The Spread of COVID-19 in London: Network Effects and Optimal Lockdowns*

Christian Julliard [†] Ran Shi [‡]

Kathy Yuan [§]

October, 2022

Abstract

We generalise a stochastic version of the workhorse SIR (Susceptible-Infectious-Removed) epidemiological model to account for spatial dynamics generated by network interactions. Using the London metropolitan area as a salient case study, we show that commuter network externalities account for about 42% of the propagation of COVID-19. We find that the UK lockdown measure reduced total propagation by 57%, with more than one third of the effect coming from the reduction in network externalities. Counterfactual analyses suggest that: i) the lockdown was somehow late, but further delay would have had more extreme consequences; ii) a targeted lockdown of a small number of highly connected geographic regions would have been equally effective, arguably with significantly lower economic costs; iii) targeted lockdowns based on threshold number of cases are not effective, since they fail to account for network externalities.

Keywords: COVID-19, networks, key players, spatial modelling, SIR model. JEL codes: I12, I18, C51, C54, D85.

^{*}We thank seminar participants at the London School of Economics and London Business School.

[†]Department of Finance, FMG, and SRC, London School of Economics, and CEPR; c.julliard@lse.ac.uk

[‡]Department of Finance, London School of Economics, r.shi1@lse.ac.uk

[§]Department of Finance, FMG, London School of Economics, and CEPR; k.yuan@lse.ac.uk

Introduction

Facing challenges of the COVID-19 pandemic, governments around the world resort to lockdown policies to contain the disease transmission. Carefully designed lockdown policies, especially ones that choosing lockdown regions selectively, can be particularly useful for preventing disease spread while controlling cutbacks on the broader economic activity. In this context, the essential trade-off between mortality and productivity points to solving the problem of minimizing the number of infectives subject to a constraint on the number of locked-down regions.

The present paper makes progress along this direction. To study the potential outcomes of different lockdown policies, we first build and estimate a statistical model for disease surveillance data to understand the COVID-19 epidemic dynamics. Our model, featuring location heterogeneity and network effects, generalizes the single-population deterministic susceptible-infectious-recovered (SIR) model developed by Kermack and McKendrick (1927)¹. We extend their model to a stochastic one that incorporates multiple subpopulations. We then estimate the model for daily COVID-19 surveillance data across London boroughs, with a special focus on the role of residential and working network links. Our estimation results confirm the essential role of network spillover effects in COVID-19 spread in London, that is, existing cases in one London borough may transmit the disease to residents of another borough through residential and working links. Specifically, on average, over forty percent of daily new COVID-19 cases can be attributed to these network effects.

Based on our estimated model, we then conduct simulations to evaluate outcomes of alternative lockdown policies. The pronounced role of network effects in the spread of COVID-19 calls for alternative views toward the design of lockdown policies.

Summary of findings: to be added

Literature.

1 A Network SIR Model

This section introduces our network SIR model. We first extend the deterministic singlepopulation SIR model to its stochastic counterpart and derive necessary distributional results for mapping the model to disease surveillance data. Then, we extend the statistical model to multiple subpopulations, with a special account for network effects, formalising a network SIR framework. Finally, we enrich the framework by introducing a parametrisation scheme tailored for analysing the spreading of COVID-19 in London.

¹The core ingredients of SIR models was first formulated by Lowell Reed and Wade Frost in the 1920s, though not published.

1.1 The deterministic SIR dynamics

We begin with notation. For a given population of fixed size N, the triplet $\{S_t, I_t, R_t\}$ represents, respectively, the *cumulative* numbers of susceptible (S), infectious (I), and recovered (R) individuals at time t. Susceptible individuals get infected through mixing with infectious individuals, featuring the dynamic

$$\dot{S}_t = -\theta_I \frac{l_t}{N} S_t,\tag{1}$$

in which the overdot notation represents derivatives with regard to time; I_t/N is the level of disease prevalence in the current population; θ_I is a parameter measuring the contact rate (times the probability of infection per contact). Infected individuals recover at a rate θ_R , which implies

$$\dot{R}_t = \theta_R I_t. \tag{2}$$

The sum $S_t + I_t + R_t$ equals the total population *N*, constant by assumption. Therefore, $\dot{S}_t + \dot{I}_t + \dot{R}_t = 0$. Plugging in (1) and (2) and rearranging terms,

$$\dot{I}_t = \left(\theta_I \frac{S_t}{N} - \theta_R\right) I_t.$$
(3)

Our analysis will focus on the dynamics of the infected population I_t . Based on this assumption, a constant, namely α , defined as $\alpha_t \triangleq \theta_I S_t / N - \theta_R$, is sufficient to capture the dynamics of the infectious and infected subpopulations in our investigation sample. The framework can be generalised to allow, as we do below, time variation in α .

1.2 The stochastic SIR model

To map the deterministic SIR model to the data, we need to introduce probabilistic "error" terms. However, arbitrarily introduced errors such as Gaussian errors may reduce statistical power due to misspecified likelihood functions. The issue is more pronounced when the number of disease incidents is relatively small (compared with the total population size), which is the case in our data (less than three detected cases per thousand people in London by the end of the lockdown period, June 2020). Thus, we aim to derive this distribution coherently from a well-defined probabilistic analogue to the deterministic SIR dynamics. The natural stochastic extension to the differential equation $dI_t/dt \approx \alpha_t I_t$ is a continuous-time Markov chain of the form:

$$\mathbb{P}\left[y \text{ new infections in } (t, t+dt) \mid I_t = x\right] = \begin{cases} \alpha_t x dt, & y = 1\\ o(dt), & y \ge 2, \\ 1 - \alpha_t x dt - o(dt), & y = 0 \end{cases}$$
(4)

where o(dt) satisfies $o(dt)/dt \rightarrow 0$ as $dt \rightarrow 0$. Interpreting this probabilistic statement is straightforward. Given the current number of infective individuals x, within an infinites-

imally short time interval, one additional person can contract the disease with probability $\alpha x dt$. This random process is a simple birth process (also known as the Yule-Furry process, see for example, Grimmett and Stirzaker (2001, p. 250)).

When working with surveillance data of infectious diseases, we only observe numbers of new cases within discrete time intervals (say, one day or one week). Based on the specification of (4), we can solve for the distribution of new case counts within a time interval of length h (instead of length dt which goes to zero), denoted by

$$p_{t \to t+h}(y \mid x) = \mathbb{P}\left[y \text{ new infections in } (t, t+h) \mid I_t = x\right].$$

Solving the implied Kolmogorov forward equation (see Grimmett and Stirzaker (2001, p. 250)), the simple birth process in (4) yields an analytical expression of the probability mass function $p_h(y \mid x)$:

$$p_{t \to t+h}(y \mid x) = \frac{\Gamma(y+x)}{\Gamma(x)\Gamma(y+1)} \left(e^{-\alpha_t h}\right)^x \left(1 - e^{-\alpha_t h}\right)^y,\tag{5}$$

where $\Gamma(\cdot)$ represents the gamma function. The above expression describes a negative Binomial distribution – a mixture of Poisson distributions with mixing of the Poisson rate driven by a gamma distribution. It provides the chance of y successes after exactly x failures in a sequence of independent Bernoulli trials, each having a probability of success $p = 1 - e^{-\alpha h}$. The negative Binomial is an appropriate representation for discrete arrival data over an unbounded positive range whose sample variance exceeds the sample mean. In such cases, the observations are overdispersed with respect to a Poisson distribution (for which the mean is equal to the variance). Since the negative binomial distribution has one more parameter than the Poisson distribution, the second parameter can be used to adjust the variance separately from the mean. Furthermore, it implies that first and second conditional moments are positively correlated – a feature consistent with epidemic dynamics. The probability $p(=1 - e^{-\alpha h})$ also has a clear interpretation in our context of disease transmission. Let y = 0 and x = 1 in (5), then p is the probability that an existing disease case infects at least one person within a period of length h (since $e^{-\alpha h} = 1 - p = p_h(0 \mid 1)$, implying that $p = 1 - p_h(0 \mid 1) = \sum_{y \ge 1} p_h(y \mid 1)$).

Denote the daily count of new infected individuals as y_t , t = 1, ..., T. Normalising the length of the time interval h to one day and denoting with \mathcal{F}_t the information available up to time t, we have

$$y_t \mid \mathcal{F}_{t-1} \sim \text{NegBinom}\left(p_t, x_{t-1}\right), \tag{6}$$

where $p_t = 1 - e^{-\alpha_t}$. Conditional expectations of new disease cases can be computed directly as

$$\mathbb{E}\left[y_t \mid \mathcal{F}_{t-1}\right] = \mu_t = a_t x_{t-1} \tag{7}$$

where $a_t = e^{\alpha_t} - 1 = p_t/(1-p_t)$ can be understood as the odds ratio comparing the probability

that an infected person does transmit the disease against that he does not. This simple linear relationship between the conditional expectations of new case numbers and existing infectious individuals is the key modelling assumption that we adopt throughout our analysis.

$$a_t = \exp\left(\theta_I \frac{S_t}{N} - \theta_R\right) - 1$$

From the negative binomial distribution, we can also calculate the conditional variance of new case counts as

$$\operatorname{var}[y_t \mid \mathcal{F}_{t-1}] = (a+1)ax_{t-1}.$$
(8)

If $a + 1 = 1/(1 - p) \rightarrow 1$, that is $p \rightarrow 0$, which means that the disease transmission rate is extremely small, the conditional variance of y_t equals its conditional mean. This corresponds to a Poisson distribution specification for y_t as $y_t \sim \text{Pois}(ax_{t-1})$. Equations (C) and (8) imply a positive correlation between the first and second moments for the number of new infected individuals – a feature consistent with the data under analysis.

The remaining issue is how to determine the number of *actively* infectious cases x_t . This issue arises because people who have been infected might recover (or they may die), as we have initially discussed in the SIR model. We adopt the following formulation for x_t :

$$x_t = \sum_{\ell=0}^{L-1} \nu(\ell) y_{t-\ell},$$
(9)

which assumes that infected individuals can transmit the disease for L periods. The decay function $\nu(\ell) \in (0,1)$ characterises "the rate of infectivity," in the language of Kermack and McKendrick (1927).² That is, for a person who has been infected for ℓ periods, the chance of transmitting the disease to another person is reduced to $100 \times \nu(\ell)$ percent of the initial level. Another way of interpreting this specification is that $100[1 - \nu(\ell)]$ percent of the infected individuals are no more infectious ℓ periods after having the contracted the disease. Taken together, equations (C) and (9) imply an autoregressive dynamic for the number of new infections.

1.3 A stochastic network SIR model

Now, we extend the stochastic model introduced above to account for multiple subpopulations connected in a network. In the context of the COVID-19 spread in London, we treat London

²The simple SIR, especially that with constant rate of recovery as in (2), is a tractable special case considered in Kermack and McKendrick (1927). In the general setup of their model, constant rate of recovered is replaced with a specification similar as in (9).

boroughs as subpopulations of constant sizes N_i , i = 1, ..., n. Time-*t* conditional expectations of new case counts are concatenated into a vector $\boldsymbol{\mu}_t = [\mu_{1t}, ..., \mu_{nt}]^{\top}$, where $\mu_{it} = \mathbb{E}[y_{it} | \mathcal{F}_{t-1}]$ is the expected number of new cases in borough *i* at time *t*. The distribution of newly infected individuals y_{it} follows the negative Binomial structure in (6), specified as follows:

$$y_{it} \sim \text{NegBinom}\left(p_{it}, x_{i,t-1}\right) \text{ where } p_{it} = \frac{\mu_{it}}{x_{i,t-1} + \mu_{it}}.$$
 (10)

To extend the conditional mean equation (C) to the multivariate case, we specify μ_t as

$$\mu_{t} = A x_{t-1} + \mu_{t}^{EN} = \underbrace{\mu_{t}^{AR}}_{\text{diag}(A)x_{t-1}} + \underbrace{\mu_{t}^{NE}}_{(A-\text{diag}(A))x_{t-1}} + \mu_{t}^{EN}$$
(11)

where $x_t = [x_{1t}, \dots x_{nt}]^\top = \sum_{\ell=0}^{L-1} \nu(\ell) y_{t-\ell}$ is the vector of infectious individuals in the London boroughs and *A* is an $n \times n$ matrix of coefficients. The first term in (11), Ax_{t-1} , and the definition of *x*, imply *vector* autoregressive dynamics in the disease propagation. The additional term μ_t^{EN} , which we call the *endemic* term, aims to capture variations in disease dynamics that are not explained by the *epidemic* component Ax_{t-1} . Such endemic forces aim to accommodate seasonality, behavioural responses, or transmission dynamics induced by external forces.³ This endemic/epidemic decomposition is commonly adopted in the empirical analysis of epidemiological surveillance data (see, e.g. Finkenstädt and Grenfell (2000), Held, Höhle, and Hofmann (2005), Lawson (2013)).

By further separating the endemic effect in equation (11) into its intra- and inter-borough elements, we have a natural decomposition of the conditional expectation of new infections into three components: the local (intra-borough) autoregressive effects μ_t^{AR} , the network (inter-boroughs) effects μ_t^{NE} , and the endemic effects μ_t^{EN} .

Autoregressive effects The local autoregressive effects capture disease dynamics as if each subpopulation were in isolation – i.e., as if new cases were driven only by infectious residents within the same borough. These are equivalent to assigning the single-population dynamic in equation (C) independently for each subpopulation. We adopt the simplest specification for the autoregressive effects by treating them as homogeneous among subpopulations. This implies that $\operatorname{diag}(A) = \gamma I$, where the coefficient γ can be understood as the odds ratio parameter *a* in equation (C).⁴

Network effects Moving from one homogeneous population to multiple subpopulations, a key new component is the network effects μ_t^{NE} , which is the main focus of our analysis.

³This is to be partially distinguished from the concept of "endemic diseases", which command stationary transmission dynamics due to the combination of slow propagation rate and temporary immunity.

⁴The specification is easily generalisable to location specific local dynamics by setting diag $(A) = \text{diag}\left(\left\{\gamma_j\right\}_{i=1}^n\right)$.

These effects are driven by a network (or multiple networks) of connections between subpopulations. Individuals contact and transmit the disease along predefined network links; the intensity of transmission is captured by the "strength" of links (i.e., varying edge values of the network). These network effects of disease transmission have long been acknowledged and were brought to light by Newman (2002). The network effects can also be interpreted as matching intensities between individuals from different subpopulations as in Acemoglu, Chernozhukov, Werning, and Whinston (2020). In particular, given an observable adjacency matrix $W = \{w_{ij}\}_{1 \le i,j \le n}$ capturing the connections among London boroughs, we model the network effects μ_t^{NE} in (11) as

$$\mu_{it}^{NE} = \phi\left(\sum_{j \neq i} w_{ij} x_{j,t-1}\right).$$
(12)

Under this formulation, for residents of borough *i*, the chance of being infected by residents from borough *j* is proportional to the strength of connections between the two boroughs, as describe by the network topology *W*. The constant ϕ determines the overall strength of these network effects. The specification in (12) can be extended to allow for multiple networks, with adjacency matrices given by $W^{(g)} = \left\{ w_{ij}^{(g)} \right\}_{1 \le i,j \le n}$, $g = 1, \ldots, G$, as follows:

$$\mu_{i,t}^{NE} = \sum_{g=1}^{G} \phi^{(g)} \left(\sum_{j \neq i} w_{ij}^{(g)} x_{j,t-1} \right).$$
(13)

Here, the topology specific coefficients $\phi^{(g)}$ capture the strength of the disease propagation though the various sets of links considered.

The above assumptions about the autoregressive and network effects directly translate into a parametric specification for matrix A in (11):

$$\boldsymbol{A} = \gamma \boldsymbol{I} + \sum_{g=1}^{G} \boldsymbol{\phi}^{(g)} \boldsymbol{W}^{(g)}.$$
(14)

Hence, our formulation implies that the conditional expectation of new infections is driven by both local autoregressive, as well as spatial, dynamics.

The set of parameters $\{\gamma, \phi^{(1)}, \ldots, \phi^{(G)}\}$, jointly with the adjacency matrices $W^{(g)}$, fully determine matrix A, which is essential for understanding the disease transmission dynamics. We expect these parameters to be non-negative for the data-generating process to be well defined, and impose this restriction by working with exponential forms. Furthermore, as the lockdown policy may affect these fundamental parameters, we model them as time varying. That is:

$$\gamma \equiv \gamma_t = \exp\left(\gamma_0 + D_t\gamma\right),\tag{15}$$

$$\phi^{(g)} \equiv \phi_t^{(g)} = \exp\left(\phi_0^{(g)} + D_t \phi^{(g)}\right), \quad g = 1, \dots, G,$$
(16)

where $D_t \in \{0, 1\}$ is a dummy variable equal to zero before the lockdown and one afterwards. The impact of the policy change is therefore quantified as $\exp(\gamma)$ for autoregressive effects and $\exp(\phi^{(g)})$ for network effects, multiplicatively. For example, if $\exp(\gamma) = 0.70$, then the policy reduces the local autoregressive effects by $100 \times (1 - 0.7)\% = 30\%$.

Endemic effects The endemic effects in (11) are specified as

$$\mu_{i,t}^{EN} = \exp\left(\mathbf{z}_t^{\top} \boldsymbol{\beta} + \eta_i\right) N_i, \qquad i = 1, \dots, n$$
(17)

which is proportional to the size of subpopulations N_i . The vector z_t contains deterministic time trends such as polynomial and trigonometrical functions of time t. Additional timevarying terms for control, such as time-varying testing intensity and positive-to-test ratio, are also included in this vector. Location-specific fixed effects η_i , are added in the endemic terms to account for static demographic and socioeconomic heterogeneities among subpopulations. In particular, control variables in the endemic term are specified as $z_t = [t, t^2, t \times D_t, t^2 \times D_t]$. Linear and quadratic terms of time are motivated by the hump-shaped disease spread trend (see, e.g., Li and Linton (2020)). These deterministic trends are allowed to change after the lockdown policy, with two interaction terms, $t \times D_t$ and $t^2 \times D_t$, added in.

Estimation of the model parameters is performed via maximum likelihood. The tuning parameter ρ of the exponential decaying kernel $\nu(\ell) = \exp(-\rho\ell)$ is selected via (maximising) the profile likelihood of each model. Standard errors are all calculated based on the outer product of the score vectors à la Berndt, Hall, Hall, and Hausman (1974).

We focus on performing inferences about parameters that determine the epidemic (autoregressive and network) effects. For parameters governing the baseline (before-policy) disease spread dynamics, we test whether their exponentials are significantly greater than zero, that is, $\exp(\gamma_0) > 0$ and $\exp\left(\phi_0^{(g)}\right) > 0$ for all $g \in \mathcal{G}$ are our null hypotheses. These tests evaluate if there are substantial autoregressive or network effects during the COVID-19 pandemic in London. For parameters quantifying the policy impacts, we ask if their exponentials are significantly smaller than one, that is, if our null hypotheses are $\exp(\gamma) < 1$ and $\exp\left(\phi^{(g)}\right) < 1$. This is because (the exponentials of) these parameters all measure (gross) percentage changes from baseline estimates after the lockdown policy is implemented.

2 Data and Estimation

This section surveys the multiple data sources, including networks of London boroughs we use in this study, and presents estimation results and their implications.

2.1 COVID-19 data and demographic information

Disease surveillance data are provided by the UK government.⁵ This database reports the number of daily new cases found in each local authority of the UK. We focus on local authorities of London, consisting of 32 London boroughs.⁶ Our sample period is from March 1 to June 4, 2020. We begin our analysis from March because the number of reported cases is small from January to February in London.⁷

We illustrate in Panel (A) of Figure 2 the total number of reported COVID-19 cases in each borough until June 4, 2020. In general, the number of cases of the pandemic in each borough correlates with subpopulation sizes, which is shown in Panel (B) of the same figure. The correlation is 0.78 throughout our sample period. This association corroborates our specification for the endemic terms in equation (17), which is proportional to the subpopulation sizes.

There are three major policy dates for COVID-19 in the UK. They are March 16 when citizens were recommended to avoid traveling and stay home, March 20 when schools and pubs were closed, and March 23 from which full nationwide lockdown measures became effective. We choose March 23 as our policy date and evaluate the impacts of nationwide lockdown measures on the autoregressive and network effects using the specifications in (15) and (16).

Additionally, for subpopulation (borough) sizes, we use records from a housing-led population projection conducted by the Greater London Authority (GLA).⁸ The GLA demography team constructs these projections based on trends in fertility, mortality and migration, and housing development on an annual basis. The starting point of these projections is dwelling records from the 2011 UK census.⁹

We use the UK nationwide testing data from the GitHub repository of Our World in Data (OWID).¹⁰ Since May 23, 2020, the UK government has stopped publishing data regarding the number of "people tested". OWID uses instead the official data for the number of "tests performed" since then. For these numbers, only official swab tests count and all serology tests are disregarded.¹¹ Available data for both series, the numbers of people tested and tests

⁵https://coronavirus.data.gov.uk/

⁶The city of London, strictly speaking, is not an official London borough (but, still a local authority), and in our dataset it is merged with Westminster. It is noteworthy that the city of London has an extremely small number of residents (less than ten thousand) and only thirteen disease cases in total (till June 4, 2020).

⁷There are only fifteen cases in total, with Southwark borough having the largest number, which is only three. ⁸Link: https://data.london.gov.uk/dataset/housing-led-population-projections

⁹Population census across all four UK countries is taken every ten years. The 2011 census is the most recent. We also use the census data for constructing our networks, as we discuss below.

¹⁰Link: https://github.com/owid/covid-19-data/tree/master/public

¹¹Official swab tests are those conducted in Public Health England (PHE) labs and National Health Service (NHS) hospitals, as well as those processed *in-person* under government guidance for a wider population. Swab tests for surveillance purpose undertaken by PHE, Office for National Statistics (ONS), Biobank, universities, and other partners do not count.

performed, are plotted in Figure 1. Within their overlapping time window (April 26 to May 22, 2020), the two series agree well.

2.2 Network construction

We construct networks connecting residents of different boroughs through commuting links. We use data from the 2011 UK Census to create a *directed* and *weighted* graph, of which 32 nodes denote the London boroughs¹². Edges of this graph quantify employment links between boroughs. For example, an edge of value 10,000 from Camden to Southwark means that there are 10,000 individuals living in Camden who go to work in Southwark. We ignore all self-pointing edges, meaning that we drop the numbers of people working in the same boroughs where they live, since the local effects are already meant to be captured by the autoregressive component μ_t^{AR} . In other words, the adjacency matrix of this graph has a diagonal of constant zeros.

This graph is visualised in Figure 3. The size of a node in this figure is proportional to its in-degree (total number of people coming to work in this borough). For clarity, an edge is drawn only if its value exceeds the 80% percentile of all edge values. Widths of these plotted edges are also proportional to their values. A clear pattern of Figure 3 is that Westminster/City of London attracts a disproportional amount of London workforce with a total in-degree that is significantly larger than any other London borough. In addition, Camden and Tower Hamlets also attract a relatively large amount of workforce.

We denote by $K = \{k_{ij}\}_{1 \le i,j \le n}$ the adjacency matrix of this graph, where *i* indexes home and *j* indexes work. We have constructed other networks based on this *K* matrix. Specifically, we consider three networks, $W^{(1)} = K$, $W^{(2)} = K^{\top}$, and $W^{(3)} = KK^{\top}$, defined respectively through adjacency matrices.¹³ These three adjacency matrices capture transmission from different contact networks: 1) "work-to-home" transmission, which measures the spreading of disease from residents in commuter's work borough to residents in commuter's home borough; 2) "home-to-work" transmission, which is from the opposite direction; and 3) "home-to-home" transmission, which is between different places of residence via a common workplace.

Generally speaking, left-multiplying the vector of active cases x_t by the matrix $W^{(1)}$ as in

¹²The 2011 census covers detailed employment information which includes office locations. Thus, for any local authority, the number of its people working in other local authorities can be calculated. Aggregate data regarding these statistics are available from https://data.london.gov.uk/dataset/place-residence-place-work-local-authority.

¹³Notations here are to provide basic ideas. In detail, when estimating the network SIR models, we always divide K and K^{\top} by the largest singular value of K to rescale their spectral norms to one. This operation improves numerical stability (of nonlinear maximum likelihood estimation) and ensures comparability across network parameters. For computing the adjacency matrix of the third network, namely $W^{(3)}$, we first normalise each row of K to sum one, then calculate the product of it with its transpose. Diagonal elements of this product are also replaced with zero. The resulting matrix is then divided by its largest singular value (which is also its largest eigenvalue because this matrix is symmetric).

equation (13) features the transmission of COVID-19 from workplaces to residential areas. To be clear, for borough *i*, the vector $\begin{bmatrix} w_{ij}^{(1)} \end{bmatrix}_{j \neq i}$ will *overweight* boroughs where more of borough *i*'s residents go to work. For example, for many London boroughs, their residents are more likely to go to work at Westminster/City of London. Thus, the propensity for contact and infection *from* Westminster/City of London is set to be higher by this network. Intuitively, larger nodes of the graph shown in Figure 3 are more emphasised in the transmission dynamics.

Similarly, left multiplying x_t by $W^{(2)}$ characterises the disease spread from residential areas to workplaces. Panel (A) of Figure 4 shows the graph defined by $W^{(2)}$. Edges are again thresholded at the 80% for presentation. Node sizes are also proportional to their total indegrees. As we can see from Panel (A) of Figure 4, network effects commanded by this graph highlight the role of boroughs such as Wandsworth, Lambeth, Lewisham, and Southwark. These boroughs are those with large numbers of residents going to work in other boroughs.

Interpreting the impact of our third network on the disease spread dynamics – the hometo-home network – is also straightforward. The adjacency matrix of this network is symmetric with element (i, j) and (j, i) defined by $\sum_{\ell=1}^{n} k_{i\ell}k_{\ell j}$. It quantifies the propensity of residents of borough *i* and *j* to contact and infect each other with the disease via common workplaces. Specifically, if work and home locations were independent, the entries of this particular adjacency matrix would be proportional to the probability of individuals from the two locations meeting each other. Panel (B) of Figure (4) visualises the graph of this network. Nodes with large in-degrees such as Westminster/City of London, Camden, and Tower Hamlets should be influential in driving the disease spread dynamics mediated through this network. These boroughs are also featured in the work-to-home network graphed in Figure 3. In addition, Wandsworth, Lambeth, and Southwark, which emerge as pivotal nodes in the home-to-work graph, are also highlighted here, meaning that boroughs with large numbers of residents working outside those boroughs also have strong network effects. An exception to this appears to be Lewisham, which does not appear to be influential in the graph.

3 Empirical Findings

The main estimation results are presented in Table 1. Five specifications are included: three specifications with one of the networks ($\mathcal{G} = (1), (2), (3)$) each; one specification containing all three networks ($\mathcal{G} = (1,2,3)$); and one specification incorporating the first two networks ($\mathcal{G} = (1,2)$). The first three specifications serve the purpose of univariate analysis. Parameter estimates and standard errors are reported in the table, as well as the pseudo- R^2 defined as

pseudo-
$$R^2 = 1 - \frac{\sum_{i=1}^n \sum_{t=1}^T (y_{it} - \hat{\mu}_{it})^2}{\sum_{i=1}^n \sum_{t=1}^T (y_{it} - \overline{y})^2},$$

where $\hat{\mu}_{it}$ is the conditional expectation of daily new COVID-19 cases calculated from the model using estimated parameters; \bar{y} stands for simple average across time and boroughs.

Before-policy estimations for the autoregressive effects, that is, estimates of $\exp(\gamma_0)$, are always significantly greater than zero. Across the five specifications, the smallest estimate is 0.252 with a standard error of 0.041. As this parameter quantifies within-borough disease transmission dynamics, this finding confirms that serious community-acquired infection has occurred during the initial outbreak of COVID-19 in London in early 2020. The nationwide lockdown policy on March 23 has reduced the infection significantly: except for the second specification (which is 5%) the magnitude ranges from 24% to 38%. The reduction is economically and statistically significant as $\exp(\gamma)$ is significantly smaller than one (again with model (2) as an exception). The analysis shows that the lockdown policy is effective in cutting down within-community transmission rates.

Estimation results from the first specification (column 1 of Table 1) show that the impact of the work-to-home network is large in magnitude: the before-policy coefficient $\exp\left(\phi_0^{(1)}\right)$ is around 0.5, which implies strong directional spillover effects from people living in major workplaces such as Westminster/City of London to residents in other boroughs via the workto-home contact network. We also find that the work-to-home disease spread is partially reduced after the lockdown. The magnitude of reduction is around 12% for the first specification and around 18% after controlling for other networks, while being not statistically different from one. The effectiveness of the lockdown policy on blocking cross-borough transmission is questionable at this level. It seems that transmitting from pivotal places such as Westminster/City of London, Camden, and Tower Hamlets to other boroughs still remains active even after the nationwide lockdown.

We find that the impact of the home-to-work network on the epidemics of COVID-19 in London is relatively smaller (column 2 of Table 1). The baseline estimation $\exp\left(\phi_0^{(2)}\right)$ is less than 0.12.¹⁴ This parameter is around 0.06 after controlling for other contact networks. Though smaller in magnitude, it is significantly greater than zero except when including the home-to-home network. We find that the lockdown policy has a stronger impact on this home-to-work spreading. After the lockdown policy is enforced, the spreading through home-to-work network is reduced by over 80% after controlling for other network effects.

We find a strong home-to-home network effect when estimating it alone (column 3 of Table 1), but it is insignificant after controlling for the work-to-home and home-to-work networks (column 4 of Table 1). Thus, we choose to use the last specification including both work-to-home and home-to-work networks as the benchmark specification (column 5 of Table 1). The maximised log-likelihood function for the fourth and the fifth specifications are -7335.229 and -7335.153, respectively. A likelihood ratio test simply fails to offer any definitive evidence on

¹⁴Our parameter estimates for different networks are comparable because all graph adjacency matrices have been rescaled by their largest singular value before feeding into our models.

including the home-to-home network. Therefore, we report further analyses of the network effects based on this benchmark specification for expositional clarity in the rest of the paper.

Table 1 also reports results for testing-related variables in vector z_t that drives the endemic effects. Regression coefficients for the number of tests (i.e. lag test) are consistently positive across specifications, being around 1.8 with standard errors around 0.3. That is, we find that greater testing capacity leads to larger endemic effects and predicts a larger number of new cases. There are two potential explanations for this result. Firstly, testing efforts are motivated by the past disease propagation trajectory, which also determines the number of new cases. The association between testing numbers and new case counts is thus due to the fact that they both relate to the current epidemic severity. The second explanation points toward under testing. If there is a serious under testing problem, increased testing capacity will artificially pick up cases that have been omitted early on, thus increasing the expected number of cases. This explanation is partly supported by the significantly negative coefficients for positive-to-test ratios in Table 1. Across models, this regression coefficient is at least -1.09with a standard error being at most 0.49. Higher positive rates predict weaker future endemic effects. Intuitively, high positive rates are more likely to be a result of serious under testing (assuming testing kit technologies remain unchanged), which artificially discounts the expected number of cases. This view has been discussed in the media, for example,

"The thinking is that higher positive test rates equate to more missed cases" (Bloomberg Opinion, July 1, 2020)

Our findings that *both* larger testing capacity and smaller positive rates are associated with more future cases are suggestive of under testing, which may cause reported cases deviating from the actual number. Adding these two test-related variables into endemic effects serves as an adjustment to this issue.

3.1 Decomposition of expected daily new cases

We now plug the estimated parameters into expected new case numbers μ_t as defined in (11) to evaluate and compare the contribution of autoregressive, network, and endemic components in explaining the observed data. We rely on parameter estimates from the benchmark specification incorporating the work-to-home and home-to-work networks (column 5 in Table 1) to perform this decomposition in this section.

Panel (A) of Figure 5 graphs the decomposition of the contribution of the autoregressive, the network, and the endemic components to the total cases aggregating across the 32 London boroughs in the sample period. That is, we plot the time-series of the three ratios $(\mathbf{1}^{\top} \hat{\mu}_t^{AR}) / (\mathbf{1}^{\top} \hat{\mu}_t)$, $(\mathbf{1}^{\top} \hat{\mu}_t^{NE}) / (\mathbf{1}^{\top} \hat{\mu}_t)$, and $(\mathbf{1}^{\top} \hat{\mu}_t^{EN}) / (\mathbf{1}^{\top} \hat{\mu}_t)$ as defined in Section 1.3. The graph demonstrates a substantial and persistent network effect, accounting for 42% of the

expected total daily new cases in London. The autoregressive and endemic effects contribute to 35% and 23%, respectively.

In Panels (B) and (C) of the same figure, we present the decomposition result for each London borough. We plot the time series of the fractions of conditional means explained by each of the three components, $\hat{\mu}_{it}^{AR}/\hat{\mu}_{it}$, $\hat{\mu}_{it}^{NE}/\hat{\mu}_{it}$, $\hat{\mu}_{it}^{EN}/\hat{\mu}_{it}$, for each borough. Light-coloured bands in these plots show 10% to 90% percentiles across all boroughs, delineating cross-sectional heterogeneity. Solid lines represent the median levels.

The left graph in Panel (B) shows community transmission (autoregressive effects) rises quickly rand remains a strong driver of the disease spread since March. The middle graph in Panel (B) shows a similar pattern for network effects but with an additional peak in mid-March. The right graph in Panel (B) shows that at the very early stage, endemic effect is a dominating component, which we interpret as infections being brought in by people coming from outside of London. Endemic effects keep dropping in early March and remain stable for a while. They are diminished to a low level in late May. Comparing all three graphs in Panel (b), we find that the fading endemic impact is largely subsumed by the network effects.

In Panel (C) of Figure 5 we further decompose the network effects by examining the contributions of the work-to-home and the home-to-work network to the total network effect across time in our sample period, respectively. Based on plots in this panel, we observe that almost all network effects in our sample period can be attributed to the work-to-home network, highlighting the importance of boroughs where many people go to work in transmitting the disease. This finding results from parameter estimates in Table 1, where the network coefficient for the second specification $\exp\left(\phi_0^{(2)}\right)$ is almost ten times smaller than its counterpart for the first specification $\exp\left(\phi_0^{(1)}\right)$. Although weaker in terms of magnitude, we also observe that the home-to-work network effects have undergone much stronger reduction after the lockdown policy (i.e. $\exp\left(\phi^{(2)}\right) < \exp\left(\phi^{(1)}\right)$).

In the graphs in all three panels, we highlight March 23, the lockdown date. Panels (B) and (C) show that the reductions in autoregressive and (work-to-home) network effect have occurred at least one week before March 23, indicating that a significant number of London residents have started observing social distancing within their communities and working from home before the official policy announcement date.

3.2 Disease R_0

Our model offers guidance on the basic reproduction number R_0 , which quantifies the expected number of new cases *directly* generated by one existing case. To be more specific, since our model features borough-level heterogeneity, the expected number of new cases varies with regard to the residence of the existing infected cases. Thus, if the "one" existing case comes from borough *i*, we have a basic reproduction number $R_0^{(i)}$. What we aim to compute here is an estimated upper bound on the maximum $\max_{i \in \{1,...,n\}} R_0^{(i)}$.

To proceed, we first calculate a simple plugging-in estimator of matrix A in equation (11), denoted by \hat{A} . For one additional case in borough i, denoted by vector e_i in which only the i-th element equals one and all other elements equal zero. This case remains contagious for L periods by assumption. The total (expected) number of new cases created directly by e_i can be estimated as $\sum_{\ell=0}^{L-1} \exp(-\hat{\rho}\ell) \hat{A}e_i$. This quantity above is uniformly bounded by $\sum_{\ell=0}^{L} \exp(-\hat{\rho}\ell) \sigma_{\max}(\hat{A})$ because $||e_i|| = 1$, in which $\sigma_{\max}(\cdot)$ represents the function that computes the largest singular value of a matrix. Thus, we have an upper bound for the estimates of R_0 as follows:

$$\widehat{R}_{0} \leq \max_{i \in \{1,\dots,n\}} \left\{ \sum_{\ell=0}^{L-1} \exp\left(-\widehat{\rho}\ell\right) \widehat{A} \boldsymbol{e}_{i} \right\} \leq \sum_{\ell=0}^{L} \exp\left(-\widehat{\rho}\ell\right) \sigma_{\max}\left(\widehat{A}\right).$$

Estimates of (upper bounds of) R_0 are presented in Table 2. According to our specification in Section 1.3, we have separate estimates of \widehat{A} before and after the lockdown policy. Thus, the table shows the largest singular values and R_0 s both before and after the policy date. The disease R_0 is around 1.4 before the lockdown policy and is around 0.8 afterwards based on our benchmark specification of incorporating work-to-home and home-to-work networks (last two columns of Table 2). The magnitude of reduction is large though not statistically significant. This reduction is due to the impacts of lockdown policies on both the autoregressive and network effects.

3.3 The network impulse response functions

To further understand how innovations in daily new COVID-19 cases propagate through networks, we define and calculate the network impulse response function (NIRF) of our model motivated by the analysis in Denbee, Julliard, Li, and Yuan (2020). For a unitary shock (or change in levels) of disease incidents in borough *i*, its impact on the expected total number of cases across *all* locations τ -period ahead is measured by

$$\operatorname{NIRF}_{i}(\tau) = \sum_{j=1}^{n} \frac{\partial \mathbb{E}\left[y_{j,t+\tau} \mid \mathcal{F}_{t}\right]}{\partial y_{it}}.$$
(18)

The empirical model we work with allows an analytical formula for the NIRF, as detailed in the Appendix.

Plots in Figure 6 illustrate NIRFs across each borough for the time horizon of one week, that is $\tau = 7$. Panel (A) shows the impulse responses before the lockdown. The Westminster/City of London subpopulation strongly dominates all other London boroughs. For one additional case that emerges in this area, three more cases are expected to occur in the whole

Greater London area, even after one week. This identifies the Westminster/City of London area as a "key player" for shock propagations through the network in the language of Denbee, Julliard, Li, and Yuan (2020). Camden, Tower Hamlets, Southwark, as well as Lambeth are among the other key players that appear to show strong network impulse responses, but the magnitude is much weaker than for Westminster/City of London. Panel (B) presents results after the lockdown. It offers another angle for us to understand the effectiveness of the lockdown policy, as there is a distinct reduction in the NIRF measure for the key areas such as Westminster/City of London and other boroughs.

Analysing network impulse responses is valuable for designing "smart" partial lockdown policies that selectively lockdown a few regions instead of deploying a full-scale lockdown. When prescribing a partial lockdown plan, the conventional wisdom is to shutdown areas that have witnessed the largest number of existing cases and are undergoing rapid growth in new cases.¹⁵ Our network impulse response analysis offers another perspective. In addition to focusing on regions that have reported severe outbreaks, lockdown policies should also target areas that are key to the disease transmission. Isolating subpopulations that are key players in the network can forestall rapid spread among the whole population, even if few cases have occurred in these areas. Optimal (partial) lockdown policies should combine both perspectives, as we demonstrate in our counterfactual analysis in the next session.

4 Counterfactual Simulations

This section presents simulation studies that evaluate counterfactual outcomes due to alternative policy interventions. We start by investigating the impacts of earlier or later nationwide lockdown measures, and compare them with actual numbers. We then compute optimal partial lockdown arrangements and their potential outcomes. We emphasize on comparing optimal policies with "naive" ones that only target areas with the largest number of existing cases. Our goal of making these comparisons is to illustrate the importance of shutting down pivotal nodes of the network in preventing disease transmission.

4.1 Alternative dates for nation-wide lockdown

The UK government implemented nation-wide lockdown on March 23. Timing of this policy has been under intensive public inspection. Dr. Neil Ferguson who, with his group's "Imperial College" model, has facilitated the lockdown decision of the UK government said:

"Had we introduced lockdown measures a week earlier, we would have reduced the final death toll by at least a half." (BBC News, June 10, 2020)

¹⁵See, for example, the lockdown of Hubei province in China and Lombardy region in Italy, both of which were the epicentres of COVID-19 outbreaks when the policies came out.

Similar arguments have been made by scientists such as Dr. James Annan and Dr. Kit Yates¹⁶. In the meantime, Dr. Yates have acknowledged that

"There had been an 'overreliance' on certain models when determining how fast the epidemic had been doubling; ...that some of the modeling groups had more influence over the consensus decisions than others." (BBC News, June 10, 2020)

We attempt the same inspection using our model, although within a limited scope by only focusing on the case of London. With our model estimates, we can change the policy indicator D_t by allowing for different policy intervention dates and simulate the model outcomes. We do exactly this exercise. Specifically, we pick the model (1,2) and plug in its parameter estimates as reported in Table 1. We consider policy dates as two weeks or one week both before and after March 23. We simulate each of the resulting models 10,000 times and then average the 10,000 paths of daily case counts across all London boroughs as our counterfactual outcomes. Throughout our simulations, we *fix* the endemic terms, although this element of the model, by definition, is also affected by the specific date of the policy due to variables such as $t \times D_t$ and $t^2 \times D_t$. We are also ignoring the impacts of alternative policies on the nationwide positiveto-test ratio¹⁷. Under this simulation design, we are indeed treating the endemic term as a deterministic force in the model. Incorporating policy-related terms into the endemic element is to statistically recover its shape and variations (across time), but not to find out how it evolves according to the state of the pandemic. The endemic effects exist only to isolate the autoregressive and network effects through controlling for variations that are not epidemic. To the extreme, we can treat the endemic effects as cross product

Simulation results are presented in Figure 7. Plots in this figure show cumulative number of all London cases from March 1 to June 4. The plot on the left panel compares the simulated outcomes from locking-down one or two weeks earlier with the actual one. Based on our results, locking down two weeks earlier translates into a reduction of 12.1% total cases (a reduction of 3,283 cases on the absolute level) in London during the period under study. The number is 9.4% (2,553 on the absolute level) if lockdown one week earlier. The plot on the right panel illustrates alternative situations in which lockdown took place one or two weeks later.

¹⁶These proponents have been broadly covered by the media. James Annan's conclusion was drawn upon his calculation made public on May 12, 2020 through a blog post, which is available at https://bskiesresearch. wordpress.com/2020/05/12/the-human-cost-of-delaying-lockdown/. Citing James Annan's calculation, Kit Yates wrote "locking-down a week earlier translates to beginning lockdown with roughly a quarter of the total cases..." in an essay to HuffPost on May 22, 2020. Details can be found at https://www.huffingtonpost. co.uk/entry/lockdown-uk-deaths_uk_5ec6efd8c5b68038a74a50ad?utm_hp_ref=uk-opinion.

¹⁷In practice, each simulation creates new time-series of London cases, which should count as part of the total positive cases in Britain, thus changing the numerator of this ratio.

4.2 Optimal partial lockdown policies

5 Conclusions

In this paper, we present and estimate a network-SIR model of the spreading of COVID-19 disease in London. Our estimates show that network play a major role in transmitting COVID-19 disease and they cannot be ignored. Based on the estimated epidermic dynamics, we investigate whether a certain target lockdown policy could contain the spread of COVID-19 disease as much as the full scale lockdown and, hence, have a lower economic cost. Our simulations show that an optimal lockdown should target areas that not only have the highest number of existing case, but also those that play a key role in transmitting disease in the contact network among the population. In our case, the contact network corresponds to the commuting network in London. In designing a lockdown policy, our finding calls for special attention to be focused on the network role of the COVID-19 transmission. These network could be train or flight networks, or any other traffic networks at the national or international level including the migration network (from hard-hit COVID-19 hotspots) identified by Coven, Gupta, and Yao (2020). As networks potentially connect regions with different jurisdictional governments, our finding indicates that coordinated regional quarantine and lockdown policies are essential in containing the spread of the COVID-19 pandemic, a conclusion echoed by the theoretical work in Chandrasekhar, Goldsmith-Pinkham, Jackson, and Thau (2020).

References

- Acemoglu, D., V. Chernozhukov, I. Werning, and M. D. Whinston (2020). A multi-risk sir model with optimally targeted lockdown. Technical report, National Bureau of Economic Research.
- Berndt, E. R., B. H. Hall, R. E. Hall, and J. A. Hausman (1974). Estimation and inference in nonlinear structural models. *Annals of Economic and Social Measurement* 3(4), 653–665.
- Chandrasekhar, A. G., P. Goldsmith-Pinkham, M. O. Jackson, and S. Thau (2020). Interacting regional policies in containing a disease. Technical report, Stanford University.
- Coven, J., A. Gupta, and I. Yao (2020). Urban flight seeded the covid-19 pandemic across the united states. Technical report, NYU Stern School of Business.
- Denbee, E., C. Julliard, Y. Li, and K. Yuan (2020). Network risk and key players: A structural analysis of interbank liquidity. *Journal of Financial Economics forthcoming*.
- Finkenstädt, B. F. and B. T. Grenfell (2000). Time series modelling of childhood diseases: a dynamical systems approach. Journal of the Royal Statistical Society: Series C 49(2), 187–205.
- Grimmett, G. R. and D. R. Stirzaker (2001). *Probability and Random Processes (Third Edition)*. Oxford University Press.
- Held, L., M. Höhle, and M. Hofmann (2005). A statistical framework for the analysis of multivariate infectious disease surveillance counts. *Statistical Modelling* 5(3), 187–199.
- Kermack, W. O. and A. G. McKendrick (1927). A contribution to the mathematical theory of epidemics. *Proceedings* of the Royal Society of London. Series A 115(772), 700–721.
- Lawson, A. B. (2013). Bayesian Disease Mapping: Hierarchical Modeling in Spatial Epidemiology. CRC press.

Li, S. and O. Linton (2020). When will the covid-19 pandemic peak? Journal of Econometrics.

Newman, M. E. (2002). Spread of epidemic disease on networks. *Physical review E* 66(1), 016128.

Table 1: Estimation results of the Network-SIR model of Section 1.3. Results in columns (1)-(3) correspond, respectively, to specifications with only one of the following transmission networks: "work-to-home", "home-to-work", and "home-to-home". Column (1,2,3) include all the three networks while column (1,2) uses only the first two networks.

Model \mathcal{G}	(1)	(2	2)	(;	3)	(1,2	2,3)	(1	,2)		
Value	est.	se.	est.	se.	est.	se.	est.	se.	est.	se.		
			Au	utoregro	essive e	ffect						
$\exp \gamma_0)$	0.381	0.033	0.456	0.034	0.424	0.032	0.365	0.036	0.344	0.038		
$\exp(\gamma)$	0.472	0.050	0.519	0.045	0.446	0.042	0.492	0.057	0.522	0.066		
Network effect												
$W^{(1)} = K$:												
$\exp\left(\phi_{0}^{(1)} ight)$	0.563	0.066					0.496	0.091	0.592	0.071		
$\exp\left(\phi^{(1)} ight)$	0.694	0.098					0.594	0.189	0.637	0.102		
$W^{(2)} = K^\top$												
$\exp\left(\phi_{0}^{(2)} ight)$			0.076	0.021			0.000	0.000	0.041	0.021		
$\exp\left(\phi^{(2)} ight)$			0.598	0.175			1.000	0.000	0.203	0.424		
$W^{(3)} = KK$	г :											
$\exp\left(\phi_{0}^{(3)} ight)$					0.280	0.035	0.056	0.048				
$\exp\left(\phi^{(3)}\right)$					0.518	0.073	0.696	0.779				
pseudo-R ²	81.9%		80.9%		81.6%		81.9%		81.	9%		
# obs.	3008		30	08	30	08	3008		3008			

Table 2: Estimation-implied upper bound for the basic reproduction number R_0 of COVID-19 in London. σ_{max} denotes the largest singular value of the estimated matrix A in equation (14).

$\operatorname{Model} \mathcal{G}$	(1)		(2)		(;	(3)		2,3)	(1,2)			
Value	est.	se.	est.	se.	est.	se.	est.	se.	est.	se.		
		Before Lockdown Policy: $D_t = 0$										
$\sigma_{\rm max}$	0.840	0.590	0.511	0.106	0.639	0.287	0.814	0.557	0.853	0.592		
R_0	1.501	1.054	0.914	0.189	1.142	0.514	1.455	0.996	1.525	1.059		
			After	Lockdo	wn Poli	$\mathbf{cy}: D_t =$	= 1					
$\sigma_{\rm max}$	0.510	0.369	0.270	0.038	0.301	0.123	0.455	0.330	0.501	0.356		
R_0	0.911	0.659	0.482	0.068	0.539	0.221	0.814	0.590	0.896	0.637		



Figure 1: UK nationwide COVID-19 tests: people tested and tests performed



(A) COVID-19 cases for boroughs in London (up till June 4, 2020)



(B) Population sizes for boroughs in London (year 2020)

Figure 2: Summary statistics: cumulative number of cases (panel A) and subpopulation sizes (panel (B)).



Figure 3: The graph of network $W^{(1)} = K$. This network is constructed using 2011 UK census data. Nodes represent London boroughs. Node sizes are proportional to their total in-degrees. Directed edges represent the amount of people going to work from one borough to another. Edge widths are proportional to their respective values in the graph. Only edges which exceed the 80% percentile are drawn.



(A) Graph of $W^{(2)}$: defined by adjacency matrix K^{\top}



(B) Graph of $W^{(3)}$: defined by adjacency matrix KK^{\top}

Figure 4: Graphs of network $W^{(2)}$ and $W^{(2)}$. Both networks are constructed based on network $W^{(1)} = K$ shown in Figure 3. Nodes represent London boroughs. Node sizes are proportional to their total in-degrees. Directed edges represent the amount of people going to work from one borough to another. Edge widths are proportional to their respective values in the graph. Only edges which exceed the 80% percentile are drawn.



(A) Decomposition of total (expected) daily new cases in London across time



(B) The autoregressive, (total) network, and endemic effects as fractions: μ_{it}^{AR}/μ_{it} , μ_{it}^{NE}/μ_{it} , μ_{it}^{EN}/μ_{it}



(C) The two network effects as fractions: $\mu_{it}^{NE,(g)}/\mu_{it},\,g\in\{1,2\}$

Figure 5: Decomposition of the conditional mean: total daily new cases in London. Effects in Panel (A) are based on summing across all London boroughs. Bands in Panel (B) are 10% to 90% percentiles across all London boroughs. Solid lines are the median level.



(B) After lockdown

Figure 6: Network impulse response functions (one-week horizon), from equation (18), before (panel (A)), and after (panel (B)), lockdown.



Figure 7: Cumulated number of cases over time (continuous black line) and counterfactual cumulated number of cases with earlier (left figure) and later (right figure) lockdown time.



Figure 8: Cumulated number of cases over time: actual (blue line); counterfactual without lockdown (red line); counterfactual with lockdown of only one borough (all cases, grey); counterfactual with lockdown of only the borough with the most cases (orange line); counterfactual with optimally chosen borough (purple line). In the left figure, the counterfactual lockdown effects are computed by severing all linkages to and from those assumed to be in lockdown, while in the right figure, instead the corresponding columns and rows of matrix W(1) and W(2) are downscaled, respectively, to 82.7% and 16.6% of their original values.



Figure 9: Cumulated number of cases over time: actual (blue line); counterfactual without lockdown (red line); counterfactual with lockdown of only two boroughs (all cases, grey); counterfactual with lockdown of only the two boroughs with most cases (orange line); counterfactual with optimally chosen two boroughs (purple line). In the left figure, the counterfactual lockdown effects are computed by severing all linkages to and from those assumed to be in lockdown, while in the right figure, instead the corresponding columns and rows of matrix W(1) and W(2) are downscaled, respectively, to 82.7% and 16.6% of their original values.



Figure 10: Cumulated number of cases over time: actual (blue line); counterfactual without lockdown (red line); counterfactual with lockdown of only two boroughs (all cases, grey); counterfactual with lockdown of only the three boroughs with most cases (orange line); counterfactual with optimally chosen three boroughs (purple line). In the left figure, the counterfactual lockdown effects are computed by severing all linkages to and from those assumed to be in lockdown, while in the right figure, instead the corresponding columns and rows of matrix W(1) and W(2) are downscaled, respectively, to 82.7% and 16.6% of their original values.

Appendix for "COVID-19 Spread in London: Network Effects and Optimal Lockdowns?"

A MLE details

Under a model \mathcal{G} , the log-likelihood function given the observed panel of daily new case numbers Y for parameters $\Theta(\mathcal{G}) = \left\{\gamma_0, \gamma, \left\{\phi_0^{(g)}\right\}_{g \in \mathcal{G}}, \left\{\phi^{(g)}\right\}_{g \in \mathcal{G}}, \beta, \left\{\eta_i\right\}_{i=1}^n\right\}$ can be written as

$$\ell(\Theta \mid \mathbf{Y}, \mathcal{G}) = \sum_{i=1}^{n} \sum_{t=1}^{T} \left\{ \log \Gamma \left(y_{it} + x_{i,t-1} \right) - \log \Gamma \left(x_{i,t-1} \right) - \log \Gamma (y_{it} + 1) \right. \\ \left. - \left(x_{i,t-1} + y_{it} \right) \log \left(1 + \mu_{it} x_{i,t-1}^{-1} \right) + y_{it} \left(\log \mu_{it} + x_{i,t-1}^{-1} \right) \right\}$$

where $\Gamma(\cdot)$ is the standard gamma function; μ_{it} is the conditional mean of daily new cases defined as

$$\mu_{it} = \underbrace{\exp\left(\gamma_0 + D_t\gamma\right)x_{i,t-1}}_{\mu_{it}^{AR}} + \sum_{g \in \mathcal{G}} \underbrace{\exp\left(\phi_0^{(g)} + D_t\phi^{(g)}\right)\left(\sum_{j \neq i} w_{ij}^{(g)}x_{j,t-1}\right)}_{\mu_{it}^{NE,(g)}} + \underbrace{\exp\left(z_t^\top \beta + \eta_i\right)N_i}_{\mu_{it}^{EN}}$$

Taking derivative, we have the following set of score functions:

$$\mathbf{S}(\Theta) = \left\{ S_{\gamma_0}, S_{\gamma}, \left\{ S_{\phi_0^{(g)}} \right\}_{g \in \mathcal{G}}, \left\{ S_{\phi^{(g)}} \right\}_{g \in \mathcal{G}}, \mathbf{S}_{\boldsymbol{\beta}}, \left\{ S_{\eta_i} \right\}_{i=1}^n \right\} (\Theta)$$

defined as

$$S_{\gamma_{0}}(\Theta) = \frac{\partial \ell}{\partial \gamma_{0}} = \sum_{i=1}^{n} \sum_{t=1}^{T} \frac{y_{it} - \mu_{it}}{1 + x_{i,t-1}^{-1} \mu_{it}} \frac{\mu_{it}^{AR}}{\mu_{it}}, \quad S_{\gamma}(\Theta) = \frac{\partial \ell}{\partial \gamma} = \sum_{i=1}^{n} \sum_{t=1}^{T} \frac{y_{it} - \mu_{it}}{1 + x_{i,t-1}^{-1} \mu_{it}} \frac{\mu_{it}^{AR}}{\mu_{it}} D_{t},$$

$$S_{\phi_{0}^{(g)}}(\Theta) = \frac{\partial \ell}{\partial \gamma} = \sum_{i=1}^{n} \sum_{t=1}^{T} \frac{y_{it} - \mu_{it}}{1 + x_{i,t-1}^{-1} \mu_{it}} \frac{\mu_{it}^{NE,(g)}}{\mu_{it}}, \quad S_{\phi^{(g)}}(\Theta) = \frac{\partial \ell}{\partial \gamma} = \sum_{i=1}^{n} \sum_{t=1}^{T} \frac{y_{it} - \mu_{it}}{1 + x_{i,t-1}^{-1} \mu_{it}} \frac{\mu_{it}^{NE,(g)}}{\mu_{it}} D_{t}$$

for any $g \in \mathcal{G}$, and

$$S_{\boldsymbol{\beta}}(\Theta) = \frac{\partial \ell}{\partial \boldsymbol{\beta}} = \sum_{i=1}^{n} \sum_{t=1}^{T} \frac{y_{it} - \mu_{it}}{1 + x_{i,t-1}^{-1} \mu_{it}} \frac{\mu_{it}^{EN}}{\mu_{it}} \boldsymbol{z}_{t}$$
$$S_{\eta_i}(\Theta) = \frac{\partial \ell}{\partial \eta_i} = \sum_{t=1}^{T} \frac{y_{it} - \mu_{it}}{1 + x_{i,t-1}^{-1} \mu_{it}} \frac{\mu_{it}^{EN}}{\mu_{it}}$$

for each $i = 1, \ldots, n$.

Our MLE estimator $\widehat{\Theta}$ solves the system of score equations, that is $S(\widehat{\Theta}) = 0$. Standard errors are then computed as $\sqrt{\operatorname{diag}\left(\left\{S(\widehat{\Theta})S(\widehat{\Theta})^{\top}\right\}^{-1}\right)}$, based on the outer product of score vectors à la Berndt, Hall, Hall, and Hausman (1974).

B Network impulse response function

Æ

To compute the NIRFs, we begin from the following fully vectorized representation. Define an $(n \times L)$ -dimension vector $\tilde{y}_t^{\top} = \left[y_t^{\top}, y_{t-1}^{\top}, \dots, y_{t-L+1}^{\top} \right]^{\top}$ which concatenates the current and lagged-(L-1) observations. Then

$$\begin{split} [\widetilde{y}_{t+1} \mid \mathcal{F}_t] &= \begin{bmatrix} Ax_t + \mu_t^{EN} \\ y_t \\ y_{t-1} \\ \vdots \\ y_{t-L+2} \end{bmatrix} \\ &= \begin{bmatrix} A\sum_{\ell=0}^{L-1} \nu(\ell) y_{t-\ell} \\ y_t \\ y_{t-1} \\ \vdots \\ y_{t-L+2} \end{bmatrix} + \begin{bmatrix} \mu_t^{EN} \\ 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix} \\ &= \begin{bmatrix} A\nu(0) \quad A\nu(1) \quad \cdots \quad A\nu(L-2) \quad A\nu(L-1) \\ I \quad 0 \quad \cdots \quad 0 \quad 0 \\ 0 \quad I \quad \cdots \quad 0 \quad 0 \\ \vdots \quad \vdots \quad \vdots \quad \vdots \quad \vdots \\ 0 \quad 0 \quad \cdots \quad I \quad 0 \end{bmatrix} \\ \widetilde{y}_t + \begin{bmatrix} \mu_t^{EN} \\ 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix} \\ \overbrace{A}(L) \end{split}$$

As a result, $\mathbb{E}\left[\widetilde{y}_{t+\tau} \mid \mathcal{F}_t\right] = \left[\widetilde{A}(L)\right]^{ au} \widetilde{y}_t + \text{const.}$, which implies that

$$\operatorname{NIRF}_{it}(\tau) = \sum_{j=1}^{n} \frac{\partial \mathbb{E}\left[y_{j,t+\tau} \mid \mathcal{F}_{t}\right]}{\partial y_{it}} = \sum_{j=1}^{n} \left[\widetilde{A}(L)\right]_{jt}^{\tau},$$

where the subscript (ji) of a matrix denotes its element on the *j*-th row and the *i*-th column. Estimations of the NIRFs are then calculated through plugging parameter estimates into the expression for matrix A.

C The network SIR model without assuming $S_t/N \approx 1$

Without assuming $S_t/N \approx 1$, the equation (C) and (C) describing the single-subject SIR model is

Table A1: Estimation results of the Network-SIR model of Section 1.3. Results in columns (1)-(3) correspond, respectively, to specifications with only one of the following transmission networks: "work-to-home", "home-to-work", and "home-to-home". Column (1,2,3) include all the three networks while column (1,2) uses only the first two networks. Test-related variables are include into the endemic terms.

Model \mathcal{G}	(1)	(2	2)	(3	3)	(1,2	2,3)	(1,	2)	
Value	est.	se.	est.	se.	est.	se.	est.	se.	est.	se.	
			A	utoregre	essive et	ffect					
$\exp\left(\gamma_{0} ight)$	0.311	0.035	0.286	0.040	0.286	0.037	0.252	0.041	0.253	0.041	
$\exp\left(\gamma ight)$	0.617	0.077	0.958	0.136	0.750	0.102	0.760	0.129	0.754	0.127	
	Network effect										
$W^{(1)} = K$:											
$\exp\left(\phi_{0}^{(1)} ight)$	0.477	0.070					0.503	0.144	0.510	0.073	
$\exp\left(\phi^{(1)} ight)$	0.885	0.136					0.821	0.169	0.827	0.127	
$W^{(2)} = K^\top$:											
$\exp\left(\phi_{0}^{(2)} ight)$			0.115	0.022			0.062	0.036	0.064	0.023	
$\exp\left(\phi^{(2)} ight)$			0.374	0.080			0.148	0.218	0.166	0.181	
$W^{(3)} = KK^{\top}$:										
$\exp\left(\phi_{0}^{(3)} ight)$					0.261	0.035	0.000	0.001			
$\exp\left(\phi^{(3)} ight)$					0.570	0.082	0.284	1.641			
			Testin	g-relate	d enden	nic effec	t				
pos-to-test	-1.37	0.47	-1.09	0.28	-1.46	0.41	-1.46	0.48	-1.46	0.49	
lag test	1.85	0.31	1.77	0.23	1.88	0.28	1.91	0.32	1.91	0.32	
pseudo-R ²	82.0)9%	82.00%		82.02%		82.12%		82.12%		
# obs.	30	08	30	08	30	08	3008		3008		

Table A2: Estimation results of the Network-SIR model of Section 1.3 adding day fixed effects. Results in columns (1)-(3) correspond, respectively, to specifications with only one of the following transmission networks: "work-to-home;" "home-to-work;" "home-to-home". Column (1,2,3) considers the three networks jointly while column (1,2) uses only the first two networks.

Model \mathcal{G}	(]	1)	(2	2)	(;	3)	(1,2	2,3)	(1	,2)	
Value	est.	se.	est.	se.	est.	se.	est.	se.	est.	se.	
			Aı	utoregro	essive e	ffect					
$\exp \gamma_0)$	0.309	0.025	0.302	0.018	0.290	0.026	0.267	0.027	0.267	0.027	
$\exp(\gamma)$	0.454	0.050	0.468	0.033	0.483	0.051	0.525	0.066	0.525	0.066	
Network effect											
$W^{(1)} = K$											
$\exp\left(\phi_{0}^{(1)} ight)$	0.291	0.042					0.285	0.044	0.285	0.044	
$\exp\left(\phi^{(1)} ight)$	0.309	0.108					0.303	0.110	0.303	0.110	
$W^{(2)} = K^{\top}$											
$\exp\left(\phi_{0}^{(2)} ight)$			0.067	0.020			0.057	0.021	0.057	0.021	
$\exp\left(\phi^{(2)} ight)$			0.000	0.000			0.000	0.000	0.000	0.000	
$W^{(3)} = KK$	Г										
$\exp\left(\phi_{0}^{(3)} ight)$					0.143	0.026	0.000	0.000			
$\exp\left(\phi^{(3)}\right)$					0.000	0.000	1.120	0.000			
pseudo-R ²	86.3%		86.3%		86.3%		86.3%		86.	3%	
# obs.	3008		30	08	30	08	3008		3008		

Table A3: Estimation results of the Network-SIR model of Section 1.3 adding week fixed effects. Results in columns (1)-(3) correspond, respectively, to specifications with only one of the following transmission networks: "work-to-home;" "home-to-work;" "home-to-home". Column (1,2,3) considers the three networks jointly while column (1,2) uses only the first two networks.

Model \mathcal{G}	(]	1)	(2	2)	(;	3)	(1,2	2,3)	(1	,2)		
Value	est.	se.	est.	se.	est.	se.	est.	se.	est.	se.		
			Au	utoregro	essive e	ffect						
$\exp(\gamma_0)$	0.302	0.033	0.323	0.020	0.274	0.032	0.259	0.036	0.260	0.035		
$\exp(\gamma)$	0.621	0.079	0.660	0.077	0.718	0.109	0.716	0.094	0.716	0.114		
Network effect												
$\boldsymbol{W}^{(1)} = \boldsymbol{K}$												
$\exp\left(\phi_{0}^{(1)} ight)$	0.351	0.062					0.362	0.065	0.355	0.064		
$\exp\left(\phi^{(1)} ight)$	0.810	0.174					0.666	0.219	0.797	0.175		
$W^{(2)} = K^{\top}$												
$\exp\left(\phi_{0}^{(2)} ight)$			0.082	0.035			0.070	0.026	0.064	0.024		
$\exp\left(\phi^{(2)} ight)$			0.102	0.213			0.000	0.000	0.000	0.000		
$W^{(3)} = KK$	Г											
$\exp\left(\phi_{0}^{(3)} ight)$					0.216	0.038	0.000	0.000				
$\exp\left(\phi^{(3)}\right)$					0.420	0.101	1.562	0.000				
$pseudo-R^2$	81.25%		80.12%		81.96%		81.42%		81.2	27%		
# obs.	3008		30	08	30	08	3008		3008			

Table A4: Sensitivity analysis of the Network-SIR model of Section 1.3. Results in columns (1)-(3) correspond, respectively, to specifications with only one of the following transmission networks: "work-to-home;" "home-to-work;" "home-to-home". Column (1,2,3) considers the three networks jointly while column (1,2) uses only the first two networks. The policy date is changed to March 16.

(1	.)	(2	2)	(3	3)	(1,2	2,3)	(1,	2)		
est.	se.	est.	se.	est.	se.	est.	se.	est.	se.		
		A	utoregro	essive ef	fect						
0.355	0.046	0.285	0.049	0.294	0.052	0.248	0.051	0.251	0.051		
0.567	0.078	0.995	0.175	0.767	0.139	0.812	0.172	0.805	0.168		
Network effect											
0.266	0.083					0.264	0.389	0.308	0.097		
1.523	0.475					1.528	2.255	1.307	0.413		
		0.197	0.042			0.117	0.095	0.130	0.043		
		0.233	0.055			0.000	0.000	0.005	0.084		
:											
				0.310	0.066	0.048	0.294				
				0.436	0.102	0.000	0.000				
		Testin	g-relate	d enden	nic effec	t					
-2.30	0.49	-1.50	0.33	-1.70	0.41	-2.19	0.48	-2.34	0.49		
2.04	0.31	2.07	0.23	2.02	0.27	2.04	0.30	2.05	0.31		
81.98% 3008		81.84% 3008		81.88% 3008		82.00% 3008		82.00% 3008			
	(1 est. 0.355 0.567 0.266 1.523 : : : -2.30 2.04 81.9 300	(1) est. se. 0.355 0.046 0.567 0.078 0.266 0.083 1.523 0.475 	(1) (2 est. se. est. 355 0.046 0.285 0.567 0.078 0.995 0.266 0.083 1.523 0.475 0.197 0.233 : Testing -2.30 0.49 -1.50 2.04 0.31 2.07 81.98% 81.8 3008 300	$\begin{array}{c c c c c c c } (1) & (2) \\ est. & se. & est. & se. \\ & & & & & & & & \\ & & & & & & & & \\ \hline & & & &$	(1) (2) (3) est. se. est. se. est. $Autoregressive$ eff 0.355 0.046 0.285 0.049 0.294 0.567 0.078 0.995 0.175 0.767 0.266 0.083 Network effect 0.266 0.083 0.266 0.083 0.266 0.083 0.266 0.083 0.266 0.083 0.197 0.042 0.197 0.042 0.310	$\begin{array}{c ccccccccccccccccccccccccccccccccccc$	$\begin{array}{c ccccccccccccccccccccccccccccccccccc$	$\begin{array}{c c c c c c c c c c c c c c c c c c c $	$\begin{array}{c ccccccccccccccccccccccccccccccccccc$		

Table A5: Sensitivity analysis of the Network-SIR model of Section 1.3. Results in columns (1)-(3) correspond, respectively, to specifications with only one of the following transmission networks: "work-to-home;" "home-to-work;" "home-to-home". Column (1,2,3) considers the three networks jointly while column (1,2) uses only the first two networks. The policy date is changed to March 20.

Model \mathcal{G}	(1	L)	(2	2)	(3	3)	(1,2	2,3)	(1,	2)	
Value	est.	se.	est.	se.	est.	se.	est.	se.	est.	se.	
			A	utoregr	essive et	ffect					
$\exp\left(\gamma_{0} ight)$	0.333	0.040	0.301	0.044	0.314	0.042	0.282	0.046	0.282	0.046	
$\exp\left(\gamma ight)$	0.596	0.078	0.946	0.140	0.703	0.098	0.709	0.122	0.708	0.120	
Network effect											
$\boldsymbol{W}^{(1)} = \boldsymbol{K}:$											
$\exp\left(\phi_{0}^{(1)} ight)$	0.362	0.078					0.377	0.230	0.384	0.085	
$\exp\left(\phi^{(1)} ight)$	1.134	0.247					1.073	0.664	1.059	0.234	
$W^{(2)} = K^{\top}:$											
$\exp\left(\phi_{0}^{(2)} ight)$			0.120	0.026			0.057	0.053	0.057	0.028	
$\exp\left(\phi^{(2)} ight)$			0.374	0.090			0.040	0.290	0.025	0.190	
$\boldsymbol{W}^{(3)} = \boldsymbol{K}\boldsymbol{K}^{\top}$:										
$\exp\left(\phi_{0}^{(3)} ight)$					0.245	0.044	0.004	0.143			
$\exp\left(\phi^{(3)} ight)$					0.621	0.116	0.104	9.481			
			Testin	g-relate	d enden	nic effec	t				
pos-to-test	-2.45	0.56	-1.79	0.35	-2.44	0.50	-2.58	0.57	-2.60	0.57	
lag test	2.06	0.32	2.05	0.24	2.09	0.29	2.08	0.32	2.08	0.32	
pseudo- R^2	81.98%		81.86%		81.87%		81.99%		81.99%		
# obs.	30	08	30	08	30	08	3008		3008		

Table A6: Sensitivity analysis of the Network-SIR model of Section 1.3. Results in columns (1)-(3) correspond, respectively, to specifications with only one of the following transmission networks: "work-to-home;" "home-to-work;" "home-to-home". Column (1,2,3) considers the three networks jointly while column (1,2) uses only the first two networks.

$\mathbf{Model}\ \mathcal{G}$	(]	1)	(2	2)	(;	3)	(1,2	2,3)	(1	,2)
Value	est.	se.	est.	se.	est.	se.	est.	se.	est.	se.
			Autore	gressive	e effect					
$\exp\left(\gamma_{0} ight)$	0.416	0.044	0.399	0.045	0.307	0.055	0.318	0.056	0.318	0.056
$\exp{(\gamma)}$: 3.16-3.20	0.793	0.215	1.073	0.120	1.315	0.299	1.039	0.318	1.003	0.320
$\exp\left(\gamma ight)$: 3.21-3.23	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
$\exp{(\gamma)}: 3.24-6.03$	0.414	0.050	0.577	0.071	0.668	0.123	0.542	0.101	0.543	0.101
			Netv	work ef	fect					
$\boldsymbol{W}^{(1)} = \boldsymbol{K}:$										
$\exp\left(\phi_{0}^{(1)} ight)$	0.477	0.079					0.542	0.089	0.542	0.089
$\exp(\phi^{(1)})$: 3.16-3.20	1.607	0.450					1.415	0.395	1.440	0.401
$\exp(\phi^{(1)})$: 3.21-3.23	2.599	0.440					1.988	0.394	2.180	0.374
$\exp(\phi^{(1)})$: 3.24-6.03	0.331	0.117					0.293	0.103	0.290	0.103
$W^{(2)} = K^\top$										
$\exp\left(\phi_{0}^{(2)} ight)$			0.160	0.043			0.115	0.039	0.115	0.039
$\exp\left(\phi^{(2)} ight)$: 3.16-3.20			0.297	0.256			0.000	0.000	0.083	0.280
$\exp\left(\phi^{(2)}\right)$: 3.21-3.23			0.996	0.352			0.000	0.000	0.466	0.568
$\exp\left(\phi^{(2)} ight)$: 3.24-6.03			0.291	0.083			0.000	0.000	0.000	0.000
$W^{(3)} = KK^{\top}$:										
$\exp\left(\phi_{0}^{(3)} ight)$					0.360	0.067	0.000	0.000		
$\exp\left(\phi^{(3)}\right)$: 3.16-3.20					0.569	0.199	0.000	0.000		
$\exp(\phi^{(3)})$: 3.21-3.23					1.028	0.208	0.000	0.000		
$\exp\left(\phi^{(3)}\right)$: 3.24-6.03					0.418	0.082	0.000	0.000		
pseudo-R ²	82.1	13%	81.74%		82.13%		82.14%		82.17%	
# obs.	30	08	30	08	30	08	30	08	30	08

Table A7: Estimation of the Network-SIR model in Section 1.3. Results in columns (1)-(3) correspond, respectively, to specifications with only one of the following transmission networks: (i) only diagonal proportional to the fraction of population, (ii) having equal off-diagonal elements (compound symmetric), (iii) off-diagonals are determined only by distances between centroids of boroughs. Column (1,2,3) considers the three networks jointly while column (1,2) uses only the first two networks. Δ AIC represents the difference between the Akaike information criteria of the estimated models and our preferred model (1,2) estimated using commuting networks.

$\operatorname{Model} \mathcal{G}$	(]	L)	(2	2)	(;	3)	(1,2	2,3)			
Value	est.	se.	est.	se.	est.	se.	est.	se.			
		Aı	utoregro	essive e	ffect						
$\exp\left(\gamma_{0} ight)$	0.684	0.014	0.343	0.033	0.349	0.033	0.332	0.086			
$\exp\left(\gamma ight)$	0.281	0.016	0.566	0.061	0.653	0.067	0.309	0.103			
Network effect											
$W^{(1)}$: diago	nal										
$\exp\left(\phi_{0}^{(1)} ight)$	0.000	0.000					0.093	2.948			
$\exp\left(\phi^{(1)} ight)$	0.000	0.000					0.000	0.000			
$W^{(2)}$: comp	ound sy	mmetr	ic								
$\exp\left(\phi_{0}^{(2)} ight)$			0.386	0.064			0.331	0.079			
$\exp\left(\phi^{(2)} ight)$			1.015	0.164			1.173	0.253			
$W^{(3)}$: squar	e expor	nential	kernel								
$\exp\left(\phi_{0}^{(3)} ight)$					0.348	0.070	0.000	0.000			
$\exp\left(\phi^{(3)}\right)$					0.994	0.190	0.820	0.029			
Δ AIC	-14	42.6	-10.9		-6	9.6	-9.1				
# obs.	30	08	30	08	30	08	3008				

Table A8: Estimation of the Network-SIR model in Section 1.3. Results in columns (1)-(3) correspond, respectively, to specifications with only one of the following transmission networks: (i) only diagonal proportional to the fraction of population, (ii) having equal off-diagonal elements (compound symmetric), (iii) off-diagonals are determined only by distances between centroids of boroughs. Column (1,2,3) considers the three networks jointly while column (1,2) uses only the first two networks. Δ AIC represents the difference between the Akaike information criteria of the estimated models and our preferred model (1,2) estimated using commuting networks.

$\operatorname{Model} \mathcal{G}$	(1,2)+	-diag.	(1,2)	+c.s.	(1,2)+s	.e. kernel	(1	,2)			
Value	est.	se.	est.	se.	est.	se.	est.	se.			
			Autoreg	gressive	e effect						
$\exp\left(\gamma_{0} ight)$	0.248	0.075	0.212	0.038	0.235	0.039	0.253	0.041			
$\exp\left(\gamma ight)$	0.774	0.235	0.893	0.167	0.813	0.142	0.754	0.127			
Network effect											
$W^{(1)} = K$:											
$\exp\left(\phi_{0}^{(1)} ight)$	0.498	0.076	0.497	0.071	0.503	0.071	0.510	0.073			
$\exp\left(\phi^{(1)} ight)$	0.830	0.131	0.272	0.124	0.625	0.120	0.827	0.127			
$W^{(2)} = K^{ op}$:											
$\exp\left(\phi_{0}^{(2)} ight)$	0.062	0.023	0.078	0.021	0.066	0.020	0.064	0.023			
$\exp\left(\phi^{(2)} ight)$	0.132	0.189	0.000	0.000	0.000	0.000	0.166	0.181			
$W^{(3)}$: diago	nal / co	mpoun	d symm	etric / s	square e	xponential	kernel				
$\exp\left(\phi_{0}^{(3)} ight)$	0.096	2.333	0.000	0.000	0.000	0.000					
$\exp\left(\phi^{(3)}\right)$	0.044	0.000	0.000	0.000	0.000	0.000					
pseudo-R ²	81.1	12%	82.24%		82	.17%	82.12%				
# obs.	3008 300			08	3	008	3008				