Global Mobility and the Threat of Pandemics: Evidence from Three Centuries

Michael A. Clemens and Thomas Ginn

Abstract

Countries restrict the overall extent of international travel and migration to balance the expected costs and benefits of mobility. Given the ever-present threat of new, future pandemics, how should permanent restrictions on mobility respond? A simple theoretical framework predicts that reduced exposure to pre-pandemic international mobility causes slightly slower arrival of the pathogen. A standard epidemiological model predicts no decrease in the harm of the pathogen if travel ceases thereafter and only a slight decrease in the harm (for plausible parameters) if travel does not cease. We test these predictions across four global pandemics in three different centuries: the influenza pandemics that began in 1889, 1918, 1957, and 2009. We find that in all cases, even a draconian 50 percent reduction in pre-pandemic international mobility is associated with 1–2 weeks later arrival and no detectable reduction in final mortality. The case for permanent limits on international mobility to reduce the harm of future pandemics is weak.

JEL: H23, I18, J68.

Center for Global Development
www.cgdev.org

Working Paper 560
December 2020

Global Mobility and the Threat of Pandemics: Evidence from Three Centuries

Michael A. Clemens
Center for Global Development
IZA

Thomas Ginn
Center for Global Development

We are grateful to John Paget, Lone Simonsen, Peter Spreeuwenberg, and Robert J. Taylor for their ideas and for sharing data from the Global Pandemic Mortality (GLaMOR) project, available online through the NIVEL Inuenza Mortality Research Project; to Emanuel Deutschmann for sharing information on the Global Transnational Mobility Dataset; and to Yasir Khan, Rose Worden, and Cassandra Zimmer for excellent research assistance. We received helpful comments from David Jaeger, Margaret Leighton, David Escamilla-Guerrero, Antonella Bancalari, Rod McCrorie, Michele Tizzoni, Sian Tsuei, Sebastian Bauhoff, Joeri Smits, Luca Braghieri, Helen Dempster, Dave Evans, Alan Gelb, Jeremy Konyndyk, Carleigh Krubiner, Mead Over, and seminar participants at the Harvard Kennedy School and the University of St. Andrews, but any errors are ours alone. The paper represents the views of the authors alone and not necessarily those of their employers, funders, or any other institutions. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript. A pre-analysis plan was irreversibly registered at https://osf.io/bn8mx.

The Center for Global Development is grateful for contributions from Open Philanthropy in support of this work.

Michael Clemens and Thomas Ginn, 2020. "Global Mobility and the Threat of Pandemics: Evidence from Three Centuries." CGD Working Paper 560. Washington, DC: Center for Global Development. https://www.cgdev.org/publication/global-mobility-and-threat-pandemics-evidence-three-centuries

Center for Global Development
2055 L Street NW
Washington, DC 20036

202.416.4000
(f) 202.416.4050

www.cgdev.org

Contents

1 Introduction

International labor mobility (worker migration) and trade in services (such as overseas tourism and study) require personal interaction. Personal interaction, however, can also spread infectious disease, causing a "manifest externality" (Bleakley 2007). Together, these imply that negative externalities from the spread of new pandemic diseases could be internalized by Pigouvian taxes or quotas on international human mobility and raise global welfare. It is well known that temporary, domestic Pigouvian restrictions on some in-person economic interactions—embraced by Pigou (1920, 169) himself—might raise welfare if the marginal economic cost of disease transmission exceeds the economic benefit of the targeted interactions (Fenichel 2013; Adda 2016).

But can *permanent* Pigouvian limits on *international* human mobility raise welfare? This question does not refer to brief emergency travel restrictions amid a raging pandemic, which have received extensive study.[1] Rather, in this paper we ask if limits on generalized international economic integration via human mobility—in normal times, before a pandemic begins—can raise welfare by internalizing the externality of transmitting new, emergent pandemics.

In principle, the degree of pre-existing global integration via international migration and travel can aggravate morbidity and mortality once a pandemic begins (Barro et al. 2020; Antràs et al. 2020).[2] This implies that countries' pre-pandemic exposure to international mobility, despite its positive effects on income per capita (Campante and Yanagizawa-Drott 2018), may be inefficiently high. The 1918 influenza pandemic killed over 50 million worldwide (Johnson and Mueller 2002). Pandemics also bring severe economic costs by restricting social interaction and exchange (Rasul 2020), with costs persisting for several decades via *in utero* morbidity (Almond 2006; Neelsen and Stratmann 2012; Guimbeau et al. 2020) and reduced capital investment (Jordà et al. 2020). The economic cost of pandemic risk per year—in *all* years—has been estimated at 0.7% of global GDP (Fan, Jamison and Summers 2016). "Smart policies have to weigh the costs and

[1] For reviews of these policies, see Ferguson et al. (2006), Mateus et al. (2014), and Ryu et al. (2020). We discuss these findings in Section 7, but our paper does not address the value of emergency travel restrictions enacted to mitigate ongoing pandemics.

[2] Barro et al. (2020, 7–8) write, "[T]he flu death rates from the Great Influenza Pandemic [of 1918] . . . likely represent the worst-case scenario today [in 2020], particularly because public-health care and screening/quarantine procedures are more advanced than they were in 1918–1920. Other factors, such as greater international travel, work in the opposite direction".

benefits of uninhibited exchanges of people," Voth (2020, 95–96) writes. "Severe restrictions may well be desirable and justifiable, bringing to an end a half-century of ever-increasing individual mobility."

In this paper, we assess the theoretical and empirical case for permanent Pigouvian limits on international mobility under the risk of emergent pandemics. We begin by presenting a standard Kermack-McKendrick (1927) SIR model of epidemic transmission, extended to include travel between a foreign country where the outbreak emerges and a home country. The base model yields three predictions: 1) lower pre-pandemic international arrivals delay the pandemic's arrival in the home country, to a small degree; 2) the date of arrival is unrelated to the overall harm; and therefore 3) the volume of pre-pandemic arrivals has *no* effect on the overall harm of the pandemic in the home country—that is, it carries no negative externality in the form of additional expected sickness or deaths during a pandemic. In brief, this is because the Poisson process representing the arrival of the first infected traveler quickly approaches certainty even under very small travel volumes, and once the infection crosses the border, the harm it causes is fully determined by within-country determinants of transmission.

An externality from pre-existing international mobility does arise in extensions of the base model. For example, externalities can arise from pre-existing international mobility if the degree of pre-existing mobility constrains the degree of feasible emergency travel restrictions during the pandemic. Theory suggests, perhaps counter-intuitively, this externality could be negative or positive, and that the magnitude would be small in a typical pandemic. Other extensions of the base model, such as rapid learning about treatments during the pandemic's initial stages or linking mobility to innovation, likewise extend the predictions. We therefore test the simple predictions of the base model across the four global pandemics for which the necessary epidemiological data exist: backwards in time, these are the influenza pandemics that began in 2009, 1957, 1918, and 1889. We test the empirical relationship between pandemic arrival delay and large variation in predetermined exposure to human mobility across both space and time. We test the empirical relationship between pandemic arrival time and mortality, and where possible, we directly test the relationship between exposure to international mobility and pandemic mortality.

We find that large changes in pre-pandemic exposure to international mobility have a small association with arrival time and no association with the harm to population health. A 50 percent lower exposure to pre-pandemic international mobility is associated with delayed pandemic arrival by roughly one week. Even 90 percent lower exposure to mobility—across countries or across time—is associated with an increased delay of 1–2 months. We find no evidence that delayed arrival correlates systematically with lower morbidity or mortality in any of the pandemics. In some cases, in fact, delayed arrival correlates with greater harm (such as in 1918) and greater exposure to mobility correlates with reduced harm (such as in 2009). These results are robust to a range of different regression specifications and different measures of mobility, mortality, and start-date. We conclude that the economic case for permanent Pigouvian restrictions on international mobility due to pandemic externalities is weak.

The contribution of this paper is twofold. First, it theoretically clarifies the conditions under which permanent Pigouvian restrictions on international mobility can improve welfare by internalizing pandemic externalities. Second, it offers empirical tests of those conditions in four real pandemics using voluminous data from a range of historical sources. It adds to the economic literature exploring the health externalities of domestic economic activity (Fenichel 2013; Goyal and Vigier 2015; Adda 2016; Glaeser et al. 2020; Milusheva 2020) and the links of international mobility and economic integration with the international spread of disease (Montalvo and Reynal-Querol 2007; Baez 2011; Horan et al. 2015; Zimmermann et al. 2020). Montalvo and Reynal-Querol (2007) and Milusheva (2020) specifically find movements – international and domestic, respectively – increase the spread of malaria. Since the threat of malaria is known and ongoing in these settings, these results are closest to our extension allowing travel to continue during a pandemic and are not in conflict with our core findings on the externality of pre-pandemic mobility. This paper builds on a theoretical literature in epidemiology and mathematical biology (in particular Ma and Earn 2006; Arino et al. 2007; and Scalia Tomba and Wallinga 2008) and complements a large literature in epidemiology focusing on emergency travel restrictions, discussed later.[3]

The economic case for generalized limits on international mobility to address pandemic exter-

[3]Our findings complement those of Lee et al. (2020) and Ahsan et al. (2020), who find that parts of migrant-origin countries with higher exposure to pre-existing migration have exhibited higher prevalence and mortality in the ongoing pandemic of coronavirus disease (COVID-19). This pattern is predicted by our model: in an ongoing pandemic, earlier arrival correlates mechanically with higher prevalence and mortality at any given moment in time. We study pandemics that have concluded, allowing tests of the theory's predictions about final size.

nalities is analogous to the standard economic case for speed limits on car travel. Society limits drivers' speed permanently—not just during emergencies—as a Pigouvian quota to internalize the external risk to others' lives caused by some drivers' benefit from fast travel (Ashenfelter and Greenstone 2004, S231). Our results, however, suggest that this analogy may be misplaced. Permanent, pandemic-related limits on international mobility, such as increased visa barriers (Czaika and Neumayer 2017), may more closely resemble speed limits for passenger airliners to avoid collisions. A policymaker might seek to address the public's fears of dying in an airliner accident by mandating a reduction in cruising-altitude airspeed from 500 to 250 miles per hour. But there is little theory or evidence to support the idea that this would reduce the risk of death (e.g. Boyd 2017). (It might even unintentionally increase the risk, by forcing planes to spend longer in the air.) As long as air travel occurs at all, its external harms are not correlated with aircraft speed over a meaningful range, and Pigouvian restrictions on speed would only reduce welfare. Similarly, basic theory and evidence from major pandemics imply that permanent limits on international mobility would have little benefit to offset their very large costs, bringing the world no closer to the Pigouvian optimum.

2 Model

We model the relationship between pre-existing international travel—travel before the initial case of a pandemic disease—and the final size of the pandemic. We incorporate travel into a Kermack-McKendrick (1927) model of disease transmission, building on Brauer and van den Driessche (2001), Ma and Earn (2006, 691), Arino et al. (2007), and Scalia Tomba and Wallinga (2008). The disease will emerge in a foreign country, and we examine how the volume of pre-existing travel to the home country affects the outbreak in the home country.

Our research question is how levels of travel *before* a pandemic affects the evolution of the pandemic. We distinguish between levels of travel before and during the pandemic, since the latter can be changed by emergency travel restrictions independent of the levels before the pandemic. We present the results under different emergency restrictions. In the base case, emergency travel restrictions are implemented when the disease arrives in the home country. We then extend the results to allow travel to continue during the pandemic. In both cases, pre-pandemic levels of

travel determine the onset date, reflecting the difficulty of immediately enacting effective emergency travel restrictions.

2.1 Base Model: Travel and Onset Date

The first outcome we examine is the expected onset date in the home country, $\tilde{t}$. Let a proportion $I^*(t)$ of foreigners in the foreign country be infected at time t. In each period, M foreigners are randomly selected to travel to the home country, and the selection is independent of their health status. The probability of drawing at least one infected traveler in M draws in period t is then[4,5]

$$\lambda(M,t) = 1 - (1 - I^*(t))^M. \tag{1}$$

The arrival of at least one infected traveler in each period can be approximated as a non-homogeneous Poisson process with the parameter $\lambda(M,t)$ as in Scalia Tomba and Wallinga (2008).[6] Let $\Lambda(M,t) = \int_0^t \lambda(M,s)\,ds$. The time of the expected first occurrence—the onset date—is represented as:

$$\tilde{t}(M) = \int_0^T t\,\lambda(M,t)\,e^{-\Lambda(M,t)}\,dt \tag{2a}$$

$$= -t\,e^{-\Lambda(M,t)}\Big|_{t=0}^{t=T} + \int_0^T e^{-\Lambda(M,t)}\,dt \tag{2b}$$

$$\approx \int_0^T e^{-\Lambda(M,t)}\,dt, \tag{2c}$$

where T represents a time when all countries have at least one case.[7] The first step comes from integration by parts and the second step approximates the first term as 0.

The first result is the inverse relationship between the onset date and the number of incoming travelers, represented by $\frac{d\tilde{t}}{dM} < 0$. Intuitively, more people traveling decreases the chance that none of them are infected in any given period.

[4] The parameter is an approximation of sampling with replacement.

[5] If only one origin country has infections, M is the number of travelers from that country. If multiple origins have different rates of infection, the parameter could be adapted to $1 - (1 - I_1^*)^{M_1}(1 - I_2^*)^{M_2}$, etc. The same modifications can be made to allow different prevalence among demographic groups in the foreign country.

[6] A non-homogeneous Poisson process is an approximation of repeated Bernoulli trials with a time-varying parameter.

[7] Equation A.2a parallels the exponential distribution which represents inter-arrival periods for a homogeneous Poisson process. Intuitively, the density function is the probability of arrival in period t, $\lambda(M,t)$, times the probability of no arrivals through period t, $e^{-\Lambda(M,t)}$. For the derivation for non-homogeneous Poisson processes, see Ma (2011).

Figure 1: Travel Quotas: International mobility and the timing of pandemic arrival

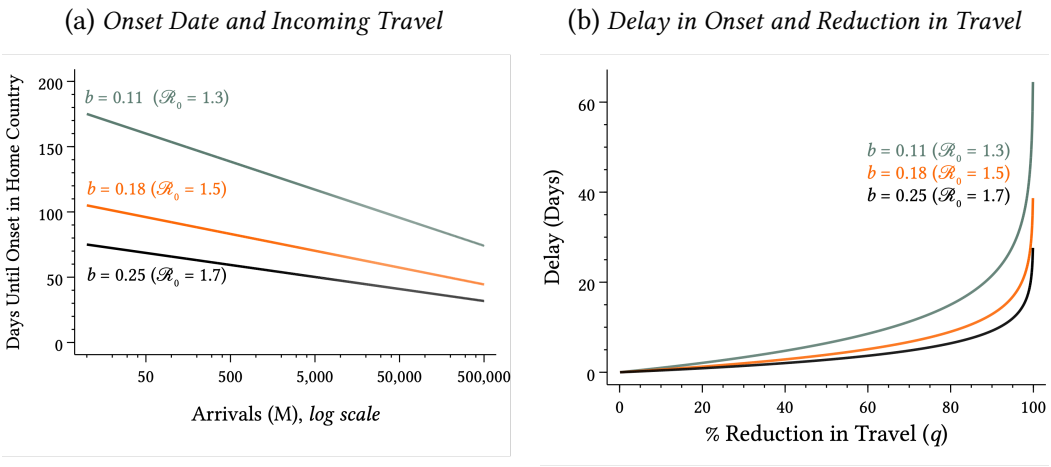


Figure 1a plots the onset date after the first global case and the total number of incoming travelers. Figure 1b plots the reduction in the expected onset date by reducing incoming arrivals, with the percent reduction on the x-axis. The values correspond to Equations 3 and 4, respectively. The lines correspond to different exponential rates of growth in the foreign country at the pandemic's onset, b, which are a function of $\mathscr{R}_0$: $b = \beta^* - \gamma = \frac{\mathscr{R}_0}{S} - \frac{1}{S}$, where β^* is the transmission rate in the foreign country, γ is the recovery rate as outlined in Equations A.1a-A.1c, and S is the generation time. The black line corresponds to $\mathscr{R}_0 = 1.7$ ($\beta^* = 0.61$) in the foreign country, the orange line to $\mathscr{R}_0 = 1.5$ ($\beta^* = 0.54$), and the teal line to $\mathscr{R}_0 = 1.3$ ($\beta^* = 0.46$), with a value of $S = 2.8$ which corresponds to $\gamma = 0.36$. The values correspond to the 75th, 50th, and 25th percentile estimates, respectively, of $\mathscr{R}_0$ for H1N1 in Biggerstaff et al. (2014)'s survey of estimates in the literature.

However, the delay from reducing the number of travelers will be minimal without drastic cuts to travel, which includes both foreigners arriving and returning citizens, since anyone can transmit spread a disease. We summarize the intuition of the closed form solution here and derive it in the appendix. First, we model $I^*(t)$, the proportion of the infected foreign population over time, to grow exponentially as in the standard Susceptible-Infectious-Recovered model at the start of the pandemic. A small number of people are infected initially, and the probability of drawing an infected traveler is close to 0. The probability increases in the number of travelers, but the marginal effect of additional travelers is small in accordance with Equation 1. As the disease grows exponentially, the number of travelers must be reduced exponentially in order to keep the probability of drawing an infected traveler small. We approximate the relationship between the onset date $\tilde{t}$ and the incoming travelers M as:

$$\tilde{t}(M) \approx \frac{1}{b}\big(ln(0.2P) - ln(M)\big), \tag{3}$$

where b is the initial exponential growth rate of the disease and P is the population size of

the foreign country.[8] The effect of an additional incoming traveler on the onset depends on the exponential growth rate of the disease in the foreign country and the log of the number of travelers.

If a quota reduces travel by some fraction $0 \leqslant q < 1$ of the previous level, the start of local transmission is delayed by

$$\tilde{t}\big((1-q)M\big) - \tilde{t}(M) \approx \frac{-ln(1-q)}{b}. \tag{4}$$

Therefore, while reducing travel can delay the onset, the delay would likely be minimal. For example, with $b = 0.18$ (corresponding to $\mathscr{R}_0 = 1.5$), a 50 percent reduction in the migration rate delays the start of local transmission in the home country by 3.9 days. A 90 percent reduction in migration provides a delay of 12.8 days, while a 99 percent reduction buys 25.6 days. Similar conclusions have been reached with alternative derivations (Scalia Tomba and Wallinga 2008).[9] We then have:

> **Prediction 1.** *The start-date of local transmission in the home country is inversely related to the prior migration rate. For a range of plausible reproduction numbers and recovery rates, a reduction in migration by an order of magnitude delays onset by 1–3 weeks.* □

This highly stylized, two-country model of arrival time serves only as a benchmark against which to compare the empirical results below. The recent epidemiology literature focuses on more sophisticated models of arrival time based on network diffusion (Colizza et al. 2006; Gautreau et al. 2008; Brockmann and Helbing 2013). For this reason we later check the empirical results for robustness to adjustments for the network centrality of travelers' origin countries.

[8] In the SIR model, $b \approx \beta^* - \gamma$, where β^* is the transmission rate in the foreign country and γ is the recovery rate in both countries as outlined in Equations A.1a-A.1c (Ma 2020).

[9] Scalia Tomba and Wallinga (2008) reach a similar relationship for the delay in median arrival time, although b in their model is the approximate exponential growth rate of $\Lambda(t) = ae^{bt}$, where ours is the approximate exponential growth rate of $I^*(t)$.

2.2 Base Model: Travel, Onset Date, and Final Size

The main outcome we examine for the home country is the overall harm. In the base model, a strict emergency travel restriction is imposed once the disease arrives in the home country that stops all international travel. The disease follows a standard Susceptible-Infectious-Recovered model (Kermack and McKendrick 1927). Under mass-action incidence — equal probability of every person within the country's borders coming into contact with one another — the fractions of susceptible (S), infectious (I), and recovered/removed (R) follow

$$\dot{S} = -S\beta I \tag{5a}$$

$$\dot{I} = S\beta I - \gamma I \tag{5b}$$

$$\dot{R} = \gamma I, \tag{5c}$$

with a dot to denote the time derivative. The transmission rate β is the probability of transmission from an infectious to susceptible person in country i per unit time, and the recovery rate γ is the reciprocal of the mean infectious duration. The basic reproduction number, the number of infections in an all-susceptible population generated by one infectious case, is $\mathscr{R}_0 \equiv \beta/\gamma$.

Let the *final size* of the local epidemic, the fraction of the home population infected over its entire course, be $F \equiv S(0) - S(\infty)$.[10] The population is fully susceptible before the outbreak ($S(0) = 1$), and the disease dies out ($I(\infty) = 0$) (Ma and Earn 2006, 693). The final size can be determined by:

$$\begin{aligned} \ln(1 - F) &= \ln(S(\infty)) \\ &= \int_0^\infty \frac{\dot{S}}{S} dt \\ &= -\beta \int_0^\infty I \, dt, \end{aligned}$$

[10] Some parts of the literature call the share F the 'attack rate', reserving 'final size' for the absolute number of cases (Andreasen 2011, 2308). Here we follow the terminology of Ma and Earn (2006). The exclusive dependence of final size on $\mathscr{R}_0$ was first derived in a different form by Kermack and McKendrick (1927) and its remarkable robustness to relaxing many of the strong assumptions here is explored inter alia by Thieme (1977), Newman (2002), Ma and Earn (2006), Arino et al. (2007), Andreasen (2011), and Miller (2012).

with the last equality from dividing Equation 5a by S and integrating. Now

$$\begin{aligned}\int_0^\infty I\,dt &= \frac{1}{\gamma}\int_0^\infty \dot{R}\,dt \\ &= \frac{1}{\gamma}(R(\infty) - R(0)) \\ &= \frac{1}{\gamma}F,\end{aligned}$$

with the first equality from integrating Equation 5b. Now

$$F(\mathscr{R}_0) = 1 - e^{-\mathscr{R}_0 F} \tag{6}$$

$$= 1 + \frac{1}{\mathscr{R}_0}\mathscr{W}\left(-\mathscr{R}_0 e^{-\mathscr{R}_0}\right), \tag{7}$$

where $\mathscr{W}$ is Lambert's product-log function (Corless et al. 1996, 337), single-valued when $\mathscr{R}_0 > 0$.

In the base case, when travel is stopped at the onset in the home country, it is apparent that the final size is independent of the onset date. The final size is only a function of β and γ, the disease parameters in the home country.

> **Prediction 2.** *Under mass-action incidence, the expected final size of any pandemic is independent of the start-date of local transmission.* □

Predictions 1 and 2 imply an additional prediction:

> **Prediction 3.** *If travel stops when the pandemic reaches the home country, then under mass-action incidence, the expected final size of any pandemic in the home country is independent of the prior immigration rate.* □

2.3 Extension: $\mathscr{R}_0$ varying over time

Several assumptions, however, are important to relax for our research question. First, in our standard model, $\mathscr{R}_0$—and therefore final size—does not depend on t or $\tilde{t}$. Additional time prior to onset could be used to set up systems of contact tracing, build hospital capacity, develop testing, discover a therapeutic, or install other measures that would decrease β or increase γ.

These would reduce $\mathscr{R}_0$ and hence the final size. Since a lower rate of travel pushes back the expected onset date, this is one channel that could overturn our second prediction.[11] However, it is important to note that $\mathscr{R}_0$ must vary with the *global* onset date. If $\mathscr{R}_0$ varies relative to the country's onset date—i.e. behavior changes once people see the implications, or a country's preparations begin at domestic onset—then a later arrival will not affect the final size.

Second, travel also has direct benefits that could reduce $\mathscr{R}_0$. Travel during, but especially before, the pandemic facilitates the spread of ideas, technology, workers, and supplies that could aid prevention and treatment (e.g. Hovhannisyan and Keller 2015; Bahar et al. 2020).[12]

The true object of interest should therefore be the total derivative of $F\big(\mathscr{R}_0(\tilde{t}(M), M)\big)$ with respect to M:

$$\underbrace{\frac{dF}{dM}}_{?} = \underbrace{\frac{\partial F}{\partial \mathscr{R}_0}}_{>0} \underbrace{\frac{\partial \mathscr{R}_0}{\partial \tilde{t}}}_{<0} \underbrace{\frac{d\tilde{t}}{dM}}_{<0} + \underbrace{\frac{\partial F}{\partial \mathscr{R}_0}}_{>0} \underbrace{\frac{d\mathscr{R}_0}{dM}}_{<0}, \tag{8}$$

which has an indeterminate sign under our formulation and hypotheses. The relationship between the final size and pre-existing travel depends on the relative effects of travel pushing the onset date forward and contributing to the preparedness and response.

2.4 Extension: Continuing Travel

In the base model, an emergency travel restriction stops all travel once the disease arrives in the home country. However, strict emergency travel restrictions may be infeasible or undesirable, and the level of continued travel may be related to the our main variable of interest, the level of travel before the pandemic. We therefore relax this emergency travel restriction and allow travel to continue during the pandemic.

Assume travelers are drawn from the foreign population as before and stay in the home country for one period. The proportion of foreigners in the home country is then $p = M/(M+N)$, where N is the total home country population. We assume foreigners and home citizens perfectly

[11] Similarly, deaths as a proportion of the infected could depend on t and therefore $\tilde{t}$.

[12] During the pandemic, travel could also increase $\mathscr{R}_0$ if the activities of travel itself – being on a plane or bus, in an airport, etc. – increase disease transmission.

mix so that all *individuals* are equally likely to interact and therefore spread the disease to one another. The probability that a given interaction in the home country is with a foreigner is then p.

Under this assumption of mass-action incidence—homogeneous and perfectly mixed populations—the fractions of home citizens who are susceptible (S), infectious (I), and recovered/removed (R) now follow

$$\dot{S} = -S\beta\big((1-p)I + pI^*\big) \tag{9a}$$

$$\dot{I} = S\beta\big((1-p)I + pI^*\big) - \gamma I \tag{9b}$$

$$\dot{R} = \gamma I, \tag{9c}$$

with a dot to denote the time derivative as before and I^* to denote the proportion of foreigners who are infected.

Define $F^* \equiv \gamma \int_0^\infty I^* dt$, the final size in the foreign country without travel.[13,14] Deriving the final size as before leads to:

$$F'(\mathscr{R}_0, p, F^*) = 1 - e^{-\mathscr{R}_0((1-p)F + pF^*)}. \tag{10}$$

If travel continues throughout the pandemic, then the final size in the home country depends on the final size in the foreign country and the number of travelers allowed in. If $F = F^*$, i.e. $\mathscr{R}_0$ is the same in both countries, then the final size in the home country is also independent of the rate of travel. If the final size is higher in the foreign country, then continuing travel will increase the final size in the home country, consistent with many people's intuitions. However, if the final size is lower in the foreign country, then continuing travel will lower the final size in the home country as the travelers are less likely to spread the disease than home citizens.

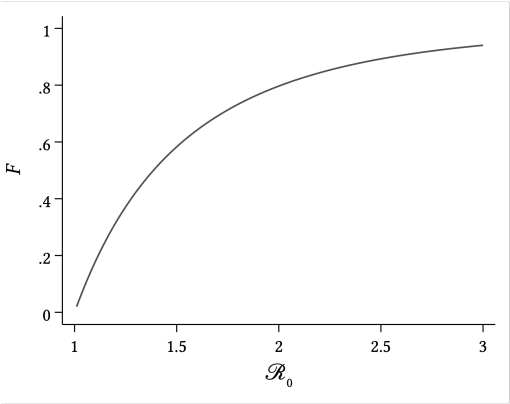
However, the effect is again likely to be small in this model. From our data on international travel

[13] We fix γ across countries, as it largely depends on biological characteristics of the disease, and allow β to vary across countries.

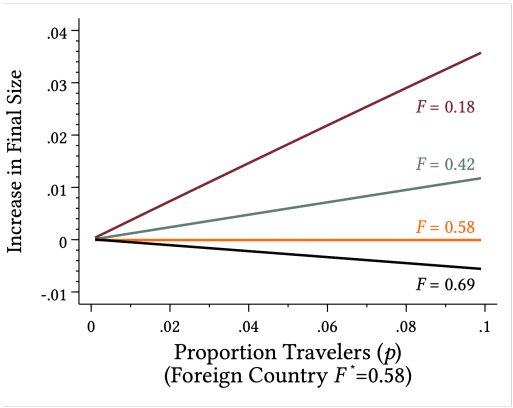
[14] The model abstracts from the effect of return travelers on the foreign country. In such a model with bilateral flows of travelers who can be infected in either country, the final sizes across countries converge, and the increases in final size are attenuated towards 0 relative to this model. The proportion of travelers in each period would determine the speed of convergence. We thank Sian Tsuei for this comment.

Figure 2: Final Size: International mobility and the total number of infected

(a) *Final Size and $\mathscr{R}_0$ with no immigration*



(b) *Changes in Final Size due to continuing travel*



Panel a shows Equation 7. Panel b shows $F'(\mathscr{R}_0, p, 1.5) - F(\mathscr{R}_0)$ on the y-axis, the increase in final size due to continued travel from a country with $\mathscr{R}_0 = 1.5$ ($F^* = 0.58$), as the level of the travel, p, varies. The lines correspond to different $\mathscr{R}_0$'s, and therefore final sizes, for the home country. The black line corresponds to $\mathscr{R}_0 = 1.7$ ($F = 0.69$) in the home country, the orange line to $\mathscr{R}_0 = 1.5$ ($F = 0.58$), the teal line to $\mathscr{R}_0 = 1.3$ ($F = 0.42$), and the maroon line to $\mathscr{R}_0 = 1.1$ ($F = 0.18$).

described below, p, the percentage of recently-arrived international travelers in the population at any given time, is about 1.5 percent in the median country, and 4 percent in the average country.[15] Thus, very quickly after the pandemic arrives, any given person's chance of acquiring the disease from an infected local greatly outstrips the chance of acquiring it from an infected traveler.[16] Differences in F and F^*, stemming from differences in $\mathscr{R}_0$, are also likely to be small since the disease is the same. Figure 2b maps these potential changes for different $\mathscr{R}_0$'s in the home country.

This simple model's ambiguous predictions highlight the importance of empirical tests, which we present below.

[15] We define recently-arrived as the total number arriving in an average 10-day window.

[16] Russell et al. (2020) estimate that during the COVID-19 pandemic, by May 2020 the fraction of incident infections from international travelers would have been less than one percent of all incident infections in most countries on earth, even if no country had imposed emergency travel restrictions during the pandemic.

3 Data on four pandemics

Testing the three hypotheses requires collecting three types of country-level data: pandemic final size, pandemic arrival time, and the international mobility of people. We require the pandemic to be over, allowing estimation of final size. We require data to be available for pandemics that caused substantial mortality in numerous countries, because estimates of incidence/prevalence alone are severely biased by cross-country differences in diagnostic methods. We also require these to be available for pandemics that spread rapidly, and thus became relevant to generalized restrictions on international migration. These criteria admit exactly four pandemics. In reverse chronological order, these are the influenza pandemics that began in 2009 ('Swine flu'), 1957 ('Asian flu'), 1918 ('Spanish flu'), and 1889 ('Russian flu'). Economic activity requiring interpersonal interaction is known to sharply increase influenza transmission (Markowitz et al. 2019). We investigate all four of these historical episodes.

Our criteria exclude a number of episodes. The requirement that the pandemic have concluded excludes COVID-19. The requirement of broad-based mortality excludes pandemics that only caused substantial mortality in a limited number of countries, such as the Severe Acute Respiratory Syndrome (SARS) outbreak of 2002, the Middle East Respiratory Syndrome (MERS) oubreak of 2012, and the Ebola virus outbreak of 2013. For these, reliable country-level estimates of final size do not exist for a large number of countries. The requirement of rapid spread excludes pandemics whose international dissemination progressed over decades. This excludes Zika virus, which spread from Uganda across Africa and Asia and then through Latin America between 1952 and 2016. Most notably, it excludes Human Immunodeficiency Virus (HIV-1 group M subtype B), which spread from Central Africa around 1930, to Haiti around 1966, to the United States around 1969, and around the world step by step from 1981 to 2001 (Gilbert et al. 2007). We likewise omit pandemics for which no country-level estimates of final size exist. This excludes several influenza pandemics before 1889 (Patterson 1985) and pandemics of yellow fever, cholera, and bubonic plague in the 19th century. It also excludes the pandemic of H3N2 influeza in 1968 ('Hong Kong flu'), for which final mortality estimates have only been computed for six countries (Viboud et al. 2005).

3.1 Final mortality

For 2009 influenza, we use two independent sources of data on final mortality, one from the literature and one that we create. We begin by using the country-level, all-age respiratory mortality estimates of Simonsen et al. (2013), which build on the earlier estimates of Dawood et al. (2012). These estimates are based on directly-reported mortality rates from 26 countries, and imputed for 185 other countries by modeling from country-level observable traits such as income per capita, population density, and physicians per capita.

While these are the best country-level estimates of 2009 influenza final mortality in the literature, they have an important limitation for our purpose: the covariates used for imputation are country-traits that we wish to control for in some regressions. For this reason, we check the robustness of all subsequent tests with alternative measures of mortality. The first is based on our own estimates of excess influenza-like illness mortality based on direct (not imputed) mortality reports to the World Health Organization, described below. The second is to restrict analysis of the Simonsen et al. sample to the 26 countries that directly reported mortality rates. The third is a reanalysis using the mortality estimates of Dawood et al. (2012), which are independently imputed by a different method.

We construct our independent measure of 2009–2010 H1N1 influenza mortality by computing excess influenza-like illness mortality as follows. From the WHO Mortality Database (December 15, 2019 update), for each year we draw the number of deaths reported by each country due to influenza-like respiratory illness[17] in people age 15–49. We use that age range because disproportionate mortality among prime-age adults strongly distinguished 2009 pandemic influenza from seasonal influenza (Shrestha et al. 2011; Simonsen et al. 2013). We then fit the trend in these mortality rates by regressing annual mortality from the four years before and after the pandemic (2005–2008 and 2011–2014) on year with Kernel Regularized Least Squares (KRLS, Hainmueller and Hazlett 2014) which simultaneously chooses parameters to maximize fit while choosing functional form to maximize smoothness. This allows a nonlinear fit in a small-N setting without arbitrary choices of functional form. The method imputes separately for each

[17] International Classification of Diseases, Tenth Revision, Clinical Modification [ICD-10-CM] codes J09–J18 and J20–J22. This includes deaths from pneumonia as Noymer (2008) recommends.

country the counterfactual on-trend mortality for 2009 and 2010. Our 2009 pandemic influenza mortality estimate is then the sum of excess mortality above trend for 2009 and 2010, for each of 99 countries. In the Appendix we plot the raw data and KRLS curvefits demonstrating that this method systematically isolates anomalous flu deaths in 2009–2010 and show that it does not do so in placebo years.

For 1957 influeza, we use the estimates of final mortality at the country level by Viboud et al. (2016). For 1918 influenza, we use the final mortality estimates for all of 1918 in Johnson and Mueller (2002, 110-114) and for the fall wave alone in Patterson and Pyle (1991, 14–15), harmonized as described in the Appendix. For 1889 influenza, first wave, we use the city-level excess mortality estimates of Valleron et al. (2010).

3.2 Time of arrival

We use date of arrival as the date of the first reported case inside a country. For 2009 influenza, we draw on the daily arrival dates in 93 countries compiled by Balcan et al. (2009), but expand this to 193 countries using the sources enumerated in the Appendix. For 1957 influenza we take monthly arrival dates from UNESCO (1958, 12–13). For 1918 influenza (overall and fall wave only) we draw many arrival dates, daily or monthly, from Patterson and Pyle (1983, 1991) and Frost and Sydenstricker (1919), but we augment their data with numerous additional countries using the sources enumerated in the Appendix. For 1889 influenza we draw many country-level arrival dates, most daily and a few monthly, from Parsons (1891b, 11–50), but augment his data with numerous additional countries using the sources enumerated in the Appendix. Also for 1889 influenza, we draw city-level weekly arrival dates from Valleron et al. (2010), alternatively defined (as in the source) as either date of first deaths or date of peak deaths.

3.3 Human mobility

For the main independent variable, our goal is to count all cross-border movements, including foreign tourism, citizens returning from abroad, and immigrants arriving across land, air, and sea ports of entry, since any could bring a disease to the destination country. We estimate these flows at the bilateral level using two main datasets on travel and migration and a simple method to

predict missing values. We use the bilateral flows to aggregate up to country-level measures and to construct a centrality measure to weight arrivals in robustness checks. We focus on the H1N1 pandemic as the only time period when these data are reliably available, and the year 2008—the year before the virus emerged in April 2009—to avoid any endogenous changes in travel patterns due to its spread.

3.3.1 Base measures of exposure to international travel and migration

We build on the methodology of the Global Transnational Mobility Dataset (GTMD), which estimates bilateral flows between 2011 and 2016 (Recchi et al. 2019). We construct our own estimates in order to expand the number of countries in the sample, estimate flows for 2008, separate non-citizen and citizen arrivals, and modify the methodology. Our estimates are highly correlated with the GTMD, and our results on the onset date and final size of the H1N1 pandemic are robust to using the GTMD estimates from 2011. We summarize our data and methodology here and provide a detailed description in the Appendix.

First, we use data on arrivals of non-resident travelers collected from destination governments and standardized by the United Nations World Tourism Organization (UNWTO 2020a). UNWTO publishes yearly arrivals, disaggregated by nationality or country of origin, for 176 countries to all ports of entry starting in 1995.[18] Most countries, however, do not report arrivals from all nationalities, and forty-two countries did not submit any arrival data for 2008. Overall, 26% of cells in the matrix of bilateral flows for 2008 can be populated by the data from UNWTO.

We therefore impute missing values using a random forest algorithm, trained with the UNWTO data and supplementary country- and bilateral-level datasets. We use 16 variables for training, including the flow of travelers in the reverse direction when available, population sizes, GDP per capita, direct flights, migration flows, geographic and linguistic distance, and historical ties. The accuracy of the prediction against a testing dataset suggests the algorithm has strong predictive power, as discussed in the Appendix. The imputations account for 4% of the global estimated

[18] These data are published by UNWTO as Outbound Tourism for each origin country, and we append and reorient for arrivals to the destination country, as they were originally supplied. Tourists are usually defined as non-residents staying at least one night and less than one year, although a few countries also count day trips without an overnight stay. Most of the data are reported by travelers' nationality, but some are supplied by country of origin. Further details are available in the Appendix.

flows, with a median of 6% and mean of 30% at the country level. Overall, while many bilateral observations are missing, most of the unreported data are likely to come from nationalities with few arrivals when the country reports any data. Our results are robust to using only the data and countries reported by the UNWTO.

The UNWTO data and imputations provide estimates of all non-resident short-term arrivals. For returning residents, we follow the GTMD methodology and assume that all these arrivals return to their country of residence in the same year. Summing resident and non-resident arrivals in each cell therefore creates a symmetric matrix of total incoming short-term flows.

Since the UNWTO data exclude migration, we also incorporate recently published estimates of migration flows (Azose and Raftery 2019). The authors use data on migrant populations from the United Nations and regional estimates of migration flows to model and estimate global flows across five-year periods, including both immigrants and returning citizens. We divide their estimates for the 2005-2010 period by five to estimate migrant flows in 2008. As migration flows are often not equal within country pairs, the final matrix of bilateral flows is asymmetric.

Our main measure of mobility is then the sum of incoming non-residents for short-term travel, residents returning from short-term travel, incoming immigrants, and citizens returning from emigration. While these are not true counts and depend on the modeling and standardization assumptions, we believe they represent the best estimates for mobility in 2008 and that the measurement error does not affect our conclusions.

The mobility data are summarized globally and for the top 20 countries in Table A1. We estimate over 2 billion international trips in 2008, or approximately 0.3 trips per capita. China, Germany, the United States, and Hong Kong recorded over 100 million international trips each, with the average country just under ten million, and the median country estimated at 1.7 million. The composition of foreign arrivals and returning citizens vary significantly; while 80% of China's international trips are foreign arrivals (notably including Hong Kong and Macao), 83% of Germany's international arrivals are German residents returning from abroad. We estimate less than 1% of global travel was for long-term migration; in most countries, incoming migrants represent a very small fraction of the total international travelers who may be carrying a disease.

3.3.2 Centrality Weights

Aggregating mobility data at the country level implicitly assumes that each traveler is equally likely to carry a disease. This is a reasonable approximation for the first period that an infected person travels internationally, when anyone could be patient zero. However, our Prediction 1 says that once the virus starts spreading abroad, travelers from countries with more incoming travel are more likely to be infected in the intermediate stages on average—regardless of where it originated—until it spreads everywhere. Our goal is therefore to weight arrivals in the empirical model by their centrality in the global network to accurately capture the potential risk each traveler poses.

We base our centrality measure on the proportion of global travel that arrives in the country, consistent with the idea that in the first stage, all travelers are equally likely to be infected. We then weight this proportion by the proportions of their origins' inflows, and weight the origins' inflows by their origins' inflows, and so on. This product of matrices converges to the eigenvector of the matrix of the proportion of global flows.[19]

However, in order to better represent the likelihood of importing the disease in the initial stages of an epidemic, we leave out each country's top origin country from this matrix of inflows.[20] Most countries' inflows are heavily concentrated from one or two neighbors. For instance, 54% of arrivals to China come from Hong Kong, and 83% of arrivals to Hong Kong come from China - countries with the most (China) and fourth-most (Hong Kong) total international inflows. Weighting inflows by the origins' inflows repeatedly counts flows within these clusters, but centrality is intended to measure global exposure beyond these regional clusters or large bilateral flows - diversity as well as volume. If most travel is within a cluster, and the disease does not originate in the cluster, this within-cluster movement represents less potential disease exposure than a hub with a similar volume of traffic from many origins.

The concentration, and therefore importance of excluding the top origin country, is evident from

[19] Banerjee et al. (2013) show that their related measure, Diffusion Centrality, also converges to eigenvector centrality.

[20] We replace the inflow with 0 and re-calculate the proportion of global travel, so that the sum of all bilateral flows is still 1. However, since the denominators are the same for every cell in the matrix, they do not affect the eigenvector measure.

the empirics. Let λ_i represent the eigenvector centrality from excluding i top origin countries and scaled to a mean of 1. When all origin countries are included (λ_0), the highest centralities are China at 65.7, Hong Kong at 61.2, Macao at 27.8, and Japan at 4.66. Leaving out the top origin country (λ_1) - our preferred centrality measure - the highest centrality weights are Germany at 23.4, France at 21.2, and Spain at 17.8.[21] The correlation between λ_0 and λ_1 is 0.12, showing the significant dependence of the measure on one origin. However, the correlation between λ_1 and λ_2 (when the top two origin countries are dropped) is 0.98, and the correlation between λ_1 and λ_5 is 0.91. This "leave-one-out" centrality (λ_1) provides a measure that is less reliant on one neighbor and better captures global exposure to an epidemic. It therefore represents our preferred centrality weight in the regressions below, and we include additional specifications with λ_0—no origin countries excluded—in the appendix.

4 International mobility and arrival delay

We first test Prediction 1, that differences in international mobility have a negative and small correlation with differences in pandemic arrival dates. We perform two separate tests, each using a different source of variation in international mobility: across countries or over time.

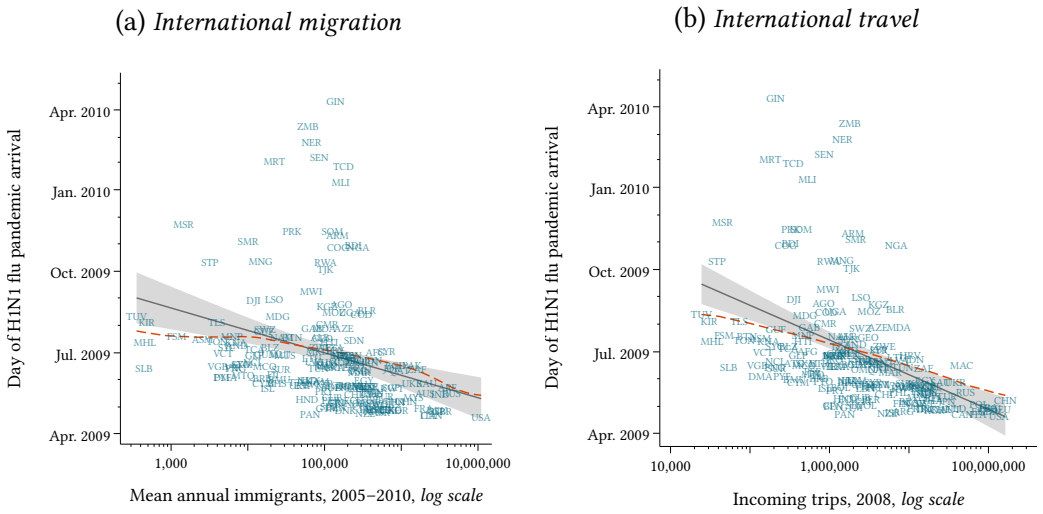
4.1 Differences in mobility across countries: Influenza 2009

For the 2009 influenza pandemic, we have reliable country-level data on relative exposure to international mobility. We can then compare this exposure to the date of pandemic arrival. Figure 3a plots each country's average immigration per capita during 2005–2010, on the horizontal axis, against the day of pandemic arrival, on the vertical axis. The solid line shows a linear OLS fit with a 95 percent confidence interval on the predicted mean. The dashed line shows a nonparametric Nadaraya-Watson (1964) kernel regression. Figure 3b repeats the exercise using incoming trips per capita.

Countries at the right edge of the figures have 1,000 times as much per-capita exposure to immigration, and 10,000 times as much exposure to incoming trips, as countries at the left edge.

[21] For comparison, China is the ninth most central at 4.2, while Hong Kong is 35th at 1.2 and Macao is 57th at 0.49.

Figure 3: INFLUENZA 2009: International mobility and the timing of pandemic arrival



Solid line shows OLS fit surrounded by 95% confidence interval of the predicted mean, in gray. Dashed line shows nonparametric Nadaraya-Watson (1964) regression, Epanechnikov kernel, bandwidth 1.5 natural log points (0.651 $\log_{10}$ points). Country labels are ISO-3166 $\alpha 3$ codes.

This is associated with a difference in pandemic arrival time of about 4 months (linear fit) or 1–2 months (nonparametric fit).

The figure implicitly treats arrivals from all countries as equally capable of transmitting disease. To relax this assumption, the top panel of Table 1 shows the linear regression coefficient estimates with and without weighting by our measure of country connectivity. We weight countries based on the proportion of global travel that arrives in the country, weighted by the proportions of their origins' inflows, and so on, as described in Section 3.3.2—given that network diffusion can affect arrival time at any give node (Gautreau et al. 2008). The coefficient estimates with and without weights are statistically indistinguishable. With these connectivity weights, the estimates imply that a 50% reduction in exposure to immigration is associated with a delay in pandemic arrival of 10.4 days (95% confidence interval 7.50–13.3 days), and a 90% reduction is associated with a delay of 34.5 days (95% c.i. 24.9–44.1 days). A 50% reduction in incoming trips per capita is associated with a delay in pandemic arrival of 11.8 days (95% c.i. 9.19–14.4 days), while a 90% reduction is associated with a delay of 39.1 days (95% c.i. 30.5–47.7 days).[22] The

[22]The general magnitude of these estimates is robust to separating the analysis by hemisphere to account for the inherent seasonality of influenza spread. These results are presented in the Appendix.

Table 1: Influenza 2009: International mobility and the timing of pandemic arrival

Dep. var.:	Day of pandemic arrival							
Mobility measure:	*ln* Immigrants				*ln* Incoming trips			
Connectivity weights?	No		Yes		No		Yes	
	est.	*s.e.*	*est.*	*s.e.*	*est.*	*s.e.*	*est.*	*s.e.*
Linear OLS								
ln Mobility	−11.0	(1.66)	−15.0	(2.13)	−16.8	(2.04)	−17.0	(1.91)
N	184		184		184		184	
R^2	0.102		0.206		0.226		0.297	
Days of delay associated with mobility reduction								
50% reduction	7.62	(1.15)	10.4	(1.48)	11.7	(1.41)	11.8	(1.33)
90% reduction	25.3	(3.83)	34.5	(4.90)	38.8	(4.69)	39.1	(4.40)
Survival regression: Loglogistic accelerated failure time								
ln Mobility	−0.120	(0.0140)	−0.149	(0.0180)	−0.158	(0.0134)	−0.159	(0.0123)
γ	0.258	(0.0166)	0.235	(0.0153)	0.233	(0.0153)	0.213	(0.0145)
N	184		184		184		184	
Days of delay associated with mobility reduction								
50% reduction	7.83	(0.951)	9.87	(1.25)	10.5	(0.937)	10.5	(0.861)
90% reduction	28.7	(3.83)	37.1	(5.28)	39.8	(4.01)	39.9	(3.69)

Observations are countries. Robust standard errors in parentheses to the right of each coefficient estimate. Constant term included but not shown. In the loglogistic survival regressions, γ is the scale parameter such that the survivor function is $S(t) \equiv (1 + (\lambda t)^{1/\gamma})^{-1}$ and $\lambda_i \equiv e^{-\beta \cdot \ln(\text{mobility})}$. 'Immigrants' is the average annual number of immigrants to each country during 2005–2010. 'Incoming trips' is the number of people arriving in each country, for any duration of stay, in 2008.

broad magnitudes of these results are consistent with the simulation in Figure 1b.

These estimates measure the relationship between overall exposure to international mobility during the interpandemic period and the arrival of a novel pandemic. They need not reflect the correlation with emergency changes to immigration or travel during the pandemic. However, their magnitude closely matches simulation-based estimates of the relationship between emergency travel restrictions and pandemic arrival using a calibrated global epidemiological-demographic-mobility model. Bajardi et al. (2011) simulate that emergency 50% cuts to incoming travel *from Mexico* at the start of the 2009 pandemic would have delayed pandemic arrival by only 4–9 days relative to no restrictions, and 90% cuts by 15–21 days.[23]

The assumption of a linear relationship in the top panel of Table 1 is restrictive. The bottom

[23]This overall delay sums the Bajardi et al. (2011) estimates for the change from 0% to 40% restriction—three days—and the simulated effects of *further* restrictions to 50% or 90% in their Figure 4. Tizzoni et al. (2012, 12) identically estimate an expected 20-day arrival delay from eliminating 90% of international travel from Mexico.

panel uses survival analysis to describe the same relationship with the loglogistic accelerated failure time model. This allows the hazard rate for arrival in a new country to vary over time, first rising then falling. The estimated changes in arrival time associated with a 50 or 90 percent reduction in exposure to international mobility are statistically indistinguishable from the values estimated in the linear specification.

We are not aware of systematic country-level data on exposure to international migration or travel before 1960. We thus do not carry out a similar exercise exploiting cross-country differences in mobility for the pandemics of 1957, 1918, and 1889. We can, however, study the earlier pandemics using a different source of variation in exposure to international mobility: over time.

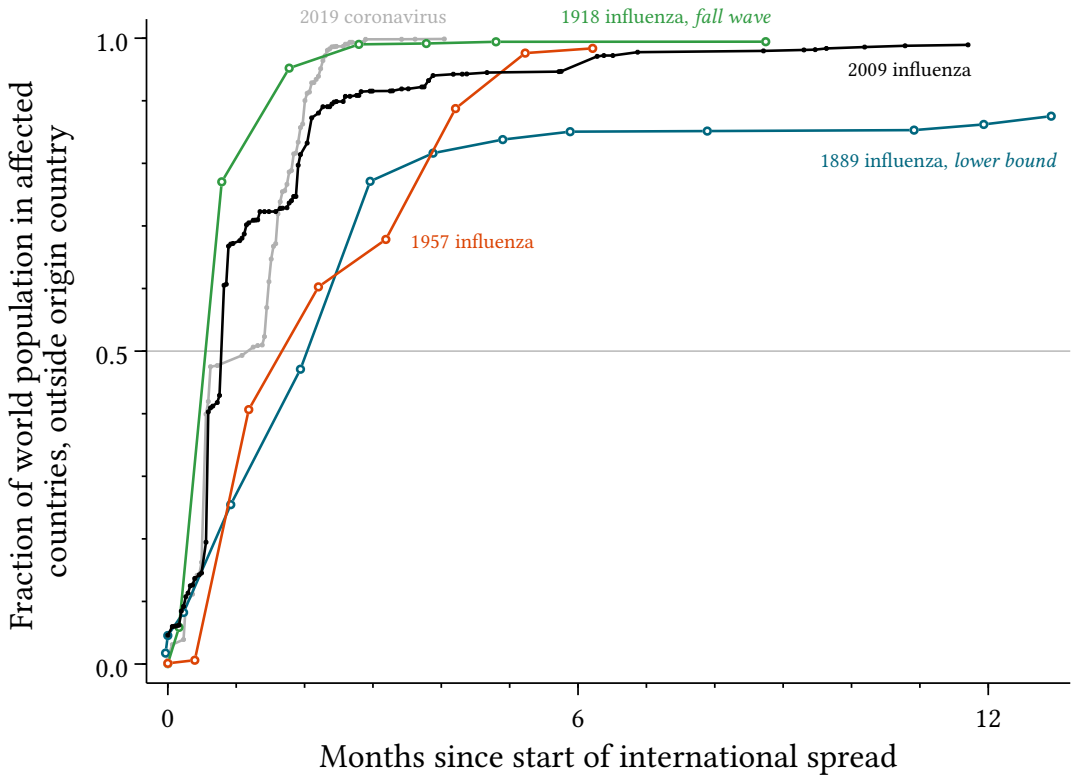
4.2 Differences in mobility over time: Pandemic spread from 1889 to 2020

By any measure, the amount and ease of international mobility has risen drastically over the past 130 years. Since 1910 alone, the global population of international migrants has risen by a factor of more than six, after accounting for changes in country borders.[24] In the United States since 1910, the number of overseas passenger arrivals rose by a factor of more than 77.[25] Global tourist arrivals rose from 25 million in 1950 (UNWTO 2018) to 1.5 billion in 2019 (UNWTO 2020b), a sixtyfold increase, while world population rose by a factor of only 3.1. From 1889 to 1997, the real cost of global freight transportation fell 94% (Shah Mohammed and Williamson 2004, 188). Since 1957 alone, the real cost of air transport per ton/kilometer has fallen by 90% (Hummels 2007).

[24] Between 1910 and 2019, the number of people living outside their country of birth or nationality rose from around 36 million to 272 million (McKeown 2004, 184; IOM 2020). This change does not account for the rise in the number of countries in the world since 1910. In the original source, Ferenczi (1937, 28) counted aliens resident in 97 countries in 1910, and by 2019 there were 193 member states in the United Nations. However, by 1930, Ferenczi's count includes aliens in 166 countries, and his estimate of total migrant population *fell* to 29 million, suggesting that even a large change in the number of countries did not have first-order effects on the measured migrant population. The events since 1930 that have added the most to the international migrant population without human mobility are the breakup of the Soviet Union and the partition of South Asia, which at the time they occurred added (respectively) about 30 million and about 8 million to the international migrant population UNDESA (2004, vii, 23). This suggests that even an estimate of global migrant population in 1910 fully adjusted for changes of borders since 1910 would be of a similar order of magnitude to that shown. An increase by a factor of six would be conservative.

[25] The number of passenger arrivals at US seaports in 1910 was 1.328 million (Source, p. 119). This excludes "travel over international land borders, crewmen, military personnel, and travelers between the United States and its outlying areas". In 2015, the number of air passenger arrivals from foreign countries was 102.3 million (Source), likewise excluding arrivals by land. Both numbers include both US citizens and foreign citizens. The ratio of these two numbers underestimates the rise because the 2015 number does not include arrivals by sea.

Figure 4: Five Pandemics: The rate of international spread to populations, 1889–2020



Populations are the number of people within the area covered by each country with borders held constant as they stood on June 1, 1991. The country of origin for each pandemic is omitted, as the goal is to compare rates of international spread. Dates for influenza 2009 and COVID-19 are daily. Dates for influenza 1889, 1918, and 1957 are unbiased estimates of start day binned by calendar month disregarding day-of-month, because day-of-month was not recorded for some countries. Thus, for all calendar months after the calendar month of first international spread, arrival dates for 1889, 1918, and 1957 are binned at the 15th of the month. Arrival dates within the first month of international spread are binned at the expected day-of-month conditional on that day being after the first day of spread. For example, in 1889, day zero of international spread is November 17, 1889, the first date the virus is recorded outside its origin country (the Russian Empire), in France. The virus arrived in both Belgium and Libya in December, and both are binned at December 15th since the day-of-month is known for Belgium (the 13th) but unknown for Libya. The day of arrival for Austria is known to be in November (the 30th), the same month as the start, and known to be after the first date of November 17th. Thus Austria is binned at November 24th—an unbiased estimate of the date if only the month were known but it were known to have followed France.

How would this affect the timing of international pandemic spread? The estimates in Table 1 can provide a naïve benchmark. If the correlations in the table reflect causal relationships, they imply that in a world with an order of magnitude less mobility, the arrival of a pandemic biologically identical to the 2009 influenza pandemic would have been delayed in the average country by one or two months.

We can compare that benchmark to the actual rate of international spread in the three earlier pandemics. Figure 4 carries out this exercise. The vertical axis shows the fraction of the world population living in countries reached by each pandemic as it progressed, holding international borders constant at their location in mid-1991. The horizontal axis is the number of days since the start of international spread—day zero is the first day the virus is reported outside its country of origin. For the historical pandemics of 1957, 1918, and 1889, the number of days is binned by month to account for the fact that arrival is reported by month only for some countries. For reference, the figure displays the same curve for the 2019 coronavirus pandemic.[26]

Consider the arrival delay for the country of the median person on earth living outside the country where each pandemic began. This is the intersection of each curve with the horizontal line in the figure, at 0.5. For the 1918 and 2009 influenza pandemics, by 2–3 weeks after the start of international spread, the median person living outside the country of pandemic origin was already living in a country reached by the virus. For the 2019 coronavirus pandemic, this occurred five weeks after the start of international spread. For the 1957 influenza pandemic it was six weeks, and for the 1889 influenza pandemic, less than eight weeks. For the median person, the delay until international spread brought the pandemic to their country varied by only six weeks even as international migration and travel exploded across 130 years of technological change.

Figure 5 presents the same data, showing the geographic extent of international spread rather than the population coverage of international spread. The country of origin of each pandemic is shown in black.[27] International borders are again fixed at their location in mid-1991. The elapsed time again takes day zero as the first day the virus is reported outside the country of origin. Again the 2019 coronavirus pandemic is juxtaposed for reference (Figure 5e). It is apparent from the four influenza pandemics that once international spread began, the virus had in each case

[26] The curve for 1889 is marked as a lower bound because the arrival date is unknown for a substantial number of countries; the true curve thus lies strictly above the one shown.

[27] For the pandemics of 2009 and 1957 the country of origin is uncontroversial. For the pandemic of 1918, we adopt the interpretation of Patterson and Pyle (1991, 5, 8) that the spring wave began in the United States on or around March 5 and the fall wave began in Brest, France, close to August 22, though particularly the spring-wave origin location and precise timing are contested (e.g. Langford 2005; Erkoreka 2009). For the pandemic of 1889, while there were reports of potentially earlier and related flu outbreaks in Greenland and Athabasca (northern Alberta, Canada), the consensus in the literature is that the most reliable report of the pandemic's origin comes from Bukhara, Russian Empire (now Uzbekistan) in the first week of June 1889 (that is, the second half of May 1889 by the 'old-style' calendar still used at the time in the Russian Empire (Parsons 1891b, 14).

circled the globe within four months. In 1889 and 1957, with the caveat of data limitations for 1889, notably fewer countries were reached within *two* months than in the other pandemics. Here again, as international mobility rose by an order of magnitude, the speed of international spread appears to shift on the order of one or two months.

This evidence would corroborate a causal interpretation of the estimates in Table 1 *if* the biological traits of the viruses in each influenza pandemic were identical. They are not identical. But they are similar. First, the four pandemics are believed to have been caused by only one type (A) and two subtypes (H1N1 and H2N2) of influenza virus. The pandemics of 1918 and 2009 were both caused by subtype H1N1, while the pandemics of 1889 and 1957 were probably both caused by subtype H2N2 (Hilleman 2002; Monto and Fukuda 2020).[28] In nature there are four types, and type A alone has 131 known subtypes (Fukuyama et al. 2020). Other types (especially B) and subtypes (especially A/H3N2 and A/H5N1) have been important in other episodes of international influenza spread (Biggerstaff et al. 2014).

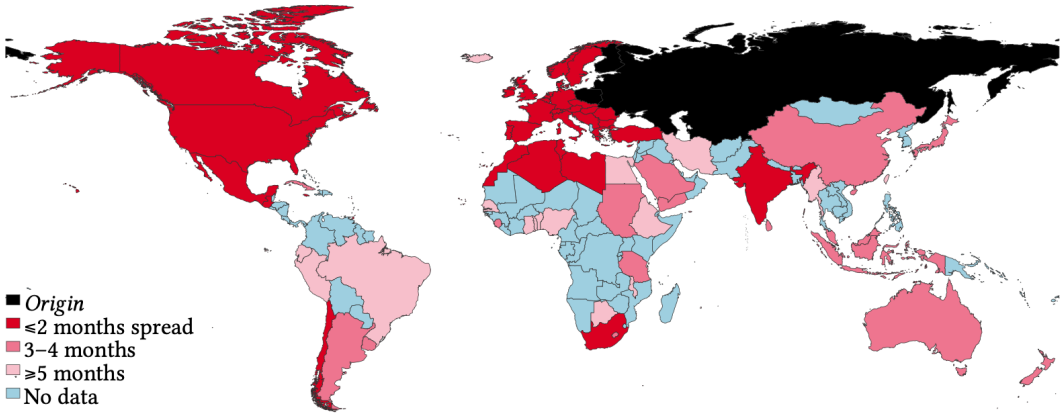
This biological similarity among the pandemics is reflected in their observed transmissibility. The basic reproduction number $\mathscr{R}_0$ lies in a relatively narrow range for all four pandemics: For the 2009 pandemic, it has been estimated at 1.2–2.3 (Boëlle et al. 2011) and 1.3–1.7 (Biggerstaff et al. 2014). For the 1957 pandemic estimates are in the range 1.5–1.7, and for 1918 in the range 1.5–2.3 (Biggerstaff et al. 2014). The two extant studies of the 1889 pandemic find $\mathscr{R}_0$ of 1.3–1.5 (Ramiro et al. 2018) and 1.9–2.4 (Valleron et al. 2010). That is, the inherent transmissibility of the 2009 virus was only "slightly smaller" than in the prior three influenza pandemics (Boëlle et al. 2011).

Controlling for this minor difference, Figure 4 suggests, would leave the above result qualitatively unchanged. Had the inherent transmissibility of the 2009 virus not been slightly smaller than in the prior pandemics, the median person's country might have been reached by roughly two weeks after the start of international spread, rather than three. In that case, the curve for 2009 would overlap with the curve for 1918, and for the median person, arrival time for all four

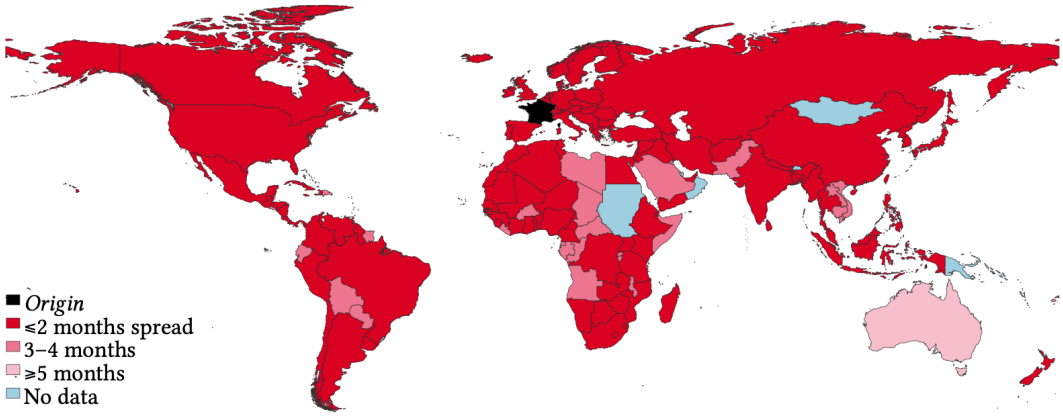
[28] No tissue samples from the 1889 pandemic survive to genetically test the hypothesis of a common subtype in the 1957 and 1889 pandemics, but exposure to the 1889 pandemic generated immune response against the 1957 pandemic (Mulder and Masurel 1958). The evidence in the literature does admit the possibility that the 1889 pandemic was caused by influenza A subtype H3N8 (e.g. Biggerstaff et al. 2014). Vijgen et al. (2005, 1603) even find it "tempting to speculate" that the 1889 virus was in fact a zoonotic coronavirus.

Figure 5: Five Pandemics: The rate of international spread to territories, 1889–2020

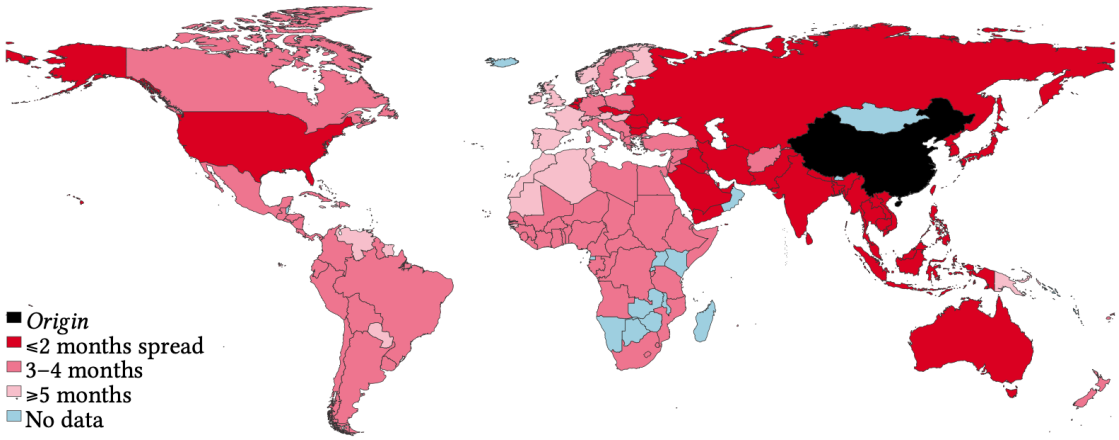
(a) *1889 influenza*



(b) *1918 influenza, fall wave*



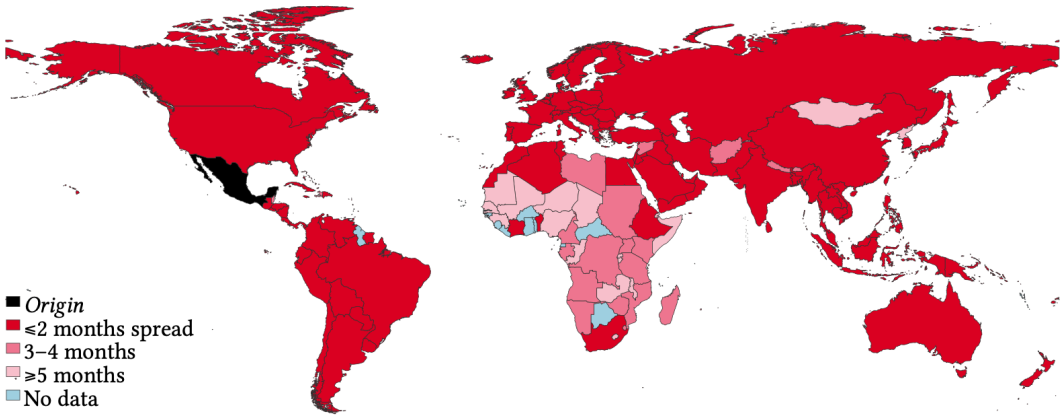
(c) *1957 influenza*



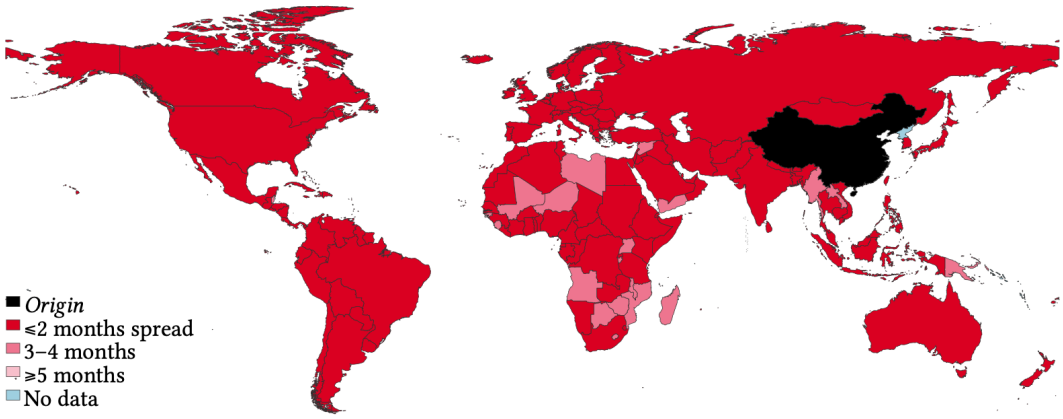
Continued on next page

Figure 5: *Continued,* Five Pandemics: The rate of international spread to territories, 1889–2020

(d) *2009 influenza*



(e) *2019 coronavirus disease*



International borders are fixed at their location on June 1, 1991 (Weidmann et al. 2010). Time zero represents the start of international transmission, the first local transmission of the virus in a country outside the country of origin. "⩽2 months" defined as the virus reported present within the territory covered by 1991 borders between 0 and 67 days in expectation (inclusive) after the beginning of international spread (the date first reported outside the country of origin). "3–4 months" defined as 68 to 128 days (inclusive) after. "⩾5 months" defined as 129 days or more.

influenza pandemics would still fall within a six-week window. The effect of an order of magnitude increase in global human mobility between the pandemics would still have an affect on the speed of international spread of roughly the same magnitude that would be predicted by a naïve extrapolation from the estimates of Table 1.[29]

Collectively, this evidence is consistent with Prediction 1: Reductions in exposure to international mobility delay pandemic arrival, but even drastic reductions delay arrival by weeks at most. Tognotti (2013, 257) states that the 1918 pandemic spread so quickly because "travel restrictions" and "border controls . . . were impractical, during a time when the movement of troops was facilitating the spread of the virus." But the evidence here indicates that any such effect was minimal: In the other pandemics where troop movements were not a major contributor, for the median person on earth, country-to-country spread was only slowed by a few weeks.

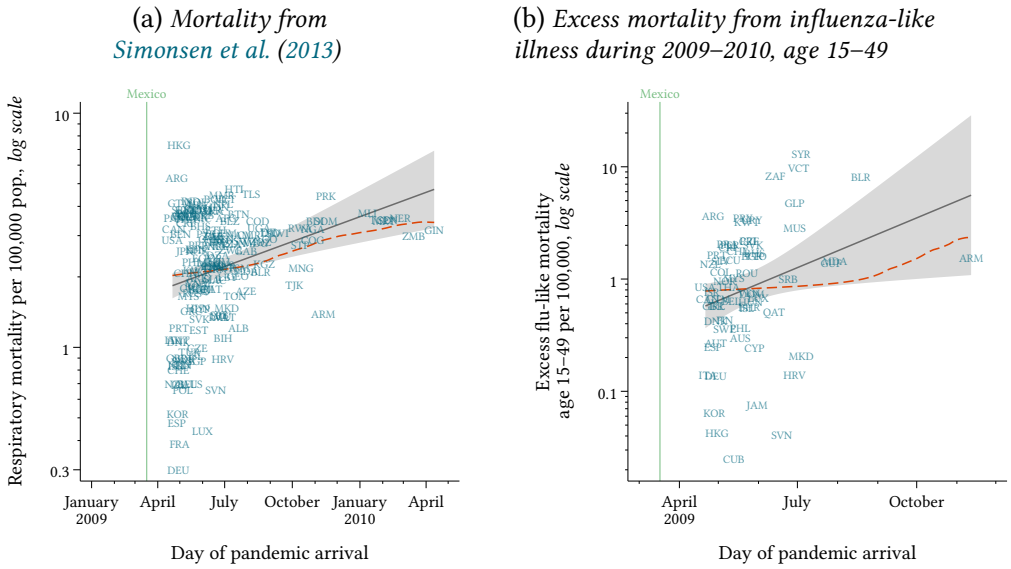
5 Arrival delay and final size

Having found a small, negative correlation between exposure to pre-pandemic international mobility and the arrival date, we now proceed to test for a relationship between arrival date and the overall harm of the pandemic. From Prediction 2, we expect no correlation between each pandemic's arrival date and its final size in each country.

These tests will be indirectly informative about the relationship between *international mobility* and final size. If we cannot detect a positive relationship between earlier arrival and larger final size, then if greater international mobility does increase final size, it would have to do so by a mechanism that does not also accelerate the arrival of the pandemic. It is not obvious what such a mechanism would be. On the other hand, a clear positive relationship between early arrival and greater harm would leave open the possibility of a positive effect of mobility on final size. Such an indirect test is feasible for all four influenza pandemics. We begin with the most recent one.

[29] In the limiting case, a more transmissible variant of the 2009 virus could not have reached the median person's country in less than a day, and even then the arrival delay for the median person would only have varied by less than eight weeks across all five pandemics.

Figure 6: Influenza 2009: Date of pandemic arrival versus final mortality



Solid line shows OLS fit surrounded by 95% confidence interval of the predicted mean, in gray. Dashed line shows nonparametric Nadaraya-Watson (1964) regression, Epanechnikov kernel, bandwidth 60 days. Country labels are ISO-3166 $\alpha 3$ codes.

5.1 Influenza 2009

Figure 6 shows the simple relationship between arrival date and final mortality for the 2009 influenza pandemic. The bivariate relationship is strong and positive, using both measures of mortality. Table 2a shows the corresponding linear regressions, and shows the effects of dropping the countries for which Simonsen et al. impute mortality, as well as including the same country-level control variables used in Table 5. The estimates in columns 2–5 are statistically indistinguishable from zero. With covariates included, the standard error on day-of-arrival is roughly equal in magnitude to the coefficient estimate.

Table 2b relaxes the assumption of a linear hazard rate and recasts the tests as survival regressions with the loglogistic accelerated failure time model allowing a time-varying hazard. Here again the coefficients in columns 2–5 cannot be statistically distinguished from zero. All of the results in Figure 6 suggest that the positive relationship seen in Figure 6 could be driven by unobserved country heterogeneity, but offer no evidence that anything causing later arrival (such as

reduced exposure to mobility) was sufficient to reduce final mortality. This indirect test corroborates the direct tests of the mobility-final size relationship in 2009 from Section 6. Collectively, these results fail to reject the null of no correlation between arrival date and final mortality in 2009.

Beyond this, the economic importance of the coefficient estimates in Table 2a is extremely small. First, most of the coefficient estimates for the relationship between delay-until-arrival and final mortality are positive. Second, in the one specification yielding a negative coefficient estimate (using the Simonsen et al. (2013) directly-reported countries), the magnitude of the statistically-insignificant coefficient is so small that the implied effect would be dwarfed by the effect of countless other public health measures. We illustrate this by translating the relationship between arrival time and final size in the table into an equivalent relationship with $\mathscr{R}_0$. Equation (10) implies that $\mathscr{R}_0\big|_{p=0} = \frac{1}{Z}\ln\left(\frac{1}{1-Z}\right)$, so

$$\frac{d\mathscr{R}_0}{dZ} = \frac{1}{Z(1-Z)} + \frac{\ln(1-Z)}{Z^2}. \tag{11}$$

Thus for example, the only negative coefficient in Table 2a, in the third regression, implies that an additional day's delay in arrival is associated with a change in mortality of −0.00311 per 100,000. Assuming an infection fatality rate of 7.6 per 100,000 infections (Riley et al. 2011), this implies that a one-week delay in arrival time—requiring a permanent 50 percent reduction in exposure to international mobility—is associated with a change in final size of −0.00286. By equation (11), this is the same change in final size that would be generated by a change in $\mathscr{R}_0$ of 0.00144.

Such an effect on $\mathscr{R}_0$ can be achieved at far less economic cost than the cost of a permanent 50 percent reduction in international mobility during the interpandemic period. Regulations on economic activity inside a country's borders, after a pandemic hits, can have multiple orders of magnitude higher impact. For example, domestic public health interventions during the 1918 pandemic in the United States (such as mask mandates, school/church closures, and case isolation) typically reduced $\mathscr{R}_0$ by 0.6–0.8 (Bootsma and Ferguson 2007).

Table 2: Influenza 2009: Start date versus final mortality

(a) *Linear OLS*

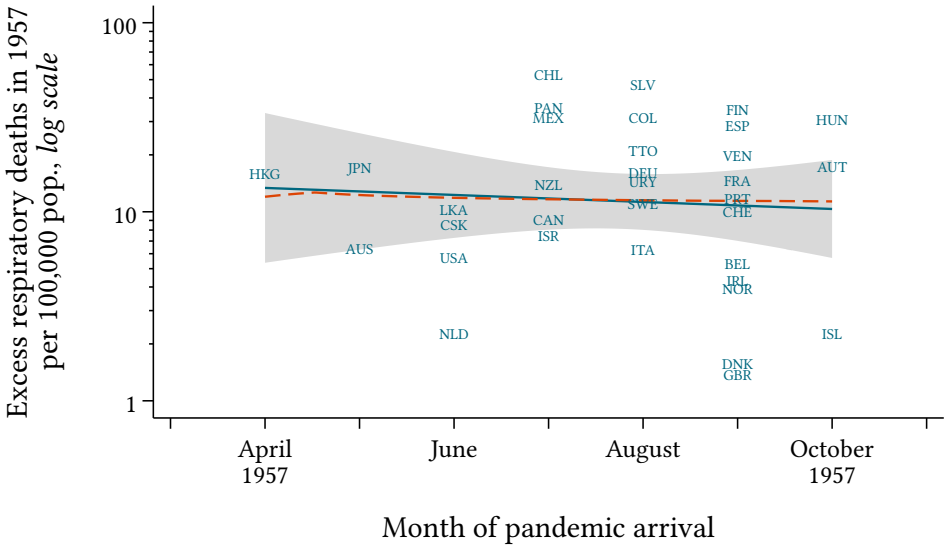
Dep. var.:	*ln* Final mortality per 100,000									
Mortality measure:	Simonsen et al. (2013)						Excess ILI mortality, 15–49			
Directly reported?	No		Yes							
	est.	*s.e.*	*est.*	*s.e.*	*est.*	*s.e.*	*est.*	*s.e.*	*est.*	*s.e.*
Day of arrival	0.00266	(0.000523)	0.00814	(0.0106)	−0.00311	(0.0122)	0.0111	(0.00481)	0.00845	(0.00478)
ln GDP/cap., PPP					−0.111	(0.634)			0.487	(0.493)
Urbanization					1.91	(1.57)			0.900	(1.26)
Healthcare access					−0.0442	(0.0237)			−0.0662	(0.0206)
Health exp./GDP					0.0148	(0.0872)			0.0527	(0.0815)
N	162		25		25		64		62	
R^2	0.0941		0.0242		0.424		0.0816		0.273	

(b) *Survival regression: Accelerated failure time, loglogistic*

Dep. var.:	Day of arrival									
Mortality measure:	Simonsen et al. (2013)						Excess ILI mortality, 15–49			
Directly reported?	No		Yes							
	est.	*s.e.*	*est.*	*s.e.*	*est.*	*s.e.*	*est.*	*s.e.*	*est.*	*s.e.*
Mortality	0.328	(0.0625)	0.0486	(0.0595)	−0.0281	(0.0365)	0.0982	(0.0400)	0.0610	(0.0401)
ln GDP/cap., PPP					0.403	(0.129)			−0.0495	(0.0932)
Urbanization					−1.13	(0.433)			−0.490	(0.361)
Healthcare access					−0.0188	(0.00452)			0.00258	(0.00502)
Health exp./GDP					−0.0434	(0.00851)			−0.047	(0.0201)
γ	0.284	(0.0181)	0.154	(0.0283)	−0.0188	(0.00452)	0.207	(0.0215)	0.00258	(0.00502)
N	162		25		25		64		62	

Robust standard errors in parentheses to the right of each coefficient estimate. Constant term included but not shown. "Healthcare access" is the Healthcare Access and Quality Index (Barber et al. 2017) for 2010, a 1–100 principal-component score derived from mortality outcomes adjusted for cause-specific risk, where 100 represents the lowest mortality. 'ILI' is influenza-like illness. 'Directly reported' countries from Simonsen et al. (2013) do not include Mexico, the pandemic-origin country. In the loglogistic survival regressions, γ is the scale parameter such that the survivor function is $S(t) \equiv (1 + (\lambda t)^{1/\gamma})^{-1}$ and $\lambda_i \equiv e^{-\beta \cdot \ln(\text{mobility})}$.

Figure 7: Influenza 1957: Date of pandemic arrival versus final mortality



Solid line shows OLS fit surrounded by 95% confidence interval of the predicted mean, in gray. Dashed line shows nonparametric Nadaraya-Watson (1964) regression, Epanechnikov kernel, bandwidth 60 days. Final mortality estimates from Viboud et al. (2016), excluding incommensurable estimates for subsets of countries. Figure for Ireland includes Northern Ireland.

5.2 Influenza 1957

The so-called 'Asian flu' pandemic of 1957 began in the Kweichow province of Southern China, now Guizhou, in February (Cox and Subbarao 2000, 413), spreading quickly around the world via Hong Kong and Singapore. We perform a corresponding test of Prediction 2 in the 1957 pandemic for the 35 countries where final mortality has been estimated by Viboud et al. (2016) at the national level.

Figure 7 shows the bivariate relationship between arrival date and final mortality across countries for the 1957 pandemic. The relationship is indistinguishable from zero, in the linear regression (solid line) and across the distribution of observed start dates in the nonparametric regression (dashed line). The corresponding linear regression estimates are shown in Table 3a, where the standard error is twice as large as the coefficient estimate. This conclusion is unchanged whether the relevant mortality rate is taken as mortality in 1957 only (column 1) or in 1957–1959 total (column 3).

As in 2009, these bivariate correlations could be confounded by unobserved country traits that mediate between final size as measured by infections and final mortality. It is possible that countries where the virus arrived early happened to be those best able to prevent infection from resulting in death, due to their greater wealth and/or stronger health systems. The table therefore includes in columns 2 and 4 the most basic controls available for this time period to capture the level of development (GDP per capita) and health conditions (the child mortality rate). If the omission of such traits were confounding a strong negative, partial correlation in columns 1 and 3, we would expect the coefficient on arrival date to become more negative when these controls are included. The opposite occurs.[30]

We conclude that in 1957 as well, no country trait causing later arrival of the pandemic (such as unobserved reduced exposure to international mobility) was sufficient to cause statistically significant reductions in final mortality. The economic magnitude of the coefficient estimates is likewise extremely small. Following again equation (11), the coefficient estimate of −0.000654 from Table 3a, column 2, implies that a seven-day delay in the 1957 pandemic arrival, assuming an infection mortality rate of 7.6 per 100,000, would correspond to a reduction in $\mathscr{R}_0$ of 0.000302 (from this pandemic's base value of 1.5–1.7 estimated by Biggerstaff et al. (2014)).

5.3 Influenza 1918, fall wave and overall

The fall wave of the 1918 pandemic most likely began in Brest, France close to August 22.[31] From there it spread very rapidly to the rest of the world via England, Boston, and Sierra Leone. We test Prediction 2 in the 1918 pandemic for the 45 countries where fall-wave final mortality has been estimated by Patterson and Pyle (1991, 14). Figure 8a shows the bivariate relationship between arrival day and final mortality in each country. In the linear and nonparametric regressions, the relationship cannot be distinguished from zero. A case in point is Mauritius, which managed to delay the arrival by eight months, but suffered high mortality nonetheless.

The genetic makeup of the fall-wave virus was different enough from the spring-wave virus to

[30] A further robustness check, dropping the linear hazard assumption in favor of survival analysis with the loglogistic accelerated failure time model, likewise gives a null results and is presented in the Appendix.

[31] "The first reports were from Brest, a major Atlantic port and landing point for American troops, on 22 August" Patterson and Pyle (1991, 8).

Table 3: Influenza 1957, 1918, and 1889: Start date versus final mortality

(a) *Influenza 1957, countries: OLS*

Dep. var.:	*ln* Final mortality per 100,000			
Mortality measure:	1957 only		1957–1959	
Date of arrival	−0.0428	−0.000654	−0.0526	−0.0398
	(0.100)	(0.0782)	(0.0987)	(0.0892)
ln GDP/cap., PPP		−0.267		−0.689
		(0.299)		(0.418)
Child mortality		0.00967		−0.0034
		(0.00296)		(0.00619)
N	34	34	35	35
R^2	0.00473	0.433	0.00887	0.139

(b) *Influenza 1918, countries: OLS*

Dep. var.:	*ln* Final mortality per 100,000			
Wave:	Fall 1918 only		Spring/Fall 1918	
Date of arrival	0.00226	0.00210	0.00520	0.00227
	(0.00348)	(0.00236)	(0.00142)	(0.00153)
ln GDP/cap., PPP		−0.516		−0.402
		(0.162)		(0.200)
Child mortality		0.00307		0.00232
		(0.00114)		(0.00135)
N	45	40	50	45
R^2	0.00822	0.643	0.147	0.488

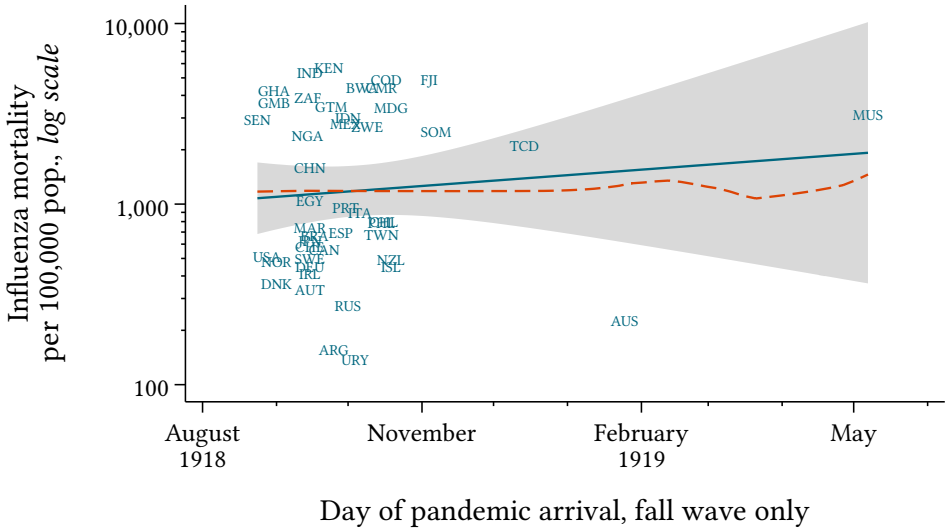
(c) *Influenza 1889, cities: OLS*

Dep. var.:	*ln* Final mortality per 100,000	
Date used:	Date of arrival	Peak mortality
Date	0.00757	−0.00383
	(0.00991)	(0.00705)
N	34	87
R^2	0.0615	0.00487

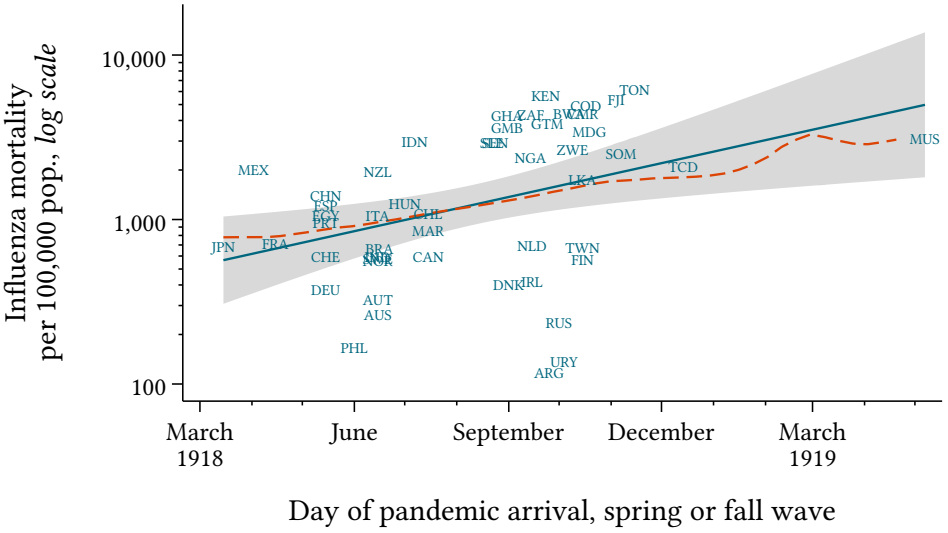
Robust standard errors in parentheses underneath each coefficient estimate. Constant term included but not shown. Covariates are predetermined (GDP per capita and child mortality are measured at 1956 for the 1957 pandemic, and at 1917 for the 1918 pandemic).

Figure 8: Influenza 1918: Date of pandemic arrival versus final mortality

(a) *Fall 1918 wave arrival and mortality only*



(b) *Spring and fall 1918 waves combined*



Solid line shows OLS fit surrounded by 95% confidence interval of the predicted mean, in gray. Dashed line shows nonparametric Nadaraya-Watson (1964) regression, Epanechnikov kernel, bandwidth 60 days. Countries where only month is known are imputed as the 15th of that month. In Figure 8a, mortality and start date are from the fall wave only. In Figure 8b, the start date is the first date that the pandemic arrived, regardless of whether that occurred in the spring or fall wave, and mortality is combined mortality from both waves.

cause far higher mortality (e.g. Sheng et al. 2011; Short et al. 2018), while similar enough that spring-wave infection conferred substantial immunity against fall-wave infection (Simonsen et al. 2018). It is therefore unclear whether the fall wave is properly considered to be a separate pandemic event. We test the robustness of the findings to relaxing that assumption.

There is likewise no sign of a negative bivariate relationship between arrival date and final mortality if we consider the pandemic of 1918 as a whole, combining the spring and fall waves in Figure 8b. Here, the start date is the first day that either the spring or fall waves reached the country, whichever was earlier, and the mortality figure combines spring- and fall-wave mortality. The sample expands by five countries for which a mortality estimate exists for all of 1918 but not for the fall wave specifically. There is a marked positive bivariate relationship between start date and final mortality: Countries that managed to escape the virus longer suffered much higher mortality.

The corresponding linear regression coefficient estimates are shown in Table 3b, in columns 1 and 3. These bivariate relationships, as above, could be confounded by unobserved country traits mediating between infection and mortality: For example, income per capita is known to explain almost half the cross-country variance in 1918 influenza mortality (Murray et al. 2006). Both of the prior estimates become statistically indistinguishable from zero with the addition of controls for the overall level of development (GDP per capita) and public health conditions (child mortality) in columns 2 and 4. The coefficient on arrival date in the fall-wave regression does not substantially change with the addition of these controls (column 2), suggesting that the bivariate correlation is not substantially confounded by the omission of the the most basic country traits that might mediate between the infection rate and resulting mortality. The coefficient on arrival date falls by half with the addition of these controls when the spring and fall waves are combined (column 4), but remains positive.

Again, we conclude that no country trait causing later arrival of the pandemic (such as unobserved reduced exposure to international mobility) was sufficient to cause detectable *reductions* in final mortality.

5.4 Influenza 1889, first wave

The 'Russian flu' of 1889 was first reliably reported in Bukhara, Russian Empire (now Uzbekistan) in June. From there is spread, via Warsaw and Saint Petersburg, rapidly around the world. We test Prediction 2 using the *city*-level final mortality figures from Valleron et al. (2010). They estimate final mortality for cities across Europe and North America, by comparing excess weekly all-cause mortality during the window October 4, 1889 to March 28, 1890 to the mortality during the first four weeks that each city is under observation. They include only large cities with anomalous mortality.[32]

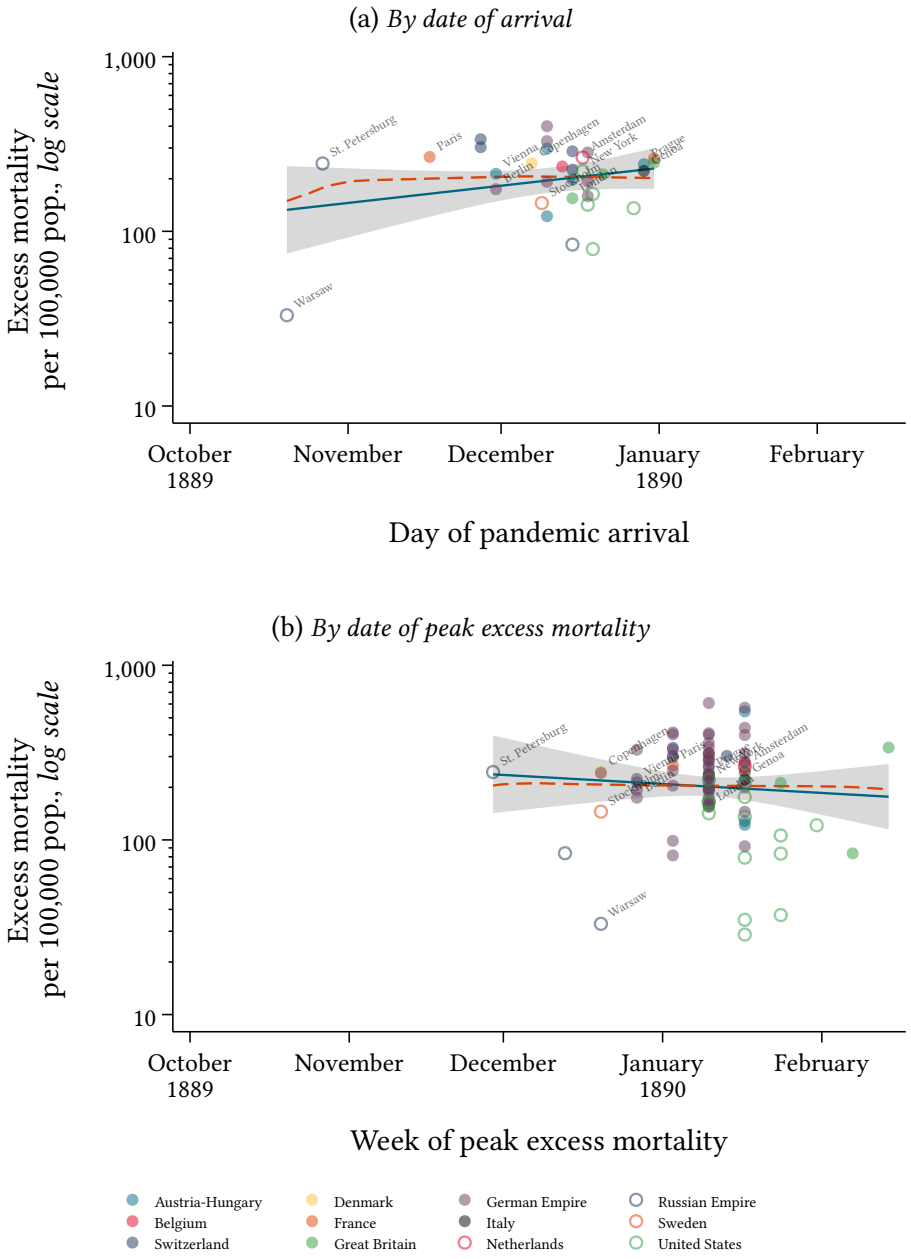
We consider only mortality in the first wave of the pandemic, because final mortality rates for the subsequent waves have not be systematically estimated. This requires stopping the analysis in mid-February 1890, when the second wave began (Parsons 1891a, 307), since the Valleron et al. data only cover the first few weeks of that wave.[33] Precise dates of pandemic arrival are only available for 34 of these cities (in Parsons 1891b and other sources discussed in the Appendix). As an alternative with more complete coverage, Valleron et al. recommend using the date of peak excess mortality as an indicator of relative timing. That is available for 87 cities meeting the above criteria.

Figure 9a shows the bivariate relationship between final mortality for the 34 cities where reliable dates of arrival are available (by day). In the country-level analysis of later pandemics above, we omitted the country of pandemic origin; in the city-level analysis here, we omit the city of pandemic origin (Bukhara, Russan Empire, now Uzbekistan). There is no relationship between final mortality and the arrival date, in the linear or nonparametric regression lines. Figure 9b

[32] Out of 172 cities with available mortality series, the source includes "the 96 cities with populations exceeding 35,000 inhabitants and a unique mortality peak during the period" (Valleron et al. 2010, 8789)

[33] To do otherwise, using the final six weeks of data collected by Valleron et al. (2010) for the initial period of the second wave would, could, for the present purpose, bias the results. For example, the city of Sheffield, England exhibited a relatively small local peak in excess deaths of 99 per 100,000 at the beginning of the second wave, just before the end of the window of observation in mid-March 1890. Including this initial period of the second wave would make Sheffield appear to exemplify a city where delayed arrival was associated with limited mortality. But this was just the beginning of the next wave, the second of several waves continuing through 1894 (Smith 1995). And in 1891, Sheffield suffered the highest influenza mortality rate seen in England during the entire pandemic (Parsons 1893, 27). Including this datum would suggest the opposite conclusion, that Sheffield was a city where delayed arrival was associated with very high mortality. Correct analysis requires considering either the entirety of subsequent waves or none of them. Unlike in the case of the 1957–1959 pandemic waves, no systematic mortality data for the second or subsequent waves of the 1889–1894 pandemic have been collected.

Figure 9: Influenza 1889: Date of pandemic arrival versus final mortality, first wave



Observations are cities. Excess mortality in each week is defined, as in Valleron et al. (2010), as all-cause all-age mortality in each week relative to baseline average mortality in the first four weeks each city is under observation (starting the week of October 4, 1889). Solid line shows OLS fit surrounded by 95% confidence interval of the predicted mean, in gray. Dashed line shows nonparametric Nadaraya-Watson (1964) regression, Epanechnikov kernel, bandwidth three weeks. Shows 1889 wave only, cut off at peak-date February 15, 1890 because the 1890 wave began in late February (Parsons 1891a, 307) and no final-mortality estimates are available for that wave. The origin of the pandemic is very likely to have been in Bukhara (now Uzbekistan) around the first week of June 1889 (that is, the second half of May 1889 by the 'old-style' calendar still used at the time in the Russian Empire (Parsons 1891b, 14))

shows the relationship between final mortality and the *peak* mortality date (by week). Again, no statistical relationship is discernible. The corresponding linear OLS regressions are shown in Table 3c. The standard errors in those regressions are comparable to or much larger than the coefficient estimates.

It is not possible for the 1889 pandemic to add control variables corresponding those used in the later pandemics (income per capita and indicators of health technology) because city-level indicators of that kind for this historical period do not exist. We note, however, that the plausible confounding effect of such unobserved traits diminishes as the geographic scope of the sample is reduced and the time period in question becomes earlier. For example, the 1957 analysis above includes Egypt, Sri Lanka, and Colombia alongside Western Europe and the United States. But these 1889 regressions include exclusively cities in relatively rich parts of Europe alongside a few in the United States, where the plausible cross-site variance in unobserved wealth and health system capability is relatively limited. Moreover, advances in health technology by 1957 made it possible for some countries to pull away from others, but we are not aware of very large differences in treatments used for influenza patients across rich European and U.S. cities in 1889. This diminishes, but need not strictly eliminate, the potential role of such traits as unobserved confounders.

We conclude that in the first wave of the 1889 pandemic, also, no country trait causing later arrival of the pandemic (such as unobserved reduced exposure to international mobility) was sufficient to cause detectable reductions in final mortality. In none of the four pandemics do we find evidence to reject Prediction 2.

6 International mobility and final size

Putting together the results of the previous two sections, we detect a small negative effect of exposure to international mobility on time-to-arrival of the four historical pandemics, and fail to detect any relationship between time-to-arrival and mortality. This implies that if exposure to international mobility does increase mortality, it does so by a mechanism that does not also accelerate the arrival of the pandemic. Because few such mechanisms are obvious, this lends

indirect support to Prediction 3: no relationship between exposure to international mobility and final size.

For the 2009 pandemic alone, we can test this prediction more directly. In that year we have measures of country-level exposure to international mobility. Prediction 3 implies the null hypothesis of zero correlation between each country's exposure to international mobility and the fraction of the population eventually infected. We do not directly observe the fraction infected, but rather mortality. For the 2009 pandemic, we can test the relationship between exposure to international mobility and pandemic mortality, conditional on key factors that causally mediate between infection and mortality.

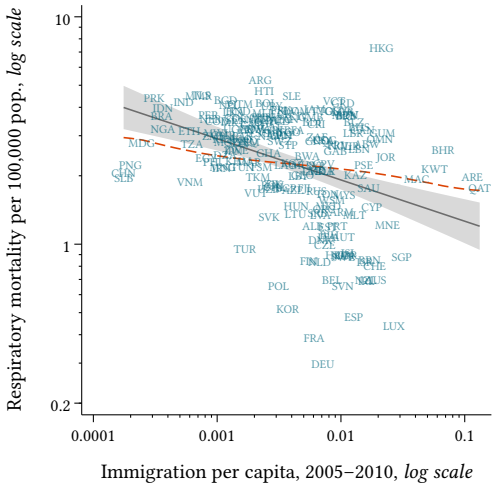
Figure 10 shows the simple cross-country correlation between exposure to international mobility and the mortality rate from the 2009 influenza pandemic. Figure 10a shows the relationship to immigrants per capita using the mortality estimates of Simonsen et al. (2013), while Figure 10b uses our independent estimates of excess prime-age mortality due to influenza-like illness in 2009–2010. Figure 10c and Figure 10d repeat the exercise for arrivals of international passengers per capita. In each figure the solid line shows a linear OLS fit surrounded by a 95 percent confidence interval on the predicted mean, while the dashed line shows a nonparametric Nadaraya-Watson (1964) regression that is robust to influential observations. There is no sign of a generalized positive relationship between exposure to mobility and pandemic mortality. If anything, the simple bivariate correlation is negative.

Table 4 shows the corresponding regression estimates. It carries out several robustness checks by presenting the results with international mobility measured as migration or travel, with mortality measured by Simonsen et al. or by excess deaths from influenza-like illness, with or without weighting countries by connectivity, with or without the countries whose mortality is imputed by Simonsen et al., and using nonparametric Spearman rank correlations rather than linear regression. The coefficient estimates remain negative in all of these permutations. The coefficient using the Simonsen et al. data is more negative when the countries without directly reported mortality data are omitted, and when our own excess mortality estimates are used.

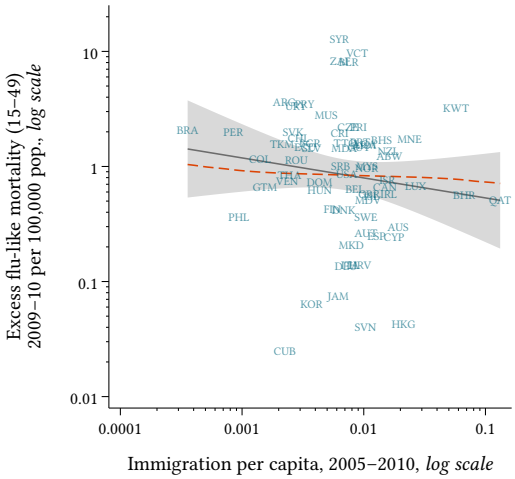
The connectivity-weighted coefficient estimates imply that a 50% reduction in exposure to either

Figure 10: Influenza 2009: International mobility and pandemic final size

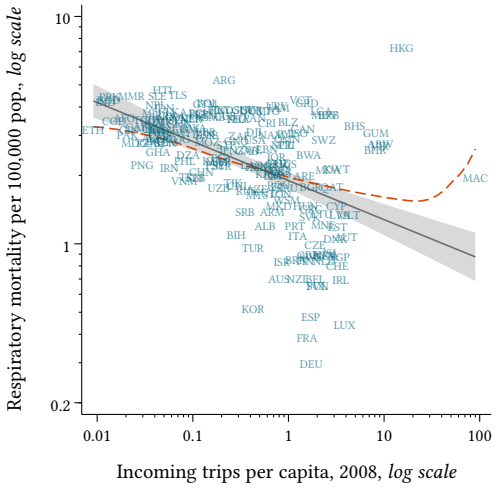
(a) *Immigration: Mortality from Simonsen et al. (2013)*



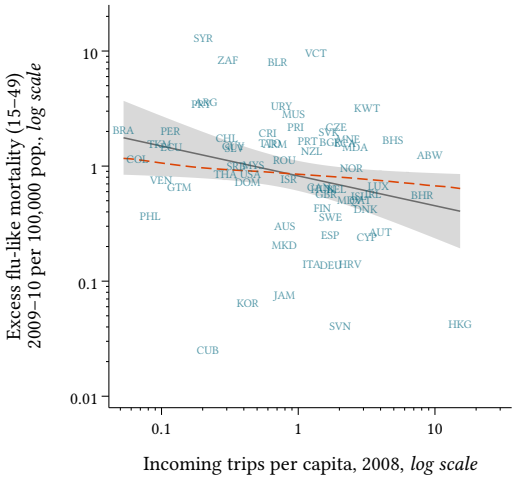
(b) *Immigration: Excess mortality from influenza-like illness 2009–2010, age 15–49*



(c) *Incoming trips: Mortality from Simonsen et al. (2013)*



(d) *Incoming trips: Excess mortality from influenza-like illness 2009–2010, age 15–49*



Solid line shows OLS fit surrounded by 95% confidence interval of the predicted mean, in gray. Dashed line shows nonparametric Nadaraya-Watson (1964) regression, Epanechnikov kernel, bandwidth 1.5 natural log points (0.651 $\log_{10}$ points). Immigration is the average annual inflow of immigrants per resident during the period 2005–2010. 'Incoming trips' is the number of people arriving in each country, for any duration of stay, per resident in 2008.

immigration or international travel is associated with an *increase* in the mortality rate of 0.277–0.325 per 100,000 used by Simonsen et al., which captures all deaths due to any respiratory condition and all ages, with a mean of 2.59. The same reduction in exposure to international mobility is associated with an increase of 0.120–0.137 in our measure of excess deaths per 100,000, which captures only prime-age (15–49) deaths due to influenza-like illness, with a mean of 0.633.

These results imply that countries more exposed to international mobility had systematically lower mortality rates in the 2009 pandemic. But this finding need not reject the null result predicted by Prediction 3, because the correlation might not estimate a strictly causal relationship. Most notably, countries with a less exposure to international mobility might be those with less wealth, higher pre-existing morbidity, and less access to advanced medical technology that could affect pandemic mortality conditional on infection (e.g. Morales et al. 2017). And international mobility by itself can cause changes in national wealth and technology (Kerr 2008; Bahar and Rapoport 2018).

We thus repeat the above regressions controlling for a set of country traits in Table 5. We omit the results using the mortality rates imputed by Simonsen et al., because their imputation model uses some of the same covariates we include here. We control for the Healthcare Access and Quality Index estimated by Barber et al. (2017). This is an index of the degree to which cause-specific mortality rates in each country are not explained by the relative prevalence of different life-threatening health conditions in that country. That is, it measures the overall ability of the healthcare system in each country to prevent death for a patient who presents with any given condition that can cause death. The index is highly correlated with other measures of health system strength, such as physicians per capita and the proportion of the population with formal health coverage. We also control for (log) real income per capita, the fraction of the population living in urban areas, and healthcare expenditure as a fraction of GDP. Of all the covariates, only the Healthcare Access and Quality Index has high and robust statistical significance throughout. Each increase of one point in that 0–100 index is conditionally correlated with 0.04–0.05 fewer all-age respiratory deaths per 100,000 (mean 2.04) and 0.05–0.06 fewer prime-age influenza-like illness deaths per 100,000 (mean 0.633). The conditional coefficient on the urbanization rate is positive but statistically insignificant in most specifications.

Table 4: Influenza 2009: International mobility and pandemic final size

Dep. var.:	*ln* Final mortality per 100,000							
Mobility measure:	*ln* Immigration per cap.				*ln* Arrivals per cap.			
Connectivity weight:	No		Yes		No		Yes	
	est.	*s.e.*	*est.*	*s.e.*	*est.*	*s.e.*	*est.*	*s.e.*
Mortality estimates from Simonsen et al. (2013)								
ln Mobility/cap.	−0.181	(0.0322)	−0.171	(0.0242)	−0.172	(0.0248)	−0.147	(0.0198)
N	187		187		187		187	
R^2	0.140		0.205		0.256		0.279	
Mean mortality	2.59	(0.0855)	2.59	(0.0855)	2.59	(0.0855)	2.59	(0.0855)
Change in mortality associated with mobility reduction								
50% reduction	0.347	(0.0654)	0.325	(0.0488)	0.327	(0.0500)	0.277	(0.0392)
90% reduction	1.34	(0.291)	1.25	(0.213)	1.25	(0.219)	1.04	(0.165)
Spearman rank correlations								
ln Mobility/cap.	−0.399	(0.0674)	−0.420	(0.0667)	−0.547	(0.0615)	−0.509	(0.0633)
Mortality estimates from Simonsen et al. (2013), directly reported countries only								
ln Mobility/cap.	−0.275	(0.113)	−0.324	(0.117)	−0.261	(0.154)	−0.243	(0.0921)
N	25		25		25		25	
R^2	0.158		0.211		0.193		0.264	
Mean mortality	2.04	(0.361)	2.04	(0.361)	2.04	(0.361)	2.04	(0.361)
Change in mortality associated with mobility reduction								
50% reduction	0.429	(0.194)	0.514	(0.208)	0.405	(0.261)	0.375	(0.154)
90% reduction	1.81	(1.00)	2.26	(1.16)	1.68	(1.32)	1.53	(0.758)
Spearman rank correlations								
ln Mobility/cap.	−0.339	(0.196)	−0.352	(0.195)	−0.479	(0.183)	−0.464	(0.185)
Excess mortality from influenza-like illness during 2009–2010, age 15–49								
ln Mobility/cap.	−0.174	(0.124)	−0.251	(0.112)	−0.259	(0.116)	−0.282	(0.0891)
N	68		68		68		68	
R^2	0.0185		0.0452		0.0660		0.118	
Mean mortality	0.633	(0.186)	0.633	(0.186)	0.633	(0.186)	0.633	(0.186)
Change in mortality associated with mobility reduction								
50% reduction	0.0810	(0.0614)	0.120	(0.0586)	0.125	(0.0610)	0.137	(0.0475)
90% reduction	0.311	(0.270)	0.496	(0.292)	0.517	(0.307)	0.578	(0.249)
Spearman rank correlations								
ln Mobility/cap.	−0.257	(0.119)	−0.232	(0.120)	−0.313	(0.117)	−0.393	(0.113)

Robust standard errors in parentheses to the right of each coefficient estimate. Constant term included but not shown.

Table 5: Influenza 2009: International mobility and pandemic final size, controlling for country traits

Dep. var.:	*ln* Final mortality per 100,000							
Mobility measure:	*ln* Immigration per cap.				*ln* Arrivals per cap.			
Connectivity weights?	No		Yes		No		Yes	
	est.	*s.e.*	*est.*	*s.e.*	*est.*	*s.e.*	*est.*	*s.e.*
Mortality estimates from Simonsen et al. (2013), directly reported countries only								
ln Mobility/cap.	−0.149	(0.138)	−0.225	(0.123)	−0.0317	(0.317)	−0.116	(0.180)
ln GDP/cap., PPP	0.163	(0.487)	0.334	(0.533)	−0.104	(0.769)	0.131	(0.567)
Urbanization	1.87	(0.979)	1.60	(0.986)	2.07	(1.57)	1.76	(1.29)
Healthcare access	−0.0459	(0.0166)	−0.0480	(0.0161)	−0.0414	(0.0181)	−0.0401	(0.0188)
Health exp./GDP	0.00928	(0.0798)	0.00175	(0.0794)	0.0168	(0.0674)	0.00902	(0.0738)
N	25		25		25		25	
R^2	0.442		0.470		0.424		0.439	
Mean mortality	2.04	(0.361)	2.04	(0.361)	2.04	(0.361)	2.04	(0.361)
Change in mortality associated with mobility reduction								
50% reduction	0.222	(0.217)	0.344	(0.203)	0.0454	(0.459)	0.171	(0.276)
90% reduction	0.837	(0.916)	1.38	(0.966)	0.155	(1.60)	0.627	(1.10)
Excess mortality from influenza-like illness (ILI) during 2009–2010, age 15–49								
ln Mobility/cap.	0.0599	(0.182)	−0.0265	(0.135)	−0.0257	(0.187)	−0.122	(0.167)
ln GDP/cap., PPP	0.267	(0.581)	0.345	(0.489)	0.357	(0.527)	0.469	(0.510)
Urbanization	0.796	(1.12)	0.689	(1.05)	0.661	(1.21)	0.292	(1.25)
Healthcare access	−0.0607	(0.0206)	−0.0601	(0.0212)	−0.0599	(0.0213)	−0.0549	(0.0230)
Health exp./GDP	0.0294	(0.0818)	0.0260	(0.0837)	0.0249	(0.0827)	0.0273	(0.0819)
N	67		67		67		67	
R^2	0.230		0.229		0.229		0.237	
Mean mortality	0.633	(0.186)	0.633	(0.186)	0.633	(0.186)	0.633	(0.186)
Change in mortality associated with mobility reduction								
50% reduction	−0.0257	(0.0768)	0.0117	(0.0602)	0.0114	(0.0834)	0.0560	(0.0796)
90% reduction	−0.0816	(0.232)	0.0399	(0.209)	0.0386	(0.289)	0.206	(0.322)

"Healthcare access" is the Healthcare Access and Quality Index (Barber et al. 2017) for 2010, a 1–100 principal-component score derived from mortality outcomes adjusted for cause-specific risk, where 100 represents the lowest mortality. Robust standard errors in parentheses to the right of each coefficient estimate. Constant term included but not shown.

With these controls, the results no longer reject the null hypothesis of no correlation between exposure to international mobility and mortality from the 2009 pandemic. In the connectivity-weighted regressions using the mortality estimates of Simonsen et al., the coefficients on migration or travel remain negative but with large standard errors. The partial correlation with the Simonsen et al. mortality estimates implies that a 50% reduction in exposure to international mobility is associated with an increase in mortality of about 0.17–0.34 per 100,000 (mean 2.04) that is not statistically significant. The partial correlation with our own mortality estimates implies that the same reduction in international mobility is associated with an increase in mortality of about 0.011–0.056 per 100,000 (mean 0.633) that is also not statistically significant.[34]

Such partial correlations have some claim to contain information about the relationship between international mobility and final size as measured by infections, since the regressions hold constant some of the clearest country-level traits that could mediate between infections and mortality. We do not interpret them as purely causal relationships. Rather, they imply that even in countries with equally effective and well-funded health systems, and equally rich and urban populations, very large differences in exposure to international mobility were not *sufficient* to cause any measurable decline in final mortality during the pandemic of 2009. This is what we would observe if changes in exposure to pre-existing international mobility had no effect on final size, as theory predicts.

Collectively, these results offer no evidence to reject the null hypothesis of no correlation between exposure to international mobility and final size, corroborating Prediction 3 for the 2009 pandemic. This supports the indirect corroboration of the same prediction for all four pandemics in Sections 4–5 considered together.

7 Caveats and interpretation

Because these results can be easily misinterpreted, we pause to discuss a series of conclusions that would not be supported by the theory and evidence discussed here.

[34] Morales et al. (2017) find no association between international air traffic and 2009 H1N1 mortality using a simulation procedure that tests whether air traffic is a causal mediator for other country traits known to be strongly associated with the H1N1 mortality rate.

First, these results do not question a role for international mobility in spreading infectious disease. Disease travels in people. Preliminary data on the ongoing coronavirus disease pandemic of 2019 show an inverse association between pre-pandemic exposure to international mobility and the delay until pandemic arrival (Keita 2020), and our findings are consistent with those. Outside of pandemic settings, some United States immigrants historically brought endemic diseases with them (Ager et al. 2020) and in modern times, forced migration has been an important channel for transmission of malaria and other endemic diseases (Montalvo and Reynal-Querol 2007; Baez 2011). But when endemic viruses, bacteria, and parasites travel with migrants, they move from areas of high endemicity to areas with low endemicity and thus low immunity. In other words, $\mathscr{R}_0$ for the arrival of endemic disease can vary greatly between countries, and by equation (10) this implies that large ongoing mobility can increase the harm to health if immunity is very low ($\mathscr{R}_0$ is very high) in the home country. New pandemic disease is distinct from endemic disease in that, by definition, immunity is low around the world. Cross-country differences in $\mathscr{R}_0$ are thus primarily determined by policy.

For similar reasons, these results do not suggest that limits on geographic mobility and personal interaction in general have no effect on the harms from infectious disease. In theory, any reduction to interpersonal interaction *inside national borders* reduces $\mathscr{R}_0$ and thus final size. Early in the COVID-19 pandemic, cities with higher exposure to domestic mobility exhibited higher prevalence (Glaeser et al. 2020), and emergency restrictions on domestic mobility reduced prevalence (Fang et al. 2020). The question considered here is whether restrictions on mobility *between countries* reduce the international externality independent of restrictions on domestic mobility. Epidemiological models have long recognized that limits on personal interaction within geographic areas ("patches") can reduce $\mathscr{R}_0$ and thus final size, while reduced (but nonzero) interaction between patches does not (e.g. Ma and Earn 2006). Theory suggests that $\mathscr{R}_0$ and thus final size within each country are strongly affected by pre-pandemic constraints to domestic mobility, and unaffected by pre-pandemic constraints to purely international mobility.

Finally, these results do not imply that emergency restrictions on international travel *during a pandemic* have no effect. The data considered here measure permanent changes in international mobility before pandemics begin, and the results do not directly test the effects of short-term emergency travel restrictions. Emergency restrictions are only relevant to our analysis in that

they present one possible mechanism for the reduced-form relationship of interest: If greater pre-pandemic mobility constrains countries from imposing sufficiently large emergency travel restrictions during the pandemic, and $\mathscr{R}_0$ abroad is sufficiently large relative to the home country (equation (10)), then this presents a mechanism by which pre-pandemic international mobility can affect final size. The mechanism requires two elements in composition: that greater pre-pandemic mobility means fewer limits to mobility during the pandemic, and that fewer limits to mobility during the pandemic have important effects on final size. Because we cannot distinguish the reduced-form relationship from zero, we fail to find evidence that this mechanism is large.

An extensive public health literature has examined one part of this composed effect: the effect of short-term emergency travel restrictions on arrival time and overall harm of pandemic disease. Most of these works find very small effects on arrival time and negligible effects on eventual prevalence. This has been the broad conclusion of works on pandemics of influenza (Cooper et al. 2006; Germann et al. 2006; Bajardi et al. 2011; Yu et al. 2012, 765; Tizzoni et al. 2012), Ebola (Otsuki and Nishiura 2016; Poletto et al. 2014), HIV (Docquier et al. 2014; Kenyon et al. 2014), and COVID-19 (Chinazzi et al. 2020, 3; Nowrasteh and Forrester 2020). A systematic review of this literature focusing on influenza by Mateus et al. (2014, 868) concludes, "Travel restrictions would make an extremely limited contribution to any policy for rapid containment of influenza at source during the first emergence of a pandemic virus". Another by Ryu et al. (2020) finds that "international travel-related NPIs [non-pharmaceutical interventions] would have limited effectiveness in controlling pandemic influenza". Ferguson et al. (2006, 448) find "that border restrictions . . . are unlikely to delay spread by more than 2–3 weeks unless more than 99% effective".

Though much of the work on emergency travel restrictions focuses on recent pandemics, these findings have been similar for past pandemics. Shortly after the 1959 influenza pandemic, an expert group convened by the World Health Organization wrote that emergency travel restrictions "were generally found to be ineffective, resulting in at best a short delay in the onset of the epidemic." Except for two countries—Israel and perhaps South Africa—"no effect was detected". The group concluded, "It seems that if such measures are to be effective, they must be very severe—so severe as seriously to interfere with international travel and traffic. This would be a high price to pay for a few additional weeks' freedom from the disease, since there is no evidence that intro-

duction can be entirely prevented. Such action could only be justified on technical grounds if the extra time permitted the application of effective specific prophylactic measures. It is, however, recognized that, confronted by a grave epidemic, health authorities might be forced by public opinion to take such action, even though it was likely to be ineffective" (WHO 1959, 19).

These findings regarding short-term reductions in international mobility are fundamentally different from ours, which relate to long-term reductions. But there are relevant to our findings in that long-term reductions in mobility would be no more effective during an eventual pandemic, but much more economically costly.

8 Conclusion

"Global commerce and travel enable infectious diseases to move around the world within days. This leads to sometimes catastrophic consequences", writes (Schloenhardt 2005, 264). "[I]t took large-scale migration and the COVID pandemic to firmly establish globalization as a widely seen threat," writes Kobrin (2020, 282), without endorsing that view. Pueyo (2020) recommends a tax on airline travel to reduce the pandemic externality. These widely-held opinions might be interpreted to mean that global integration through international mobility raises the expected harm from new pandemics in the future. We find little support for such a conclusion in epidemiological theory or the empirical traits of major recent pandemics.

Moreover, we find weak evidence of the opposite pattern: that in some pandemics, less exposure to international mobility was associated with greater harm—particularly in the 2009 influenza pandemic. This could arise because more isolated countries with less frequent exposure to a variety of pathogens develop less cross-immunity to reduce the harm from a new pandemic. This "hygiene hypothesis" appears to have shaped the mortality experience of isolated island-states in the 1918 pandemic (Shanks et al. 2012; Shanks and Brundage 2013) and simulations suggest it could have important effects today as well (Thompson et al. 2019; Sehrawat and Rouse 2020). It could also occur because exposure to international mobility gives countries higher incomes, stronger health systems, and greater capacity for innovation (e.g. Hovhannisyan and Keller 2015), all of which could reduce the harm inflicted by new pandemics. Separately, such a

relationship could arise because isolation complicates globally coordinated surveillance, and "the prevention of disease is inextricably linked to international cooperation and rights protections" (Meier et al. 2020, 1436). Our analysis cannot distinguish or test for such mechanisms.

We do confirm that greatly reduced exposure to international mobility can slightly delay pandemic arrival, with a halving of exposure to international mobility producing a delay on the order of one week. In principle, even a brief delay of that kind could reduce pandemic harm by "buying time to coordinate an appropriate public health response" (Wells et al. 2020), and could "provide a small but important delay in the spread of a pandemic, especially if other disease control measures are implemented during the afforded time" (Epstein et al. 2007). Our results fail to reject zero correlation between large differences in arrival time and the overall harm in four pandemics. This suggests that in practice, small changes in pandemic arrival time are not typically a binding constraint on the ability or willingness of countries to take measures that reduce final size. Most recently, in the COVID-19 pandemic, the *New England Journal of Medicine* wrote of the United States, "We had ample warning, but when the disease first arrived, we were incapable of testing effectively and couldn't provide even the most basic personal protective equipment to health care workers and the general public" (NEJM 2020).

The literature contains numerous policies that could effectively reduce $\mathscr{R}_0$ and thus the final size of new pandemics. Any costs associated with those policies could in principle play the role of a welfare-enhancing Pigouvian tax. These policies during pandemics include interventions inside national borders, such as internationally-assisted containment in the country of origin (Hollingsworth et al. 2006, 498), mass screening tests (Atkeson et al. 2020), and measures to prevent interpersonal transmission within other countries. They furthermore include interventions *at* the border, such as public information campaigns for travelers at entry points, quarantines, isolation, and contact tracing (Huizer et al. 2015, 19; Selvey et al. 2015, 197). But the theoretical and empirical case for Pigouvian policy before the next pandemic begins—such as taxes or quotas on international mobility in general—remains weak.

References

Adda, Jérôme, "Economic activity and the spread of viral diseases: Evidence from high frequency data," *Quarterly Journal of Economics*, 2016, *131* (2), 891–941.

Ager, Philipp, James J Feigenbaum, Casper Worm Hansen, and Hui Ren Tan, "How the Other Half Died: Immigration and Mortality in US Cities," Working Paper 27480, National Bureau of Economic Research July 2020.

Ahsan, Reshad, Kazi Iqbal, Mahreen Khan, Ahmed Mushfiq Mobarak, and Abu Shonchoy, "Using Migration Patterns to Predict COVID-19 Risk Exposure in Development Countries," New Haven: Yale School of Management 2020.

Almond, Douglas, "Is the 1918 influenza pandemic over? Long-term effects of in utero influenza exposure in the post-1940 US population," *Journal of Political Economy*, 2006, *114* (4), 672–712.

Andreasen, Viggo, "The final size of an epidemic and its relation to the basic reproduction number," *Bulletin of Mathematical Biology*, 2011, *73* (10), 2305–2321.

Antràs, Pol, Stephen J Redding, and Esteban Rossi-Hansberg, "Globalization and Pandemics," Working Paper 27840, National Bureau of Economic Research September 2020.

Arino, Julien, Fred Brauer, P Van Den Driessche, James Watmough, and Jianhong Wu, "A final size relation for epidemic models," *Mathematical Biosciences and Engineering*, 2007, *4* (2), 159–175.

Ashenfelter, Orley and Michael Greenstone, "Using mandated speed limits to measure the value of a statistical life," *Journal of Political Economy*, 2004, *112* (S1), S226–S267.

Atkeson, Andrew, Michael C Droste, Michael Mina, and James H Stock, "Economic Benefits of COVID-19 Screening Tests," Working Paper 28031, National Bureau of Economic Research October 2020.

Azose, Jonathan J and Adrian E Raftery, "Estimation of emigration, return migration, and transit migration between all pairs of countries," *Proceedings of the National Academy of Sciences*, 2019, *116* (1), 116–122.

Baez, Javier E, "Civil wars beyond their borders: The human capital and health consequences of hosting refugees," *Journal of Development Economics*, 2011, *96* (2), 391–408.

Bahar, Dany and Hillel Rapoport, "Migration, knowledge diffusion and the comparative advantage of nations," *Economic Journal*, 2018, *128* (612), F273–F305.

_ , Prithwiraj Choudhury, and Do Yoon Kim, "Innovation on Wings: Flight connectedness and the diffusion of knowledge," Working Paper. Washington, DC: Brookings Institution 2020.

Bajardi, Paolo, Chiara Poletto, Jose J Ramasco, Michele Tizzoni, Vittoria Colizza, and Alessandro Vespignani, "Human mobility networks, travel restrictions, and the global spread of 2009 H1N1 pandemic," *PloS One*, 2011, *6* (1), e16591.

Balcan, Duygu, Hao Hu, Bruno Goncalves, Paolo Bajardi, Chiara Poletto, Jose J Ramasco, Daniela Paolotti, Nicola Perra, Michele Tizzoni, Wouter Van den Broeck et al., "Seasonal transmission potential and activity peaks of the new influenza A (H1N1): a Monte Carlo likelihood analysis based on human mobility," *BMC medicine*, 2009, *7* (1), 45.

Banerjee, Abhijit, Arun G Chandrasekhar, Esther Duflo, and Matthew O Jackson, "The diffusion of microfinance," *Science*, 2013, *341* (6144).

Barber, Ryan M, Nancy Fullman, Reed JD Sorensen, Thomas Bollyky, Martin McKee, Ellen Nolte, Amanuel Alemu Abajobir, Kalkidan Hassen Abate, Cristiana Abbafati, Kaja M Abbas et al., "Healthcare Access and Quality Index based on mortality from causes amenable to personal health care in 195 countries and territories, 1990–2015: a novel analysis from the Global Burden of Disease Study 2015," *The Lancet*, 2017, *390* (10091), 231–266.

Barro, Robert J, José F Ursúa, and Joanna Weng, "The Coronavirus and the Great Influenza Pandemic: Lessons from the ?Spanish Flu? for the Coronavirus?s Potential Effects on Mortality and Economic Activity," Working Paper 26866, National Bureau of Economic Research March 2020.

Biggerstaff, Matthew, Simon Cauchemez, Carrie Reed, Manoj Gambhir, and Lyn Finelli, "Estimates of the reproduction number for seasonal, pandemic, and zoonotic influenza: a systematic review of the literature," *BMC infectious diseases*, 2014, *14* (1), 480.

Bleakley, Hoyt, "Disease and development: evidence from hookworm eradication in the American South," *Quarterly Journal of Economics*, 2007, *122* (1), 73–117.

Boëlle, Pierre-Yves, Severine Ansart, Anne Cori, and Alain-Jacques Valleron, "Transmission parameters of the A/H1N1 (2009) influenza virus pandemic: a review," *Influenza and other respiratory viruses*, 2011, *5* (5), 306–316.

Bolt, Jutta, Robert Inklaar, Herman de Jong, and Jan Luiten van Zanden, "Rebasing 'Maddison': New income comparisons and the shape of long-run economic development," Maddison Project Working paper 10. Groningen Growth and Development Centre, Faculty of Economics and Business. University of Groningen 2018.

Bootsma, Martin CJ and Neil M Ferguson, "The effect of public health measures on the 1918 influenza pandemic in US cities," *Proceedings of the National Academy of Sciences*, 2007, *104* (18), 7588–7593.

Boyd, Douglas D, "A review of general aviation safety (1984–2017)," *Aerospace Medicine and Human Performance*, 2017, *88* (7), 657–664.

Brauer, Fred and Pauline van den Driessche, "Models for transmission of disease with immigration of infectives," *Mathematical Biosciences*, 2001, *171* (2), 143–154.

Brockmann, Dirk and Dirk Helbing, "The Hidden Geometry of Complex, Network-Driven Contagion Phenomena," *Science*, 2013, *342* (6164), 1337–1342.

Campante, Filipe and David Yanagizawa-Drott, "Long-range growth: economic development in the global network of air links," *Quarterly Journal of Economics*, 2018, *133* (3), 1395–1458.

Chinazzi, Matteo, Jessica T. Davis, Marco Ajelli, Corrado Gioannini, Maria Litvinova, Stefano Merler, Ana Pastore y Piontti, Kunpeng Mu, Luca Rossi, Kaiyuan Sun, Cécile Viboud, Xinyue Xiong, Hongjie Yu, M. Elizabeth Halloran, Ira M. Longini, and Alessandro Vespignani, "The effect of travel restrictions on the spread of the 2019 novel coronavirus (COVID-19) outbreak," *Science*, 2020, *368* (6489), 395–400.

Colizza, Vittoria, Alain Barrat, Marc Barthélemy, and Alessandro Vespignani, "The role of the airline transportation network in the prediction and predictability of global epidemics," *Proceedings of the National Academy of Sciences*, 2006, *103* (7), 2015–2020.

Cooper, Ben S, Richard J Pitman, W John Edmunds, and Nigel J Gay, "Delaying the international spread of pandemic influenza," *PLoS Med*, 2006, *3* (6), e212.

Corless, Robert M, Gaston H Gonnet, David EG Hare, David J Jeffrey, and Donald E Knuth, "On the Lambert *W* function," *Advances in Computational mathematics*, 1996, *5* (1), 329–359.

Cox, Nancy J and K Subbarao, "Global epidemiology of influenza: past and present," *Annual review of*

medicine, 2000, *51* (1), 407–421.

Czaika, Mathias and Eric Neumayer, "Visa restrictions and economic globalisation," *Applied Geography*, 2017, *84*, 75–82.

Dawood, Fatimah S, A Danielle Iuliano, Carrie Reed, Martin I Meltzer, David K Shay, Po-Yung Cheng, Don Bandaranayake, Robert F Breiman, W Abdullah Brooks, Philippe Buchy et al., "Estimated global mortality associated with the first 12 months of 2009 pandemic influenza A H1N1 virus circulation: A modelling study," *The Lancet Infectious Diseases*, 2012, *12* (9), 687–695.

Docquier, Frédéric, Ch Vasilakis, and D Tamfutu Munsi, "International migration and the propagation of HIV in sub-Saharan Africa," *Journal of Health Economics*, 2014, *35*, 20–33.

Epstein, Joshua M, D Michael Goedecke, Feng Yu, Robert J Morris, Diane K Wagener, and Georgiy V Bobashev, "Controlling pandemic flu: the value of international air travel restrictions," *PloS One*, 2007, *2* (5), e401.

Erkoreka, Anton, "Origins of the Spanish Influenza pandemic (1918–1920) and its relation to the First World War," *Journal of molecular and genetic medicine: an international journal of biomedical research*, 2009, *3* (2), 190.

Fan, Victoria Y, Dean T Jamison, and Lawrence H Summers, "The Inclusive Cost of Pandemic Influenza Risk," Working Paper 22137, National Bureau of Economic Research March 2016.

Fang, Hanming, Long Wang, and Yang Yang, "Human Mobility Restrictions and the Spread of the Novel Coronavirus (2019-nCoV) in China," Working Paper 26906, National Bureau of Economic Research March 2020.

Fenichel, Eli P, "Economic considerations for social distancing and behavioral based policies during an epidemic," *Journal of Health Economics*, 2013, *32* (2), 440–451.

Ferenczi, Imre, *World Statistics of Aliens: A Comparative Study of Census Returns 1910–1920–1930*, London: International Labour Office (League of Nations), 1937.

Ferguson, Neil M, Derek AT Cummings, Christophe Fraser, James C Cajka, Philip C Cooley, and Donald S Burke, "Strategies for mitigating an influenza pandemic," *Nature*, 2006, *442* (7101), 448–452.

Frost, W. H. and Edgar Sydenstricker, "Epidemic Influenza in Foreign Countries," *Public Health Reports (1896–1970)*, 1919, *34* (25), 1361–1376.

Fukuyama, Hidehiro, Ryo Shinnakasu, and Tomohiro Kurosaki, "Influenza vaccination strategies targeting the hemagglutinin stem region," *Immunological Reviews*, 2020, *forthcoming*.

Gautreau, Aurélien, Alain Barrat, and Marc Barthélemy, "Global disease spread: Statistics and estimation of arrival times," *Journal of Theoretical Biology*, 2008, *251* (3), 509–522.

Germann, Timothy C., Kai Kadau, Ira M. Longini, and Catherine A. Macken, "Mitigation strategies for pandemic influenza in the United States," *Proceedings of the National Academy of Sciences*, 2006, *103* (15), 5935–5940.

Gilbert, M. Thomas P., Andrew Rambaut, Gabriela Wlasiuk, Thomas J. Spira, Arthur E. Pitchenik, and Michael Worobey, "The emergence of HIV/AIDS in the Americas and beyond," *Proceedings of the National Academy of Sciences*, 2007, *104* (47), 18566–18570.

Glaeser, Edward L, Caitlin S Gorback, and Stephen J Redding, "How Much does COVID-19 Increase with Mobility? Evidence from New York and Four Other U.S. Cities," Working Paper 27519, National Bureau of Economic Research July 2020.

Goyal, Sanjeev and Adrien Vigier, "Interaction, protection and epidemics," *Journal of Public Economics*,

2015, *125*, 64–69.

Guimbeau, Amanda, Nidhiya Menon, and Aldo Musacchio, "The Brazilian Bombshell? The Long-Term Impact of the 1918 Influenza Pandemic the South American Way," Working Paper 26929, National Bureau of Economic Research April 2020.

Hainmueller, Jens and Chad Hazlett, "Kernel regularized least squares: Reducing misspecification bias with a flexible and interpretable machine learning approach," *Political Analysis*, 2014, *22* (2), 143–168.

Hilleman, Maurice R, "Realities and enigmas of human viral influenza: pathogenesis, epidemiology and control," *Vaccine*, 2002, *20* (25-26), 3068–3087.

Hollingsworth, T Déirdre, Neil M Ferguson, and Roy M Anderson, "Will travel restrictions control the international spread of pandemic influenza?," *Nature Medicine*, 2006, *12* (5), 497–499.

Horan, Richard D., Eli P. Fenichel, David Finnoff, and Christopher A. Wolf, "Managing dynamic epidemiological risks through trade," *Journal of Economic Dynamics and Control*, 2015, *53*, 192 – 207.

Hovhannisyan, Nune and Wolfgang Keller, "International business travel: an engine of innovation?," *Journal of Economic Growth*, 2015, *20* (1), 75–104.

Huizer, Y.L., C.M. Swaan, K.C. Leitmeyer, and A. Timen, "Usefulness and applicability of infectious disease control measures in air travel: A review," *Travel Medicine and Infectious Disease*, 2015, *13* (1), 19 – 30.

Hummels, David, "Transportation costs and international trade in the second era of globalization," *Journal of Economic Perspectives*, 2007, *21* (3), 131–154.

IOM, *World Migration Report 2020*, Geneva: United Nations, 2020.

Johnson, Niall PAS and Juergen Mueller, "Updating the accounts: global mortality of the 1918-1920" Spanish" influenza pandemic," *Bulletin of the History of Medicine*, 2002, pp. 105–115.

Jordà, Òscar, Sanjay R Singh, and Alan M Taylor, "Longer-run Economic Consequences of Pandemics," Working Paper 26934, National Bureau of Economic Research April 2020.

Keita, Sekou, "Air passenger mobility, travel restrictions, and the transmission of the covid-19 pandemic between countries," *CEPR Covid Economics*, 2020, *9*, 77–96.

Kenyon, Chris, Robert Colebunders, Helene Voeten, and Mark Lurie, "Migration intensity has no effect on peak HIV prevalence: an ecological study," *BMC infectious diseases*, 2014, *14* (1), 350.

Kermack, William Ogilvy and Anderson G McKendrick, "A contribution to the mathematical theory of epidemics," *Proceedings of the Royal Society of London. Series A*, 1927, *115* (772), 700–721.

Kerr, William R., "Ethnic scientific communities and international technology diffusion," *Review of Economics and Statistics*, 2008, *90* (3), 518–537.

Kobrin, Stephen J, "How globalization became a thing that goes bump in the night," *Journal of International Business Policy*, 2020, *3* (3), 280–286.

Langford, Christopher, "Did the 1918-19 Influenza Pandemic Originate in China?," *Population and Development Review*, 2005, *31* (3), 473–505.

Lee, Jean N., Mahreen Mahmud, Jonathan Morduch, Saravana Ravindran, and Abu S. Shonchoy, "Migration, Externalities, and the Diffusion of COVID-19 in South Asia," *Journal of Public Economics*, 2020, *forthcoming*.

Ma, Dan, "The hazard rate function," Applied Probability and Statistics 2011.

Ma, Junling, "Estimating Epidemic Exponential Growth Rate And Basic Reproduction Number," *Infec-*

tious Disease Modelling, 2020, *5*, 129–141.

__ **and David JD Earn**, "Generality of the final size formula for an epidemic of a newly invading infectious disease," *Bulletin of Mathematical Biology*, 2006, *68* (3), 679–702.

Markowitz, Sara, Erik Nesson, and Joshua J. Robinson, "The effects of employment on influenza rates," *Economics & Human Biology*, 2019, *34*, 286–295.

Mateus, Ana LP, Harmony E Otete, Charles R Beck, Gayle P Dolan, and Jonathan S Nguyen-Van-Tam, "Effectiveness of travel restrictions in the rapid containment of human influenza: a systematic review," *Bulletin of the World Health Organization*, 2014, *92*, 868–880D.

McKeown, Adam, "Global Migration, 1846-1940," *Journal of World History*, 2004, *15* (2), 155–189.

Meier, Benjamin Mason, Roojin Habibi, and Y. Tony Yang, "Travel restrictions violate international law," *Science*, 2020, *367* (6485), 1436–1436.

Miller, Joel C, "A note on the derivation of epidemic final sizes," *Bulletin of Mathematical Biology*, 2012, *74* (9), 2125–2141.

Milusheva, Sveta, "Using Mobile Phone Data to Reduce Spread of Disease," Policy Research Working Paper 9198, The World Bank 2020.

Mohammed, Saif I Shah and Jeffrey G Williamson, "Freight rates and productivity gains in British tramp shipping 1869–1950," *Explorations in Economic History*, 2004, *41* (2), 172–203.

Montalvo, Jose G and Marta Reynal-Querol, "Fighting against malaria: prevent wars while waiting for the "miraculous" vaccine," *Review of Economics and Statistics*, 2007, *89* (1), 165–177.

Monto, Arnold S and Keiji Fukuda, "Lessons from influenza pandemics of the last 100 years," *Clinical Infectious Diseases*, 2020, *70* (5), 951–957.

Morales, Kathleen F, John Paget, and Peter Spreeuwenberg, "Possible explanations for why some countries were harder hit by the pandemic influenza virus in 2009: A global mortality impact modeling study," *BMC infectious diseases*, 2017, *17* (1), 642.

Mulder, J and N Masurel, "Pre-Epidemic Antibody against 1957, Strain of Asiatic Influenza in Serum of Older People living in the Netherlands.," *The Lancet*, 1958, *271* (7025), 810–14. Originally published as Volume 1, Issue 7025.

Murray, Christopher JL, Alan D Lopez, Brian Chin, Dennis Feehan, and Kenneth H Hill, "Estimation of potential global pandemic influenza mortality on the basis of vital registry data from the 1918–20 pandemic: a quantitative analysis," *The Lancet*, 2006, *368* (9554), 2211–2218.

Nadaraya, Elizbar A, "On estimating regression," *Theory of Probability & Its Applications*, 1964, *9* (1), 141–142.

Neelsen, Sven and Thomas Stratmann, "Long-run effects of fetal influenza exposure: Evidence from Switzerland," *Social Science & Medicine*, 2012, *74* (1), 58–66.

NEJM, "Dying in a Leadership Vacuum," *New England Journal of Medicine*, 2020, *383* (15), 1479–1480. PMID: 33027574.

Newman, M. E. J., "Spread of epidemic disease on networks," *Phys. Rev. E*, Jul 2002, *66*, 016128.

Nowrasteh, Alex and Andrew C Forrester, "How US Travel Restrictions on China Affected the Spread of COVID-19 in the United States," Cato Working Paper 58. Washington, DC: Cato Institute 2020.

Noymer, Andrew, "Influenza analysis should include pneumonia," *American Journal of Public Health*, 2008, *98* (11), 1927–1928.

Otsuki, Shiori and Hiroshi Nishiura, "Reduced risk of importing Ebola virus disease because of travel restrictions in 2014: a retrospective epidemiological modeling study," *PLoS One*, 2016, *11* (9), e0163418.

Parsons, H Franklin, "The influenza epidemics of 1889-90 and 1891, and their distribution in England and Wales," *British Medical Journal*, 1891, *2* (1597), 303–308.

__ , *Report on the influenza epidemic of 1889–90 by Dr. Parsons, with an introduction by the Medical Officer of the local government board*, London: Her Majesty's Stationery Office, 1891.

__ , *Further report and papers on epidemic influenza 1889–92*, Local Government Board Report. London: HM Stationery Office, 1893.

Patterson, K David, "Pandemic and epidemic influenza, 1830–1848," *Social science & medicine*, 1985, *21* (5), 571–580.

__ **and Gerald F Pyle**, "The diffusion of influenza in sub-Saharan Africa during the 1918–1919 pandemic," *Social Science & Medicine*, 1983, *17* (17), 1299–1307.

__ **and** __ , "The Geography and Mortality of the 1918 Influenza Pandemic," *Bulletin of the History of Medicine*, 1991, *65* (1), 4–21.

Pigou, A. C., *The Economics of Welfare*, London: MacMillan & Co., 1920.

Poletto, Chiara, MF Gomes, A Pastore y Piontti, Luca Rossi, Livio Bioglio, Dennis L Chao, IM Longini Jr, M Elizabeth Halloran, Vittoria Colizza, and Alessandro Vespignani, "Assessing the impact of travel restrictions on international spread of the 2014 West African Ebola epidemic," *Eurosurveillance*, 2014, *19* (42), 20936.

Pueyo, Salvador, "Jevons' paradox and a tax on aviation to prevent the next pandemic," May 2020.

Ramiro, Diego, Sara Garcia, Yolanda Casado, Laura Cilek, and Gerardo Chowell, "Age-specific excess mortality patterns and transmissibility during the 1889–1890 influenza pandemic in Madrid, Spain," *Annals of Epidemiology*, 2018, *28* (5), 267–272.

Rasul, Imran, "The Economics of Viral Outbreaks," *AEA Papers and Proceedings*, May 2020, *110*, 265–68.

Recchi, Ettore, Emanuel Deutschmann, and Michele Vespe, "Estimating transnational human mobility on a global scale," *Robert Schuman Centre for Advanced Studies Research Paper No. RSCAS*, 2019, *30*.

Riley, Steven, Kin O Kwok, Kendra M Wu, Danny Y Ning, Benjamin J Cowling, Joseph T Wu, Lai-Ming Ho, Thomas Tsang, Su-Vui Lo, Daniel KW Chu et al., "Epidemiological characteristics of 2009 (H1N1) pandemic influenza based on paired sera from a longitudinal community cohort study," *PLoS Med*, 2011, *8* (6), e1000442.

Russell, Timothy W., Joesph Wu, Samuel Clifford, John Edmunds, Adam J Kucharski, and Mark Jit, "The effect of international travel restrictions on internal spread of COVID-19," *medRxiv*, 2020.

Ryu, Sukhyun, Huizhi Gao, Jessica Y Wong, Eunice YC Shiu, Jingyi Xiao, Min Whui Fong, and Benjamin J Cowling, "Nonpharmaceutical measures for pandemic influenza in nonhealthcare settings—international travel-related measures," *Emerging Infectious Diseases*, 2020, *26* (5), 961.

Scalia Tomba, Gianpaolo and Jacco Wallinga, "A simple explanation for the low impact of border control as a countermeasure to the spread of an infectious disease," *Mathematical Biosciences*, 2008, *214* (1-2), 70–72.

Schloenhardt, Andreas, "From black death to bird flu: infectious diseases and immigration restrictions in Asia," *New Eng. J. Int'l & Comp. L.*, 2005, *12*, 263.

Sehrawat, Sharvan and Barry T Rouse, "Does the hygiene hypothesis apply to COVID-19 susceptibil-

ity?," *Microbes and Infection*, 2020, *forthcoming*.

Selvey, Linda A, Catarina Antão, and Robert Hall, "Evaluation of border entry screening for infectious diseases in humans," *Emerging infectious diseases*, 2015, *21* (2), 197.

Shanks, G Dennis and John F Brundage, "Pacific islands which escaped the 1918–1919 influenza pandemic and their subsequent mortality experiences," *Epidemiology & Infection*, 2013, *141* (2), 353–356.

__ , Tracy Hussell, and John F Brundage, "Epidemiological isolation causing variable mortality in island populations during the 1918–1920 influenza pandemic," *Influenza and Other Respiratory Viruses*, 2012, *6* (6), 417–423.

Sheng, Zong-Mei, Daniel S Chertow, Xavier Ambroggio, Sherman McCall, Ronald M Przygodzki, Robert E Cunningham, Olga A Maximova, John C Kash, David M Morens, and Jeffery K Taubenberger, "Autopsy series of 68 cases dying before and during the 1918 influenza pandemic peak," *Proceedings of the National Academy of Sciences*, 2011, *108* (39), 16416–16421.

Short, Kirsty R, Katherine Kedzierska, and Carolien E van de Sandt, "Back to the future: Lessons learned from the 1918 influenza pandemic," *Frontiers in Cellular and Infection Microbiology*, 2018, *8*, 343.

Shrestha, Sundar S, David L Swerdlow, Rebekah H Borse, Vimalanand S Prabhu, Lyn Finelli, Charisma Y Atkins, Kwame Owusu-Edusei, Beth Bell, Paul S Mead, Matthew Biggerstaff et al., "Estimating the burden of 2009 pandemic influenza A (H1N1) in the United States (April 2009–April 2010)," *Clinical Infectious Diseases*, 2011, *52* (suppl_1), S75–S82.

Simonsen, Lone, Gerardo Chowell, Viggo Andreasen, Robert Gaffey, John Barry, Don Olson, and Cécile Viboud, "A review of the 1918 herald pandemic wave: importance for contemporary pandemic response strategies," *Annals of epidemiology*, 2018, *28* (5), 281–288.

__ , Peter Spreeuwenberg, Roger Lustig, Robert J Taylor, Douglas M Fleming, Madelon Kroneman, Maria D Van Kerkhove, Anthony W Mounts, W John Paget et al., "Global mortality estimates for the 2009 Influenza Pandemic from the GLaMOR project: a modeling study," *PLoS Medicine*, 2013, *10* (11).

Smith, F. B., "The Russian Influenza in the United Kingdom, 1889–1894," *Social History of Medicine*, 04 1995, *8* (1), 55–73.

Thieme, HR, "A model for the spatial spread of an epidemic," *Journal of Mathematical Biology*, 1977, *4* (4), 337–351.

Thompson, RN, CP Thompson, Omri Pelerman, Sunetra Gupta, and Uri Obolski, "Increased frequency of travel in the presence of cross-immunity may act to decrease the chance of a global pandemic," *Philosophical Transactions of the Royal Society B*, 2019, *374* (1775), 20180274.

Tizzoni, Michele, Paolo Bajardi, Chiara Poletto, José J Ramasco, Duygu Balcan, Bruno Gonçalves, Nicola Perra, Vittoria Colizza, and Alessandro Vespignani, "Real-time numerical forecast of global epidemic spreading: case study of 2009 A/H1N1pdm," *BMC medicine*, 2012, *10* (1), 165.

Tognotti, Eugenia, "Lessons from the history of quarantine, from plague to influenza A," *Emerging infectious diseases*, 2013, *19* (2), 254.

UNDESA, *World Economic and Social Survey 2004: International Migration*, New York: United Nations, 2004.

UNESCO, "How Asian 'Flu Swept around the World," *The UNESCO Courier*, 1958, *11* (5), 12–13.

UNWTO, "Tourism Highlights 2017," Madrid: United Nations World Tourism Organization 2018.

__ , "Methodological Notes to the Tourism Statistics Database," United Nations World Tourism Organization 2020. Accessed May 30, 2020.

__ , "World Tourism Barometer," 18 (1). Madrid: United Nations World Tourism Organization 2020.

Valleron, Alain-Jacques, Anne Cori, Sophie Valtat, Sofia Meurisse, Fabrice Carrat, and Pierre-Yves Boëlle, "Transmissibility and geographic spread of the 1889 influenza pandemic," *Proceedings of the National Academy of Sciences*, 2010, *107* (19), 8778–8781.

Viboud, Cécile, Lone Simonsen, Rodrigo Fuentes, Jose Flores, Mark A Miller, and Gerardo Chowell, "Global mortality impact of the 1957–1959 influenza pandemic," *Journal of infectious diseases*, 2016, *213* (5), 738–745.

__ **, Rebecca F Grais, Bernard AP Lafont, Mark A Miller, and Lone Simonsen**, "Multinational impact of the 1968 Hong Kong influenza pandemic: evidence for a smoldering pandemic," *Journal of infectious diseases*, 2005, *192* (2), 233–248.

Vijgen, Leen, Els Keyaerts, Elien Moës, Inge Thoelen, Elke Wollants, Philippe Lemey, Anne-Mieke Vandamme, and Marc Van Ranst, "Complete genomic sequence of human coronavirus OC43: molecular clock analysis suggests a relatively recent zoonotic coronavirus transmission event," *Journal of Virology*, 2005, *79* (3), 1595–1604.

Voth, Joachim, "Trade and travel in the time of epidemics," in Richard Baldwin and Beatrice Weder di Mauro, eds., *Economics in the Time of COVID-19*, London: Centre for Economic Policy Research, 2020, pp. 93–96.

Watson, Geoffrey S., "Smooth Regression Analysis," *Sankhyā: The Indian Journal of Statistics, Series A (1961-2002)*, 1964, *26* (4), 359–372.

Weidmann, Nils B., Doreen Kuse, and Kristian Skrede Gleditsch, "The Geography of the International System: The CShapes Dataset," *International Interactions*, 2010, *36* (1), 86–106.

Wells, Chad R, Pratha Sah, Seyed M Moghadas, Abhishek Pandey, Affan Shoukat, Yaning Wang, Zheng Wang, Lauren A Meyers, Burton H Singer, and Alison P Galvani, "Impact of international travel and border control measures on the global spread of the novel 2019 coronavirus outbreak," *Proceedings of the National Academy of Sciences*, 2020, *117* (13), 7504–7509.

WHO, *Expert Committee on Respiratory Virus Diseases: First Report*, WHO Technical Report Series, No. 170. Geneva: World Health Organization, 1959.

Yu, Hongjie, Simon Cauchemez, Christl A Donnelly, Lei Zhou, Luzhao Feng, Nijuan Xiang, Jiandong Zheng, Min Ye, Yang Huai, Qiaohong Liao et al., "Transmission dynamics, border entry screening, and school holidays during the 2009 influenza A (H1N1) pandemic, China," *Emerging infectious diseases*, 2012, *18* (5), 758.

Zimmermann, Klaus F., Gökhan Karabulut, Mehmet Hüseyin Bilgin, and Aslı Cansın Doker, "Inter-country distancing, globalisation and the coronavirus pandemic," *The World Economy*, 2020, *43* (6), 1484–1498.

Online Appendix

Global Mobility and the Threat of Pandemics: Evidence from Three Centuries

Michael A. Clemens and Thomas Ginn — December 2020

A1 Expected Onset Date Derivation

The disease emerges in the foreign country, which can intuitively represent all countries besides the home country. It follows the standard SIR model outlined in Equations 5a-5c:

$$\dot{S^*} = -S^*\beta^* I^* \tag{A.1a}$$

$$\dot{I^*} = S^*\beta^* I^* - \gamma I^* \tag{A.1b}$$

$$\dot{R^*} = \gamma I^*, \tag{A.1c}$$

with asterisks to indicate values in the foreign country. Figure A1b represents the evolution of $I^*(t)$ with an initial value in the foreign country of 1 in 7 billion, $\beta^* = 0.58$, and $\gamma = 0.38$, leading to $\mathscr{R}_0^* = 1.5$.

In each period, M foreigners are randomly selected to travel to the home country, and the selection is independent of their health status. The probability of drawing at least one infected traveler in M draws in period t is then $\lambda(M,t) = 1-(1-I^*(t))^M$.[35,36] The arrival of at least one infected traveler in each period can be approximated as a non-homogeneous Poisson process with the parameter $\lambda(M,t)$ as in Scalia Tomba and Wallinga (2008).[37] Let $\Lambda(M,t) = \int_0^t \lambda(M,s)\,ds$. The time of the expected first occurrence—the onset date—is represented as:

$$\tilde{t}(M) = \int_0^T t\,\lambda(M,t)\,e^{-\Lambda(M,t)}\,dt \tag{A.2a}$$

$$= -t\,e^{-\Lambda(M,t)}\Big|_{t=0}^{t=T} + \int_0^T e^{-\Lambda(M,t)}\,dt \tag{A.2b}$$

$$\approx \int_0^T e^{-\Lambda(M,t)}\,dt, \tag{A.2c}$$

where T represents a time when all countries have at least one case.[38] The first step comes from integration by parts and the second step approximates the first term as 0.

Figure A1c shows $\lambda(5,000,t)$ and $\lambda(500,000,t)$, the probability of at least one positive case arriving to a country with 5,000 and 500,000 average daily arrivals, respectively. These values reflect travel patterns in our data on 2008, when the average country received about 4,500 incoming travelers, and the country

[35] The parameter is an approximation of sampling with replacement.

[36] If only one origin country has infections, M is the number of travelers from that country. If multiple origins have different rates of infection, the parameter could be adapted to $1-(1-I_1^*)^{M_1}(1-I_2^*)^{M_2}$, etc. The same modifications can be made to allow different prevalence among demographic groups in the foreign country.

[37] A non-homogeneous Poisson process is an approximation of repeated Bernoulli trials with a time-varying parameter.

[38] Equation A.2a parallels the exponential distribution which represents inter-arrival periods for a homogeneous Poisson process. Intuitively, the density function is the probability of arrival in period t, $\lambda(M,t)$, times the probability of no arrivals through period t, $e^{-\Lambda(M,t)}$. For the derivation for non-homogeneous Poisson processes, see Ma (2011).

receiving the most, China, received approximately 445,000 per day.

At the onset of the pandemic, $I^*(t)$ is small enough that $\lambda(M,t)$, and therefore $\Lambda(M,t)$, is close to 0. Even with a large number of travelers, the disease is sparse enough in the foreign population that there is almost no chance of drawing a positive case as traveler. However, $I^*(t)$ is growing exponentially in the initial phase; $I^*(t) \approx ae^{bt}$ where $a = I^*(0)$ and $b = \beta^* - \gamma^*$ (Ma 2020). With exponential growth and thousands of draws, $\lambda(M,t)$ quickly increases from 0 to 1 once $I^*(t)$ exceeds some inflection point at a time $\tau(M)$.[39] Separating A.2c at $\tau(M)$:

$$\tilde{t}(M) = \underbrace{\int_0^{\tau(M)} e^{-\Lambda(M,t)}\,dt}_{\Lambda(M,t)\approx 0 \implies e^{-\Lambda(M,t)}\approx 1} + \underbrace{\int_{\tau(M)}^{T} e^{-\Lambda(M,t)}\,dt}_{\Lambda(M,t)>0 \implies e^{-\Lambda(M,t)}\approx 0} \tag{A.3}$$

$$\approx \tau(M). \tag{A.4}$$

Intuitively, the expected onset date can be approximated by the number of periods that the probability of infected travelers arriving is close to 0. Using Figures A1a and A1c to estimate $\lambda(M, \tau(M)) \approx 0.2$ for all M, solving for $\tau(M)$ yields:

$$1 - (1 - I^*(\tau(M)))^M = 0.2 \tag{A.5}$$

$$ae^{b\tau(M)} = 1 - 0.8^{1/M}, \tag{A.6}$$

using the exponential approximation of $I^*(t)$ during the initial phase of the pandemic described above.[40] Approximating $0.8^{1/M} \approx 1 - \frac{0.2}{M}$, setting the initial infected population $a = 1/P$ where P is the foreign population, and setting $b = \beta^* - \gamma$ yields:

$$\tilde{t}(M) \approx \frac{1}{\beta^* - \gamma}\big(ln(0.2P) - ln(M)\big). \tag{A.7}$$

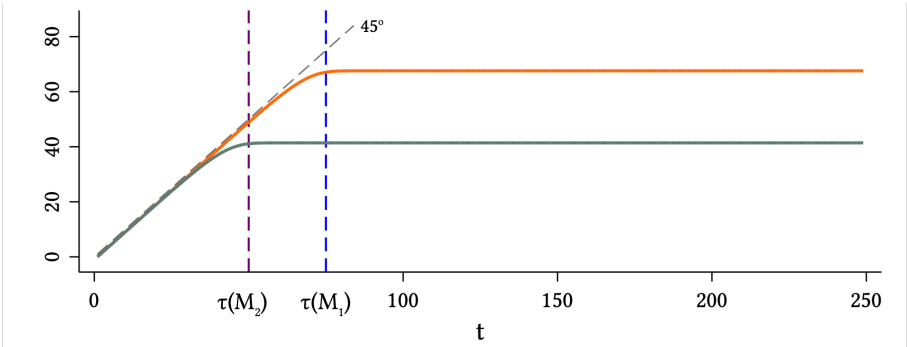
[39] We can define $\tau(M)$ as the time when $\int_{\tau(M)}^{T} e^{-\Lambda(M,t)}\,dt = 1$, or the time when the expected onset date converges to within 1 day of its limiting value. Intuitively, the probability of the *first* arrival occurring after $\tau(M)$ is small and approaching zero.

[40] To approximate the relationship between $\tau(M)$ and M, we find the time when expected waiting time converges in Figure A1a, and the corresponding probability of at least one infected traveler in Figure A1c. This results in Equation A.5. The probability of at least one infected traveler per period is 0.2, and the probability that no infected travelers have arrived quickly approaches 0. The 0.2 estimate will vary slightly with M, β^*, and γ, though further numerical simulations and Equation 3 suggest this impact is small.

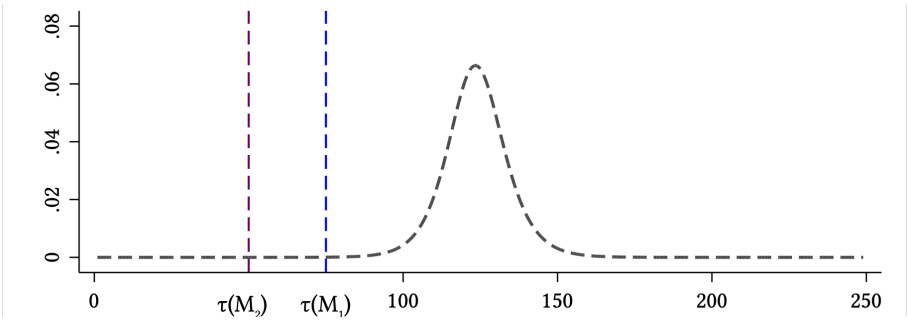
Appendix Figure A1: Approximating $t(\tilde{M})$

Orange: $M_1 = 5,000$ Green $M_2 = 500,000$

(a) *Expected Waiting Time Through Time t:* $\int_0^t e^{-\int_0^t \lambda(M,s)ds} dt$



(b) *Proportion Infected in Foreign Country:* $I^*(t)$



(c) *Probability* ≥ 1 *Infected Traveler in t:* $\lambda(M,t) = 1 - (1 - I^*(t))^M$

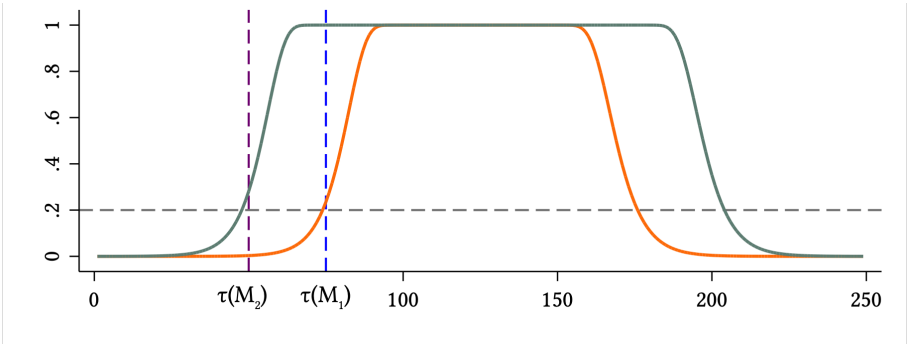


Figure A1a shows the expected waiting time for the disease to arrive through t. Figure A1b shows Equation A.1b for $\beta^* = 0.58$, $\gamma^* = 0.38$, $p = 0$, and an initial value of $I^*(0) = 1/7,000,000,000$. The y-axis is the proportion of the population that is infected at time t. Figure A1c shows the corresponding probability of drawing at least one infected traveler in period t among 5,000 and 500,000 incoming travelers on the orange and green curves, respectively. The y-axis is the probability of drawing at least one infected traveler. The y-axis is the number of periods in the same unit as the x-axis and β (usually days). See the text and corresponding footnotes for explanations of $\tau(M)$ and the blue and purple dashed lines.

A2 Summary statistics

Appendix Table A1: Mobility in 2008: Global Travel and the Top 20 Countries

Trips Rank	Country	Annual Incoming Trips	Trips Per Capita	% Trips Foreign	% Trips Immigrant	Centrality Rank
	World	2,083,384,334	0.31	50.4	0.8	
	Average Country	9,556,809	2.34	55.3	1.9	
	Median Country	1,687,618	0.57	55	0.9	
1	China	163,323,823	0.12	79.7	0.1	9
2	Germany	142,315,188	1.73	18.5	0.4	1
3	United States	136,484,770	0.45	44.5	1.6	6
4	Hong Kong	106,415,446	15.29	16.7	0.1	35
5	France	101,315,253	1.57	61	0.4	2
6	United Kingdom	99,226,661	1.61	32.8	0.7	4
7	Spain	78,789,732	1.71	73	0.7	3
8	Poland	76,163,646	2	78.7	0.2	12
9	Italy	73,962,717	1.26	58.4	0.6	5
10	Russia	51,774,794	0.36	46.8	1.8	14
11	Canada	47,125,763	1.42	37.4	1.1	23
12	Macao	46,395,134	90.05	49.5	0.1	57
13	Netherlands	38,972,675	2.37	26.1	0.3	7
14	Ukraine	38,720,080	0.84	66.4	0.7	31
15	Mexico	37,605,013	0.34	57.4	0.9	39
16	Austria	33,191,256	3.99	65.1	0.3	11
17	Japan	30,107,584	0.24	28.6	0.6	13
18	Turkey	29,837,421	0.42	83.3	0.4	17
19	Saudi Arabia	25,156,819	0.97	60.4	1.7	42
20	Switzerland	24,869,726	3.25	34.6	0.6	8

Trips are the total number of incoming foreigners, incoming immigrants, returning citizens, and returning emigrants. It is estimated based on data from UNWTO (2020a), Azose and Raftery (2019), and the random forest model as described in the text. The percentage of foreigners assumes all immigrants from Azose and Raftery (2019) are foreigners, though some are returning emigrants. Azose and Raftery (2019) estimate migration over the five-year period between mid-2005 and mid-2010; we divide these estimates by 5 for an annual estimate for 2008. Centrality rank is based on the eigenvector centrality of the global flows matrix that leaves out the top origin for each destination country.

A3 Detailed data sources

A3.1 Connectivity

A3.1.1 Global mobility estimates

As discussed in the main text, the mobility data is based on the United Nations World Tourism Organization's (UNWTO) measure of outbound flows UNWTO (2020a), drawing on the methodology used to construct the Global Transnational Mobility Dataset (Recchi et al. 2019), and migration data from Azose and Raftery (2019). The UNWTO data consist of administrative reports of arrivals from each destination

country. The arrivals are listed by the origin country, although some countries use aggregations, meaning these data do not capture the total flows. Since each country measures arrivals differently, the measurements are categorized into the eight types listed in Table A2. Of the 17,216 bilateral routes reported, 29% report more than one type. In these cases, we follow the GTMD ranking of types, which mostly corresponds to the type's prevalence in the overall dataset.

Appendix Table A2: UNWTO Data Types

Rank	Type	Routes Total	Routes Used	% Total Flows
1	TFR: Arrivals of non-resident tourists at national borders, by country of residence	3,896	3,896	34%
2	TFN: Arrivals of non-resident tourists at national borders, by nationality	3,561	3,368	13%
3	VFR: Arrivals of non-resident visitors at national borders, by country of residence	2,724	2,370	13%
4	VFN: Arrivals of non-resident visitors at national borders, by nationality	2,999	2,248	27%
5	TCER: Arrivals of non-resident tourists in all types of accommodation establishments, by country of res.	1,355	798	9%
6	TCEN: Arrivals of non-resident tourists in all types of accommodation establishments, by nationality	498	333	1%
7	THSR: Arrivals of non-resident tourists in hotels and similar establishments, by country of residence	1,520	253	2%
8	THSN: Arrivals of non-resident tourists in hotels and similar establishments, by nationality	665	108	1%
	Total	17,216	13,374	

We check the measurement error induced by the different data types across countries by estimating the variation where the same route is reported in multiple categories. Using within-route fixed effects to weight each data type, we find that the correlation between the weighted and unweighted sum of arrivals at the country level is 0.98, and therefore argue that this measurement error is unlikely to change our conclusions.

Across 218 countries, and therefore 218^2-218 = 47,306 bilateral routes in total, the UNWTO data provides counts for 12,195 bilateral routes in 2008, or 26% of the total cells. We estimate the missing cells in multiple stages. First, we use estimates from 2007 (966 additional routes) and 2006 (213 additional routes) when available, which yields 13,374 total routes. Next, since zeroes were otherwise not recorded in the UNWTO data, we impute a flow of 0 if all of the following are true:

- the UNWTO estimate (2006-08) is missing for the country pair, but the destination reports estimates from other origins,
- the UNWTO estimate (2006-08) is also missing in the reverse direction, but the origin reports estimates from other destinations,
- the estimated migration flow between 2005 and 2010 is less than one, and

- the Global Transnational Mobility Dataset (GTMD) estimate for 2011 (the earliest year available) is less than one. The GTMD filled in missing values from the UNWTO data using private data on air passenger flows between airports, so this condition in effect states that the air passenger data used in the GTMD is zero in 2011.

This captures 2,749 total cells where zero is imputed.

In the next stage, we predict the bilateral flows that are still missing with random forest models. We train the model on the route's characteristics, combining the following data sources:

- direct flights: the number of nonstop routes in 2009 between the countries, using data from OpenFlights.org (2009),
- migration: bilateral migration flows estimated by Azose and Raftery (2019) in 5-year intervals, dividing the 2005-10 estimate by 5 for an estimate of 2008,
- gravity data: origin and destination populations, GDP per capita (which come from the World Bank's World Development Indicators), the distance between the most populated cities, difference of time zones in hours , an indicator for a common official language, a language spoken by at least 9% of the population, contiguity, a common colonizer, currency, or religion, all from CEPII's GeoDist and Gravity datasets Mayer and Zignago (2011) and Head et al. (2010)
- UN sub-regional and European Union (EU) classifications, and a dummy for whether they're in the same sub-region or both in the EU.

We first train a model that includes the flow in the reverse direction when available. This model is used to impute all cells with one direction missing and one direction measured or imputed up to this point in our methodology. The parameters of the model (number of trees and variables) are tuned by training the model on a random half of the data, testing on the other half, and minimizing the validation root mean squared error. A second random forest is trained using all of the values imputed to this point and used to predict the remaining missing cells. The dependent variable in the model is the natural log of the estimates greater than zero and zeroes remaining.

We assess our method in multiple ways. We estimate alternative predictions using an OLS gravity model on the same independent variables used to train the random forest, as well as origin and destination fixed effects. We also try a lasso specification that selects the independent variables used for predictions. In Figure A2, we compare these estimates to the true values for the 26% of cells in the bilateral matrix where data is available. Panel a shows that the random forest best approximates the true distribution; the mean estimates from the OLS and lasso exceed the true means. Panel b shows that the random forest also predicts the true values better than the OLS across the distribution.

A3.2 Influenza 2009

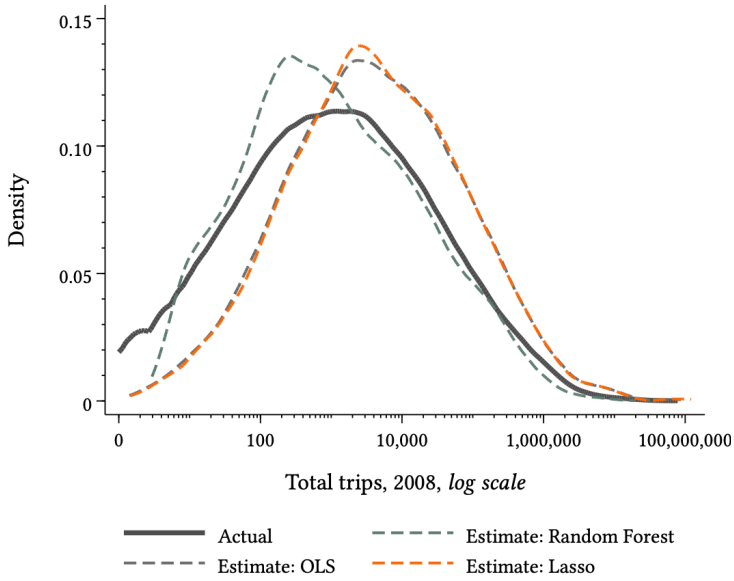
2009 final mortality estimates, as noted in the main text, are drawn in part from Simonsen et al. (2013) and Dawood et al. (2012). Mortality rates per 100,000 are estimated using 2009 country population from the World Bank *World Development Indicators* (SP.POP.TOTL).

The Simonsen et al. (2013) dataset is freely available online at https://www.nivel.nl/en/influenza-mortality (Accessed November 23, 2020), part of the Influenza Mortality Research Project at the Netherlands Institute for Health Services Research (NIVEL).

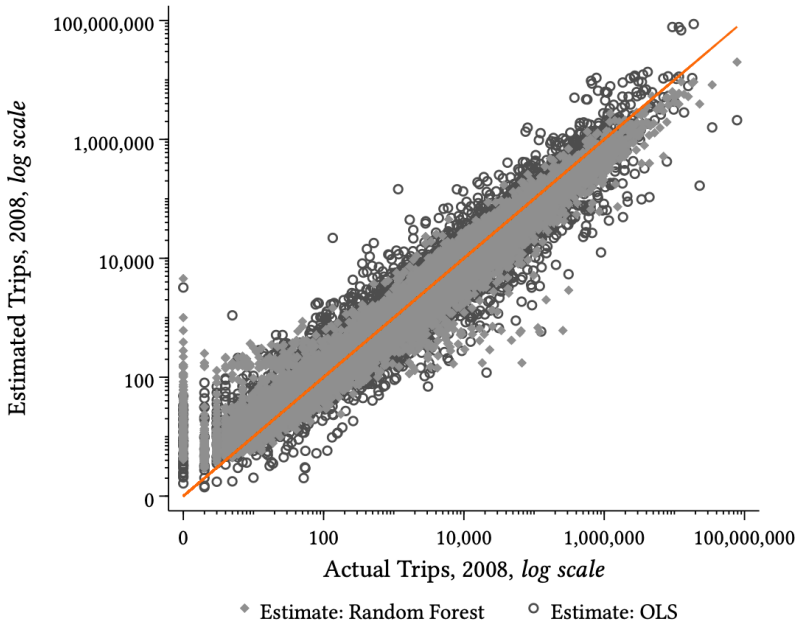
Arrival dates for the 2009 pandemic are collected as follows. We begin with the dates for 93 countries collected by Balcan et al. (2009): Argentina, Australia, Austria, Bahamas, Bahrain, Bangladesh, Barbados, Belgium, Bermuda, Bolivia, Brazil, British Virgin Islands, Bulgaria, Canada, Cayman Islands, Chile,

Appendix Figure A2: Mobility: 2008 Actual vs. Predicted Bilateral Flows (26% of Cells)

(a) *Kernel Densities*



(b) *Correlations*



Panel (a) shows the kernel densities of four estimates of bilateral trip flows in 2008. The dark gray line shows the actual data from UNWTO. The dashed lines show the predicted values across the same 26% of cells. The teal dashed line shows the predicted values from the random forest described in the text. The gray dashed line shows the predicted values from an ordinary least squares specification, and the orange dashed line shows the predicted values from a lasso algorithm. The kernel function is the Epanechnikov kernel and other default Stata settings are used. Panel (b) plots the random forest and OLS estimates against the actual values, with the orange line representing 45° and an estimate that equals the measured value.

China (Mainland), Colombia, Côte d'Ivoire, Cuba, Cyprus, Czech Republic, Denmark, Dominica, Dominican Republic, Ecuador, Egypt, El Salvador, Estonia, Finland, France, French Polynesia, Germany, Greece, Guatemala, Honduras, Hong Kong, Hungary, Iceland, India, Ireland, Israel, Italy, Jamaica, Japan, Jordan, Kuwait, Laos, Lebanon, Luxembourg, Macau, Malaysia, Morocco, Netherlands, Netherlands Antilles, New Zealand, Nicaragua, Norway, Oman, Panama, Papua New Guinea, Paraguay, Peru, Philippines, Poland, Portugal, Puerto Rico, Qatar, Romania, Russia, Saudi Arabia, Singapore, Slovakia, Solomon Islands, South Africa, South Korea, Spain, Sri Lanka, Suriname, Sweden, Switzerland, Thailand, Trinidad and Tobago, Turkey, Ukraine, United Arab Emirates, United Kingdom, United States, Uruguay, Venezuela, Vietnam, Western Sahara, and Western Samoa. From the World Health Organization *Influenza A(H1N1) Emergency Updates* of June 17, June 19, July 6, July 27, August 12, August 19, and November 13, 2009 we take arrival dates for Andorra, Burundi, Chad, Haiti, Kiribati, Marshall Islands, Martinique, Saint Kitts and Nevis, Saint Martin (France), Saint Vincent and the Grenadines, Seychelles, Sint Eustatius (Netherlands), Sint Maarten (Netherlands), Timor-Leste, Tonga, Turks and Caicos Islands (UK), Tuvalu, and Wallis and Futuna (France).

For 72 other countries we manually collected reports of the first case from individual sources, predominantly archived press reports. These are: Afghanistan from WHO EMRO. Afghanistan, Surveillance, forecasting, and response, Albania from Reuters. Albania confirms first cases of H1N1 flu virus. July 21, 2009, American Samoa (US) from PAHO. Update Influenza A (H1N1) Regional Report (17 Jul 2009). July 17, 2009, Angola from ANGOP. "Angola reports first H1N1 cases." August 25, 2009, Antigua and Barbuda from Associated Press. "Boy, 9, becomes first swine flu case in Antigua". Taiwan News. June 20,2009 (archived), Armenia from WHO Pandemic (H1N1) 2009 - update 76. November 27, 2009, Azerbaijan from "Reuters. Azerbaijan reports first two H1N1 flu cases. July 30, 2009", Belarus from RIANovosti. Belarus registers first case of swine flu - Health Ministry. August 19, 2009. (archived), Belize from Caribbean360.com. July 8: Belize (BZ): first case confirmed. July 8, 2009. (archived), Benin from The Times. "Swine flu confirmed in Benin," May 2, 2009. (archived), Bhutan from Wangchuk S, et al. Influenza surveillance from November 2008 to 2011; including pandemic influenza A(H1N1)pdm09 in Bhutan. May 13, 2013, Bosnia and Herzegovina from Mondo. Novi grip stigao u BiH!. June 29, 2009. (archived), Chad from WHO Pandemic (H1N1) 2009 - Update 86. February 5, 2010, Costa Rica from Avalos, Angela. "Confirmada primera tica con fiebre porcina." Nacion.com. April 28, 2009. (archived), Democratic Republic of the Congo from Independent Online. DRC First Swine Flu Case. August 16, 2009. (archived), Djibouti from WHO Pandemic (H1N1) 2009 - update 64. August 30, 2009, Ethiopia from Reuters. "Ethiopia confirms first cases of H1N1". June 19, 2009. (archived), Gabon from WHO Pandemic (H1N1) 2009 - Update 60. July 31, 2009, Georgia from Reuters. Georgia reports its first case of H1N1 flu. June 18, 2009, Greenland (Denmark) from Focus Information Agency. "Swine flu hits Greenland as toll rises in Europe". November 11, 2009, Guam (US) from Karrigan, Kevin. First Case of Swine Flu Confirmed On Guam. PNC News. July 1, 2009 (archived), Guinea from WHO Pandemic (H1N1) 2009 - Update 96. April 16, 2010, Indonesia from Reuters. Indonesia confirms first cases of H1N1 flu virus, June 24, 2009, Iraq from Kami, Aseel. Reuters. "Iraq confirms first cases of H1N1 flu virus", Isle of Man (UK Crown Dependency) from BBC News. First Isle of Man Swine Flu Case. June 12, 2009, Kenya from Reuters. Kenya confirms first case of H1N1 flu virus.June 29, 2009, Kosovo from Reuters. "Kosovo registers first case of H1N1 flu virus." July 27,2009, Kyrgyzstan from Trend News Agency. "Swine flu hits Kyrgyzstan." August 24, 2009, Latvia from Reuters. Latvia has first confirmed case of H1N1 flu. June 23, 2009, Lesotho from IOL. Lesotho Confirms first Cases of Swine Flu. Spetember 9, 2009, Libya from Reuters. CORRECTED: Libya reports first case of new H1N1 flu. July 5, 2009, Lithuania from Reuters. Lithuania confirms first case of H1N1 flu. JUNE 26, 2009, Macedonia from Reuters. Macedonia confirms first two cases of H1N1 flu virus. July 4, 2009, Madagascar from Rajatonirina et al. "The Spread of Influenza A(H1N1)pdm09 Virus in Madagascar Described by a Sentinel Surveillance Network", Malawi from Associated Press. Taiwan News. September 10, 2009, Mali from The New Humanitarian. "Lab Confirms H1N1." January 13, 2010, Malta from The Malta Independent. "Maltese Swine flu cases confirmed" July 3, 2009, Mauritania from WHO Pandemic (H1N1) 2009 - Update 86. February 5, 2010, Mauritius from Reuters. Mauritius confirms first of H1N1 flu virus. June 30, 2009, Mexico from Roos,

Robert. University of Minnesota CIDRAP. May 1, 2009, Micronesia from Palikir, Pohnpei. "First Reported case of Pandemic Influenza A/H1N1 in the Federated States of Micronesia." College of Micronesia-FSM. July 21, 2009, Moldova from WHO Pandemic (H1N1) 2009 - Update 60. July 31, 2009, Monaco from Reuters. "Mexican swine flu victims were young." June 29, 2009, Mongolia from MonInfo. "Breaking news: Swine flu outbreak in Mongolia." October 13, 2009, Montserrat from GIU.gov.ms. Montserrat reports first h1n1 flu case. November 24, 2009. (archived), Mozambique from CM Journal. Gripe A: Moçambique regista primeiro caso. August 18, 2009, Myanmar (Burma) from Reuters. Myanmar reports first case of H1N1 flu. June 27, 2009, Namibia from SABCNews. First two cases of swine flu confirmed in Namibia. July 20, 2009, Nepal from ExpressBuzz. Swine flu strikes Nepal. June 29, 2009 (archived), Niger from WHO Pandemic (H1N1) 2009 - Update 89. February 26, 2010, Nigeria from Reuters. "Nigeria records first swine flu case in U.S. girl." November 5, 2009, North Korea from Sue-young, Kim. The Korea Times. "Worker at Gaeseong Site Diagnosed With H1N1 Flu." November 16, 2009, Northern Mariana Islands from Radio New Zealand International. "Guam reports its first swine flu-related death." July 21, 2009. (archived), Pakistan from "Nishtar, Sania, 2010, H1N1 Outbreak in Pakistan: Lessons Learnt, NTS Working Paper Series No. 4, Singapore: RSIS Centre for Non-Traditional Security (NTS) Studies", Palestine from Reuters. "Palestinians report first case of H1N1 flu." June 10, 2009, Republic of Congo from IOL. "Congo students diagnosed with swine flu." October 28, 2009, Rwanda from Wane et al. 2009 Pandemic Influenza A (H1N1) Virus Outbreak and Response – Rwanda, October, 2009–May, 2010, San Marino from Radio e Televisione della Repubblica di San Marino. "Influenza A/H1N1: primo caso in Repubblica." November 5, 2009. (archived), São Tome and Príncipe from WHO Pandemic (H1N1) 2009 - Update 71. October 17, 2009, Senegal from WHO Pandemic (H1N1) 2009 - Update 87. February 12, 2010, Serbia from Sekularac, Ivana. Reuters. "Serbia confirms first case of H1N1 flu". June 24, 2009, Slovenia from Reuters. "Slovenia confirms first H1N1 case." June 19, 2009, Somalia from The New Humanitarian. "WHO confirms first cases of H1N1." November 16, 2009, Sudan from Heavens, Andrew. Reuters. "Sudan reports first H1N1 flu case." July 16, 2009, Swaziland from WHO Pandemic (H1N1) 2009 - Update 60. July 31, 2009, Syria from The Times of India. Syria confirms first swine flu case. July 4, 2009. (archived), Taiwan from Jennings, Ralph. Reuters. Taiwan reports first case of H1N1 flu. May 19, 2009, Tajikistan from WHO Pandemic (H1N1) 2009 - update 69. October 4, 2009, Tanzania from University of Minnesota CIDRAP. H1N1 FLU BREAKING NEWS. July 9, 2009, U.S. Virgin Islands from The St. John Source. "Swine Flu Case Confirmed on St. Thomas." June 16, 2009, Yemen from Reuters. Jordan, Qatar, Yemen identify first H1N1 flu cases. June 16, 2009, Zimbabwe from University of Minnesota CIDRAP. H1N1 FLU BREAKING NEWS. July 9, 2009.

For 18 countries we were not able to locate an arrival date in health agency reports or press reports, so we use the date that each country reported its first confirmed pandemic-strain influenza test (A/H1N1pdm09) to the World Health Organization's *FluNet* database, between March 2009 and March 2010: Algeria, Anguilla (UK), Aruba (Netherlands), Cambodia, Cameroon, Croatia, Fiji, French Guiana (France), Guadeloupe (France), Iran, Kazakhstan, Maldives, Montenegro, New Caledonia, Saint Lucia, Tunisia, Uganda, and Zambia (accessed April 22, 2020).

GDP per capita at purchasing power parity and urbanization rate for the year 2008 are from the World Bank *World Development Indicators.* That source lists no 2008 GDP per capita values for Cuba or Syria; these, respectively $6,336 and $5,307, are taken from Bolt et al. (2018). The Healthcare Access and Quality Index for the year 2010 is from Barber et al. (2017). Current health expenditure as a fraction of GDP for the year 2009 is from the World Health Organization's Global Health Expenditure Database, Version December 2019 (WHO 2019), with the following exceptions: The WHO reports no value for Hong Kong, which is instead taken from the Hong Kong *Domestic Health Accounts 2009–2010 Summary Report*, Food and Health Bureau, Government of Hong Kong Special Administrative Region, p. 1 (5.2 percent). The WHO reports no 2009 value for Montenegro, which was still in the process of separating from Serbia at the time, so the value for Serbia is used for Montenegro (9.3 percent). The WHO reports no values for Puerto Rico; the 2009 value (11.9 percent) is from *Informe de la Salud en Puerto Rico 2014*, Secretaría Auxiliar de Planificación y Desarrollo, Departamento de Salud, Estado Libre Asociado de Puerto Rico,

2014, p. 9.

Latitude of the capital city in each country is from the GEO CEPII database (Mayer and Zignago 2011).

A3.3 Influenza 1957

For final mortality we use the annual respiratory death rate per 10,000 estimated by Viboud et al. (2016), converted to a rate per 100,000 for comparability with the other pandemics. We omit their estimates for Scotland, England & Wales, and West Berlin, because these are subsets of other geographic areas in the dataset. We combine their estimates for the Republic of Ireland and Northern Ireland into a single mortality rate for the island of Ireland, weighting by the relative populations reported by Viboud et al. (2016). The pandemic arrival months for almost all countries are taken from UNESCO (1958). For a few countries, UNESCO (1958) does not report an arrival month so we obtained them elsewhere: Germany and New Zealand from Dunn (1958), and Iceland from Sigurjónsson et al. (1959).

GDP per capita for 1956 (1990 U.S. dollars at purchasing power parity) is from the Maddison Project (Bolt et al. 2018). Child mortality (under age 5, per 1,000 live births) in 1956 is from Gapminder (2020), which for this historical period draws on UNIGME, a data collaboration of UNICEF, WHO, the UN Population Division and the World Bank. The source contains estimates for Czechia and Slovakia but not for Czechoslovakia. We estimate the rate for Czechoslovakia as a population-weighted average of the source's figures for Czechia and Slovakia, where the relative populations in 1960 are from the Clio Infra database of historical population with modern borders (Fink-Jensen 2015).

A3.4 Influenza 1918

Final mortality for the fall wave is from Patterson and Pyle (1991, 14–15) and mortality for the combined spring and fall waves is from Johnson and Mueller (2002, 110-114).

Fall-wave arrival dates for the geographic areas corresponding to mid-1991 international borders are from Patterson and Pyle (1991, 8–12) for most countries: Albania, Algeria, Andorra, Antigua and Barbuda, Armenia, Australia, Austria, Azerbaijan, Bangladesh, Barbados, Belarus, Belgium, Bosnia and Herzegovina, Bulgaria, Burundi, Cambodia, China, China, Hong Kong, Colombia, Costa Rica, Croatia, Cyprus, Czechoslovakia, D.P.R. of Korea, Ecuador, Egypt, El Salvador, Equatorial Guinea, Estonia, Finland, Former Yugoslavia, France, Gambia, Georgia, Germany, Ghana, Greece, Grenada, Guinea-Bissau, Honduras, India, Ireland, Japan, Kuwait, Laos, Latvia, Libya, Lithuania, Luxembourg, Maldives, Moldova, Monaco, Montenegro, Morocco, Myanmar, Netherlands, Norway, Pakistan, Panama, Paraguay, Peru, Philippines, Republic of Korea, Romania, Rwanda, San Marino, São Tome and Príncipe, Saudi Arabia, Serbia, Slovenia, Sri Lanka, Swaziland, Sweden, Switzerland, Syrian Arab Republic, Tanzania (Mainland), Thailand, Tunisia, Turkey, United Kingdom, Venezuela, and Vietnam. Arrival dates for most of Africa are from Patterson and Pyle (1983, 1300–1304): Angola, Benin, Botswana, Burkina Faso, Cabo Verde, Cameroon, Central African Republic, Chad, Congo, Côte d'Ivoire, D.R. Congo, Djibouti, Eritrea, Ethiopia, Gabon, Guinea, Kenya, Lesotho, Liberia, Malawi, Mali, Mauritania, Mozambique, Niger, Nigeria, Senegal, Sierra Leone, Somalia, South Africa, Togo, Uganda, Zambia, and Zimbabwe, as well as providing a date for Yemen (Aden). From Vollmer and Wójcik (2017) we take fall-wave arrival dates for Bolivia, Guatemala, Guyana, Haiti, Hungary, Iraq, Jamaica, Namibia, Nepal, Nicaragua, and Poland. From Killingray (1994) we take fall-wave arrival dates for Bahamas, Belize, Dominica, Dominican Republic, Saint Lucia, St. Kitts and Nevis, St. Vincent and the Grenadines, Suriname, Trinidad and Tobago, and Puerto Rico. From Frost and Sydenstricker (1919, 1361–1363) we take fall-wave arrival dates for Canada, Cuba, Italy, Madagascar, Mexico, Spain, United States, and Uruguay. From Vaughan (1921, 66) we take fall-wave arrival dates for Denmark, New Zealand, Portugal, and Russia. From Langfeldt Lind (2012, 10) we take fall-wave arrival dates for Iran, Israel/Palestine, Jordan, and Lebanon. We take fall-wave arrival dates for Argentina and Brazil from Jordan (1925, 946), Fiji and Tonga from Herda (2000, 135), and Bahrain and the United Arab Emirates (Afkhami 2003). Finally, we collect fall-wave arrival dates for Chile (Chowell et al. 2014, 1805),

Iceland (Gottfredsson 2008), Indonesia (Chandra et al. 2013, 186), Mauritius (Gealogo 2009, 271), Samoa (Tomkins 1992), and Taiwan (Hsieh and Chan 2011, 2). We impute the arrival month for Afghanistan from that of Northern India, because in Afghanistan the epidemic "peaked about the same time as did the epidemic in Northern India" (Kohn 2007, 1–2). Shanks and Brundage (2013) state that Papua New Guinea was never hit by the pandemic.

Spring-wave arrival dates are from Patterson and Pyle (1991, 5–8), except the following: Malaysia and Singapore from Liew (2007), Mexico and the United States from Frost and Sydenstricker (1919, 1361). When none of the sources offers a separate spring-wave arrival date, we consider the date of first arrival in the combined spring-fall regressions to be the arrival date of the fall wave. When the sources provide an exact date we use it. When they provide a partial date we create an unbiased estimate of the date conditional on the information given: "October" is estimated as October 15, "early October" is estimated as October 7, "first week of October" is estimated as October 3, and so on.

GDP per capita for 1917 (1990 U.S. dollars at purchasing power parity) is from the Maddison Project (Bolt et al. 2018), supplemented by the estimates of Prados de La Escosura (2012) for several countries in Africa at 1913 where no Maddison estimate is available: Belgian Congo, Botswana, Cameroon, Chad, Gambia, Kenya, Madagascar, Mauritius, Nigeria, Senegal, Sierra Leone, Somalia, and Southern Rhodesia. Child mortality (under age 5, per 1,000 live births) in 1917 is from Gapminder (2020), which for this historical period draws on Vladimir Shkolnikov's Human Mortality Database and B.R. Mitchell's *International Historical Statistics*. The source contains estimates for Czechia and Slovakia but not for Czechoslovakia. We estimate the rate for Czechoslovakia as a population-weighted average of the source's figures for Czechia and Slovakia, where the relative populations in 1920 are from Rothenbacher (2002, 145–6).

A3.5 Influenza 1889

Final mortality estimates by city are from Valleron et al. (2010), calculated as in the source: flu mortality is sum-total excess all-cause mortality during the period of observation October 4, 1889 to March 28, 1890, with average mortality during the first four weeks of this period taken as the baseline. Most arrival dates are from Parsons (1891b), with the exception of the Swiss cities (Basel, Bern, Gens/Geneva, and Zürich) from Schmid (1895), and Warschau/Warsaw from Ruhemann (1891). Peak dates are the week of highest excess mortality in the Valleron et al. (2010) data during the first wave (up to and including the first half of February 1890).

A3.6 Coronavirus 2019

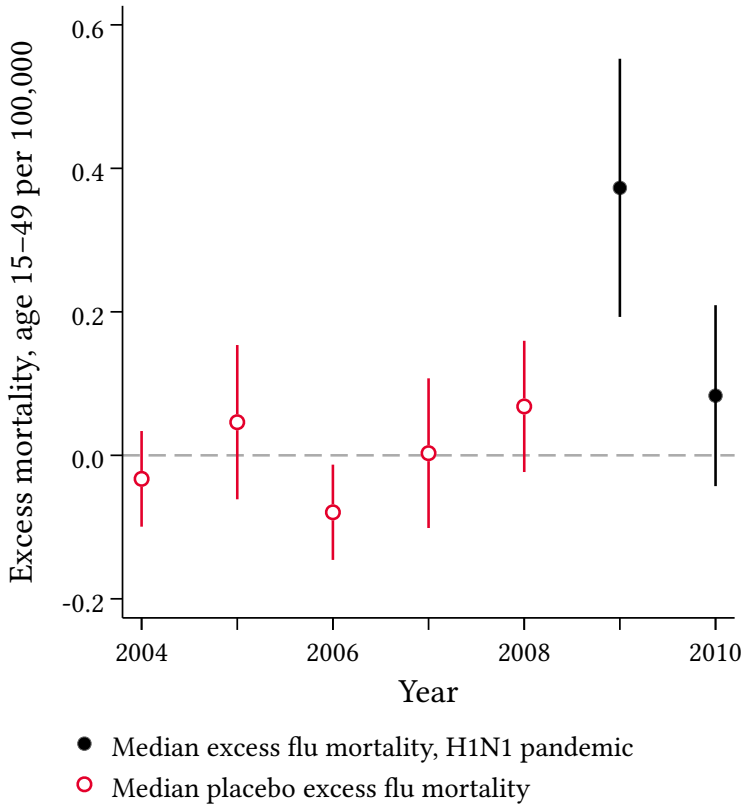
The date of arrival of the novel 2019 coronavirus in each country and country population are from Our World in Data at Oxford University (Ritchie 2020), as of June 21, 2020.

A3.7 Historical population within modern borders

For all years (1889, 1918, 1957, and 2009) we begin with the population estimates from the Maddison Project (Bolt et al. 2018), which use constant 1990 borders. For countries with no Maddison estimate, we use the estimates of the Clio Infra database (Fink-Jensen 2015) linearly interpolated between the decadal estimates in the source. Clio Infra uses 2012 borders, but for none of the relevant countries were there substantial changes in borders between 1990 and 2012. Both Maddison and Clio Infra omit estimates that could be used to reconstruct the population within the 1991-borders area of the former Soviet Union in either 1889 or 1918; we take those from Vinogradov (2016, 10).

For the 2009 pandemic, borders are resolved to their mid-1991 location by summing the populations of successor states into units representing the Former Soviet Union, Former Yugoslavia, and Former Czechoslovakia.

Appendix Figure A3: Placebo tests for excess mortality estimation



Excess mortality estimates are actual Influenza-Like Illness mortality (age 15–49) compared to counterfactual mortality predicted for 2009 and 2010 by Kernel Regularized Least Squares (Hainmueller and Hazlett 2014) fit to the data for the four years before and after the pandemic (2005–2008 and 2011–2014), pooled. Placebo estimates fit the counterfactual curve with the four pre-pandemic years *except the placebo year* closest to the pandemic, along with four post-pandemic years. For example, the placebo estimate for 2007 is actual mortality in 2007 compared to counterfactual mortality predicted by a fit to the data from 2004–2006, 2008, and 2011-2014 pooled. Vertical spikes show bootstrapped 95% confidence intervals on the median, 1,000 draws each.

A4 Excess mortality estimates, Influenza-Like Illness age 15–49

We construct our measure of 2009–2010 H1N1 influenza mortality by computing excess influenza-like illness mortality as follows. We begin with annual prime-age influenza-like illness (ILI) deaths for each country. From the World Health Organization Mortality Database (December 15, 2019 update), for each year we draw the number of deaths reported by each country due to influenza-like respiratory illness. This is the sum of deaths in the International Classification of Diseases, Tenth Revision, Clinical Modification [ICD-10-CM] codes J09–J18 and J20–J22. This includes deaths from pneumonia as Noymer (2008) recommends. We restrict the deaths to people age 15–49, because disproportionate mortality among prime-age adults strongly distinguished the 2009 pandemic strain of H1N1 influenza from seasonal influenza (Shrestha et al. 2011; Simonsen et al. 2013). Due to a large number of missing values for country population in the WHO Mortality Database, we estimate rates per 100,000 population age 15-49, using

age-specific population in each year linearly interpolated from the quinquennial estimates in the United Nations *World Populaton Prospects 2019* (POP/DB/WPP/Rev.2019/POP/F07-1).

We estimate counterfactual prime-age influenza-like illness deaths—in the absence of the pandemic—as follows. We fit the trend in these mortality rates by regressing annual mortality from the four years before and after the pandemic (2005–2008 and 2011–2014) on year with Kernel Regularized Least Squares (KRLS, Hainmueller and Hazlett 2014) to make minimal assumptions on functional form. With the KRLS fit, we impute separately for each country the counterfactual on-trend mortality for 2009 and 2010 combined. Our 2009 pandemic influenza mortality estimate is then the sum of excess mortality above trend for 2009 and 2010, for each of 99 countries.

We check this method with a set of placebo tests. Estimating excess mortality in this way should not spuriously generate excess deaths in non-pandemic years. Figure A3 shows this test. The solid dots (black) are the cross-country median of actual prime-age ILI mortality in 2009 and 2010 in excess of the KRLS counterfactual trend when the counterfactual is estimated using 2005–2008 and 2011–2014. The hollow (red) dots are placebo excess mortality: the median of actual mortality in excess of trend when the trend is estimated with a KRLS fit using the four years before but closest to the actual pandemic—except the placebo year—and the four years after the actual pandemic. For example, the counterfactual for placebo year 2008 is estimated with a KRLS fit of 2004–2007 and 2011–2014, the counterfactual for placebo year 2007 is 2004–2006, 2008, and 2011–2014, and so on. 95% confidence intervals on the median in each year are bootstrapped with 1,000 draws in each year. Median placebo excess mortality is indistinguishable from zero in each pre-pandemic year (2004–2008). It is highly statistically significant in the real pandemic year (2009). It is elevated but not statistically distinguishable from zero in 2010. This evidence is consistent with the ability of this excess mortality method to detect the occurrence of the pandemic.

A transparent check on the method is simply to plot the raw data and the KRLS counterfactual fits, in Figure A4. In the figure, countries are ordered from the country with the largest percent rise in excess mortality in 2009 to the smallest. Only the first 48 of the 99 countries are shown, that is, the 48 countries with the largest percent rise in excess mortality at the time of the pandemic. Visual inspection clearly shows the arrival of the pandemic in countries known to be hit by it, such as Mexico, the United States, Australia, Argentina, and many others. Statistical noise is evident in the figure, but overall the figure corroborates the ability of this excess mortality method to extract a meaningful signal of pandemic mortality.

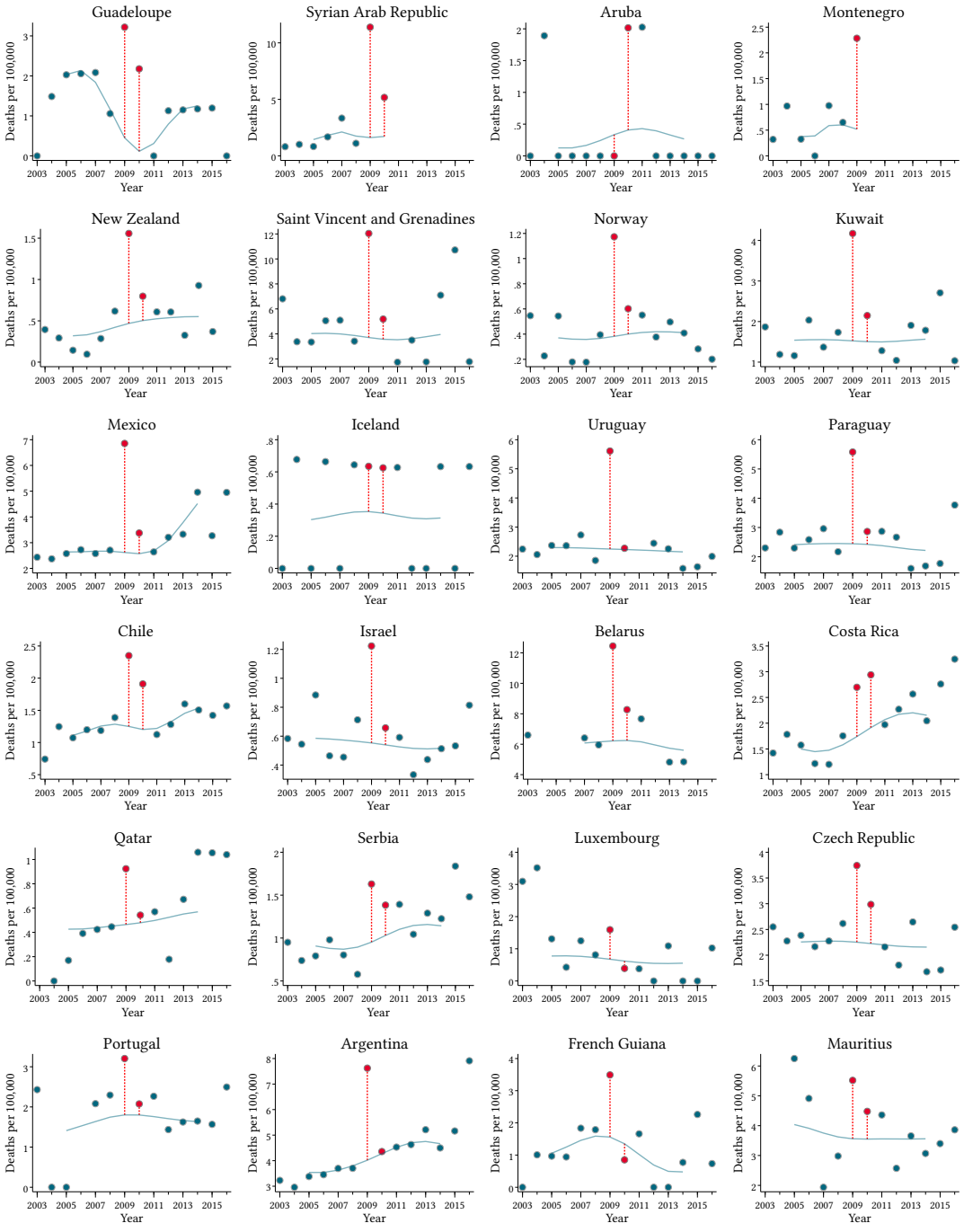
A5 Heterogeneity by hemisphere

Table A3 tests whether the relationship between international mobility and timing of the 2009–2010 H1N1 influenza pandemic arrival are heterogeneous by hemisphere, given that the timing of influenza spread often exhibits seasonal patterns. We separate countries according to whether their capital city is located north of the Tropic of Cancer, south of the Tropic of Capricorn, or between the Tropics, following e.g. Simonsen (1999).

A6 Arrival time regressions with per-capita mobility

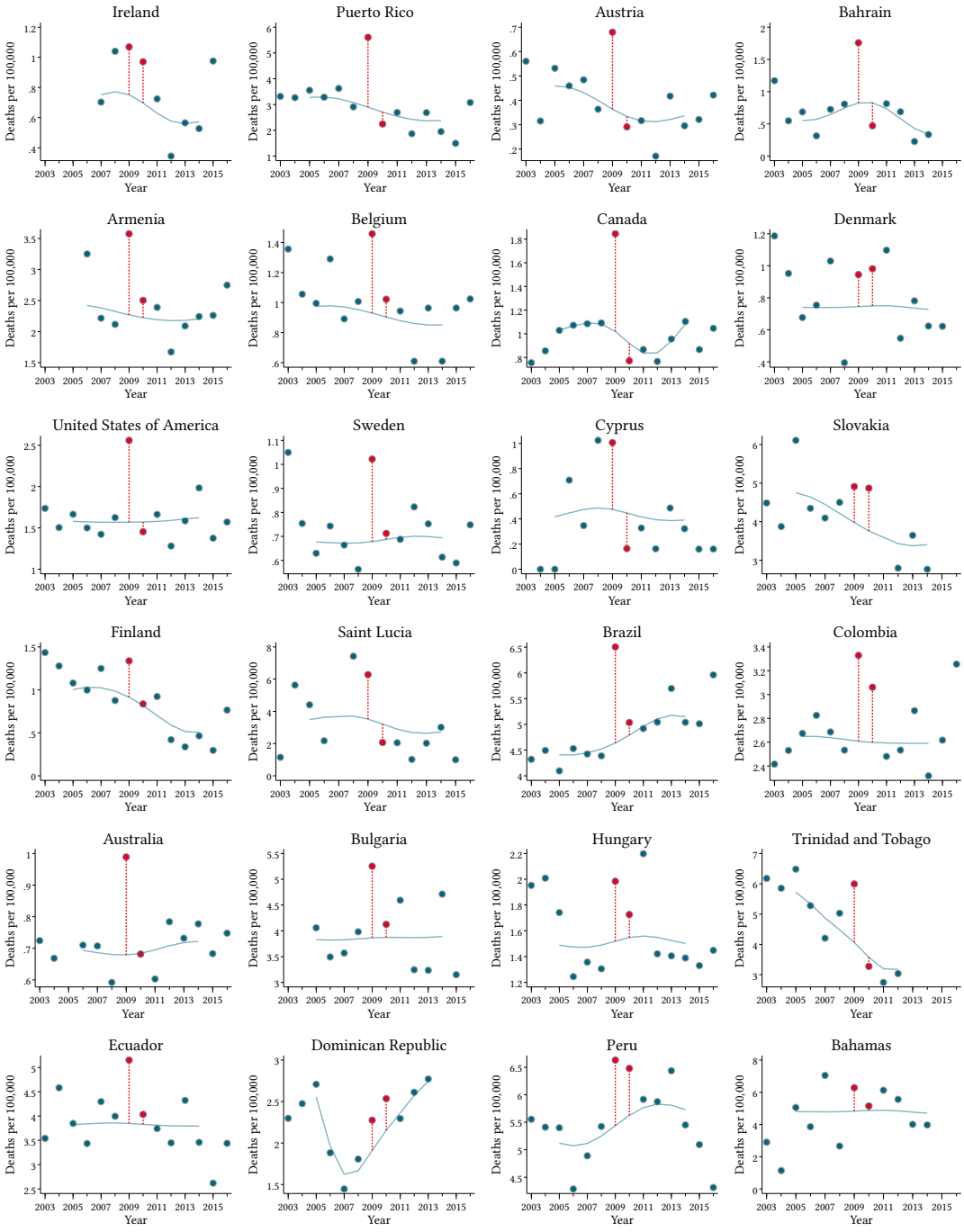
Table A4 reanalyzes main-text Table 1, using per-capita measures of immigration and incoming trips rather than absolute numbers of immigrants and trips.

Appendix Figure A4: Excess mortality, first 24 of top 48 countries by relative rise in 2009



Dots show raw data for the age-specific mortality rate per 100,000 due to Influenza-Like Illness, age 15-49, in the WHO Mortality Database, 15 December 2019 update. Regression curve is Kernel Regularized Least Squares (Hainmueller and Hazlett 2014) fit to the data for the four years before and after the pandemic (2005–2008 and 2011–2014) pooled. Dotted vertical red spikes in 2009 and 2010 show the values summed to arive at excess mortality estimates.

Appendix Figure A4: Excess mortality, next 24 of top 48 countries by relative rise in 2009



Dots show raw data for the age-specific mortality rate per 100,000 due to Influenza-Like Illness, age 15-49, in the WHO Mortality Database, 15 December 2019 update. Regression curve is Kernel Regularized Least Squares (Hainmueller and Hazlett 2014) fit to the data for the four years before and after the pandemic (2005–2008 and 2011–2014) pooled. Dotted vertical red spikes in 2009 and 2010 show the values summed to arrive at excess mortality estimates.

Appendix Table A3: Influenza 2009: International mobility and the timing of pandemic arrival, heterogeneity by latitude

Dep. var.:	Day of pandemic arrival							
Mobility measure:	*ln* Immigrants				*ln* Incoming trips			
Weight by connectivity?	—	Yes	—	Yes	—	Yes	—	Yes
	est.	*s.e.*	*est.*	*s.e.*	*est.*	*s.e.*	*est.*	*s.e.*
Linear OLS: Countries with capital city north of the Tropic of Cancer								
ln Mobility/cap.	−13.9	(2.40)	−15.7	(2.33)	−16.5	(2.23)	−14.6	(1.74)
N	88		88		88		88	
R^2	0.280		0.353		0.373		0.416	
Linear OLS: Countries with capital city south of the Tropic of Capricorn								
ln Mobility/cap.	−16.6	(6.87)	−16.9	(6.06)	−19.7	(15.7)	−22.3	(10.5)
N	10		10		10		10	
R^2	0.339		0.513		0.130		0.390	
Linear OLS: Countries with capital city between the Tropics of Cancer and Capricorn								
ln Mobility/cap.	−3.70	(3.33)	−11.9	(3.64)	−15.8	(3.77)	−20.1	(3.92)
N	86		86		86		86	
R^2	0.00739		0.0829		0.104		0.197	

Robust standard errors in parentheses to the right of each coefficient estimate.

A7 Arrival time regressions with alternative, all-origins weights on incoming trips

Table A5 reanalyzes the weighted 'incoming trips' regressions in main-text Table 1 and Appendix Table A4 using an alternative weight. The preferred weights in Table 1 and Appendix Table A4 use the "leave-one-out" centrality (λ_1), omitting the top origin country from the eigenvector calculation in order to reduce the dependence on travel with one neighbor. The alternative measure calculates the eigenvector centrality using all origins, including the top origin country (λ_0). The results are qualitatively similar across the different versions of the weights.

A8 Alternative 2009 H1N1 influenza mortality estimates by Dawood et al.

Table A6 repeats the analysis of the bivariate relationship between mobility measures and 2009 H1N1 influenza mortality using the estimates of Dawood et al. (2012). The results are qualitatively similar to those in the main text using the mortality estimates of Simonsen et al. (2013): Greater exposure to international mobility is strongly associated with lower final mortality. Estimates using the two different measures of mortality are not strictly comparable because Dawood et al. (2012) estimate not only mortality due to respiratory illness but also to cardiovascular illness.

A9 Survival analysis for pandemics of 1957, 1918, and 1889

Table A7 shows regression tables for the relationship between final mortality and arrival date for the pandemics of 1957, 1918, and 1889, where the linear regression specifications in the main text are replaced

Appendix Table A4: Influenza 2009: International mobility and the timing of pandemic arrival, with per-capita mobility exposure

Dep. var.:	Day of pandemic arrival							
Mobility measure:	*ln* Immigrants/capita				*ln* Incoming trips/capita			
Connectivity weights?	No		Yes		No		Yes	
	est.	*s.e.*	*est.*	*s.e.*	*est.*	*s.e.*	*est.*	*s.e.*
Linear OLS								
ln Mobility/cap.	−11.6	(3.56)	−13.8	(3.23)	−11.6	(3.21)	−11.3	(2.64)
N	179		179		179		179	
R^2	0.0533		0.126		0.102		0.146	
Days of delay associated with mobility reduction								
50% reduction	8.01	(2.47)	9.54	(2.24)	8.03	(2.22)	7.83	(1.83)
90% reduction	26.6	(8.2)	31.7	(7.44)	26.7	(7.39)	26.0	(6.07)
Survival regression: Loglogistic accelerated failure time								
ln Mobility/cap.	−0.0725	(0.0264)	−0.0976	(0.0249)	−0.0735	(0.0229)	−0.0836	(0.0197)
γ	0.293	(0.0165)	0.283	(0.0154)	0.289	(0.0158)	0.281	(0.0149)
N	179		179		179		179	
Days of delay associated with mobility reduction								
50% reduction	4.64	(1.73)	6.30	(1.66)	4.71	(1.51)	5.37	(1.3)
90% reduction	16.3	(6.47)	22.7	(6.45)	16.6	(5.63)	19.1	(4.96)

Observations are countries. Robust standard errors in parentheses to the right of each coefficient estimate. Constant term included but not shown. In the loglogistic survival regressions, γ is the scale parameter such that the survivor function is $S(t) \equiv (1 + (\lambda t)^{1/\gamma})^{-1}$ and $\lambda_i \equiv e^{-\beta \times \ln(\text{mobility/cap.})}$. 'Immigrants' is the average annual number of immigrants to each country during 2005–2010. 'Incoming trips' is the number of people arriving in each country, for any duration of stay, in 2011.

Appendix Table A5: Influenza 2009: International mobility and the timing of pandemic arrival, with all-origins weights

Dep. var.:	Day of pandemic arrival			
Connectivity weights?	Yes			
Mobility measure:	*ln* Incoming trips		*ln* Incoming trips/cap.	
	est.	*s.e.*	*est.*	*s.e.*
Linear OLS				
ln Mobility	−14.1	(1.75)	−11.1	(2.77)
N	184		179	
R^2	0.225		0.124	
50% reduction	9.80	(1.22)	7.68	(1.92)
90% reduction	32.5	(4.04)	25.5	(6.38)

with survival analysis—the loglogistic accelerated failure time model. The specification in the main text implicitly assumes a constant hazard rate; these regressions in the Appendix allow for a rise then fall in the hazard rate over time. With this change, the findings remain qualitatively identical to those in the main text. There is no statistically significant bivariate relationship between final mortality and arrival

Appendix Table A6: Influenza 2009: Mortality: linear, Dawood et al. mortality

Dep. var.:	*ln* Final mortality per 100,000							
Mobility measure:	*ln* Immigration per cap.				*ln* Arrivals per cap.			
Connectivity weight:	No		Yes		No		Yes	
	est.	*s.e.*	*est.*	*s.e.*	*est.*	*s.e.*	*est.*	*s.e.*
ln Mobility/cap.	−0.0846	(0.0307)	−0.0758	(0.021)	−0.0839	(0.0185)	−0.0603	(0.0142)
N	185		185		185		185	
R^2	0.0561		0.0763		0.108		0.0863	
Mean mortality	3.31	(0.114)	3.31	(0.114)	3.31	(0.114)	3.31	(0.114)
50% reduction	0.200	(0.0746)	0.178	(0.0509)	0.198	(0.0450)	0.141	(0.0340)
90% reduction	0.711	(0.284)	0.631	(0.191)	0.704	(0.171)	0.492	(0.124)

Robust standard errors in parentheses to the right of each coefficient estimate.

time in the pandemics of 1957, 1918 (fall wave) and 1889. When the spring and fall waves of 1918 are considered as a single pandemic event, as in the linear regressions in the main text, longer delays in arrival are associated with higher final mortality with high statistical significance.

Appendix Table A7: Influenza 1957, 1918, and 1889: Start date versus final mortality, survival analysis specification

(a) *Influenza 1957, countries: Accelerated failure time, loglogistic*

Dep. var.:	Date of arrival			
Mortality measure:	1957 only		1957–1959	
	est.	*s.e.*	*est.*	*s.e.*
ln Mortality	−0.0204	(0.0460)	−0.0243	(0.0545)
γ	0.167	(0.0295)	0.165	(0.0272)
N	34		35	

(b) *Influenza 1918, countries: Accelerated failure time, loglogistic*

Dep. var.:	Date of arrival			
Wave:	Fall 1918 only		Spring/Fall 1918	
	est.	*s.e.*	*est.*	*s.e.*
ln Mortality	0.0573	(0.0981)	0.200	(0.0692)
γ	0.424	(0.0688)	0.305	(0.0487)
N	45		50	

(c) *Influenza 1889, cities: Accelerated failure time, loglogistic*

Dep. var.:	Date of arrival			
Date:	Start date		Peak mortality date	
	est.	*s.e.*	*est.*	*s.e.*
ln Mortality	0.00741	(0.0486)	−0.00951	(0.00819)
γ	−3.25	(0.251)	−3.76	(0.132)
N	34		87	

Robust standard errors in parentheses to the right of each coefficient estimate. Constant term included but not shown.

Appendix References

Afkhami, Amir, "Compromised Constitutions: The Iranian Experience with the 1918 Influenza Pandemic," *Bulletin of the History of Medicine*, 2003, *77* (2), 367–392.

Azose, Jonathan J and Adrian E Raftery, "Estimation of emigration, return migration, and transit migration between all pairs of countries," *Proceedings of the National Academy of Sciences*, 2019, *116* (1), 116–122.

Balcan, Duygu, Hao Hu, Bruno Goncalves, Paolo Bajardi, Chiara Poletto, Jose J Ramasco, Daniela Paolotti, Nicola Perra, Michele Tizzoni, Wouter Van den Broeck et al., "Seasonal transmission potential and activity peaks of the new influenza A (H1N1): a Monte Carlo likelihood analysis based on human mobility," *BMC medicine*, 2009, *7* (1), 45.

Barber, Ryan M, Nancy Fullman, Reed JD Sorensen, Thomas Bollyky, Martin McKee, Ellen Nolte, Amanuel Alemu Abajobir, Kalkidan Hassen Abate, Cristiana Abbafati, Kaja M Abbas et al., "Healthcare Access and Quality Index based on mortality from causes amenable to personal health care in 195 countries and territories, 1990–2015: a novel analysis from the Global Burden of Disease Study 2015," *The Lancet*, 2017, *390* (10091), 231–266.

Bolt, Jutta, Robert Inklaar, Herman de Jong, and Jan Luiten van Zanden, "Rebasing 'Maddison': New income comparisons and the shape of long-run economic development," Maddison Project Working paper 10. Groningen Growth and Development Centre, Faculty of Economics and Business. University of Groningen 2018.

Chandra, Siddharth, Eva Kassens-Noor, Goran Kuljanin, and Joshua Vertalka, "A geographic analysis of population density thresholds in the influenza pandemic of 1918–19," *International Journal of Health Geographics*, 2013, *12* (1), 9.

Chowell, Gerardo, Lone Simonsen, Jose Flores, Mark A Miller, and Cécile Viboud, "Death patterns during the 1918 influenza pandemic in Chile," *Emerging infectious diseases*, 2014, *20* (11), 1803.

Dawood, Fatimah S, A Danielle Iuliano, Carrie Reed, Martin I Meltzer, David K Shay, Po-Yung Cheng, Don Bandaranayake, Robert F Breiman, W Abdullah Brooks, Philippe Buchy et al., "Estimated global mortality associated with the first 12 months of 2009 pandemic influenza A H1N1 virus circulation: A modelling study," *The Lancet Infectious Diseases*, 2012, *12* (9), 687–695.

Dunn, Frederick L, "Pandemic influenza in 1957: review of international spread of new Asian strain," *Journal of the American Medical Association*, 1958, *166* (10), 1140–1148.

Fink-Jensen, Jonathan, "Total Population," Clio Infra database, IISH Dataverse. Amsterdam: International Institute of Social History, Royal Netherlands Academy of Arts and Sciences 2015.

Frost, W. H. and Edgar Sydenstricker, "Epidemic Influenza in Foreign Countries," *Public Health Reports (1896–1970)*, 1919, *34* (25), 1361–1376.

Gapminder, "Child Mortality Estimates, Version 11," Stockholm: Gapminder 2020.

Gealogo, Francis A., "The Philippines in the World of the Influenza Pandemic of 1918–1919," *Philippine Studies*, 2009, *57* (2), 261–292.

Gottfredsson, Magnus, "The Spanish flu in Iceland 1918. Lessons in medicine and history," *Laeknabladid*, November 2008, *94* (11), 737—745.

Hainmueller, Jens and Chad Hazlett, "Kernel regularized least squares: Reducing misspecification bias with a flexible and interpretable machine learning approach," *Political Analysis*, 2014, *22* (2), 143–168.

Head, Keith, Thierry Mayer, and John Ries, "The erosion of colonial trade linkages after indepen-

dence," *Journal of International Economics*, 2010, *81* (1), 1–14.

Herda, Phyllis S, "Disease and the colonial narrative-The 1918 influenza pandemic in Western Polynesia," *New Zealand Journal of History*, 2000, *34* (1), 133–144.

Hsieh, Ying-Hen and Chi-Ho Chan, "Excess healthcare burden during 1918-1920 influenza pandemic in Taiwan: implications for post-pandemic preparedness," *BMC public health*, 2011, *11* (1), 41.

Johnson, Niall PAS and Juergen Mueller, "Updating the accounts: global mortality of the 1918-1920" Spanish" influenza pandemic," *Bulletin of the History of Medicine*, 2002, pp. 105–115.

Jordan, Edwin O, "The present status of the influenza problem," *American Journal of Public Health*, 1925, *15* (11), 943–947.

Killingray, David, "The influenza pandemic of 1918–1919 in the British Caribbean," *Social History of Medicine*, 1994, *7* (1), 59–87.

Kohn, George C, *Encyclopedia of Plague and Pestilence: From Ancient Times to the Present*, third ed., New York: Facts on File, 2007.

Liew, Kai Khiun, "Terribly Severe Though Mercifully Short: The Episode of the 1918 Influenza in British Malaya," *Modern Asian Studies*, 2007, *41* (2), 221–252.

Lind, Kjell Jostein Langfeldt, "The Impact of the 1918 Influenza Pandemic on Greater Syria," M.A. Dissertation. London: School of Oriental and African Studies 2012.

Mayer, Thierry and Soledad Zignago, "Notes on CEPII's distances measures: The GeoDist database," 2011.

Noymer, Andrew, "Influenza analysis should include pneumonia," *American Journal of Public Health*, 2008, *98* (11), 1927–1928.

OpenFlights.org, "OpenFlights Airports Database," 2009.

Parsons, H Franklin, *Report on the influenza epidemic of 1889–90 by Dr. Parsons, with an introduction by the Medical Officer of the local government board*, London: Her Majesty's Stationery Office, 1891.

Patterson, K David and Gerald F Pyle, "The diffusion of influenza in sub-Saharan Africa during the 1918–1919 pandemic," *Social Science & Medicine*, 1983, *17* (17), 1299–1307.

__ and __, "The Geography and Mortality of the 1918 Influenza Pandemic," *Bulletin of the History of Medicine*, 1991, *65* (1), 4–21.

Prados de La Escosura, Leandro, "Output per head in pre-independence Africa: quantitative conjectures," *Economic History of Developing Regions*, 2012, *27* (2), 1–36.

Ritchie, Hannah, "Our World in Data COVID-19 dataset," June 1 update. Oxford: Oxford Martin School at the University of Oxford 2020.

Rothenbacher, Franz, *The European Population 1850–1945*, London: Palgrave Macmillan UK, 2002.

Ruhemann, Jacob, *Die Influenza in dem Winter 1889–90*, Leipzig, Thieme, 1891.

Schmid, Friedrich, *Die Influenza in der Schweiz in den Jahren 1889–1894: Auf Grund amtlicher Berichte und sonstigen Materials*, Komissionsverl. d. Buchhandlung Schmid, Francke, 1895.

Shanks, G Dennis and John F Brundage, "Pacific islands which escaped the 1918–1919 influenza pandemic and their subsequent mortality experiences," *Epidemiology and Infection*, 2013, *141* (2), 353–356.

Shrestha, Sundar S, David L Swerdlow, Rebekah H Borse, Vimalanand S Prabhu, Lyn Finelli, Charisma Y Atkins, Kwame Owusu-Edusei, Beth Bell, Paul S Mead, Matthew Biggerstaff et al.,

"Estimating the burden of 2009 pandemic influenza A (H1N1) in the United States (April 2009–April 2010)," *Clinical Infectious Diseases*, 2011, *52* (suppl_1), S75–S82.

Sigurjónsson, Júlíus, Björn Sigurdsson, and Halldór Grímsson, "Experience with influenza vaccination in Iceland, 1957," *Bulletin of the World Health Organization*, 1959, *20* (2-3), 401.

Simonsen, Lone, "The global impact of influenza on morbidity and mortality," *Vaccine*, 1999, *17*, S3–S10.

__ , Peter Spreeuwenberg, Roger Lustig, Robert J Taylor, Douglas M Fleming, Madelon Kroneman, Maria D Van Kerkhove, Anthony W Mounts, W John Paget et al., "Global mortality estimates for the 2009 Influenza Pandemic from the GLaMOR project: a modeling study," *PLoS Medicine*, 2013, *10* (11).

Tomkins, Sandra M., "The Influenza Epidemic of 1918-19 in Western Samoa," *Journal of Pacific History*, 1992, *27* (2), 181–197.

UNESCO, "How Asian 'Flu Swept around the World," *The UNESCO Courier*, 1958, *11* (5), 12–13.

UNWTO, "Methodological Notes to the Tourism Statistics Database," United Nations World Tourism Organization 2020. Accessed May 30, 2020.

Valleron, Alain-Jacques, Anne Cori, Sophie Valtat, Sofia Meurisse, Fabrice Carrat, and Pierre-Yves Boëlle, "Transmissibility and geographic spread of the 1889 influenza pandemic," *Proceedings of the National Academy of Sciences*, 2010, *107* (19), 8778–8781.

Vaughan, Warren Taylor, *Influenza: An Epidemiologic Study* number 1, Baltimore: The American Journal of Hygiene, 1921.

Viboud, Cécile, Lone Simonsen, Rodrigo Fuentes, Jose Flores, Mark A Miller, and Gerardo Chowell, "Global mortality impact of the 1957–1959 influenza pandemic," *Journal of infectious diseases*, 2016, *213* (5), 738–745.

Vinogradov, A.G., *The population of Russia and the USSR from ancient times to the present*, IPGEB, 2016.

Vollmer, Sebastian and Juditha Wójcik, "The long-term consequences of the global 1918 influenza pandemic: A systematic analysis of 117 IPUMS international census data sets," Working Paper, University of Mainz 2017.

WHO, "Methodology for the update of the Global Health Expenditure Database, 2000–2017," Geneva: World Health Organization 2019.