and 6 million euros.

Table VII displays the costs that we consider in the simulations. We distinguish three age groups as the diseases have very different effects by age, especially when we compare the elderly to the other age groups.[24]

For children, the cost of flu-like illnesses are not predominantly medical. Treatments are cheap and complications such as otitis media or pneumonia, although expensive, are rare. Death rates due to influenza are very small, about 1 per 100,000. The cost of these diseases for this age group comes from the loss of human capital, a fact that is often neglected in the

---

[24]Data on costs and healthcare use are taken from Prosser et al. (2006) for children, from Nichol (2001) for adults and from Molinari et al. (2007) for the elderly. Medical costs are weighted by the probability of health care usage. Data on mortality from influenza by age group comes from the National Vital Statistics Report 2011. Data on wages are taken from INSEE, "Revenus salariaux médians des salariés de 25 à 55 ans selon le sexe en 2011" (http://www.insee.fr/fr/themes/tableau.asp?reg_id=0&ref_id=NATnon04146). Labor market participation data comes from OECD skill data set. All US dollars converted into euros with an exchange rate of 0.8. Loss of human capital is costed using a return to schooling of 5 percent, median wages by sex and average labor market participation by sex over a period of 42 years.

epidemiological literature. We assume that children who experience a bout of flu-like illness loose 3 days of schools. This reduces their human capital in that year by about 0.8 percent. We assume a return to schooling of 5 percent per year, and we compute the net present value of earnings over the life-cycle (42 years), with an annual discount factor equal to 0.95. We also take into account the effect on retirement benefits and assume that individuals live 15 years in retirement. This yields a loss of about 100 euros per illness episode. Another large cost is the loss in productivity as an adult usually stays home to supervise the child. We assume that this person is the mother and we impute a cost based on the median female income, weighted by the labor market participation of women. This induces a cost of about 100 euros as well.

For adults, the medical costs are also small compared to the overall cost. The main cost comes from the risk of death. Although this risk is low (about 0.4 cases per 100,000 individuals), the value of a statistical life is high. The second largest cost for adults is the loss of productivity as they stay home for an average of 2 days, and they suffer an additional loss in productivity for 0.7 days, where they work at 50 percent of their capacity. These numbers are taken from Nichol (2001) and we use the average wage by sex, combined with average labor market participation by sex, assuming that men and women have equal probability of contracting the disease.

Finally, the costs for the elderly come mainly from death as the probability of death is about 100 per 100,000 cases. Medical costs are also much higher than for younger individuals.

Table VIII presents similar statistics for acute diarrhea.[25] Costs are lower for children than for flu-like illnesses, as the disease is more short-lived and requires less health care use, except for newborns. Mortality risks are also lower, especially for adults and the elderly.

---

[25]Data on costs and healthcare use are taken from Van Den Brandhof et al. (2004). Medical costs are weighted by the probability of health care usage. Data on mortality from acute diarrhea by age group comes from the National Vital Statistics Report 2011. Data on wages are taken from INSEE, "Revenus salariaux médians des salariés de 25 à 55 ans selon le sexe en 2011" (http://www.insee.fr/fr/themes/tableau.asp?reg_id=0&ref_id=NATnon04146). Labor market participation data comes from OECD skill data set. All costs are adjusted for inflation and in 2005 euros. US dollars converted into euros with an exchange rate of 0.8. Loss of human capital is costed using a return to schooling of 5 percent, median wages by sex and average labor market participation by sex over a period of 42 years.

Table IX displays the costs induced by the policy. The school closure policy leads to 2 weeks of loss of schooling over the year, which we value in the same way as in Tables VII and VIII. The effect of closing public transportation is a priori more difficult to evaluate. We rely on evaluations made after the French strike of 1995, which lasted for 22 days and blocked the entire public transportation grid. Dubois et al. (2007) evaluate the effect from an aggregate perspective, looking at sectoral GDP effects at a monthly frequency. They find a loss of about 400 million euros per week, with the transportation and manufacturing sectors bearing the largest costs. Alternatively, Coindet (1998), using micro data for the same period, calculates a cost which is equivalent, based on individual's self-declared work behavior during that period.

Figure VIII presents the cost of each policy, taking into account their joint effect on both flu-like illnesses and acute diarrhea. The graph is computed with a value of statistical life of 5 million euros regardless of age. We present robustness calculations with respect to various statistical life values in Appendix D. A policy that closes down schools is never cost effective. It represents a cost varying between 16 and 37 euros per capita, depending on when it is implemented in the year. As seen in Tables VII and VIII the cost of diseases is high for children. Adults and the elderly would actually benefit from such a policy as they would not suffer from a loss of human capital. Alternatively, we evaluate a policy where schools would be open for an additional two weeks in early July to compensate for the loss of human capital during the closure when prevalence is high. This policy yields positive benefits, equal to a maximum of 6 euros per capita if implemented in mid December or early January. This policy is similar to the one implemented in Japan or Bulgaria.

Figure VIIIb shows the effect of a public transportation shutdown. Such a policy is not beneficial either, although the cost is less than for school closures, varying between 8 and 19 euros per capita.

Given that mortality is an important component of the overall cost of the diseases we consider, we simulate counterfactuals where we vary the probability of death. Mortality rates for influenza vary depending on the type (strain) of the virus. Excess mortality can vary by a factor of 1 to 5 in post-war pandemics (see Simonsen et al. (1998)). Taking drastic measures such as the closure of all schools, may become important in years when the virus

31

is particularly deadly. We therefore explore what level of mortality risk would justify closing down schools. Fixing the value of a statistical life at 5 million euros, we vary the mortality rate from influenza from its baseline value in Table VII. We find that closing down schools or public transportations would be cost effective for a death rates about 2.1 times higher than the baseline. In the case of acute diarrhea, mortality rates and costs are low. Closing down schools would only be efficient if the probability of death is at least 40 times higher.

VI. The Health Cost of Transportation Infrastructure

The economic literature has evaluated investments on transportation infrastructures through their effect on productivity and growth, as in Fernald (1999), Donaldson (2010), Duranton and Turner (2012) or Banerjee et al. (2012), or through their effect on alleviating congestion as in Duranton and Turner (2011). While this literature has found mostly beneficial effects of expanding transportation infrastructure, the literature has not considered the cost of those infrastructures in the dissemination of diseases, leading to health care costs and lost productivity. We investigate in more detail the health costs induced by the expansion of high-speed rail lines in France. We use the results in Section 3. (and displayed in Table VI) and simulate a counterfactual where we take out one high-speed link, to evaluate the difference in the aggregate and annual prevalence of flu-like illnesses and of acute diarrhea. We perform Monte Carlo simulations, where we eliminate randomly one link each time - out of the 44 possible links that existed in France in 2009, in order to evaluate the average effect of connecting two regions. We find that one high-speed link contributes to less than 1 percent of the annual prevalence of both flu-like illnesses and acute diarrhea. Using the costs displayed in Tables VII and VIII, we find a cost of this high-speed link of about 80 million euros per year, 80 percent of which is due to flu-like illnesses. To put this number in perspective, this is lower than the gain in terms of time saved by the introduction of high speed trains, evaluated at the average hourly wage, calculated to be in the order of about 100 million

euros per year and per link.[26] However, the gain from high-speed trains is likely higher, as it may also have increased productivity of firms and offered more opportunities to workers.

VII. CONCLUSION

This paper studies the economic determinants of the spread of viruses across time and space, both in the short and longer-run. In the short-run, it assesses the effect of potential public health measures aimed at reducing interpersonal contacts, such as the closure of schools, of public transportations or preventing larger gatherings of people. In the longer-run, it explores the role of the transportation network , as well as the role of economic cycles and increases in trade within and across regions in disseminating diseases. We show that the expansion of transportation networks and the growth in trade can explain part of the rise in the transmission rate of viral diseases that we document during the last quarter of a century.

We find that measures that prevent interpersonal contacts, such as limiting travel or closing down schools work well, for certain type of diseases and certain age groups. However, their efficacy depends crucially on disease characteristics such as their incubation period. We find that such measures benefit mostly younger people, who see a reduction in the incidence of these diseases. These measures impose costs, which makes them not necessarily desirable as a tool to contain seasonal flu or gastroenteritis epidemics. In particular, the most popular one, which consists of closing down schools comes as a net costs, especially for young individuals, unless closures are offset by an equivalent opening of schools during the summer. Our results therefore support school closure policies such as the ones in place in Japan. However, school closures or public transportation shutdowns would become cost effective for flu epidemics in instances when their death rate is above average. We also find that such measures can have adverse effects on the elderly, which is a population that is most at risk of complications due to those diseases.

Viruses represented a major threat to health in the past century. Looking into the future, it is unclear whether developed countries will suffer less from viral diseases. While

---

[26]Traffic on high-speed trains in France carried the equivalent of 49 billion person-kilometers in 2009, at an average speed of 250km per hour, a 60 percent improvement on car transportation or regular trains. The average hourly wage is equal to 14 euros.

better use and development of vaccines can help protect from outbreaks, new emerging diseases regularly make the headlines. They may be on the rise because of closer contacts between men and wild habitats in Asia and Africa, and because of closer economic ties between continents. Another reason can be the increase in global temperatures. Our results show that warmer temperature help containing epidemics of familiar viruses responsible for influenza or gastroenteritis. New viruses may not be as sensitive to temperature and there is also a fear that older viruses preserved in ice or in the permafrost may come back when such regions warm up due to global warming.[27] This is why it is essential to better understand whether less conventional methods to curb epidemics are efficient and when to optimally use them. It is also important to take into account how economic growth and development of infrastructure shape disease transmission in the long-run.

---

[27]See Legendre et al. (2014).

References

ADAMS, P., M. D. HURD, D. MCFADDEN, A. MERRILL, AND T. RIBEIRO (2003): "Healthy, wealthy, and wise? Tests for direct causal paths between health and socioeconomic status," *Journal of Econometrics, 112, 3–56.*

ADDA, J. *(2007): "Behavior Towards Health Risks: An Empirical Study Using the "Mad Cow" Crisis as an Experiment," Journal of Risk and Uncertainty, 35, 285–305.*

ADDA, J., J. BANKS, AND H.-M. V. GAUDEKER *(2009): "The Impact of Income Shocks on Health: Evidence from English Cohorts," Journal of the European Economic Association, 7, 1361–1399.*

ALMOND, D. *(2006): "Is the 1918 influenza pandemic over? Long-term effects of in utero influenza exposure in the post-1940 US population," Journal of Political Economy, 114, 672.*

ANDERSON, R. M. AND R. M. MAY *(1991): Infectious Diseases of Humans: Dynamics and Control, Oxford, UK: Oxford University Press.*

ASHENFELTER, O. AND M. GREENSTONE *(2004): "Using Mandated Speed Limits to Measure the Value of a Statistical Life," Journal of Political Economy, 112, S226–S267.*

AULD, C. *(2006): "Estimating Behavioral Response to the AIDS Epidemic," The B.E. Journal of Economic Analysis & Policy. Contributions in Economic Analysis & Policy, 5.*

BANERJEE, A., E. DUFLO, AND N. QIAN *(2012): "On the Road: Access to Transportation Infrastructure and Economic Growth in China," NBER Working Paper No. 17897.*

CAUCHEMEZ, S., M. D. V. KERKHOVE, B. N. ARCHER, M. CETRON, B. J. COWLING, P. GROVE, D. HUNT, M. KOJOUHAROVA, P. KON, K. UNGCHUSAK, H. OSHITANI, A. PUGLIESE, C. RIZZO, G. SAOUR, T. SUNAGAWA, A. UZICANIN, C. WACHTEL, I. WEISFUSE, H. YU, AND A. NICOLL *(2014): "School closures during the 2009 influenza pandemic: national and local experiences," BMC Infectious Diseases, 14, 1–11.*

CAUCHEMEZ, S., A.-J. VALLERON, P.-Y. BOELLE, A. FLAHAULT, AND N. M. FERGUSON *(2008): "Estimating the impact of school closure on influenza transmission from Sentinel data," Nature, 452, 750–755.*

CHAN, T. Y., B. H. HAMILTON, AND N. W. PAPAGEORGE *(2015): "Health, Risky Behavior and the Value of Medical Innovation for Infectious Disease," Mimeo.*

COINDET, J.-P. *(1998): "Home-to-Work Trips During the Transportation Strikes in Ile-de-France at the End of 1995," Journal Of Transportation And Statistics, 43–51.*

CUTLER, D., A. DEATON, AND A. LLERAS-MUNEY *(2006): "The determinants of mortality." Journal of Economic Perspectives, 20, 97–120.*

DE PAULA, A., G. SHAPIRA, AND P. E. TODD *(2013): "How Beliefs about HIV Status Affect Risky Behaviors: Evidence from Malawi," Journal of Applied Econometrics.*

DONALDSON, D. *(2010): "Railroads of the Raj: Estimating the Impact of Transportation Infrastructure," American Economic Review (forthcoming).*

Dubois, E., P.-O. Beffy, F. Arnaud, and B. Ourliac (2007): "Note de Conjoncture," Tech. rep., INSEE.

Dupas, P. (2011): "Do Teenagers Respond to HIV Risk Information? Evidence from a Field Experiment in Kenya," American Economic Journal: Applied Economics, 3, 1–34.

Duranton, G. and M. A. Turner (2011): "The Fundamental Law of Road Congestion: Evidence from US Cities," The American Economic Review, 101, 2616–2652.

——— (2012): "Urban Growth and Transportation," Review of Economic Studies, 79, 1407–1440.

Epstein, J. M., D. M. Goedecke, F. Yu, R. J. Morris, and D. K. Wagener (2007): "Controlling Pandemic Flu: The Value of International Air Travel Restrictions," PLoS ONE, 2, e401.

Ettner, S. L. (1996): "New evidence on the relationship between income and health," Journal of Health Economics,, 15, 67–85.

Evans, W. N. and T. J. Moore (2012): "Liquidity, Economic Activity, and Mortality," Review of Economics and Statistics, 94, 400–418.

Ferguson, N. M., D. A. T. Cummings, C. Fraser, J. C. Cajka, P. C. Cooley, and D. S. Burke (2006): "Strategies for mitigating an influenza pandemic," Nature, 442, 448–452.

Fernald, J. G. (1999): "Roads to Prosperity? Assessing the Link between Public Capital and Productivity," The American Economic Review, 89, 619–638.

Flahault, A., T. Blanchon, D. Y, T. L, V. J. F, and V. A. J (2006): "Virtual surveillance of communicable diseases: a 20-year experience in France." Statistical Methods in Medical Research, 15, 413–421.

Fogli, A. and L. Veldkamp (2013): "Germs, Social Networks and Growth," .

Geoffard, P.-Y. and T. Philipson (1996): "Rational Epidemics and Their Public Control," International Economic Review, 37, 603–624.

Greenwood, J., P. Kircher, C. Santos, and M. Tertilt (2013): "An Equilibrium Model of the African HIV/AIDS Epidemic," .

Guris, D., A. Jumaan, L. Mascola, B. Watson, J. Zhang, S. Chaves, P. Gargiullo, D. Perella, R. Civen, and J. F. Seward (2008): "Changing varicella epidemiology in active surveillance sites - United States, 1995-2005." J Infect Dis, 197, S71–5.

Hufnagel, L., D. Brockmann, and T. Geisel (2004): "Forecast and control of epidemics in a globalized world," Science, 101, 15124–15129.

InVS (2006): "Bilan annuel du réseau," Tech. rep., Institut de Veille Sanitaire.

Kelly, E. (2011): "The Scourge of Asian Flu In utero Exposure to Pandemic Influenza and the Development of a Cohort of British Children," The Journal of Human Resources, 46, 669–694.

Kermack, W. O. and A. G. McKendrick (1927): "A contribution to the mathematical theory of epidemics," Proc. R. Soc. Lond A, 115, 700–721.

KREMER, M. (1996): "Integrating Behavioral Choice into Epidemiological Models of AIDS," *The Quarterly Journal of Economics*, 111, 549–573.

LAKDAWALLA, D., N. SOOD, AND D. GOLDMAN (2006): "HIV Break- throughs and Risky Sexual Behavior," *Quarterly Journal of Economics*, 121, 1063–1102.

LEGENDRE, M., J. BARTOLI, L. SHMAKOVA, S. JEUDY, K. LABADIE, A. ADRAIT, M. LESCOT, O. POIROT, L. BERTAUX, C. BRULEY, Y. COUTÉ, E. RIVKINA, C. ABERGEL, AND J.-M. CLAVERIE (2014): "Thirty-thousand-year-old distant relative of giant icosahedral DNA viruses with a pandoravirus morphology," *PNAS*.

LEMPEL, H., J. M. EPSTEIN, AND R. A. HAMMOND (2009): "Economic Cost and Health Care Workforce Effects of School Closures in the U.S." *Nature*, 5, RRN1051.

LOWEN, A. C., S. MUBAREKA, J. STEEL, AND P. PALESE (2007): "Influenza Virus Transmission Is Dependent on Relative Humidity and Temperature," *PLoS Pathog*, 3, e151.

MOE, K. AND J. A. SHIRLEY (1982): "The effects of relative humidity and temperature on the survival of human rotavirus in faeces," *Archives of Virology*, 72, 179–186.

MOLINARI, N.-A. M., I. R. ORTEGA-SANCHEZ, M. L. MESSONNIER, W. W. THOMPSON, P. M. WORTLEY, E. WEINTRAUB, AND C. B. BRIDGES (2007): "The annual impact of seasonal influenza in the US: Measuring disease burden and costs," *Vaccine*, 25, 5086–5096.

MORETTI, E. AND M. NEIDELL (2011): "Pollution, Health, and Avoidance Behavior Pollution, Health, and Avoidance Behavior Evidence from the Ports of Los Angeles," *Journal of Human Resources*, 46, 154–175.

MURPHY, K. M. AND R. TOPEL (2006): "The Value of Health and Longevity," *Journal of Political Economy*, 114, 871–904.

NEIDELL, M. (2009): "Information, Avoidance Behavior, and Health The Effect of Ozone on Asthma Hospitalizations," *The Journal of Human Resources*, 44, 450–478.

NICHOL, K. L. (2001): "Cost-Benefit Analysis of a Strategy to Vaccinate Healthy Working Adults Against Influenza," *JAMA Internal Medicine*, 161, 749–759.

OSTER, E. (2005): "Sexually Transmitted Infections, Sexual Behavior, and the HIV/AIDS Epidemic," *The Quarterly Journal of Economics*, 120, 467–515.

——— (2012): "Routes of Infection: Exports and HIV Incidence in Sub-Saharan Africa," *Journal of the European Economic Association*, 10, 1025–1058.

PHILIPSON, T. (2000): "Economic Epidemiology and Infectious Diseases," in *Handbook of Health Economics*, New York: A. Culyer and J. Newhouse, vol. Volume 1B, 1761–1797, elsevier science ed.

PROSSER, L. A., C. B. BRIDGES, T. M. UYEKI, V. L. HINRICHSEN, M. I. MELTZER, N.-A. M. MOLINARI, B. SCHWARTZ, W. W. THOMPSON, K. FUKUDA, , AND T. A. LIEU (2006): "Health Benefits, Risks, and Cost-Effectiveness of Influenza Vaccination of Children," *Emerging Infectious Diseases*, 12, 1548–1558.

RUHM, C. (2000): "Are Recessions Good for Your Health?" *Quarterly Journal of Economics, 115,* 617–650.

——— (2005): "Healthy Living in Hard Times," *Journal of Health Economics, 24, 341.*

RUHM, C. J. (2003): "Good Times Make You Sick," *Journal of Health Economics, 22, 637–658.*

SIMONSEN, L., M. J. CLARKE, L. B. SCHONBERGER, N. H. ARDEN, N. J. COX, AND K. FUKUDA (1998): "Pandemic versus Epidemic Influenza Mortality: A Pattern of Changing Age Distribution," *The Journal of Infectious Diseases, 178, 53–60.*

SUHRCKE, M., D. STUCKLER, J. E. SUK, M. DESAI, M. SENEK, M. MCKEE, S. TSOLOVA, S. BASU, I. ABUBAKAR, AND P. HUNTER (2011): "The Impact of Economic Crises on Communicable Disease Transmission and Control: A Systematic Review of the Evidence," *PLoS ONE, 6, e20724.*

TANK, K., J. B. WIJNGAARD, G. P. KONNEN, R. BOHM, G. DEMAREE, A. GOCHEVA, M. MILETA, S. PASHIARDIS, L. HEJKRLIK, C. KERN-HANSEN, R. HEINO, P. BESSEMOULIN, G. MULLER-WESTERMEIER, M. TZANAKOU, S. SZALAI, T. P. ALSDOTTIR, D. FITZGERALD, S. RUBIN, M. CAPALDO, M. MAUGERI, A. LEITASS, A. BUKANTIS, R. ABERFELD, A. F. V. V. ENGELEN, E. FORLAND, M. MIETUS, F. COELHO, C. MARES, V. RAZUVAEV, E. NIEPLOVA, T. CEGNAR, J. A. LOPEZ, B. DAHLSTROM, A. MOBERG, W. KIRCHHOFER, A. CEYLAN, O. PACHALIUK, L. V. ALEXANDER, AND P. PETROVIC (2002): "Daily dataset of 20th-century surface air temperature and precipitation series for the European Climate Assessment." *International Journal of Climatology, 22, 1441–1453.*

TAUBENBERGER, J. AND D. MORENS (2006): "1918 Influenza: the Mother of All Pandemics," *Emerging Infectious Diseases, 12.*

THORNTON, R. (2008): "The Demand for and Impact of Learning HIV Status: Evidence from a Field Experiment," *American Economic Review, 98, 1829–1863.*

VAN DEN BRANDHOF, W. E., G. A. DE WIT, M. A. S. DE WIT, AND Y. T. H. P. VAN DUYNHOVEN (2004): "Costs of gastroenteritis in The Netherlands," *Epidemioly and Infection, 132, 211–221.*

VISCUSI, W. K. AND J. E. ALDY (2003): "The Value of a Statistical Life: A Critical Review of Market Estimates throughout the World," *Journal of Risk and Uncertainty, XXVII, 5–76.*

TABLE I: Characteristics of viral diseases

	Influenza	Gastroenteritis	Chickenpox
Incubation time (days)	1-4	0-2	10-20
Start of contagious phase	-1	0	-2
Length of symptoms (days)	7-10	2-5	6-10
Acquired immunity	No	No	Yes
Vaccine exists	Yes	Yes	Yes
Date vaccine conceived	1940	2006	1988
Systematic vaccination	Partly	No	No
	Average yearly incidence (% of total population)		
Europe and USA	5-20%	20-40%	0.1-0.5%

Notes: Incidence rates are collected for the US and particular European countries, usually around 2008. For chickenpox, the range corresponds to the US (lower range) where immunization is widespread and England and France where no immunization exists. In these countries the incidence rate is almost equal to the percentage of newborns.

TABLE II: Average annual percentage increase in incidence

	Children	Adults	Elderly	Time period
Flu-like Illness	-0.3%	-2.1%	-5.1%***	1984-2010
	(2.0)	(1.5)	(1.3)	
Acute diarrhea	2.4%***	2.5%***	0.8%	1990-2010
	(0.9)	(0.7)	(1.0)	
Chickenpox	-0.9%	-	-	1990-2010
	(0.9)			

Notes: This table displays estimates of a regression of the log average yearly incidence of a given disease per 100,000 on a linear trend. Each entry is obtained from a separate regression. Standard errors corrected for heteroskedasticity are in parenthesis. ***, ** and * denotes significance at 1%, 5%, and 10% level.

TABLE III: Time (weeks) to reach the annual peak of infection

	Flu-like illness	Acute diarrhea
Trend (years)	-.07***	-.40***
	(0.01)	(0.03)
Adults	0.65***	-.08
	(0.1)	(0.15)
Elderly	0.8***	1.05***
	(0.16)	(0.33)
Const.	28.07***	38.74***
	(0.18)	(0.45)
Obs.	1,573	1,259

Notes. This table displays estimates of equation (III) in the text. Each column presents estimates from a separate regression. The dependent variable is the time to reach the seasonal peak of infection, in weeks, - defined as the date when the cumulative density of infected individuals within each season reaches the 80% threshold. The explanatory variables are a linear time trend (years), age group indicators as well as region indicators (not shown). Standard errors corrected for heteroskedasticity in parenthesis. ***, ** and * denotes significance at 1%, 5%, and 10% level.

TABLE IV: Propagation rates of viruses: OLS and IV results

| | Flu-like illness | | | Acute diarrhea | | | Chickenpox |
	Children	Adults	Elderly	Children	Adults	Elderly	Children
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Panel A: OLS estimates, within region transmission only							
Within region (α_{within})	0.280***	0.455***	0.120***	0.182***	0.260***	0.034***	0.182***
Panel B: OLS estimates, within and between region transmission							
Within region (α_{within})	0.243***	0.412***	0.103***	0.179***	0.250***	0.032***	0.182***
	(0.008)	(0.008)	(0.004)	(0.005)	(0.007)	(0.002)	(0.023)
Between region ($\alpha_{between}$)	0.005***	0.006***	0.003***	0.002**	0.006***	0.001***	0.001
	(0.001)	(0.001)	(0.000)	(0.001)	(0.001)	(0.000)	(0.003)
Panel C: IV estimates, within and between region transmission							
Within region (α_{within})	0.205**	0.290***	0.111***	0.275***	0.502***	0.059***	1.901***
Between region ($\alpha_{between}$)	0.012***	0.016***	0.005***	0.008**	0.001	0.003**	-0.001
Observations	25,005	25,005	25,005	18,753	18,753	18,753	19,038
Autocorrelation (AR1)	0.34	0.17	0.37	0.11	0.09	0.19	0.33
F test (endogeneity)	4.430	10.730	9.690	18.770	21.760	12.020	30.760
P-value	0.109	0.005	0.008	00	00	0.002	0.0

Notes. This table displays estimates of equations (4), and (5) in the text. Each column/panel presents estimates from a separate regression. The dependent variable is the incidence of a given disease in a particular age group. All regressions include region, year and week fixed effects. Standard errors are corrected for serial and spatial correlation, using a a Prais–Winsten regression, where a region specific AR(1) process is assumed. Regressions in Panel C use lagged temperature and precipitation in all regressions to instrument the endogenous effects within and between region. Lagged birth cohort size is also used as an instrument for the incidence of chickenpox. The F-tests in Panel C is a joint Hausman test for the endogeneity of the regressors. The first stage is displayed in Table V. ***, ** and * denotes significance at 1%, 5%, and 10% level.

41

TABLE V: First stage results: F-tests, p-values and R-square

		Children	Adults	Elderly
		(1)	(2)	(3)
Within region (dependent variable: $I_{rt-\tau}S_{rt-\tau}$)				
Flu-like illness	F-test	3.73	3.71	2.74
	p-val	(.008)	(.008)	(.03)
	R^2	0.272	0.270	0.25
Acute diarrhea	F-test	5.53	5.53	5.53
	p-val	(.001)	(.001)	(.001)
	R^2	0.55	0.55	0.55
Chickenpox	F-test	4.08		
	p-val	(.002)		
	R^2	0.32		
Across regions (dependent variable: $\sum_{c \in R\backslash r} I_{ct-\tau}S_{rt-\tau}$)				
Flu-like illness	F-test	612.30	558.04	501.93
	p-val	(0.0)	(0.0)	(0.0)
	R^2	0.417	0.416	0.40
Acute diarrhea	F-test	1284.12	1258.88	1309.37
	p-val	(0.0)	(0.0)	(0.0)
	R^2	0.8	0.8	0.8
Chickenpox	F-test	56.52		
	p-val	(0.0)		
	R^2	0.67		

Notes. The table reports the F test of joint significance of the instruments, with p-values in parenthesis as well as the R-square of the regression. The dependent variables are defined in equation (5). Each entry in the table corresponds to a separate regression. The instruments are lagged cumulative temperature and precipitations within a year (defined as May to May). For flu-like illnesses and acute diarrhea, the lag is one week. For chickenpox, the instruments are lagged three weeks, and include also the size of the birth cohorts at regional level, within the last six years. All regressions include region, year, and week fixed effects. Robust standard errors are clustered at region level.

TABLE VI: Propagation rates of viruses: fully interacted model

| | Flu-like illness | | | Acute diarrhea | | | Chickenpox |
	Children	Adults	Elderly	Children	Adults	Elderly	Children
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Panel A. Interpersonal contacts							
School holiday, within	-.145***	-.101***	-.006	-.071***	0.014	0.01***	-.197**
	(0.019)	(0.019)	(0.008)	(0.009)	(0.013)	(0.004)	(0.078)
School holiday, between	-.003***	-.0009	0.0001	0.003***	-.004***	-.0007***	0.005
	(0.0007)	(0.0008)	(0.0003)	(0.0007)	(0.0008)	(0.0003)	(0.004)
Transportation strike, within	-.002	-.056**	-.036***	0.038	0.016	0.017	-.060
	(0.026)	(0.024)	(0.011)	(0.026)	(0.039)	(0.012)	(0.154)
Transportation strike, between	-.008**	-.010***	-.004*	-.002	-.003	-.001	0.009
	(0.004)	(0.004)	(0.002)	(0.003)	(0.003)	(0.001)	(0.013)
High-speed train	0.005**	0.014***	0.007***	0.01***	0.006***	0.0008	0.001
	(0.003)	(0.002)	(0.001)	(0.002)	(0.002)	(0.0006)	(0.016)
Adverse weather, within	-.050***	-.035**	0.008	-.033***	-.006	-.0002	-.019
	(0.015)	(0.016)	(0.006)	(0.008)	(0.01)	(0.003)	(0.087)
Panel B: Trade							
Within region	0.078	-.092	0.008	-1.669	-1.381*	-.159	
	(0.088)	(0.075)	(0.036)	(4.315)	(0.714)	(0.81)	
Between regions	0.035***	0.049***	0.009**	0.018***	0.029***	0.002	0.026
	(0.008)	(0.007)	(0.004)	(0.006)	(0.007)	(0.002)	(0.029)
Panel C: Unemployment							
Within region	-.012**	-.008	0.002	0.005	-.014***	-.001	0.004
	(0.005)	(0.005)	(0.003)	(0.003)	(0.004)	(0.001)	(0.02)
Between regions	-.014***	0.006*	0.009***	0.01***	0.005	-.002**	0.018
	(0.003)	(0.003)	(0.002)	(0.003)	(0.004)	(0.001)	(0.02)

Table continued on next page.

43

TABLE VI: (CONTINUED)

| | Flu-like illness | | | Acute diarrhea | | | Chickenpox |
| | Children | Adults | Elderly | Children | Adults | Elderly | Children |
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Panel D. Within region characteristics							
Temperature (in C)	-.016***	-.027***	-.004***	-.005***	-.005***	-.0004	-.004
	(0.002)	(0.002)	(0.0009)	(0.0009)	(0.001)	(0.0003)	(0.006)
Population density	1.173***	0.701***	-.252**	0.407***	-.164	-.170***	4.440**
	(0.176)	(0.178)	(0.123)	(0.118)	(0.157)	(0.055)	(1.762)
Panel E. Between region characteristics							
Distance (1000km)	0.0008	0.005	0.004*	-.004	-.011***	-.005***	0.019
	(0.003)	(0.003)	(0.002)	(0.002)	(0.003)	(0.001)	(0.09)
Population ratio	0.002**	-.0002	0.0009**	0.005***	0.01***	0.002***	0.025***
	(0.001)	(0.001)	(0.0004)	(0.001)	(0.001)	(0.0004)	(0.006)
Log GDP ratio	-.009	0.018**	-.0007	-.0001	0.066***	0.018***	-.051
	(0.01)	(0.009)	(0.004)	(0.008)	(0.009)	(0.003)	(0.042)
Temperature Difference (in C)	0.0002*	0.0009***	0.0002***	0.0001	0.0006***	0.0002***	0.002**
	(0.0001)	(0.0001)	(0.00006)	(0.0001)	(0.0001)	(0.00003)	(0.0007)
Prevalence in Paris	0.046***	0.043***	0.011***	-.007**	-.003	0.003**	0.023
	(0.006)	(0.006)	(0.003)	(0.003)	(0.004)	(0.001)	(0.016)
Observations	24,407	24,407	24,407	18,537	18,537	18,537	17,602
R-squared	0.57	0.73	0.37	0.54	0.67	0.27	0.24

Notes. This table displays estimates of the parameters α_{within}^k and $\alpha_{between}^k$ defined in equation (6) in the text. Each column represents a separate regression with incidence rates as the dependent variables. Panels are organised for ease of exposition, but all the variables listed in the table are included in the regression. In addition, the regressions control for region, year and week fixed effects in levels and region fixed effects, quarterly dummies and a linear trend interacted with transmission rates. Lagged temperature and precipitation are used as instruments in all regressions to instrument the endogenous effects within and between region. Lagged birth cohort size is also used as an instrument for the incidence of chickenpox. Standard errors are corrected for serial and spatial correlation, using a a Prais-Winsten regression, where a region specific AR(1) process is assumed. ***, ** and * denotes significance at 1%, 5%, and 10% level.

44

TABLE VII: Costs of disease per case, influenza

Children	
GP visit (32% chance)	6.68 €
Otitis media (0.28% chance)	17.38 €
Pneumonia (12% chance)	16.45 €
Hospitalisation (0.07% chance)	2.45 €
Hospitalisation (sequelae pneumonia 0.7 per 100,000)	3.61 €
Loss of human capital (3 days off school, 5% return)	99 €
Parent stays home (50% of time, labor market particip. 0.65)	102 €
Value of statistical life	1.3-6 million €
Probability of death	0.7 per 100,000
Cost of death	9-42 €
Adults	
Absent from work (2 days of work at average wage)	78.90€
Reduced productivity (0.7 days at 50%)	13.80 €
GP visit (45% chance)	9.45 €
Hospitalisation (0.04% chance)	1.80 €
Value of statistical life	1.3-6 million €
Probability of death	4 per 100,000
Cost of death	52-240 €
Elderly	
Outpatient visit	219 €
Hospital	476 €
Value of statistical life	1.3-6 million €
Probability of death	102 per 100,000
Cost of death	1326-6120 €

Notes. Data on costs and healthcare use are taken from Prosser et al. (2006) for children, from Nichol (2001) for adults and from Molinari et al. (2007) for the elderly. These studies weight medical costs by the probability of health care usage. Data on mortality from influenza by age group comes from the National Vital Statistics Report 2011. Data on wages are taken from INSEE, "Revenus salariaux médians des salariés de 25 à 55 ans selon le sexe en 2011" (http://www.insee.fr/fr/themes/tableau.asp?reg_id=0&ref_id=NATnon04146). Labor market participation data comes from OECD skill data set. All US dollars converted into euros with an exchange rate of 0.8. Loss of human capital is costed using a return to schooling of 5 percent, median wages by sex and average labor market participation by sex over a period of 42 years. Net present value numbers are displayed, calculated with a discount factor equal to 0.95.

TABLE VIII: Costs of disease per case, acute diarrhea

Children	
Medication, GP visits	3.16 €
Hospitalisation	5.4 €
Loss of human capital (2 days off school, 5% return)	66 €
Absence of other person for informal care	57 €
Value of statistical life	1.3-6 million €
Probability of death	0.4 per 100,000
Cost of death	5.2-24 €
Adults	
Work absence	105.8 €
Medication, GP visits	2.61 €
Hospitalisation	5.15 €
Value of statistical life	1.3-6 million €
Probability of death	0.5 per 100,000
Cost of death	6-30 €
Elderly	
Medication, GP visits	16.3 €
Hospitalisation	83 €
Value of statistical life	1.3-6 million €
Probability of death	22 per 100,000
Cost of death	286-1320 €

Notes. Data on costs and healthcare use are taken from Van Den Brandhof et al. (2004). Medical costs are weighted by the probability of health care usage. Data on mortality from acute diarrhea by age group comes from the National Vital Statistics Report 2011. Data on wages are taken from INSEE, "Revenus salariaux médians des salariés de 25 à 55 ans selon le sexe en 2011" (http://www.insee.fr/fr/themes/tableau.asp?reg_id=0&ref_id=NATnon04146). Labor market participation data comes from OECD skill data set. All costs are adjusted for inflation and in 2005 euros. US dollars converted into euros with an exchange rate of 0.8. Loss of human capital is costed using a return to schooling of 5 percent, median wages by sex and average labor market participation by sex over a period of 42 years.Net present value numbers are displayed, calculated with a discount factor equal to 0.95.

TABLE IX: Cost of policy, euros per capita

School closure	
10 days off school, return to schooling, 5% per year	350 €
Parent stays home (50% of time, labor market particip. 0.65)	150 €
Public transportation closure	
€400 million per week	19 €

Notes. Data on wages are taken from INSEE, "Revenus salariaux médians des salariés de 25 à 55 ans selon le sexe en 2011" (http://www.insee.fr/fr/themes/tableau.asp?reg_id=0&ref_id=NATnon04146). Labor market participation data comes from OECD skill data set. Cost of public transportation closure comes from Dubois et al. (2007).

Flu–like Illnesses

Acute Diarrhea

Chickenpox
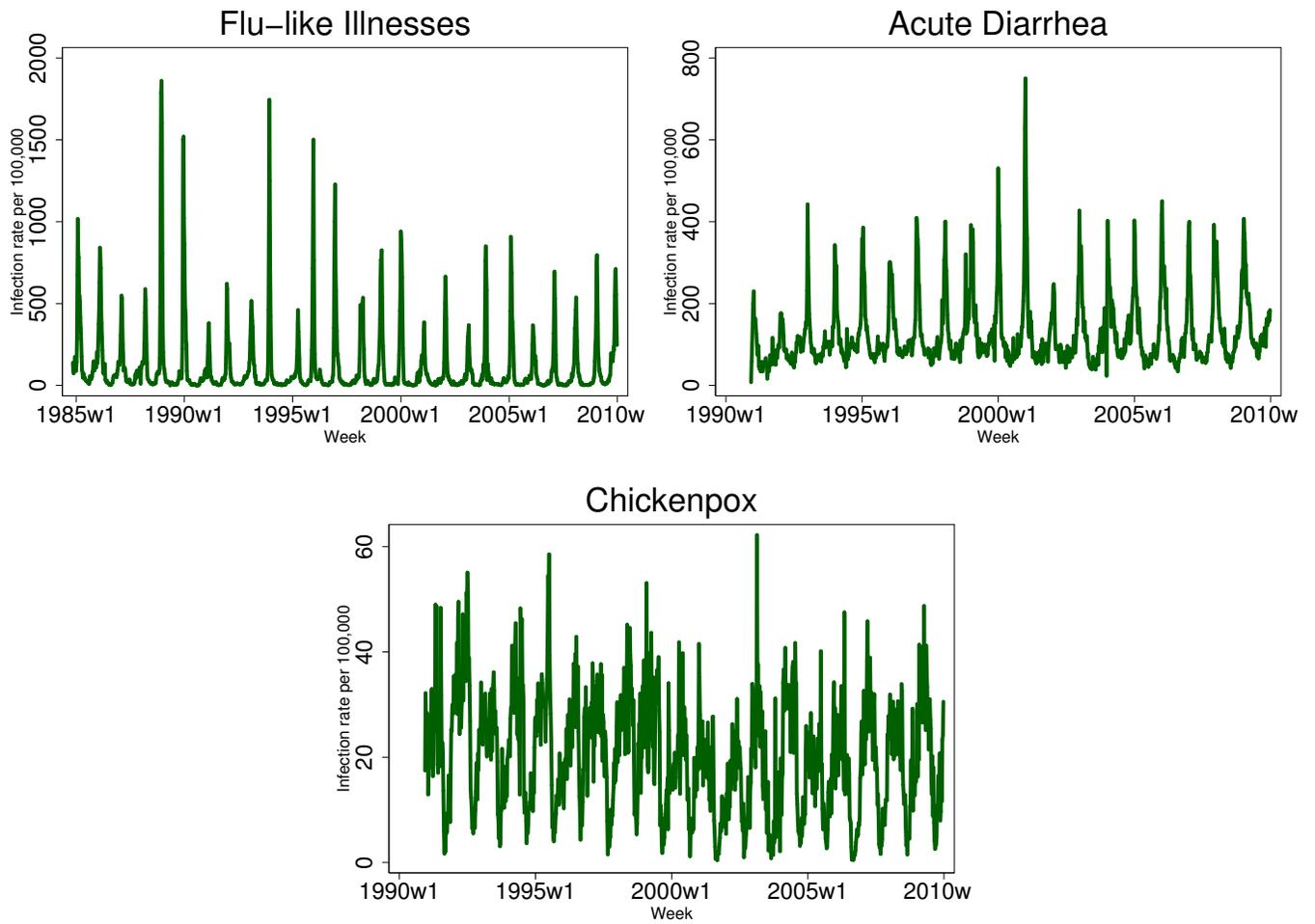
Notes. Data from InVS. Infection rates per 100,000 individuals.

FIGURE II: Prevalence rates over calendar year, by age. Average 1984-2010

Flu–Like Illnesses

Infection rate per 100,000

Children — Adults
Elderly

Acute Diarrhea

Infection rate per 100,000

Children — Adults
Elderly

Chickenpox

Infection rate per 100,000

Children — Adults
Elderly
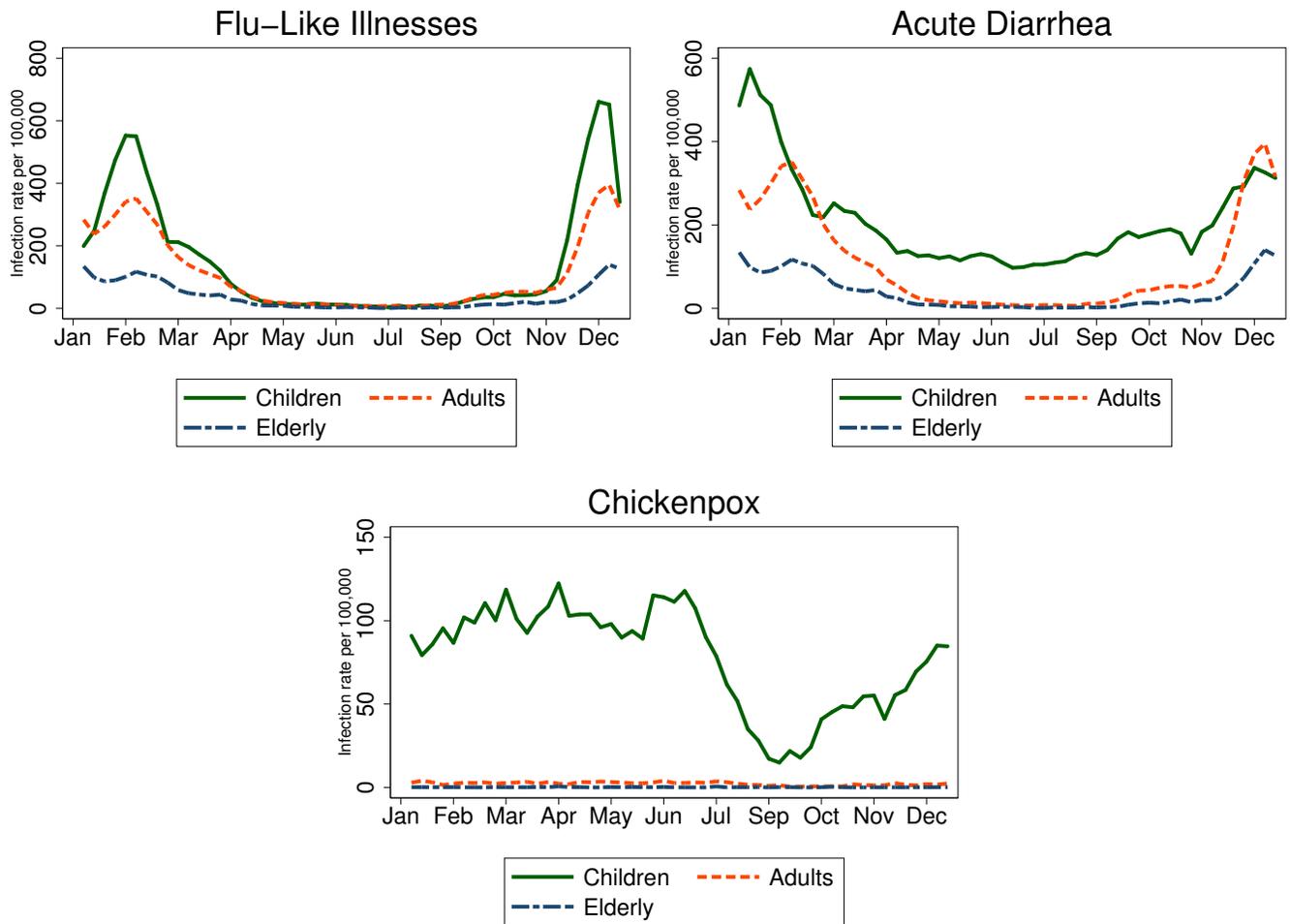
Notes. Data from InVS. Weekly infection rates per 100,000 individuals.

49

FIGURE III: Public Transportation Strikes, France, 1984-2010

(a) Frequency within Year

(b) Frequency by Duration

Average Duration: 12.3 days; Median duration 8 days.
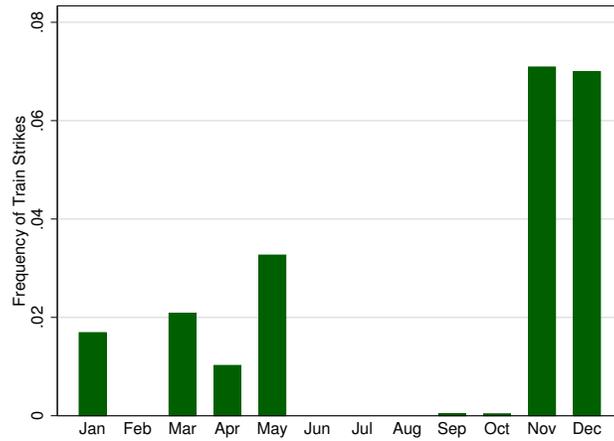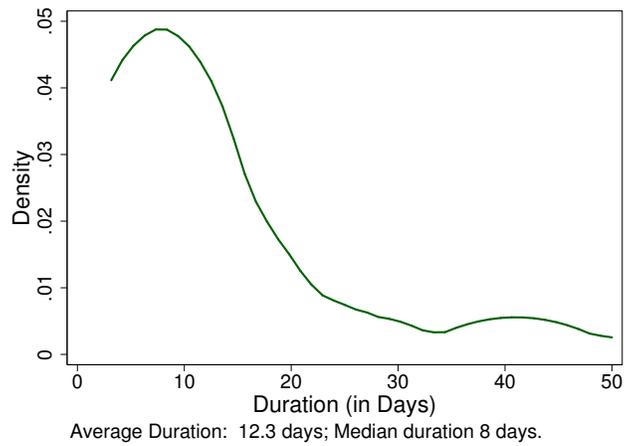
Notes. Data collected by the author using the LexisNexis search of the French written press. All occurrences of a public transportation strike lasting three days at least, either nationally or regionally were recorded from the national press (Le Monde, Le Figaro).

FIGURE IV: Expansion of the high speed railway
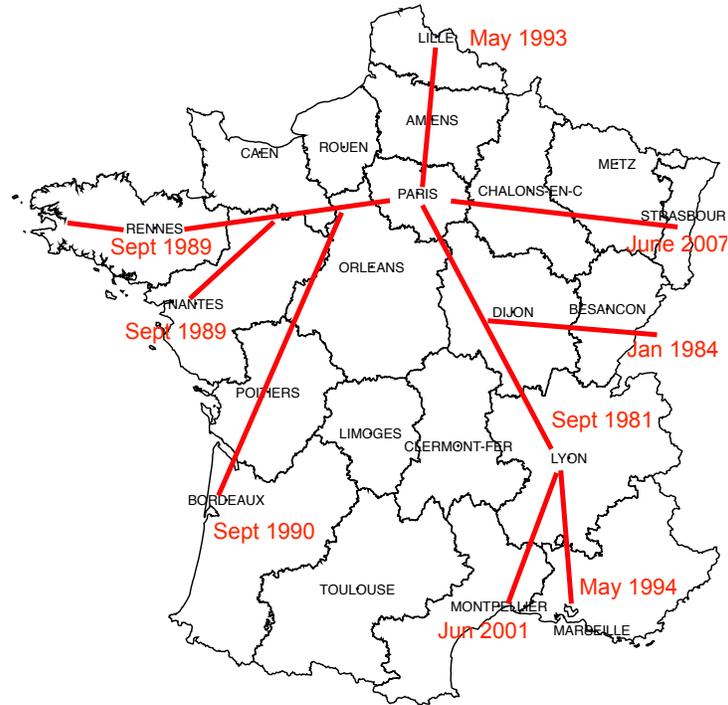
Notes. The map displays the major connections of the high-speed rail link in France and the date of opening. To avoid cluttering, only direct links are displayed. The analysis also takes into account opening of links between two cities where a stop in Paris was no longer necessary, e.g. Marseille - Lille without change of trains in Paris.

FIGURE V: Event Analysis: School Closures
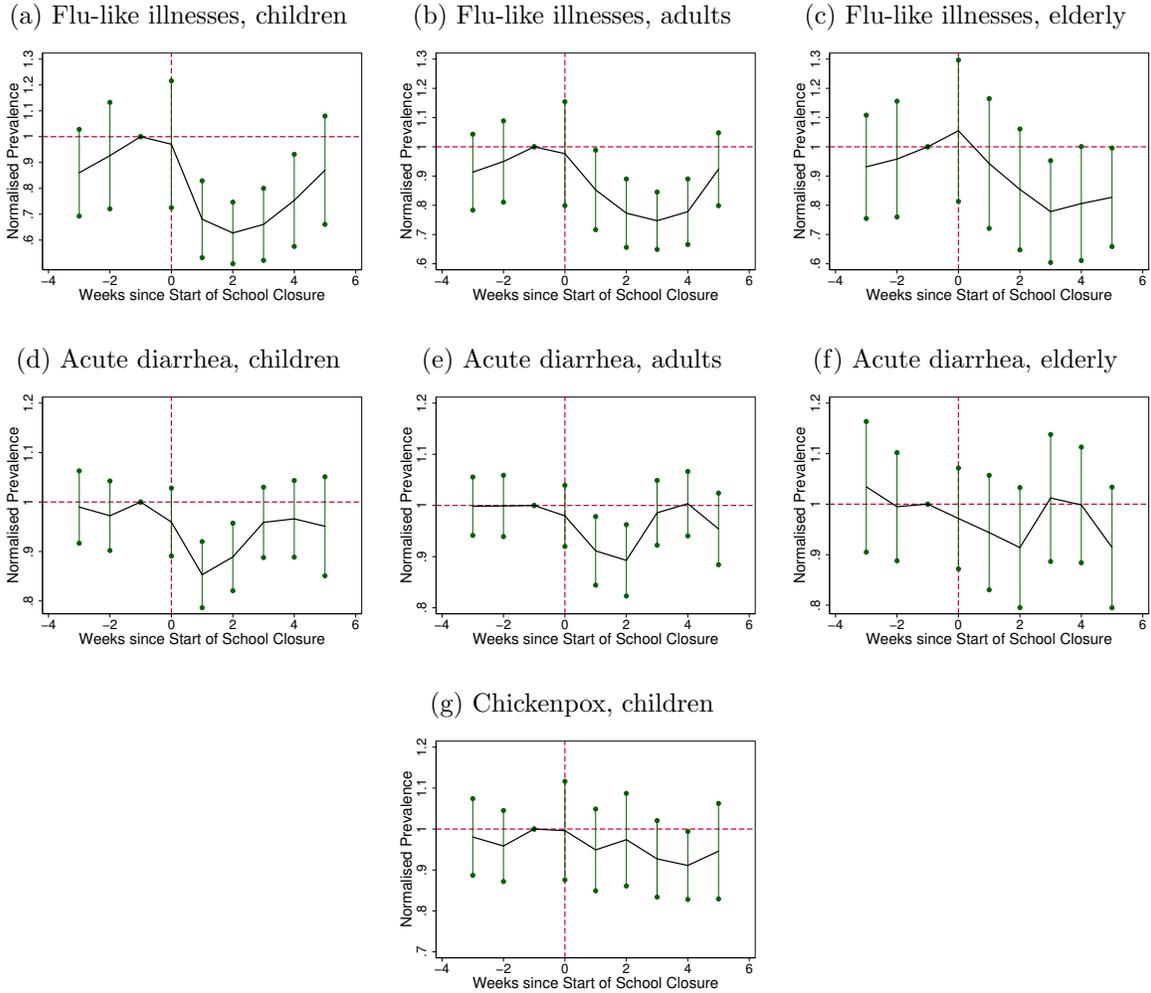
(a) Flu-like illnesses, children

(b) Flu-like illnesses, adults

(c) Flu-like illnesses, elderly

(d) Acute diarrhea, children

(e) Acute diarrhea, adults

(f) Acute diarrhea, elderly

(g) Chickenpox, children

The different panels display the estimates of equation (2) and show the weekly incidence of the disease within a region as a function of the time since the first week of school holidays. Results obtained from a regression controlling also for region, year and week effects. Standard errors are clustered by region. The incidence is normalised by the average incidence one week before the strike (labelled as -1). School closures last typically for one or two weeks, except for the 8 week summer break. Solid line is the mean effect by week, 95 percentiles are shown as vertical bars.

52

FIGURE VI: Event Analysis: Public Transportation Strike
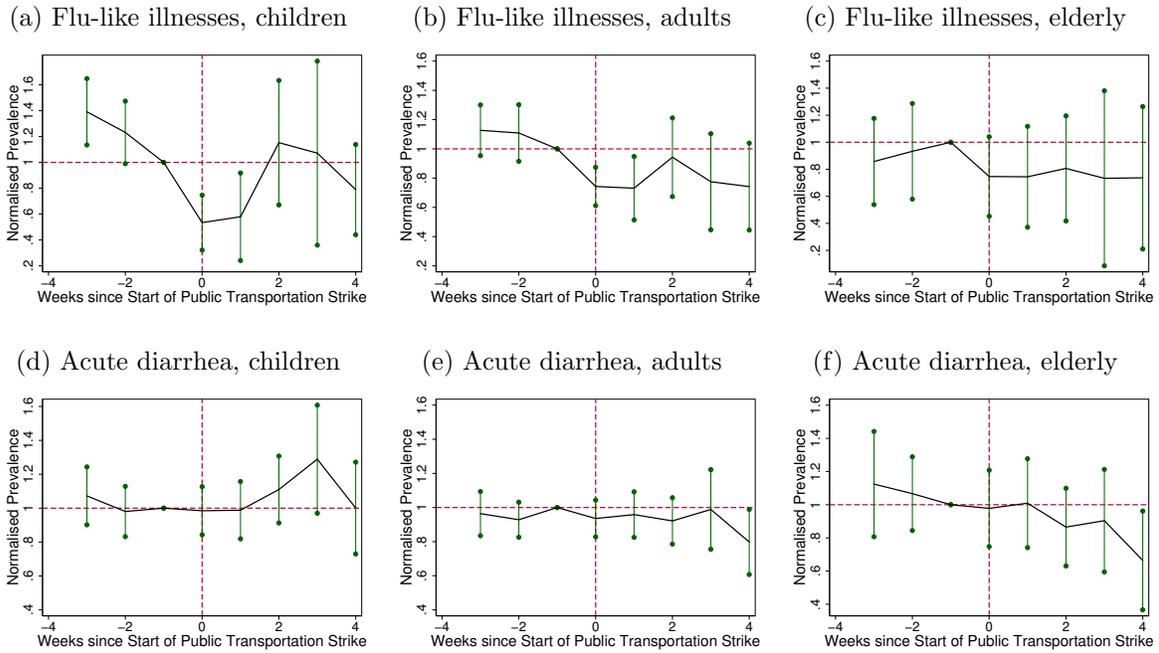
(a) Flu-like illnesses, children   (b) Flu-like illnesses, adults   (c) Flu-like illnesses, elderly

(d) Acute diarrhea, children   (e) Acute diarrhea, adults   (f) Acute diarrhea, elderly

The different panels display the estimates of equation (2) and show the weekly incidence of the disease within a region as a function of the time since the first week of a public transportation strike. Results obtained from a regression controlling also for region, year and week effects. Standard errors are clustered by region. The incidence is normalised by the average incidence one week before closure (labelled as -1). Solid line is the mean effect by week, 95 percentiles are shown as vertical bars. Chickenpox is omitted as transportation strikes do not have statistically significant effects.

FIGURE VII: Effect of policy measures on annual disease prevalence, by week of imple-
mentation

(a) Flu-like illness and school closure

(b) Acute diarrhea and school closure

(c) Flu-like illness and public
transportation closure

(d) Acute diarrhea and public trans-
portation closure
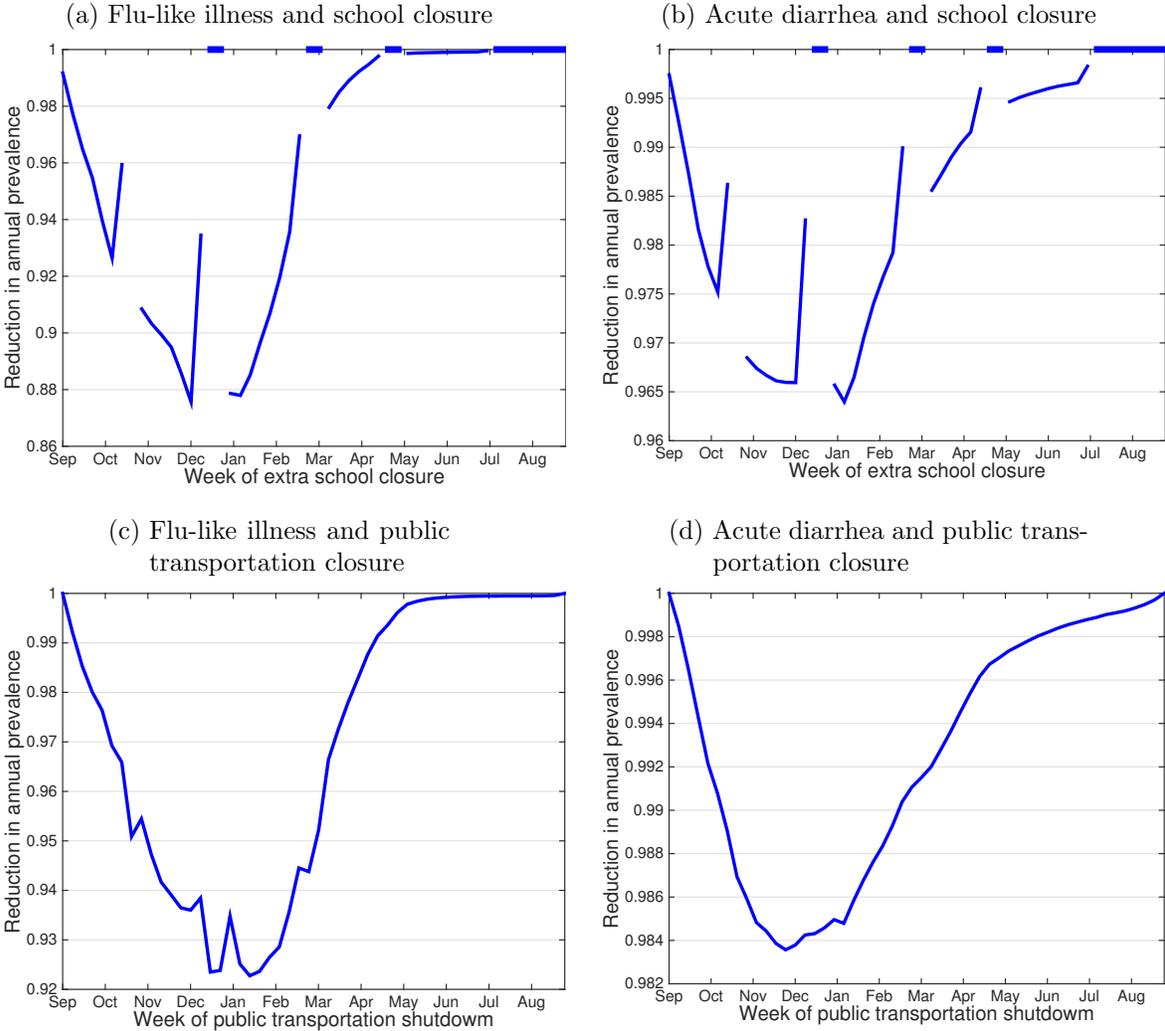
Notes. The different panels display the annual prevalence of diseases as a function of the week
of implementation of the policy. The prevalence are in deviation from a baseline scenario, which
includes school closures due to regular holidays. School closures lasts for two weeks and public
transportation closures for one week. The discontinuous lines in Panels (a) and (b) are due to
periods of holidays, when schools are closed anyway. The results are based on estimated coefficients
of equation (6) and shown in Table VI. The results were averaged over 500 replications, drawing
realisations for the shock in equation (6) and for the estimated coefficients from their asymptotic
distribution.

FIGURE VIII: Overall gain of policy measures closures, by week of implementation

(a) School closure   (b) Public transportation closure
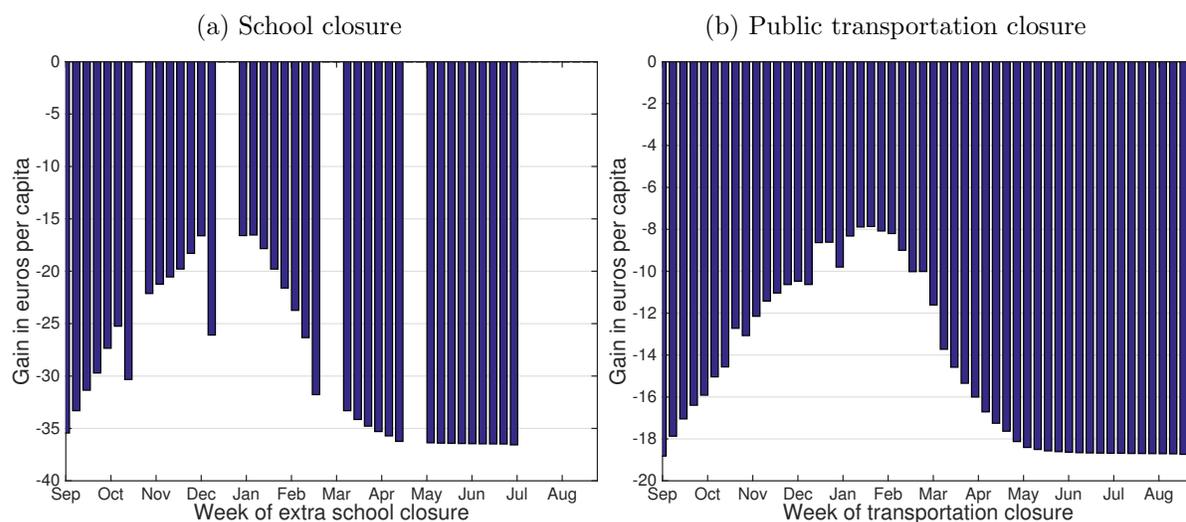
Notes. The different panels display the gain in per capita euros of the policy compared to a baseline scenario, which includes school closures due to regular holidays. The calculations use a value of statistical life equal to 5 million euros regardless of age. The results are based on estimated coefficients of equation (6) and cost measures displayed in Tables VII, VIII and IX. The results were averaged over 500 replications, drawing realisations for the shock in equation (6) and for the estimated coefficients from their asymptotic distribution.

Appendix

A   Supplementary Data Description

Data on viral diseases   The data on viral diseases is collected by the Institut de Veille Sanitaire (InVS) on a continuous basis from 1984. It relies on information sent by a network of GPs across the whole country. The network represents about 1-2 percent of all French GPs over the period we analyze. These doctors are representative of the population of active GPs. Figure I depicts the geographical dispersion of these GPs in 2006. Each dot indicates the localisation of a particular GP. The network is more dense in the Paris area, as well as along the Mediterranean coast, which corresponds to areas with higher population density.
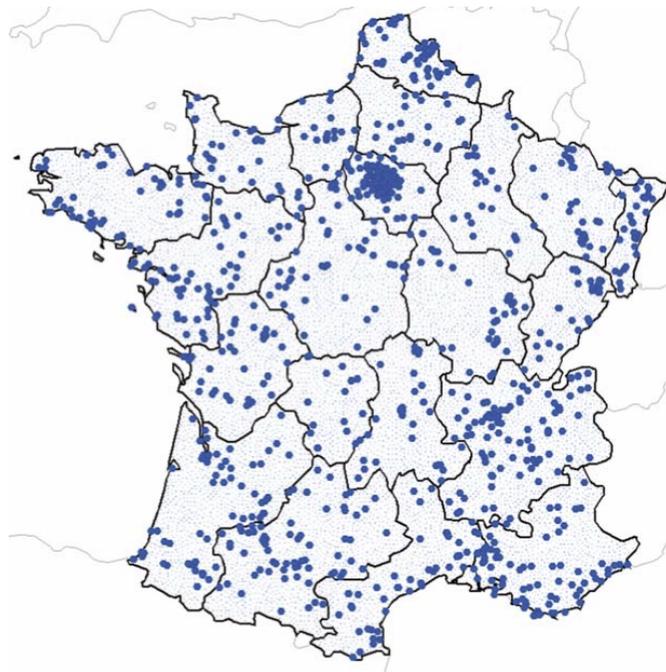
FIGURE I: Density of General Practitioners belonging to the Reseau Sentinelles, 2006
Source: InVS (2006)

The doctors report all cases of particular diseases, on a daily to weekly basis through a web/electronic interface. Prior to the existence of the internet and computers, doctors were equipped with an electronic terminal (Minitel) allowing them to send data. The data are then compiled by the InVS to produce incidence rates by region and

by age groups, using additional data on the number of general practitioners and the size of demographic groups by area derived from the French census. The French censuses took place in 1982, 1990 and 1999 in an exhaustive way. Since 2004, the census is updated every year, based on a random survey. The incidence rates are calculated, assuming that conditional on observables, the doctors are randomly selected. The data and information on how the data is constructed is available from the InVS (http://www.invs.sante.fr) and is described in more detail in Flahault et al. (2006).

Weather Data: We obtained daily temperatures from the European Climate Assessment and Dataset (Tank et al. (2002)) for the period 1984-2010. We aggregate the temperature at a weekly level to make it consistent with the data on disease incidences. The data was obtained for major cities in France and we allocate each city to a region. In the case where more than one reading was obtained for a region, we took the data for the most populous city. In case there is no reading for a particular region, we linearly interpolated the data between the two nearest stations.

Population Data: The data were extracted from the French census conducted by INSEE for each region and linearly interpolated for the missing years. Data on cohort size was gathered from INSEE, for each region and year between 1975 and 2010.

Trade data: The data come from the French Comissariat Général au Développement Durable and describe the intra and inter-regional transportation of goods, expressed in metric tons per year. For each region, the data report the quantities transported within the region and to all other regions by road, rail or boat (a number of regions are connected by rivers or canals). The data span the period 1975 to 2011, except for rail transportation which only goes to 2006. To get consistent measures across years, we use the sum of quantities transported by road and boat and discard the information on rail transport. Rail transportation of goods represents about 20% of all goods transported. Nevertheless, the measure we use is highly correlated with the total transportation measure for the years 1975-2006, where we have the data for all means of transportation. The overall correlation is equal to 0.997 and the correlations within origin-destination region pairs are also very close to one (the median correlation among the 21*21 measures is 0.98).

B   COMPUTATION OF THE SUSCEPTIBLE POPULATION IN THE CASE OF CHICKENPOX

We denote nb_{rt} the size of the cohort born in year t in region r. We abstract from infant mortality as well as regional migration, so that the cohort does not vary (substantially and systematically) in size over the first J years. We denote by ρ_{rt} the fraction of children who contract chickenpox in year t and region r out of the pool of children at risk. The susceptible population consist of the sum of all children for each birth cohort that have not yet been infected, starting with the newborns, up to the children of age J.[28]

$$S_{rt} = nb_{rt} + (1 - \rho_{rt-\tau})nb_{rt-\tau} + \ldots + \prod_{c=1}^{J}(1 - \rho_{rt-c})nb_{rt-J}. \tag{1}$$

We assume that at age $J+1$, all children have been exposed to the disease. The number of new cases in year t and region r is:

$$I_{rt} = \rho_{rt}S_{rt}.$$

We compute the size of the susceptible population using the history of disease incidence as well as data on the size of the birth cohorts, by region and year. The computation proceeds iteratively. Note that we can write the propensity to become infected as:

$$\rho_{rt} = \frac{I_{rt}}{nb_{rt} + (1 - \rho_{rt-\tau})nb_{rt-\tau} + \ldots + \prod_{c=1}^{J}(1 - \rho_{rt-c})nb_{rt-J}} \tag{2}$$

We first focus on the stationary equilibrium which is used as an initial guess in the iterative procedure. With a stable population and no variations in the infectiousness of the disease, one can solve for the stationary solution ρ as a function of the number of cases and the size of the birth cohort:

$$\bar{\rho}_r = 1 - \left(1 - \frac{\bar{I}_r}{\bar{nb}_r}\right)^{1/J+1}$$

---

[28]In practice, we set this parameter to 16 for two reasons. First, the disease is predominant in young children, with a median age of about 6 years old (Guris et al. (2008)) and the incidence of diseases is extremely small for an adult population (see Figure II). Second, the data on the size of the birth cohorts starts in 1975, and setting $J = 16$ allows us to calculate the population of susceptible individuals from the start of our data set describing the incidence in 1991.

This value can be used to compute a first guess of the susceptible population using equation (1), combining data on incidence and size of various cohorts, for all periods in the data set. At the nth iteration, the set of propensities to become infected verifies:

$$\rho_{rt}^n = \frac{I_{rt}}{S_{rt}(\rho_{rt}^{n-1})}$$

We repeat this procedure until convergence, which, practically, occurs very quickly. We then check that $I_{rt}(\rho_{rt}^n)$ is indeed close to the observed incidence of the disease in that region and time period. We are able to match these data very closely, with a correlation equal to 0.993.

C MEASUREMENT ERROR

We show in this section how measurement error affects the estimation of equation (5). Denote the observed incidence rate as: $\tilde{I}_{rt} = I_{rt} + \varepsilon_{rt}$, where ε_{rt} is the error, arising from the fact the incidence of diseases are estimated from the number of cases seen by a network of general practitioners, which is subject to sampling error. This error may not be classical. It is possible that the incidence rate is more likely measured with error when the incidence is very high, and patients have to queue to see a doctor. The measurement error on the incidence rate also affects the measurement of the proportion of susceptible individuals. Taking into account the measurement error, we can rewrite equation (5) expressed with observed variables as

$$\tilde{I}_{rt} = \alpha_{within}\tilde{I}_{rt-\tau}\tilde{S}_{rt-\tau} + \alpha_{between}\sum_{c\in R\backslash r}\tilde{I}_{ct-\tau}\tilde{S}_{rt-\tau} + X_{rt}\delta + \tilde{\eta}_{rt} \tag{3}$$

The error term depends on the disease into consideration as it implies different definitions of the susceptible population. In the simple case of gastroenteritis, where we define the susceptible population as $S_{rt} = N_{rt} - I_{rt-\tau}$, i.e. the total population minus those who were sick the week before, the error term is expressed as:

$$\begin{aligned}\tilde{\eta}_{rt} &= \eta_{rt} + \epsilon_{rt} - \alpha_{within}\tilde{S}_{rt-\tau}\epsilon_{rt-\tau} + \alpha_{within}\tilde{I}_{rt-\tau}\varepsilon_{rt-2} - \alpha_{within}\varepsilon_{rt-\tau}\varepsilon_{rt-2} \\ &\quad + \alpha_{between}\left(\varepsilon_{rt-2}\sum_{c\in R\backslash r}\left(\tilde{I}_{ct-\tau} - \varepsilon_{ct-\tau}\right) - \tilde{S}_{rt-\tau}\sum_{c\in R\backslash r}\varepsilon_{ct-\tau}\right)\end{aligned} \tag{4}$$

In the case of influenza and chickenpox, the error term is more complex as the susceptible population consists of an estimate of all those who have been infected during

the flu season or in the preceding years. For these two diseases, the error term contains measurement error terms over a longer period than gastroenteritis. In all cases, measurement error implies a complex structure for the error term, which exhibit serial correlation and heteroskedasticity. Importantly, given this structure, we cannot exclude the possibility that $cov(\tilde{I}_{rt-1}\tilde{S}_{rt-1}, \tilde{\eta}_{rt}) \neq 0$, which leads to biased inference of the parameters of interest.

D ROBUSTNESS OF RESULTS AND FURTHER EVIDENCE

In this section, we explore additional departures from our model (6). We investigate heterogeneous effects with respect to the lagged incidence of the disease and we test for interactions between diseases.

DA. Non-linear effects

We first look at non-linear effects of lagged incidence rate by looking at various quantiles. The standard epidemiological model interacts the lagged incidence (I_{rt-1} in our notations in equation (6)) with the size of the susceptible population (S_{rt-1}). There is a priori no reason for this particular matching function to be expressed as the product of these two variables, in a geometric way. We explore whether transmission rates vary with the level of the lagged incidence rate, which is equivalent to postulating that the matching function can be expressed as $f(I_{rt-1})S_{rt-1}$, which generalises our model. Another reason for a non-linear relationship can be due to behavioral adjustment. Susceptible individuals may observe incidence rates and avoid getting into contact with potentially sick individuals. In this case the function $f()$ is decreasing in the incidence rate.

We allow the parameter (or matrix of parameters) α to vary, and we discretise the lagged incidence into 6 intervals, denoted $[I^q, I^{q+1}]$. We denote $1_{I_{rt-\tau}\in[I^q,I^{q+1}]}$ an indicator variable equal to one when the incidence in region r and period $t-\tau$ is within the corresponding interval.

$$I_{rt} = \sum_{q=1}^{6} 1_{I_{rt-\tau}\in[I^q,I^{q+1}]} \left[\alpha_{within}^q I_{rt-\tau}S_{rt-\tau} + \alpha_{between}^q \sum_{c\in R\backslash r} I_{ct-\tau}S_{rt-\tau} \right] + X_{rt}\delta + \eta_{rt} \quad (5)$$

We are particularly interested in episodes with high incidence, which are more relevant for policy. We display in the appendix Table II, Panel A the intercept for the transmission within a region. We find some evidence that the transmission rate is higher, the higher the lagged incidence rate of the disease, especially for increasing

ages. For children, however, the confidence intervals tend to overlap, suggesting that the baseline model fits the data well. We find similar patterns for transmissions across regions, displayed in Table II, Panel B. These results do not offer much support for behavioral adjustment, where the transmission rate would be lower the higher the past incidence. However, to fully investigate this aspect, more precise data on actual behavior would be required.

DB. Interactions across diseases

We estimate a model where we allow for the incidence rate of acute diarrhea to influence the effect on how flu-like illnesses are transmitted and vice-versa. There can be two opposite effects, an incapacitation one, which limits the propagation of the other disease as sick individuals tend to stay home, and a weakening effect that enhances the effect of a given disease. The results are displayed in the appendix Table III and the results are to be interpreted as the effect of a standard deviation increase in the incidence of a disease on the transmission rate of the other disease. Sadly, the weakening effect appears to dominate in the case of flu-like illnesses which propagate faster when more individuals suffer from acute diarrhea, at least for adults and the elderly. We also find similar effects for all age groups in the case of transmission of acute diarrhea during flu-like illnesses outbreak. No such effect can be seen in the transmission rate of chickenpox. We find therefore similar interactions effects across diseases as in Oster (2005), although the channel would be quite different. From a public health point of view, this is bad news as one epidemic tend to reinforce the other.

DC. Cost of policy measures

We provide further detail on the cost of policy measures studied in Section V.. The appendix Table IV assess the robustness of the results to the value of statistical life. We use values ranging from 1.3 to 6 million euros, which correspond to the range found in the literature. Our results are broadly insensitive to the different values, except for very low values for flu-like illnesses in the case of public transportation closures.

TABLE I: First Stage Results, Within Region Variables

| | Flu-like illness | | | Acute diarrhea | | | Chickenpox |
	Children	Adults	Elderly	Children	Adults	Elderly	Children
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Weeks with temp < 0C	223.665***	210.473***	109.109***	83.327	83.488	83.609	-11.534
	(55.083)	(50.863)	(28.418)	(53.997)	(54.019)	(54.105)	(11.208)
Weeks with temp < 5C	-37.777*	-32.975*	-4.659	-31.785	-31.890	-32.008	-1.788
	(20.246)	(19.047)	(13.193)	(20.269)	(20.297)	(20.347)	(3.177)
Weeks with temp < 10C	99.446**	91.200**	30.488	40.847**	41.079**	41.192**	-1.370
	(41.980)	(39.042)	(21.844)	(18.572)	(18.579)	(18.642)	(4.074)
Weeks with temp > 15C	-14.431	-19.766	-28.083	89.177**	89.433**	89.712**	5.454
	(53.001)	(50.294)	(24.962)	(31.924)	(31.956)	(32.014)	(4.512)
Weeks with temp > 20C	-39.658	-34.258	5.471	4.560	4.368	4.342	2.100
	(29.938)	(29.087)	(23.519)	(48.104)	(48.190)	(48.285)	(3.370)
Weeks with temp > 25C	-75.856	-74.756	39.482	120.152	120.399	120.949	6.290
	(132.059)	(128.372)	(93.359)	(129.444)	(129.729)	(130.099)	(11.957)
Rainfall in last 4 weeks	-157.950*	-142.275**	-65.663*	-74.804***	-74.726***	-75.053***	3.972*
	(72.614)	(66.801)	(31.548)	(24.878)	(24.917)	(24.967)	(2.175)
Rainfall in last 8 weeks	-57.777	-63.217	-47.487	133.778**	133.637**	134.092**	-7.968*
	(63.741)	(58.981)	(32.820)	(51.331)	(51.439)	(51.594)	(3.856)
Newborns in region, year t							-0.000
							(0.004)
Newborns in region, year t-1							0.001
							(0.004)
Newborns in region, year t-2							0.001
							(0.004)
Newborns in region, year t-3							-0.009
							(0.016)
Newborns in region, year t-4							0.021
							(0.017)
Newborns in region, year t-5							-0.034***
							(0.011)
Newborns in region, year t-6							0.040***
							(0.008)
Observations	25710	25710	25710	19146	19146	19146	19770
R-squared	0.272	0.270	0.251	0.547	0.547	0.547	0.324

Table continued on next page.

62

TABLE I: (CONTINUED), Between Region Variables

	Flu-like illness			Acute diarrhea			Chickenpox
	Children	Adults	Elderly	Children	Adults	Elderly	Children
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Regions with temp< 0	5197.422***	4861.747***	3035.855***	644.641***	641.847***	643.693***	21.508***
	(92.207)	(89.323)	(51.171)	(33.036)	(32.647)	(32.310)	(2.705)
Regions with temp< 10	717.391***	675.270***	404.442***	363.595***	365.400***	365.959***	0.202
	(31.895)	(30.018)	(15.502)	(12.273)	(12.284)	(12.316)	(2.061)
Regions with temp< 20	138.320***	144.402***	277.030***	-3.917	-4.133	-4.350	-67.470***
	(27.593)	(25.164)	(14.319)	(16.073)	(16.114)	(16.133)	(4.123)
Weeks temperature < 0C	2898.468***	2752.972***	1183.457**	149.746	152.689	154.802	-20.295
	(850.651)	(789.656)	(450.970)	(152.895)	(153.027)	(153.441)	(42.014)
Weeks temperature < 5C	93.986	113.636	357.999*	-112.485	-113.314	-114.792	24.841
	(340.467)	(308.929)	(193.777)	(116.153)	(116.183)	(116.162)	(20.766)
Weeks temperature < 10C	324.209	277.901	-211.724	-304.009**	-301.942**	-301.735**	51.100*
	(346.128)	(325.039)	(181.006)	(135.142)	(134.904)	(135.025)	(25.473)
Weeks temperature > 15C	-768.818*	-806.184*	-656.123***	1108.189***	1109.753***	1113.363***	96.934***
	(403.075)	(399.076)	(227.492)	(147.357)	(146.793)	(146.769)	(22.586)
Weeks temperature > 20C	-293.220	-223.112	365.054	814.980***	814.545***	814.854***	68.499**
	(409.859)	(393.466)	(232.349)	(126.865)	(127.356)	(127.518)	(29.447)
Weeks temperature > 25C	-3771.850***	-3608.756***	-1493.348***	-478.493	-478.492	-474.643	-49.846
	(659.066)	(631.153)	(367.159)	(353.371)	(352.091)	(353.393)	(41.877)
Rainfall in last 4 weeks	-1482.096*	-1286.211*	-398.239	-608.933***	-607.601***	-611.145***	43.341***
	(717.245)	(668.429)	(348.319)	(149.681)	(150.343)	(150.478)	(10.399)
Rainfall in last 8 weeks	-2216.681***	-2218.766***	-1358.273***	1312.478***	1311.016***	1316.546***	-29.057
	(696.852)	(676.010)	(375.703)	(433.502)	(433.653)	(434.552)	(30.646)
Newborns in region, year t							-0.003
							(0.013)
Newborns in region, year t-1							0.012*
							(0.006)
Newborns in region, year t-2							0.042***
							(0.009)
Newborns in region, year t-3							-0.125***
							(0.032)
Newborns in region, year t-4							0.104**
							(0.043)
Newborns in region, year t-5							0.004
							(0.022)
Newborns in region, year t-6							-0.088***
							(0.023)
Observations	25710	25710	25710	19146	19146	19146	19811
R-squared	0.417	0.416	0.397	0.796	0.797	0.797	0.668

Notes. This table displays estimates of the first stage regressions for the within and the between region variables, as in equation (5). Each column represents a separate regression with incidence times the number of susceptible individuals as the dependent variables. The instruments relating to weather are lagged, with a lag of one week for flu-like illnesses and acute diarrhea and three weeks in the case of chickenpox. All regressions include region, year and week fixed effects. Standard errors are clustered at regional level. ***, **, and * denotes significance at 1%, 5%, and 10% level.

TABLE II: Non-linear effects of lagged incidence, by quantiles of incidence, within region

	Flu-like illness			Acute diarrhea			Chickenpox
	Children	Adults	Elderly	Children	Adults	Elderly	Children
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Panel A: within region effects (α_{within}^q)							
Quantile [0-25]	0.3	0.186	0.025	0.252***	0.31***	0.033	1.584***
	(0.184)	(0.19)	(0.098)	(0.055)	(0.068)	(0.022)	(0.371)
Quantile [25-50]	0.303***	0.278***	0.081*	0.244***	0.309***	0.027	1.714***
	(0.094)	(0.1)	(0.044)	(0.054)	(0.066)	(0.021)	(0.364)
Quantile [50-75]	0.359***	0.327***	0.044	0.222***	0.32***	0.035*	1.603***
	(0.069)	(0.074)	(0.031)	(0.053)	(0.064)	(0.021)	(0.363)
Quantile [75-90]	0.502***	0.427***	0.075***	0.237***	0.382***	0.044**	1.675***
	(0.063)	(0.067)	(0.028)	(0.054)	(0.065)	(0.021)	(0.365)
Quantile [90-95]	0.542***	0.424***	0.101***	0.249***	0.437***	0.061***	1.562***
	(0.062)	(0.067)	(0.027)	(0.054)	(0.066)	(0.021)	(0.368)
Quantile [95-100]	0.46***	0.36***	0.116***	0.245***	0.476***	0.081***	1.612***
	(0.062)	(0.067)	(0.027)	(0.054)	(0.066)	(0.021)	(0.367)

Table continued on next page.

64

TABLE II: (CONTINUED)

| | Flu-like illness | | | Acute diarrhea | | | Chickenpox |
	Children	Adults	Elderly	Children	Adults	Elderly	Children
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Panel B: between region effects ($\alpha^q_{between}$)							
Quantile [0-25]	0.007	0.003	-.003	0.009*	0.013**	0.006***	0.013
	(0.031)	(0.035)	(0.015)	(0.005)	(0.006)	(0.002)	(0.01)
Quantile [25-50]	0.004	0.006	0.002	0.01**	0.012**	0.005***	0.012
	(0.011)	(0.012)	(0.005)	(0.004)	(0.005)	(0.002)	(0.01)
Quantile [50-75]	0.002	0.008	0.004*	0.011***	0.012**	0.005***	0.017*
	(0.005)	(0.006)	(0.002)	(0.004)	(0.005)	(0.002)	(0.01)
Quantile [75-90]	0.002	0.007*	0.003**	0.013***	0.011**	0.005***	0.009
	(0.003)	(0.004)	(0.001)	(0.004)	(0.005)	(0.002)	(0.01)
Quantile [90-95]	0.002	0.007**	0.003***	0.013***	0.009*	0.004**	0.011
	(0.003)	(0.003)	(0.001)	(0.004)	(0.005)	(0.002)	(0.011)
Quantile [95-100]	0.003	0.008**	0.003***	0.013***	0.007	0.003**	0.01
	(0.003)	(0.003)	(0.001)	(0.004)	(0.005)	(0.002)	(0.011)

Notes. This table displays estimates of the parameters α^q_{within} and $\alpha^q_{between}$ defined in equation (5) in the Appendix. Each column represents a separate regression with incidence rates as the dependent variables. All regressions include region, year and week fixed effects. IV regressions using lagged temperature and precipitation as instruments. Standard errors are corrected for serial and spatial correlation, using a a Prais-Winsten regression, where a region specific AR(1) process is assumed. The model also includes school holidays, transportation strikes, quarterly dummies and linear trend interacted with the transmission rates, which are not displayed. ***, ** and * denotes significance at 1%, 5%, and 10% level.

TABLE III: Cross effects between diseases

	Children	Adults	Elderly
		Flu-like illness	
Acute diarrhea	-.001	0.009***	0.007***
	(0.002)	(0.003)	(0.001)
		Acute diarrhea	
Flu-like illness	0.005***	0.009***	0.001**
	(0.001)	(0.002)	(0.0005)
		Chickenpox	
Flu-like illness	-.012	-	-
	(0.008)		

Notes. This table displays estimates of the effect of the incidence of other diseases on the incidence of a given disease (parameters α_{within}^k and $\alpha_{between}^k$ defined in equation (6) in the text). Each entry presents a separate regression with incidence rates as the dependent variables. All the variables listed in Table VI are also included in the regression but not shown. Standard errors are corrected for serial and spatial correlation, using a a Prais-Winsten regression, where a region specific AR(1) process is assumed. ***, ** and * denotes significance at 1%, 5%, and 10% level.

TABLE IV: Maximum per capita gain of policy and value of statistical life

Value of statistical life (millions of euros)	Flu-like illnesses		Acute diarrhea	
	School closure	Public transportation closure	School closure	Public transportation closure
1.3	-23 €	-14.1 €	-23 €	- 18.7 €
3	-19 €	-11.2 €	-22.8 €	-18.7 €
5	-14 €	-7.8. €	-22.6 €	-18.7 €
6	-12 €	-6.1 €	-22.6 €	-18.6€

Notes. Simulations performed using the cost parameters in Tables VII and VIII and the estimated parameters in Tables VI. The table displays the cost taking into account the impact of the policy on a given disease.