DISCUSSION PAPER SERIES

DP14750

A DYNAMIC STRUCTURAL MODEL OF VIRUS DIFFUSION AND NETWORK PRODUCTION: A FIRST REPORT

Victor Aguirregabiria, Jiaying Gu, Yao Luo and Pedro Mira

INDUSTRIAL ORGANIZATION

ISSN 0265-8003

A DYNAMIC STRUCTURAL MODEL OF VIRUS DIFFUSION AND NETWORK PRODUCTION: A FIRST REPORT

Victor Aguirregabiria, Jiaying Gu, Yao Luo and Pedro Mira

Discussion Paper DP14750
Published 12 May 2020
Submitted 11 May 2020

Centre for Economic Policy Research
33 Great Sutton Street, London EC1V 0DX, UK
Tel: +44 (0)20 7183 8801
www.cepr.org

This Discussion Paper is issued under the auspices of the Centre's research programmes:

- Industrial Organization

A DYNAMIC STRUCTURAL MODEL OF VIRUS DIFFUSION AND NETWORK PRODUCTION: A FIRST REPORT

Abstract

This paper presents a dynamic structural model to evaluate economic and public health effects of the diffusion of COVID-19, as well as the impact of factual and counterfactual public policies. Our framework combines a SIR epidemiological model of virus diffusion with a structural game of network production and social interactions. The economy comprises three types of geographic locations: homes, workplaces, and consumption places. Each individual has her own set of locations where she develops her life. The combination of these sets for all the individuals determines the economy's production and social network. Every day, individuals choose to work and consume either outside (with physical interaction with other people) or remotely (from home, without physical interactions). Working (and consuming) outside is more productive and generates stronger complementarities (positive externality). However, in the presence of a virus, working outside facilitates infection and the diffusion of the virus (negative externality). Individuals are forward-looking. We characterize an equilibrium of the dynamic network game and present an algorithm for its computation. We describe the estimation of the parameters of the model combining several sources of data on COVID-19 in Ontario, Canada: daily epidemiological data; hourly electricity consumption data; and daily cell phone data on individuals' mobility. We use the model to evaluate the health and economic impact of several counterfactual public policies: subsidies for working at home; testing policies; herd immunity; and changes in the network structure. These policies generate substantial differences in the propagation of the virus and its economic impact.

JEL Classification: C57, C73, L14, L23, I18

Keywords: COVID-19, Virus diffusion, dynamics, Production and social networks, Production externalities, Disease-Specific Public Health Interventions

Victor Aguirregabiria - victor.aguirregabiria@utoronto.ca
University of Toronto and CEPR

Jiaying Gu - jiaying.gu@utoronto.ca
University of Toronto

Yao Luo - yao.luo@utoronto.ca
University of Toronto

Pedro Mira - mira@cemfi.es
CEMFI

A Dynamic Structural Model of Virus Diffusion and Network Production: A First Report*

Victor Aguirregabiria, Jiaying Gu, Yao Luo, and Pedro Mira

May 11, 2020

Abstract

This paper presents a dynamic structural model to evaluate economic and public health effects of the diffusion of COVID-19, as well as the impact of factual and counterfactual public policies. Our framework combines a SIR epidemiological model of virus diffusion with a structural game of network production and social interactions. The economy comprises three types of geographic locations: *homes*, *workplaces*, and *consumption places*. Each individual has her own set of locations where she develops her life. The combination of these sets for all the individuals determines the economy's production and social network. Every day, individuals choose to work and consume either outside (with physical interaction with other people) or remotely (from home, without physical interactions). Working (and consuming) outside is more productive and generates stronger complementarities (positive externality). However, in the presence of a virus, working outside facilitates infection and the diffusion of the virus (negative externality). Individuals are forward-looking. We characterize an equilibrium of the dynamic network game and present an algorithm for its computation. We describe the estimation of the parameters of the model combining several sources of data on COVID-19 in Ontario, Canada: daily epidemiological data; hourly electricity consumption data; and daily cell phone data on individuals' mobility. We use the model to evaluate the health and economic impact of several counterfactual public policies: subsidies for working at home; testing policies; herd immunity; and changes in the network structure. These policies generate substantial differences in the propagation of the virus and its economic impact.

Keywords: COVID-19; Virus diffusion; Dynamics; Production and social networks; Production externalities; Public health.

JEL classifications: C57, C73, L14, L23, I18

Victor Aguirregabiria. University of Toronto and CEPR. Address: 150 St. George Street. Toronto, ON, M5S 3G7, Canada. E-mail: victor.aguirregabiria@utoronto.ca

Jiaying Gu. University of Toronto. Address: 150 St. George Street. Toronto, ON, M5S 3G7, Canada. E-mail: jiaying.gu@utoronto.ca

Yao Luo. University of Toronto. Address: 150 St. George Street. Toronto, ON, M5S 3G7, Canada. E-mail: yao.luo@utoronto.ca

Pedro Mira. CEMFI. Casado del Alisal 5, 28014 Madrid, Spain. Email: mira@cemfi.es

*We would like to thank comments from Connor Campbell, Marc-Antoine Chatelain, Matt Mitchell, and Adonis Yatchew.

1 Introduction

The purpose of this paper is to develop a framework that combines an epidemiological model of COVID-19 diffusion with a structural game of network production and social interactions. The model emphasizes several aspects which are typically absent from epidemiological models but are work horses in economic models: (i) individuals make choices to maximize their own welfare and respond to incentives affecting this welfare; (ii) individuals interact with each other and their choices and contributions to economic and social outcomes depend on the behavior of coworkers, suppliers, clients, family members, and friends. The production (social) system is a network. Importantly, production and social networks also determine physical links between individuals that can facilitate infection and the diffusion of a virus. In this paper, we develop a structural econometric model that emphasizes the relationship between the production/social network in an economy and the diffusion of COVID-19 and its economic impact. The model can be used to evaluate the impact that different public policies have on the propagation of a virus and its economic effects.

Our model incorporates the following features.

Production and social network. The economy comprises a set of geographic locations and a set of individuals. We distinguish three types of locations: *homes*, *workplaces*, and *consumption places*. Each individual has her own set of locations where she develops her life: her *home(s)*, *workplaces*, and *consumption places*. The combination of these sets for all the individuals determines the economy's production and social network. The structure of the network is determined by data on individuals' mobility in the absence of COVID-19.

Endogenous individuals' choices. Every day, individuals choose to work and consume either outside (with physical interaction with other people) or remotely (from home, without physical interactions). Working (consuming) *outside* is more productive and generates complementarities. Therefore, in the absence of a virus, working outside generates a positive externality.

Epidemiological model. In the presence of a virus transmitted through physical contact, working or consuming *outside* facilitates the diffusion of the virus (negative externality). The epidemiological part of our model incorporates substantial extensions with respect to standard SIR models. First, the production/social network is an important component of our epidemiological model. Second, the probability of infection is endogenous as it depends on working and consumption decisions of the own individual and of other people in her production and social network.[1] Third, the probability

[1] Other types of individuals' choices – which are important for the spread of the virus and have economic implications – are hygiene and keeping a distance from others in personal interactions. So far, we have focused on working and consumption decisions because they have strong economic implications and we have mobility data that measures

of infection varies across local regions in the network. The model provides a landscape of the probability of infection over a city or region, and this landscape evolves endogenously.

Information structure and testing. A special feature of COVID-19 is that asymptomatic individuals, some of which may never develop symptoms, are infectious. In the absence of testing, an individual without symptoms does not know whether she is healthy (noninfectious), or infected asymptomatic, or even already recovered without having developed symptoms. This incomplete information facilitates the diffusion of the virus and is an important element in our model. In this context, testing asymptomatic individuals can reduce this uncertainty, both for the tested individual and economy-wide. The degree of testing is a policy variable chosen by the government.

We characterize an equilibrium of the dynamic game and present an algorithm for its computation. We describe the estimation of the parameters of the model combining several sources of data on COVID-19 in Ontario, Canada, with a daily frequency and at postal code level: epidemiological data; individuals' mobility data; and electricity consumption commercial and for residential customers. In this first report, we calibrate the model and use it to evaluate the health and economic impact of factual and counterfactual public policies: subsidies for working at home; more testing; herd immunity; and changes in the structure of the production/social network.

This paper tries to contribute to a rapidly growing economic literature on the diffusion of COVID-19 and its economic impact. Our paper is closely related to the economic literature on *rational epidemics* that extends the SIR epidemiological model (Kermack and McKendrick, 1927) to take into account how individuals react to changes in prevalence: see Kremer (1996), Geoffard and Philipson (1996), Auld (2003), Chan, Hamilton, and Papageorge (2016), or Greenwood et al. (2019), among others. Most of this literature has focused on the HIV epidemics. Recent papers apply this approach to study COVID-19. Alvarez, Argente, and Lippi (2020), Eichenbaum, Rebelo, and Trabandt (2020), Hall, Jones, and Klenow (2020) combine an epidemiological model with a macro equilibrium model where individuals make consumption and labor supply decisions and are (intertemporal) utility maximizers. Within this framework, Jones, Philippon, and Venkateswaran (2020) study the tradeoffs faced by a social planner who tries to mitigate the spread of COVID-19. They show that the social planner's solution implies a much more drastic reduction in consumption and output than in a decentralized equilibrium. In a similar vein, Acemoglu et al. (2020) study Pareto optimal lockdown policies for COVID-19 – where the key tradeoff is between deaths and economic loss. They show that the Pareto frontier can be substantially improved if lockdown

them.

policies apply differently across age groups.[2]

Our model has several features that we have not seen in the rational epidemics literature in general or in its application to COVID-19. In our model, individuals are affected by their own production/social group and not only by the aggregate. They have information not only about health states at the aggregate level but also about individuals' in their teams. Furthermore, their working and consumption decisions are the result of a network game with complementarities between the players.

Our paper is also related to the literature of learning in social network games. Computational tractability is a fundamental issue in this literature. Models with fully rational players with perfect Bayesian updating beliefs are intractable except in very stylized cases, i.e., very few players and simple exogenous networks. [3] Authors have proposed different forms of adaptive or naive learning from neighbors (Bala and Goyal, 1998; Golub and Jackson, 2010). We follow this approach. More specifically, our assumptions on agents' information structure and beliefs updating are in the spirit of Acemoglu et al. (2011, 2014) and Mossel et al. (2020). In our model, agents combine local and economy-wide information and use an adaptive rule to update their beliefs about health and probability of infection. In contrast, to most models in this literature, agents in our model are forward-looking and solve a dynamic programming problem. However, for tractability we need to impose restrictions on their beliefs about other agents in the game.

In contrast to the standard SIR model, the infection rate in our model is endogenous and heterogeneous across workplaces and consumption places. Pichler (2015) proposes a model of endogenous sickness absences to study their procyclical behavior. Using state-level aggregate data from Germany, he finds evidence that the probability that a sick individual goes to work is higher in boom than in a bust, and this implies a broader spread of a virus during periods of economic expansion. In our model, the risk of infection depends on the number of infected coworkers infected who decide to work in the workplace and not at home. This is an endogenous decision. There is complementarity in the production function between coworkers' choices of working in-site or at home. This implies that an increase in the risk of infection of an individual has a multiplier effect on coworkers' decision of working at home.

In our model, the local structure of the production/social network plays a key role in the diffusion of the virus in a local community and across communities. Measures of social connectivity and

[2] Berger, Herkenhoff, and Mongey (2020) and Piguillem and Shi (2020) study the value of information from testing.

[3] See the discussion on this issue in the recent paper by Mossel et al. (2020; pages 1235-1236), and their citation to Gale and Kariv (2003): *"The computational difficulty of solving the model is massive even in the case of three persons"*.

mobility are important. Kuchler, Russel, and Stroebel (2020) use data from Facebook to measure the degree of social connectivity in Italy and in US. They present evidence on the relationship between an index of social connectivity and the density of COVID-19 cases. In an influential paper, Adda (2016) uses detailed weekly data on disease incidence in France covering a period of 25 years to measure how exogenous changes in social distancing – public transportation strikes, opening of new railway lines, school closure due to holidays – affect the probability of infection.

Individuals' endogenous mobility choices are also an important component of our model. Engle, Stromme, and Zhou (2020) use daily data on COVID-19 cases, mobility data from the market research company *Unacast* (from mobile phones), and the timing of social-distancing policies in New York. Based on their empirical analysis, the authors conclude that both increasing risk of infection and policy restrictions have had a significant effect in reducing mobility.

A motivation for our paper is to provide a structural framework to evaluate the economic impact of factual and counterfactual public policies to mitigate the spread of COVID-19. Recent papers present evidence for Japan, Italy, and France, respectively. Inoue and Todo (2020) use a large dataset with information from more than 1.6 million firms and almost 6 million supply-chain links in Tokyo to quantify the economic impact of a hypothetical lockdown policy in this city. Their estimates and experiments show a huge production loss of 309 billion yen per day. This effect would quickly spread to the whole Japanese economy such that in one month total output would be reduced by 86%. Boeri, Caiumi, and Paccagnella (2020) study the impact of COVID-19 on employment and on the type of jobs in Italy. Barrot, Grassi, and Sauvagnat (2020) propose measures on the degree of remote working for different industries in France and use these measures to estimate how social-distancing policies have affected production. They conclude that a six weeks confinement reduces GDP by approximately 5.6%. The upstream sectors are the most negatively affected. The analysis emphasizes the importance of industrial composition for the aggregate economic impact.

A rapidly growing literature presents reduced form evidence on the socioeconomics of COVID-19. Borjas (2020) shows how socioeconomic characteristics have a significant impact on the probability of being tested and on the result of the test for COVID-19 in New York city. Persons living in poorer, black, and immigrant neighborhoods were less likely to be tested and had a higher probability of a positive result conditional on testing. Fang, Yang and Wang (2020) study the causal effect of Wuhan lockdown on human mobility. Using data on real-time location of smart phones and a Differences-in-Differences approach, they find that the lockdown reduced inflow into Wuhan by 76%, outflows from Wuhan by 56%, and within-Wuhan mobility by 54%.

The rest of this paper is organized as follows. Section 2 presents the model. Section 3 describes data sources and provides a brief discussion on estimation. Section 4 presents a calibration of the model and our policy experiments. We summarize and conclude in section 5.

2 Model

2.1 The network

The economy consists of a set of L geographic locations, $\mathcal{L} = \{1, 2, ..., L\}$, and a set of N individuals, $\mathcal{I} = \{1, 2, ..., N\}$. We index individuals by i and locations by ℓ. Time is discrete and indexed by $t \in \{0, 1, ...\}$. One period is one day. There are three types of locations: *homes*, *workplaces*, and *consumption places*. Each individual has her own set of locations where she develops her life: her home(s), $\mathcal{L}_i^H$, workplaces, $\mathcal{L}_i^W$, and consumption places, $\mathcal{L}_i^C$. Each of these individual-specific sets may contain one or multiple locations.[4]

An individual's *household* consists of all the other individuals who share the same home: that is, the set $\mathcal{H}_i \equiv \{j : \mathcal{L}_i^H \cap \mathcal{L}_j^H \neq \varnothing\}$. Similarly, an individual's *production (consumption) team* consists of all the other people who share a workplace (consumption place) with her: that is, the set $\mathcal{W}_i \equiv \{j : \mathcal{L}_i^W \cap \mathcal{L}_j^W \neq \varnothing\}$ for production, and the set $\mathcal{C}_i \equiv \{j : \mathcal{L}_i^C \cap \mathcal{L}_i^C \neq \varnothing\}$ for consumption.

The combination of all these sets, $\{\mathcal{L}_i^H, \mathcal{L}_i^W, \mathcal{L}_i^C : i \in \mathcal{I}\}$ or equivalently $\{\mathcal{H}_i, \mathcal{W}_i, \mathcal{C}_i : i \in \mathcal{I}\}$, describes the *network* in this economy. This network is an exogenous primitive in the model. The network can vary across economies because industrial composition, geography, transportation infrastructures, culture, etc. In this model, the network not only describes social and economic interlinks but also physical contacts. We assume that an individual with the virus can infect other individual only if they share a common location, either home, or workplace, or consumption place.

A network can be represented by a graph consisting of nodes and edges. Nodes correspond to the set of individuals while edges represent who they are connected to, either through homes, workplaces or consumption places. The *degree of a node*, d_i, is defined as the number of neighbors it has. In our model, d_i is the cardinality (number of elements) of the set $\{\mathcal{H}_i, \mathcal{W}_i, \mathcal{C}_i\}$. There are three interesting properties that can be used to describe a graph: (i) the distribution of degrees d_i, which measures the heterogeneity of individual's connectivity; (ii) the *clustering coefficient*, which measures how often a triple of nodes forms a triangle; and (iii) the *average path length*, where a

[4] The specific definition of *home*, *workplace*, or *consumption place* depends on the data. We use mobility data based on the geographic location of an individual's cell phone. In our data, an individual's home is defined as the location where the phone sits between 6pm to 8am. The set of workplaces consists of the locations – other than home– where the phone is located on working days from 9am to 5pm. The set of consumption places are those visited during weekends and holidays between 9am and 5pm.

path length is defined as the smallest number of edges one needs to travel to connect two nodes. This statistic measures how connected the network is. A *regular graph* is a graph where each node has the same degree: the distribution of degrees is degenerate. Social networks in reality have a non-uniform distribution of degrees, with some individuals having more connections than others.

2.2 Health and diagnosis states and transitions

Variable $x_{it} \in \mathcal{X}$ describes the health and diagnosis state of individual i at period t. It can take ten possible values: $\mathcal{X} \equiv \{H, AU, AD, SU, SD, RAU, RSU, RAD, RSD, Death\}$.

State H (for *Healthy*) means that the individual has not been infected with the virus. States AU, AD, SU, and SD represent infected individuals at different states depending on the development of symptoms and on the existence of diagnosis. State AU (for *infected Asymptomatic Undiagnosed*) represents an individual who is infected and can transmit the virus but has not developed symptoms yet and has not been diagnosed.[5] State AD (for *infected Asymptomatic Diagnosed*) means that the individual is infected and asymptomatic but she has been diagnosed. State SU (for *infected Symptomatic Undiagnosed*) represents an infected individual who has developed symptoms but has not been diagnosed. State SD (for *infected Symptomatic Diagnosed*) represents an infected individual who has developed symptoms and has been diagnosed.

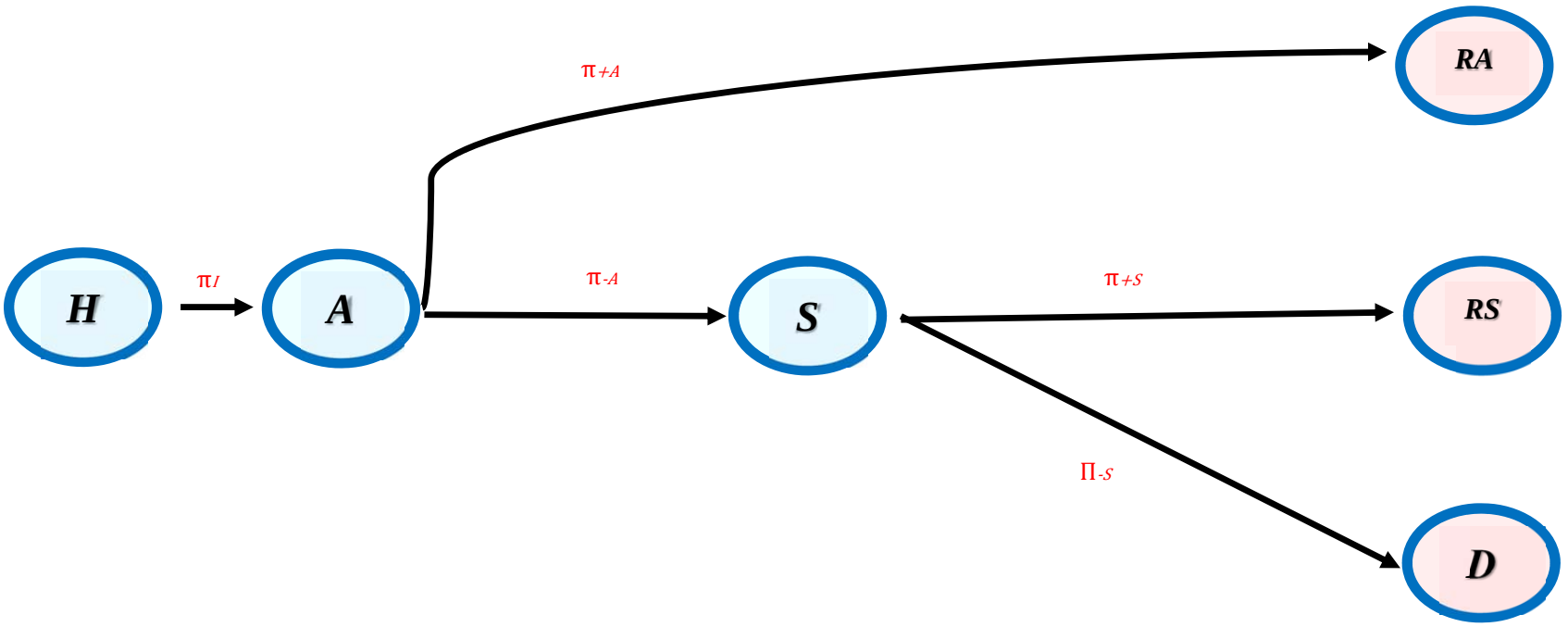
States RAU, RSU, RAD, and RSD represent recovered individuals who are at different states depending on whether they developed symptoms and whether they were diagnosed.[6] States RAU and RSU represent recovered individuals who had never been diagnosed – asymptomatic and symptomatic, respectively – such that the individual does not know if she has been infected. In contrast, states RAD and RSD represent recovered individuals who had been previously diagnosed such that they know they were infected and now are recovered. Finally, $Death$ means death because of the virus.

Transitions between states are based on two types of shocks: health shocks, and testing shocks. Figure 1 presents a flow diagram of a simplified version of our model with only four health states. Figure 2 presents the flow diagram of our model.

[5] Throughout the paper, we use the term infected as synonymous of infectious. In reality, this is not exactly the case. A virus needs to replicate itself sufficiently in a person's body before this person becomes infectious. Our model can be trivially extended to include an additional state between states H and AU such that the model would distinguish between infected and infectious. This additional state – say E form "Exposed" – would represent an individual who had the virus but has not become infectious yet.

[6] We assume that "Recovered" implies not infectious and immune.

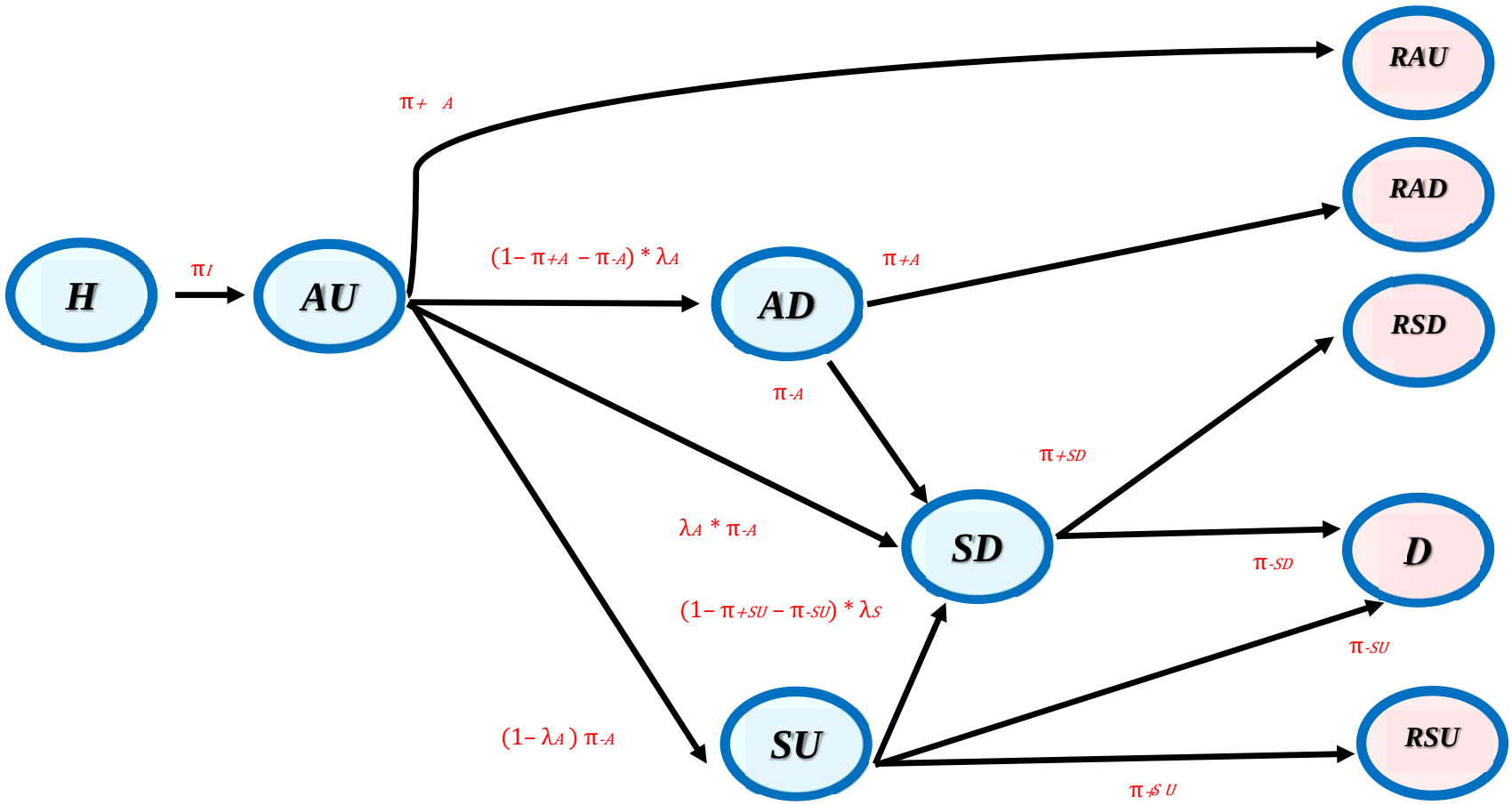
Figure 1: Flow Diagram of States and Transitions in Simplified Model



States: ***H*** *(Healthy);* ***A*** *(Infected Asymptomatic);* ***S*** *(Infected Symptomatic);* ***RA*** *(Recovered from Asymptomatic);* ***RS*** *(Recovered from Symptomatic);* ***D*** *(Death).*

Probabilties: π_I *=Infection;* π_{+A} *=Positive health shock (Recovery) for asymptomatic.;* π_{-A} *=Negative health shock (symptoms) for asymptomatic.;* π_{+S} *=Positive health shock (Recovery) for symptomatic.;* π_{-S} *=Negative health shock (Death) for symptomatic.*

Figure 2: Flow Diagram of States and Transitions



States: ***H*** *(Healthy);* ***AU*** *(Infected Asymptomatic Undiagnosed);* ***AD*** *(Infected Asymptomatic Diagnosed);* ***SU*** *(Infected Symptomatic Undiagnosed);* ***SD*** *(Infected Symptomatic Diagnosed);* ***RAU*** *(Recovered from AU);* ***RSU*** *(Recovered from SU);* ***RAD*** *(Recovered from AD);* ***RSD*** *(Recovered from SD);* ***D*** *(Death).*

Probabilties: π_I *=Infection;* λ_A *=Test to asymptomatic;* λ_S *=Test to symptomatic;* π_{+A} *=Positive health shock (Recovery) for asymp.;* π_{-A} *=Negative shock (symptoms) for asymp.;* π_{+SU} *=Positive shock (Recovery) for symp-undiag.;* π_{-SU} *=Negative shock (Death) for symp-undiag.;* π_{+SD} *=Positive shock (Recovery) for symp-diag.* π_{-SD} *=Negative shock (Death) for symp-diag.*

(i) Infection [Transition $H \rightarrow AU$]. Every day t, a healthy individual (H) can become infected with a probability π_{it}^{I}. This probability is endogenous. It depends on the individual's behavior (confinement or not) and on the behavior of other individuals in her social group and in the rest of the society. We describe the form of this endogenous probability in section 2.5 below. Individuals in state H can be randomly selected to be tested. However, we assume that there are not false positives such that the result of the test is negative and the individual remains in state H.

(ii) Transitions from AU. Every period, an individual in state AU (infected asymptomatic undiagnosed) receives a health shock and a testing shock. The health shock can take three possible values: positive (with probability π_{+A}), negative (with probability π_{-A}), or neutral. If the shock is positive, the individual recovers and becomes immune. If the shock is negative, the individual develops symptoms. We assume that individuals cannot transition within one day from asymptomatic to death: they need to develop symptoms before dying.

The testing shock is independent of the health shock and it determines whether the individual is tested for the virus (with probability λ_A) or not (with probability $1 - \lambda_A$). We assume that tests are accurate such that they cannot provide neither false positives nor false negatives. We also assume that an individual in AU cannot be tested positive on the same day that she receives a positive health shock and recovers.[7]

Under these conditions, there are five possible transitions from state AU. (1) Neutral health shock and no testing – with probability $(1-\pi_{+A}-\pi_{-A})\,(1-\lambda_A)$ – implies that the individual remains in the same state AU. (2) Neutral health shock and testing – with probability $(1 - \pi_{+A} - \pi_{-A})$ λ_A – implies that she remains asymptomatic but now is diagnosed: she moves to state AD. (3) Regardless testing, a positive health shock – with probability π_{+A} – means that she recovers and has not been diagnosed such that she does not know that she was infected. This corresponds to state RAU. (4) Negative health shock and no testing – with probability $\pi_{-A}\,(1 - \lambda_A)$ – implies that the individual develops symptoms but is still undiagnosed: she moves to state SU. (5) Finally, with a negative health shock and testing – with probability $\pi_{-A}\,\lambda_A$ – the individual moves to state SD where she is both symptomatic and diagnosed.

There are two relevant extensions of the model regarding the probability λ_A. First, it is interesting to allow for false negatives in the results of the test. In this extension, the probability λ_A could be interpreted as the product of two probabilities: the probability of being selected for

[7] We are assuming that all the testing is PCR testing. A PCR test detects the infection while it is active. For a PCR test, individuals in states A and S should test positive, and individuals in states H and R negative. For this test, the rates of false positives or false negatives are very low. An interesting extension of the model would be to include other tests that are emerging, including testing for antibodies and immunity.

testing times the probability of a positive result of test conditional on infection, i.e., a *true positive*. In that model, parameter λ_A measures the government testing effort in two different dimensions: the number of tests and the quality of the testing procedure. However, this extension of the model requires also some non-trivial changes in individuals' beliefs about their actual health status.

Other relevant extension is to allow the probability that an asymptomatic individual is selected for testing to depend on the number of members in her social group who are diagnosed as infected.

(iii) Transitions from SU. The transitions from state SU are similar to those from AU, but a main difference is that a negative health shock implies death. Furthermore, the probabilities of a positive and a negative health shock for a symptomatic individual – π_{+SU} and π_{-SU}, respectively – are different that for an asymptomatic. The government policy on testing can be different for symptomatic and asymptomatic, such that probability λ_S is different to λ_A. Similarly as for state AU, we assume that the individual cannot recover and test positive on the same day.

Health and testing shocks determine four possible transitions from state SU. (1) Under neutral health shock and no testing, the individual remains in state SU – with probability $(1-\pi_{+SU}-\pi_{-SU})$ $(1-\lambda_S)$. (2) With neutral health shock and testing, she becomes diagnosed and moves to state SD – with probability $(1-\pi_{+SU}-\pi_{-SU})\,\lambda_S$. (3) A positive health shock – regardless testing – means that she recovers and remains undiagnosed: she moves to state RSU with probability π_{+SU}. (4) Finally, regardless testing, with a negative health shock – with probability π_{-SU} – the individual dies.

(iv) Transitions from AD and from SD. For diagnosed individuals, testing does not matter and only health shocks determine the transitions from these states. Again, we consider that health shocks can take three values, positive, negative, and neutral.

For an individual in state AD, the probability distribution of the health shocks is the same as under state AU. That is, we assume that diagnosis does not affect the health transition when the individual is asymptomatic. There are three possible transitions. Under a positive health shock, the individual recovers and arrives to state RAD – with probability π_{+A}. Under a negative health shock, she develops symptoms and moves to state SD – with probability π_{-A}. And with a neutral shock, she stays in state AD – with probability $1-\pi_{+A}-\pi_{-A}$.

Being diagnosed can affect the distribution of health shocks if an individual is symptomatic: that is, SD individuals are more likely to receive some treatment than SU individuals. Therefore, the probabilities π_{-SD} and π_{+SD} can be different than the probabilities π_{-SU} and π_{+SU}.

There are three possible transitions under state SD. Given a positive health shock, the in-

dividual recovers and arrives to state RSD – with probability π_{+SD}. Under a negative health shock – with probability π_{-SD} – she dies. Finally, with a neutral shock she stays in state SD with probability $1 - \pi_{+SD} - \pi_{-SD}$.

Finally, we assume that all the recovered states – RAU, RAD, RSU, RSD – are absorbing states. Individuals in states RAU and RSU can be subject to random testing, but the test will be negative and individuals remain in the same state.

2.3 Individual decisions

Every period t, individuals make two decisions: working at home or outside, and consuming at home or outside. For the rest of this model section, we describe a simplified version where the only decision is working either at home or outside. We represent this decision using the binary variable a_{it}, where $a_{it} = 0$ means working outside, and $a_{it} = 1$ means *confinement* at home. We also abstract from the home and consumption teams – they only include the own individual – and focus on the workplace team $\mathcal{W}_i$ that has size $|\mathcal{W}_i|$.

The set of feasible choices for an individual depends on her current state. In particular, diagnosed individuals have mandatory confinement. We use $\mathcal{A}(x_{it})$ to represent the choice set under state x_{it} such that $\mathcal{A}(AD) = \mathcal{A}(SD) = \{1\}$, $\mathcal{A}(Death) = \varnothing$, and at any other state $\mathcal{A}(x_{it}) = \{0, 1\}$. For simplicity, we assume that confinement means that the individual does not have physical relationship with any other member of the society.

The assumption that individuals who are undiagnosed or recovered have the freedom to decide to work at home or outside deserves some explanation. One may be concerned that this decision is taken by the firm's manager or, in the case of mandatory confinement policies, by the government. These are important concerns that we take into account.

We incorporate government confinement policies in the model. We take into account that these policies can be applied with very different degrees of flexibility and not uniformly in all the sectors and regions of the economy. Though we can evaluate an hypothetical policy where the government has the ability to lockdown every individual at home, we are interested in more realistic policies that consist of penalties for working or consuming outside or subsidies for confinement at home. These penalties and subsidies may vary across industries and/or geographic locations.

In our data, we cannot distinguish between managers, self-employed, and salaried workers. If we had that information, we could take into account that a manager's decision affects and limits the possible choices of her workers.

2.4 Information structure

The assumptions about individuals' information are important for the model predictions on individual behavior and diffusion of the virus.

(i) Information about the network. An individual knows the identity of the members of her social group but she does not have information about the structure of the network outside of her own units, e.g., friends of friends, etc. According to this condition, we assume below that individuals have only information about members of her social group and economy-wide aggregate information provided by government and media.

(ii) Information about own health status. We consider that, without a test, an individual cannot distinguish between being healthy (H), infected asymptomatic undiagnosed (AU), and recovered after being asymptomatic undiagnosed (RAU). For the sake of notational simplicity, we use $\widetilde{H}$ to represent the union of these three states: $\widetilde{H} \equiv H \cup AU \cup RAU$. We assume that an individual's information about her own health status is captured by the variable $\widetilde{x}_{it}$ such that:

$$\widetilde{x}_{it} = \begin{cases} \widetilde{H} & \text{if} \quad x_{it} \in \{H,\, AU,\, RAU\} \\ \\ x_{it} & \text{if} \quad x_{it} \notin \{H,\, AU,\, RAU\} \end{cases} \tag{1}$$

For an individual in $\widetilde{H}$, it is important to know the likelihood of being in state H, or AU, or RAU. In particular, her confinement decision can have implications on her future health only if she is in state H, but it is completely irrelevant if she is already in states AU or RAU. Therefore, an individual in state $\widetilde{H}$ form beliefs about the probability of being in each of the three specific states. We represent these beliefs as the probabilities $B_{it}^{H|\widetilde{H}}$, $B_{it}^{AU|\widetilde{H}}$, and $B_{it}^{RAU|\widetilde{H}}$ such that $B_{it}^{H|\widetilde{H}} + B_{it}^{AU|\widetilde{H}} + B_{it}^{RAU|\widetilde{H}} = 1$. These beliefs are part of the individual's information set at period t. In section 2.8 below, we describe our assumptions about the initial value and the updating rule of these beliefs.

An individual in state RSU knows that she has experienced symptoms in the past and now does not have those symptoms, but – similarly as someone in state RAU – she does not know that is immune because she has not been diagnosed. From the point of view of an individual's information, state RSU is different to $\widetilde{H}$ only if the symptoms from COVID-19 are different to those from other diseases, like the common flu. For instance, if COVID symptoms were clearly distinguishable, then state RSU would be equivalent to state RSD. At the other extreme, if the symptoms were the same as those from a common flu, then state RSU would be part of $\widetilde{H}$. More generally, we can have a probabilistic belief that captures the informative content of COVID symptoms. In our numerical experiments in section 4, we have assumed that state RSU is equivalent to RSD.

(iii) Information about health statuses of team members. An individual knows the value $\widetilde{x}_{jt}$ for any other individual in her social group. For computational simplicity, we assume that individuals do not use information on team members' health history (before period t).[8]

(iv) Information about health statuses of individuals outside the own team. An individual does not know the health status of individuals outside her team. However, she has information at the aggregate level for the whole economy. In particular, for every state $x \in \mathcal{X}$, she knows the proportion of individuals in state x at period t. We represent this aggregate shares as $S_t(x)$, and $\mathbf{S}_t$ is the vector $\{S_t(x) : x \in \mathcal{X}\}$. The implicit assumption is that the Health Ministry collects this information and communicates it to the citizenship.

(v) Aggregate probability of confinement. An individual has rational beliefs on the equilibrium probability of confinement at period t for each state $\widetilde{x}$. We use $Q_t(\widetilde{x})$ to represent these average probabilities, and $\mathbf{Q}_t$ to represent the vector $\{Q_t(\widetilde{x}) : \widetilde{x} \in \widetilde{\mathcal{X}}\}$.

(vi) Previous day's own decision of confinement. Individual i knows her own choice at previous period, $a_{i,t-1}$. As we explain below in the description of the utility function, lagged choices are payoff relevant because there are costs of changing the form of working – outside or remotely. For computational simplicity, we assume that individuals do not use information in the lagged actions of team members.

(vii) Private information productivity shocks. Finally, individuals are subject to productivity shocks which are their own private information and are independently distributed across individuals and over time. We represent those shocks as $\varepsilon_{it}(0)$ – if working outside – and $\varepsilon_{it}(1)$ – if working at home.

Summarizing, the information set of individual i at period t is:

$$\Omega_{it} = (\widetilde{\mathbf{x}}_{it},\, \mathbf{S}_t,\, \mathbf{Q}_t,\, \varepsilon_{it}(0), \varepsilon_{it}(1)) \tag{2}$$

where we use $\widetilde{\mathbf{x}}_{it}$ to represent in a compact form the vector of state variables $(\widetilde{x}_{it},\, B_{it}^{H|\widetilde{H}},\, B_{it}^{AU|\widetilde{H}},\, a_{i,t-1},\, \{\widetilde{x}_{jt} : j \in \mathcal{W}_i\})$.

2.5 Probability of infection

Let $n_{it}^{(x,0)}$ be the number of members in i's team who are in state x and choose to be work outside. The probability of infection is an increasing function of the number of infected people that individual i interacts with at period t: that is, a function of $n_{it}^{(AU,0)} + n_{it}^{(SU,0)}$. Let π_{it}^{I} be the probability of

[8] Team members' health history may be informative about an individuals' own health. Suppose that yesterday individuals i and j were in state $\widetilde{H}$ and they were working together, and today individual j is in state SD. This information from period $t-1$ contributes to update upwards individual i's belief of being infected at period t.

infection and let $\Lambda(.)$ be the logistic function. Then,

$$\pi_{it}^{I} = \begin{cases} 0 & \text{if } a_{it} = 1 \\ 1 - (1 - \rho_I)^{n_{it}^{(AU,0)} + n_{it}^{(SU,0)}} & \text{if } a_{it} = 0 \end{cases} \tag{3}$$

where the parameter $\rho_I \in (0, 1)$ measures the probability of getting infected from one infectious teammate. The expression assumes independence (and homogeneity) between the events of getting infected from each sick member. Note that π_{it}^{I} is zero if there are not infected team members.

The probability of infection depends on variables which are not part of the information set of individual i. In particular, individual i does not know the number $n_{it}^{(AU,0)} + n_{it}^{(SU,0)}$ because: (i) she cannot distinguish team members who are healthy from those who are infected but undiagnosed; and (ii) she does not know their current confinement decisions a_{jt}. Given her information set Ω_{it}, individual i forms expectations about her infection probability π_{it}^{I}. We describe these beliefs in section 2.8.1.

2.6 Production function

The amount of output generated by an individual depends on her own health status and confinement choice, and on the health statuses and confinement choices of her coworkers. If an individual is diagnosed with infection, she is isolated and does not participate in production such that her output is zero. Therefore, we have that $Y_{it} = 0$ if $x_{it} \in \{AD,\ SD,\ Death\}$. For the other states, the production function is:

$$Y_{it} = \alpha(a_{it}) + \beta(a_{it}, 0)\ n_{it}^{(a=0)} + \beta(a_{it}, 1)\ n_{it}^{(a=1)} + \gamma(a_{it})\ \overline{Q}_t \tag{4}$$

where $\alpha(0)$, $\alpha(1)$, $\beta(0,0)$, $\beta(0,1)$, $\beta(1,0)$, $\beta(1,1)$, $\gamma(0)$, and $\gamma(1)$ are structural parameters, and $n_{it}^{(a=0)}$ and $n_{it}^{(a=1)}$ are the numbers of other individuals in the production team who decide to work at the workplace and remotely from home, respectively.

Parameter $\alpha(a)$ represents the output of an individual when nobody else in the production unit works and her confinement choice is a. We expect $\alpha(0) > \alpha(1)$ since confinement reduces an individual's feasible actions.

Parameter $\beta(a, a')$ measures the contribution of a coworker to the output of an individual when the coworker's confinement choice is a' and the individual's choice is a. We expect $\beta(a, 0) > \beta(a, 1)$ and $\beta(0, a') > \beta(1, a')$. Furthermore, we expect to have complementarity (supermodularity) between the confinement decisions of coworkers such that:

$$\beta(0,0) - \beta(0,1) - \beta(1,0) + \beta(1,1) > 0 \tag{5}$$

Finally, $\overline{Q}_t$ is the proportion of individuals confined at the economy level – i.e., $\overline{Q}_t \equiv \sum_{\widetilde{x}} S_t(\widetilde{x})$ $Q_t(\widetilde{x})$ – and the term $\gamma(a_{it})\,\overline{Q}_t$ accounts for the possible dependence of productivity on the aggregate level of confinement. This effect may be different when working at home or outside. For instance, complementarity in working at home at the aggregate level means that $\gamma(1) - \gamma(0) > 0$. This parameter can vary substantially across industries: see Barrot, Grassi, and Sauvagnat, (2020) for related evidence.

2.7 Preferences

An individual's utility depends on the utility from consumption, $u(C_{it})$, plus the utility from her health status, $\phi(x_{it})$, and minus adjustment costs $\omega(a_{it}, a_{i,t-1})$.[9] In this version of the model we do not consider intertemporal consumption smoothing. Consumption is equal to output minus net taxes (taxes minus subsidies), $\tau_i(a_{it}, x_{it})$. These taxes may depend on the individual's confinement decision and on her health/diagnosis state.[10] Therefore, consumption is $C_{it}(a_{it}) = Y_{it}(a_{it}) - \tau_i(a_{it}, x_{it})$, and the utility function is:

$$U_{it}(a_{it}) \quad = \quad u(Y_{it}(a_{it}) - \tau_i(a_{it}, x_{it})) + \phi(x_{it}) - \omega(a_{it}, a_{i,t-1}) + \varepsilon_{it}(a_{it}) \tag{6}$$

where $\{\phi(x) : x \in \mathcal{X}\}$, $\omega(1,0)$, and $\omega(0,1)$ are parameters, and $\varepsilon_{it}(0)$ and $\varepsilon_{it}(1)$ are private information shocks in individual i's utility of working outside and confined, respectively, and they are independently and identically distributed across individuals and over time with an extreme value type I distribution. We assume that $u(.)$ is a linear function: $u(C) = C$.[11]

For the utility from health status, we assume that $\phi(Death) = 0$ and:

$$\phi(x) = \begin{cases} \phi_{alive} + \phi_{health} & \text{for } x \in \{H, AU, AD, RAU, RSU\} \\ \phi_{alive} + \phi_{health} + \phi_{immu} & \text{for } x \in \{RAD, RSD\} \\ \phi_{alive} & \text{for } x \in \{SU, SD\} \end{cases} \tag{7}$$

Parameter ϕ_{alive} represents the flow utility from being alive. Parameter ϕ_{health} represents the extra utility from being (or feeling) healthy. Since an individual cannot distinguish between states H, AU, or RAU, we assume that these states report the same utility. Parameter ϕ_{immu} captures the additional utility from the knowledge of being recovered and immune.

[9] The cost of no change is zero, such that $\omega(0,0) = \omega(1,1) = 0$.

[10] For instance, we may think in different tax/subsidy policies for immune individuals.

[11] Alternatively, our specification can be interpreted as one where the utility function is logarithmic, $u(C) = \ln(C)$, the production function is Cobb-Douglas, $Y_{it} = \exp\{\alpha(a_{it}) + \beta(a_{it}, 0)\, n_{it}^{(a=0)} + \beta(a_{it}, 1)\, n_{it}^{(a=1)} + \gamma(a_{it})\, \overline{Q}_t\}$, and taxes are proportional, i.e., $C_{it} = Y_{it}(1 - \tau_{it})$.

Changing the location for working involves adjustment costs. Parameter $\omega(1,0)$ is the cost of moving from working outside to working at home; similarly, $\omega(0,1)$ is the cost of moving from working at home to working outside. They capture actual sunk investment costs as well as habits. These costs can play an important role to explain persistence in individual behavior and slow transitions at the aggregate level.

2.8 Beliefs and equilibrium

To maker her choice, an individual needs to form beliefs about different objects. First, if an individual is in state $\widetilde{H}$, she needs to form beliefs about the probability of her actual health status, H, AU, or RAU. We represent these beliefs using the probabilities $B_{it}^{H|\widetilde{H}}$, $B_{it}^{AU|\widetilde{H}}$, and $B_{it}^{RAU|\widetilde{H}}$, where, by definition, we have that $B_{it}^{H|\widetilde{H}} + B_{it}^{AU|\widetilde{H}} + B_{it}^{RAU|\widetilde{H}} = 1$. Second, individuals have beliefs about the probability distribution of their own health at $t+1$ given their information and their own decision at period t. We denote these beliefs as $B_i^i(x_{i,t+1}|\widetilde{x}_{it}, a_{it})$. Third, they have beliefs about the current health status and confinement choices of coworkers, that we denote as $B_{it}^{n(x,a)}$, for $x \in \widetilde{\mathcal{X}}$ and $a = 0, 1$. Finally, individuals have beliefs about the probability distribution of next period health status of their team members – that we represent as $B_i^j\left(x_{jt+1} \mid \Omega_{it}\right)$ – and about the evolution of the state variables at the aggregate level – that we denote as $B_i^{\mathbf{S}}(\mathbf{S}_{t+1}, \mathbf{Q}_{t+1} \mid \Omega_{it})$. Section 2.8.1 describes our conditions on all these beliefs.

Let $B_i(\Omega_{it})$ represent – in a compact form – all the above individual i's *beliefs* given her information set Ω_{it}. Given these beliefs, an individual's best response is the solution of a dynamic programming problem with the following Bellman equation:

$$V^{B_i}(\Omega_{it}) = \max_{a_{it} \in \mathcal{A}(\widetilde{x}_{it})} \mathbb{E}\left[u(C_{it}(a_{it})) + \phi(x_{it}) - \omega(a_{it}, a_{i,t-1}) + \varepsilon_{it}(a_{it}) + \delta V^{B_i}(\Omega_{it+1}) \mid a_{it}, B_i, \Omega_{it}\right] \tag{8}$$

where $\delta \in (0,1)$ is the time discount factor. The expectation operator in this Bellman equation involves the distribution of Ω_{it+1} conditional on Ω_{it}, as well as the distribution of $n_{it}^{(x,0)}$ – which affect individual i's output and probability of infection — conditional on Ω_{it}.

2.8.1 Beliefs

To make our equilibrium concept relatively simple to compute, we introduce several conditions on beliefs. This section describes these conditions.

(i) Current health status and confinement choices of other individuals in the team. For any state x different to $\widetilde{H}$, individual i knows the number of coworkers in that state – $n_{it}^{(x)}$ – but she does

not know their current confinement choices: i.e., she does not know $n_{it}^{(x,0)}$ and $n_{it}^{(x,1)}$. We assume that individuals use the aggregate frequencies in $Q_t(\widetilde{x})$ to form probabilistic beliefs about the choices of team members in state $\widetilde{x}$. Individuals believe that $n_{it}^{(x,0)}$ has a Binomial distribution with arguments $n_{it}^{(x)}$ and $p_{it}^{(x,0)} = 1 - Q_t(x)$. We represent the density function of this Binomial distribution as $B_{it}^{n(x,0)}(n)$.

An individual also knows the number of coworkers in state $\widetilde{H}$, that we denote as $n_{it}^{(\widetilde{H})}$. But she does not the values of $n_{it}^{(x,a)}$ for $x = H$, AU, RAU and $a = 0, 1$. We assume that individuals use the aggregate frequencies in $\mathbf{S}_t$ to form her beliefs about the actual health status of a team member at state $\widetilde{H}$. Therefore, for $x \in \widetilde{H}$, individuals believe that variable $n_{it}^{(x,a)}$ has a Binomial distribution with arguments $n_{it}^{(\widetilde{H})}$ and $p_{it}^{(x,a)|\widetilde{H}}$, where:

$$p_{it}^{(x,a)|\widetilde{H}} \equiv \frac{S_t(x)}{S_t(\widetilde{H})} \; (1 - Q_t(\widetilde{H}))^{1-a} \; Q_t(\widetilde{H})^a. \tag{9}$$

We represent the density function of this Binomial distribution as $B_{it}^{n(x,a)|\widetilde{H}}(n)$.

(ii) Expected probability of infection. Conditional on working choice a_{it}, the expected probability of infection for individual i – that we denote as $\overline{\pi}_{it}^{I,own}(a_{it})$ – has the following expression:

$$\overline{\pi}_{it}^{I,own}(a_{it}) \equiv (1 - a_{it}) \sum_{n} \sum_{n'} B_{it}^{n(AU,0)|\widetilde{H}}(n) \; B_{it}^{n(SU,0)}(n') \; \left[1 - (1 - \rho_I)^{[n+n']}\right] \tag{10}$$

(iii) Next period own health status, and updating of beliefs under state $\widetilde{H}$. Let $B_{it}^{i}(x'|\widetilde{x}, a)$ be the belief that individual i has at period t about the probability of $x_{it+1} = x'$ given that $\widetilde{x}_{it} = \widetilde{x}$ and $a_{it} = a$. For states x_{it} other than $\widetilde{H}$, the individual knows her current health status x_{it} such that the expression for $B_{it}^{i}(x'|\widetilde{x}, a)$ is straightforward and is given by the transition rules that we have described in section 2.2 above. Note that all the transitions from states different to $\widetilde{H}$ do not depend on the individual's choice a_{it}: once the individual is infected, her confinement choices do not have any effect of her own confinement choices.

Here we focus on beliefs on the distribution of x_{it+1} given that the current state is $\widetilde{H}$. An individual in this state uses probabilistic beliefs about her actual status, that we denote as $B_{it}^{H|\widetilde{H}}$, $B_{it}^{AU|\widetilde{H}}$, and $B_{it}^{RAU|\widetilde{H}}$. The only transition probabilities that depend on the individual's choice are the probabilities of $x_{it+1} = H$ and $x_{it+1} = AU$. They have the following form:

$$\begin{cases} B_{it}^{i}(H \mid \widetilde{H}, a_{it}) &= B_{it}^{H|\widetilde{H}} \; \left(1 - \overline{\pi}_{it}^{I,own}(a_{it})\right) \\ \\ B_{it}^{i}(AU \mid \widetilde{H}, a_{it}) &= B_{it}^{AU|\widetilde{H}} \, (1 - \pi_{+A} - \pi_{-A}) \, (1 - \lambda_A) + B_{it}^{H|\widetilde{H}} \; \overline{\pi}_{it}^{I,own}(a_{it}) \end{cases} \tag{11}$$

The first equation says that an individual is healthy at $t+1$ if she was healthy at period t – that has subjective belief $B_{it}^{H|\widetilde{H}}$ – and was not infected during that period – that has subjective belief $1-\overline{\pi}_{it}^{I,own}(a_{it})$. The second equation establishes that she arrives to state AU at $t+1$ either if she was at state AU at period t and she gets a neutral health shock and no testing, or shes was at state H at period t and gets infected. The rest of the transition probabilities from state $\widetilde{H}$ do not depend on a_{it}. They are:

$$\begin{cases} B_{it}^{i}(RAU \mid \widetilde{H}, a_{it}) &= B_{it}^{RAU|\widetilde{H}} + B_{it}^{AU|\widetilde{H}}\ \pi_{+A} \\ B_{it}^{i}(AD \mid \widetilde{H}, a_{it}) &= B_{it}^{AU|\widetilde{H}}\ (1-\pi_{+A}-\pi_{-A})\ \lambda_A \\ B_{it}^{i}(SU \mid \widetilde{H}, a_{it}) &= B_{it}^{AU|\widetilde{H}}\ \pi_{-A}\ (1-\lambda_A) \\ B_{it}^{i}(SD \mid \widetilde{H}, a_{it}) &= B_{it}^{AU|\widetilde{H}}\ \pi_{-A}\ \lambda_A \end{cases} \tag{12}$$

At period $t+1$, if the individual is in state $\widetilde{H}$, the beliefs about her actual health status are updated using the following natural formula. For $x = H,\ AU,\ RAU$:

$$B_{i,t+1}^{x|\widetilde{H}} = \frac{B_{it}^{i}(x \mid \widetilde{H}, a_{it})}{B_{it}^{i}(H \mid \widetilde{H}, a_{it}) + B_{it}^{i}(AU \mid \widetilde{H}, a_{it}) + B_{it}^{i}(RAU \mid \widetilde{H}, a_{it})} \tag{13}$$

Note that this updating rule depends on the individual's previous confinement choice and on previous infection probabilities. In particular, $B_{i,t+1}^{H|\widetilde{H}}$ is greater (and $B_{i,t+1}^{AU|\widetilde{H}}$ is smaller) with $a_{it}=1$ than with $a_{it}=0$.

(iv) Next period health statuses of other individuals in the production unit. Let $B_{it}^{j}(x'|\widetilde{x})$ be the belief that individual i has at period t about the probability of $x_{jit+1}=x'$ given that $\widetilde{x}_{jt}=\widetilde{x}$ where j is a team member. For states x_{jt} other than $\widetilde{H}$, individual i knows coworker j's health status x_{jt} such that the expression for $B_{jt}^{j}(x'|\widetilde{x})$ is straightforward and is given by the transition rules in section 2.2 above. To predict next period's health of team members currently in state $\widetilde{H}$, an individual needs to form beliefs about team members' current health, probability of infection, and confinement choice. Beliefs on current health are based on aggregate frequencies: $S_t(H)/S_t(\widetilde{H})$, $S_t(AU)/S_t(\widetilde{H})$, and $S_t(RAU)/S_t(\widetilde{H})$. Her belief on the probability of confinement is also the aggregate probability $Q_t(\widetilde{H})$. Similarly, she also uses aggregate frequencies to predict the probability of infection of other team members. That is,

$$\overline{\pi}_{t}^{I,other}(\mathbf{Q}_t) = (1-Q_t(\widetilde{H}))\ \left[1-(1-\rho_I)^{|\mathcal{W}|[S_t(AU)\ (1-Q_t(AU))+S_t(SU)\ (1-Q_t(SU))]}\right], \tag{14}$$

where $|\mathcal{W}|$ is the average number of members in a social group. This implies the following beliefs

for the transition of health status for team members who are currently at state $\widetilde{H}$:

$$\begin{cases} B_{it}^{j}\left(H \mid \widetilde{H}\right) &= \dfrac{S_t(H)}{S_t(\widetilde{H})}\left[1-\overline{\pi}_t^{I,other}(\mathbf{Q}_t)\right] \\ \\ B_{it}^{j}\left(AU \mid \widetilde{H}\right) &= \dfrac{S_t(AU)}{S_t(\widetilde{H})}\left(1-\pi_{+A}-\pi_{-A}\right)\left(1-\lambda_A\right)+\dfrac{S_t(H)}{S_t(\widetilde{H})}\overline{\pi}_t^{I,other}(\mathbf{Q}_t) \end{cases} \tag{15}$$

Note that the expressions in equation (15) are equivalent to those in (11) but replacing $\overline{\pi}_{it}^{I,own}(a_{it})$ with $\overline{\pi}_t^{I,other}(\mathbf{Q}_t)$ and $B_{it}^{x|\widetilde{H}}$ with $S_t(x)/S_t(\widetilde{H})$. Similarly, beliefs $B_{it}^{j}(x_{jt+1} \mid \widetilde{H})$ for states $x_{jt+1} \in \{RAU, RAD, AD, SU, SD\}$ are the same as the beliefs $B_{it}^{i}(x_{it+1} \mid \widetilde{H})$ in equation (12) but replacing the own $B_{it}^{x|\widetilde{H}}$ with $S_t(x)/S_t(\widetilde{H})$.

(vi) Adaptive beliefs about next day aggregate states. For beliefs about next day realization of aggregate variables – $\mathbf{Q}_{t+1}$ and $\mathbf{S}_{t+1}$ – we assume that individuals have adaptive beliefs:

$$B_i(\mathbf{S}_{t+1}, \mathbf{Q}_{t+1} \mid \Omega_{it}) = 1\left\{\mathbf{S}_{t+1} = \mathbf{S}_t,\ \mathbf{Q}_{t+1} = \mathbf{Q}_t\right\} \tag{16}$$

2.8.2 Equilibrium

Given her information and beliefs, an individual's best response is the solution to the Bellman equation in (8). Following Aguirregabiria and Mira (2007), the solution to this dynamic programming problem can be described as a conditional choice probability satisfying a best response condition:

$$P\left(\widetilde{\mathbf{x}}_{it}, \mathbf{S}_t, \mathbf{Q}_t\right) = \Lambda\left(v\left(1; \widetilde{\mathbf{x}}_{it}, \mathbf{S}_t, \mathbf{Q}_t\right) - v\left(0; \widetilde{\mathbf{x}}_{it}, \mathbf{S}_t, \mathbf{Q}_t\right)\right) \tag{17}$$

where $P\left(\widetilde{\mathbf{x}}_{it}, \mathbf{S}_t, \mathbf{Q}_t\right)$ is the conditional probability of confinement; $v\left(a; \widetilde{\mathbf{x}}_{it}, \mathbf{S}_t, \mathbf{Q}_t\right)$ is the expected intertemporal utility of choosing alternative a; and $\Lambda(.)$ is the logistic function – that is, the distribution of $\varepsilon_{it}(0) - \varepsilon_{it}(1)$.

Let $\widetilde{\mathbf{x}}_t$ be the vector $(x_{it},\ a_{i,t-1},\ B_{it}^{H|\widetilde{H}},\ B_{it}^{AU|\widetilde{H}} : i = 1, 2, ..., N)$. That is, $\widetilde{\mathbf{x}}_t$ contains health status, current beliefs about status conditional on $\widetilde{H}$, and last period choice of every individual in the economy.[12]

DEFINITION. Given $\widetilde{\mathbf{x}}_t$ (and the corresponding $\mathbf{S}_t$), an equilibrium is a conditional choice probability function of confinement, $P\left(\widetilde{\mathbf{x}}_{it}, \mathbf{S}_t, \mathbf{Q}_t\right)$, and an aggregate vector of probabilities of confinement, $\mathbf{Q}_t = \{Q_t(\widetilde{x}) : \widetilde{x} \in \widetilde{\mathcal{X}}\}$, that satisfy the following conditions: (i) $P\left(\widetilde{\mathbf{x}}_{it}, \mathbf{S}_t, \mathbf{Q}_t\right)$ is the best response probability function of the dynamic programming problem in equation (8); and (ii) for any $\widetilde{x} \in \widetilde{\mathcal{X}}$, $Q_t(\widetilde{x})$ is the probability that results from the aggregation of individual choice probabilities:

$$Q_t(\widetilde{x}) = \frac{\sum_{i \in \mathcal{I}} 1\{\widetilde{x}_{it} = \widetilde{x}\}\ P\left(\widetilde{x}, \widetilde{\mathbf{x}}_{-it}, \mathbf{S}_t, \mathbf{Q}_t\right)}{\sum_{i \in \mathcal{I}} 1\{\widetilde{x}_{it} = \widetilde{x}\}} \qquad \blacksquare \tag{18}$$

[12] Since $\mathbf{S}_t$ is the vector of frequencies of each health status, we have that $\mathbf{S}_t$ is a deterministic function of $\widetilde{\mathbf{x}}_t$.

Remark 1. Equilibrium existence. The best response conditional choice probability in equation (17) is a well-defined function that is continuously differentiable in all its arguments. Equilibrium conditions (18) define a fixed point mapping in $\mathbf{Q}_t$. This mapping is continuos from $[0,1]^{|\widetilde{\mathcal{X}}|}$ into $[0,1]^{|\widetilde{\mathcal{X}}|}$. By Brower's theorem, an equilibrium exists.

Remark 2. Strategic complementarity vs. substitutability in individuals' confinement decisions. If an individual's propensity to confinement increases with the aggregate probability of confinement – if $P\left(\widetilde{\mathbf{x}}_{it}, \mathbf{S}_t, \mathbf{Q}_t\right)$ increases with $\mathbf{Q}_t$ –, then we say that confinement decisions are *strategic complements* and we have a coordination game. Otherwise, confinement decisions are *strategic substitutes* and we have an entry game.

Depending on the values of the parameters, this model can generate either complementarity or substitutability. The complementarity between individuals' confinement decisions in the production function can generate strategic complementarity in this game. However, individuals' concern for their future health can generate strategic substitutability. The larger the proportion of confined individuals, the lower the probability of getting infected and the smaller the expected health benefits of current confinement.

Strategic complementarity or substitutability in confinement decisions can have important policy implications. Under complementarity, small incentives to confinement may generate large changes in the aggregate probability $\mathbf{Q}_t$. In contrast, under substitutability, it may be difficult to achieve a high aggregate probability of confinement.

Remark 3. Multiplicity. Under substitutability, the model has a unique equilibrium $\mathbf{Q}_t$. Under complementarity, the game can have multiple equilibria.

Remark 4. Comparison to Markov Perfect Equilibrium. The concept of Markov Perfect equilibrium (MPE) is commonly used in the literature of dynamic games – especially, in industrial organization – (Maskin and Tirole, 1988; Ericson and Pakes, 1995). The equilibrium concept that we use has clear similarities with MPE but it has an important differences. Individuals do not have rational expectations about the future evolution of the aggregate state variables $\{\mathbf{Q}_{t+s}, \mathbf{S}_{t+s} : s \geq 1\}$. We avoid this assumption of rational expectations mostly because of its computational burden for the solution (and estimation) of a network game. Note that the stochastic process for $\mathbf{S}_t$ and $\mathbf{Q}_t$ depends – in a complicated way – on the network structure and the current state $\mathbf{x}_t$ of the whole economy. It seems unrealistic to assume that individuals know the whole network structure. Furthermore, our adaptive expectations assumption is plausible in such a complex environment and at the daily frequency.

Remark 5. Endogenous stochastic process of $\{\widetilde{\mathbf{x}}_t, \boldsymbol{a}_t : t \geq 1\}$. As defined above, the equilibrium concept that we use takes $\widetilde{\mathbf{x}}_t$ as given and it applies to one period. However, this equilibrium concept implies a stochastic process for the vectors of state and decisions variables. Given $\widetilde{\mathbf{x}}_t$, the equilibrium at period t implies and aggregate $\mathbf{Q}_t$ and the corresponding CCPs for every individual i: $P_i(\widetilde{\mathbf{x}}_t) \equiv P\left(\widetilde{\mathbf{x}}_{it}, \mathbf{S}_t, \mathbf{Q}_t\right)$. These probabilities define the distribution of the vector of choices $\boldsymbol{a}_t$ conditional on $\widetilde{\mathbf{x}}_t$. Then, the transition probabilities of the health state variable and the updating rule of beliefs define the probability distribution of $\widetilde{\mathbf{x}}_{t+1}$ conditional on $\widetilde{\mathbf{x}}_t$ and $\boldsymbol{a}_t$.

3 Data, Identification, and Estimation

3.1 Data

Ideally, the model could be estimated using individual level panel data with a daily frequency, with information on state and decision variables for the own individual and for members of her social group. As far as we know, this type of data – at the individual level – is not available yet for COVID-19.[13] In this section, we discuss several sources of data on COVID-19 that can be combined to estimate the model parameters. We focus on data available for Ontario, Canada.

(i) Basic clinical and epidemiological data on COVID-19. Medical research provides measures of the distribution of the incubation period, the time to develop symptoms, recovery time, basic reproduction number R_0, and death rate. This information can be found in medical journals. We use these data to calibrate some of the epidemiological parameters of the model, i.e., the π parameters.

(ii) Epidemiological data from Public Health Ontario (PHO). This dataset contains information on every identified case of COVID-19. For each case, it provides the following variables: month and day of detection, gender, age group, transmission through travel dummy, hospitalization, admission to intensive care unit (ICU), death, recovery; with the dates of each of these events. Importantly, this data file also provides information on the exact geographic location of the physician that reported the case: name of hospital/clinic, postal code, and exact longitude and latitude of the hospital. A separate data file provides information on testing at the aggregate Ontario level: number of tests, and positive and negative results.

We use these data to measure the state variables $S_t(x)$ as well as testing probabilities λ_{At} and λ_{St}. The geographic and demographic dimension of these data implies that we can measure the

[13] Several government and private institutions are currently collecting survey data on individuals' symptoms, health status, and mobility choices. An example is the *COVID-19 Symptoms & Social Distancing Web Survey* from the Program on the Global Demography of Aging at Harvard University: https://www.hsph.harvard.edu/pgda/covid/

shares S_t at a more disaggregate level that the province of Ontario. We can measure shares S_{gmt} where g is an index for gender-age group, and m is an index for geographic region, e.g., postal code, or a geographic level that represents the region of attraction of a hospital/clinic.

On the negative side, these epidemiological data contain information on the shares $S_t(x)$ only for some states x or grouped states. More specifically, we observe $S_t(\widetilde{x})$ for the following states $\widetilde{x}$: $Death$; recovered-diagnosed, $\widetilde{RD} \equiv RAD \cup RSD$; and nonrecovered-diagnosed, $\widetilde{D} \equiv AD \cup SD$. As a residual, we have the share of the remaining states: $\widetilde{H} \equiv H \cup AU \cup RAU$, and symptomatic-undiagnosed both recovered and nonrecovered, $\widetilde{SU} \equiv SU \cup RSU$.

As for data on testing, publicly available information from PHO does not distinguish between asymptomatic and symptomatic, such that we have only a measure λ_t. In summary, the epidemiological data from PHO can be summarize as follows:

$$\textit{Epid. Data} = \left\{ S_{gmt}(Death),\, S_{gmt}(\widetilde{RD}),\, S_{gmt}(\widetilde{D}),\, S_{gmt}(\widetilde{H} \cup \widetilde{SU}),\, \lambda_t : g,\ m, t \right\} \tag{19}$$

Figure 3 presents daily time series of: (a) number of new cases; (b) new deaths; (c) new tests; (d) new recovered; (e) change in the number of hospitalized; (f) change in the number in ICU; (g) change in the number in ICU with ventilator; and (h) logarithm of cumulative cases. In these figures, we highlight the date of March 17 when the Ontario government declared a state of emergency. For series (d) to (g), the first date with reported data was March 29.[14] These figures show a rapid growth in the number of new cases and deaths. There is also fast growth in the number of new tests. Figure 3(h), for the logarithm of cumulative cases, shows three periods in terms of the average growth rate.

In figure 4, we present a sequence of maps to illustrate the geographic diffusion of COVID-19 in Southern Ontario – that accounts for 94% of the Ontario population. Each region represents a public health unit based on the boundaries defined by Public Health Ontario. The colors represent different levels of the (cumulative) number of confirmed cases per thousand population in a public health unit. We present maps for eight different days with a weekly frequency, from March 17 to May 5. The diffusion of the virus is very heterogeneous across geographic regions. Simcoe and Brockville were the first regions to achieve relatively high levels of infection per capita. Toronto (population 5.5M) and Ottawa (1M) have also reached similar levels. In contrast, the third and fourth most populated cities in Ontario – Hamilton and London, with 0.7M and 0.4M people, respectively – have reached so far a much smaller number of cases per capita.

[14] For those series, the first observation is not the daily change but the level on that date.

Figure 3: Time Series of the Evolution of Covid-19 in Ontario

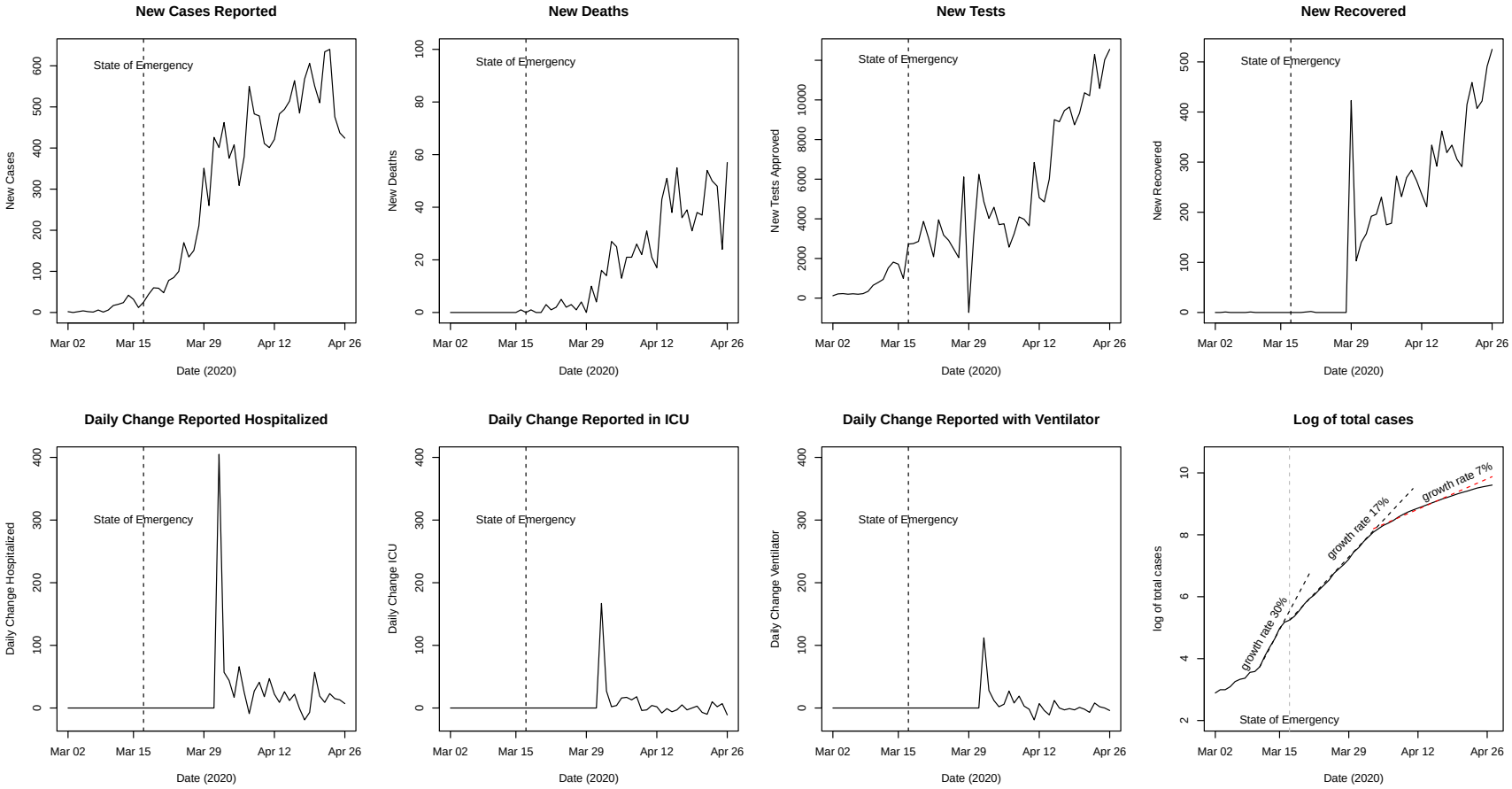
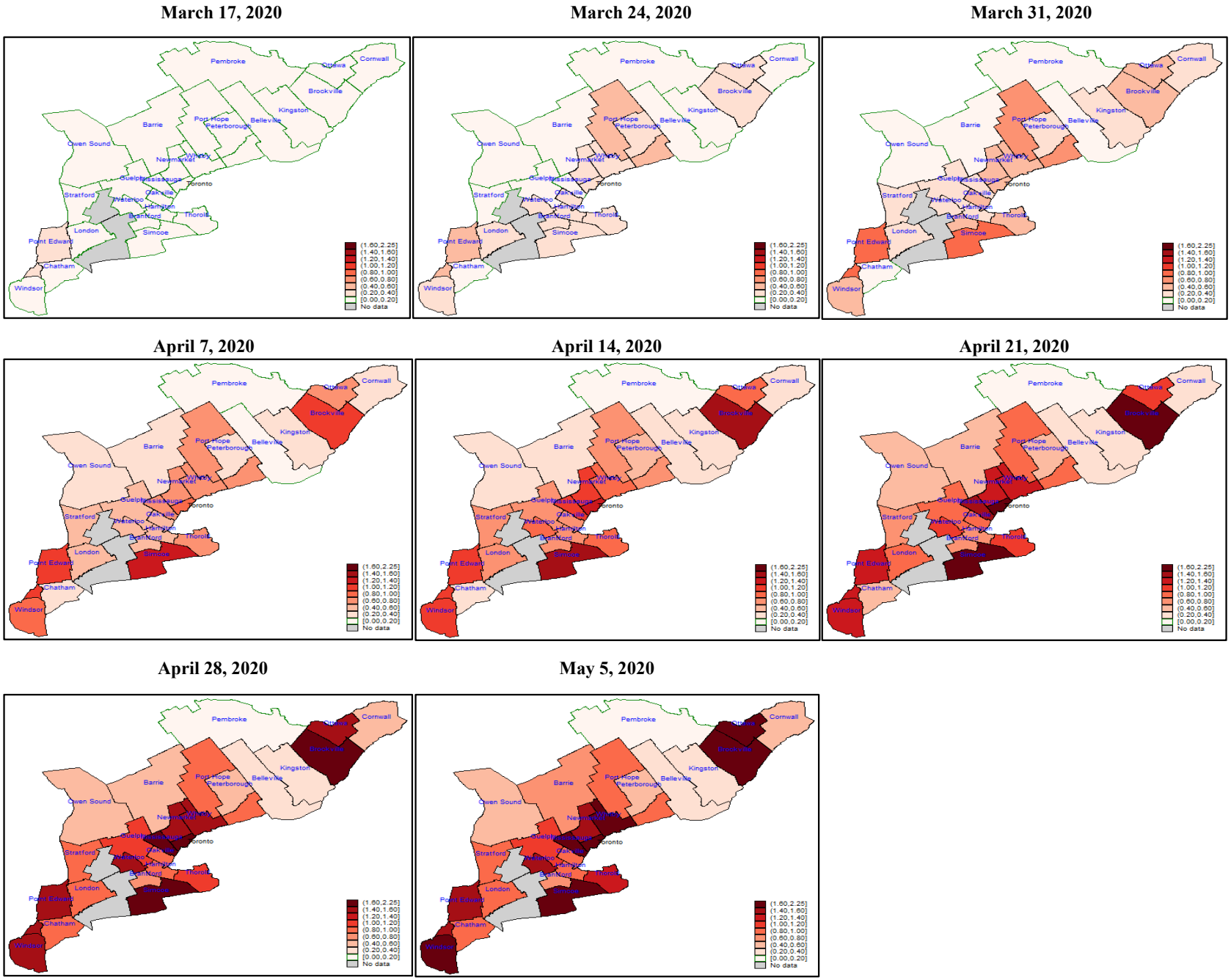


Figure 4: Maps of the Evolution of Covid-19 in Southern Ontario



(iv) Data on social network. Different marketing companies, as well as Google, collect real time information on the geographic location of mobile phones and other mobile devices. Using these data before COVID-19, it is possible to assign each mobile phone a home location $\mathcal{H}_i$ (i.e., the place where the phone sits at night), a workplace $\mathcal{W}_i$ (i.e., the place where the phone is most frequently found during working time); and consumption places $\mathcal{C}_i$.

This type of data is proprietary and it is quite expensive at the individual cell phone level. However, it is possible to get this type of information at an aggregate geographic level and in a probabilistic form. Here we describe a type of data that some market-data companies (e.g., *Safe Graph* in US, or *Environics Analytics* in Canada) have made available to academic researchers.[15]

Let $\mathcal{M} = \{1, 2, ..., M\}$ be the set of geographic locations in Ontario, e.g., set of postal codes. For each triple $(m_{\mathcal{H}}, m_{\mathcal{W}}, m_{\mathcal{C}}) \in \mathcal{M}^3$, define $G_{m_{\mathcal{H}},m_{\mathcal{W}},m_{\mathcal{C}}}$ as the proportion of individuals (mobile phones) that have $m_{\mathcal{H}}$ as home, $m_{\mathcal{W}}$ as workplace, and $m_{\mathcal{C}}$ as consumption place.[16] Of course, most of the entries in the array $\{G_{m_{\mathcal{H}},m_{\mathcal{W}},m_{\mathcal{C}}}\}$ are zero. Note that this array represents individual mobility choices before the spread of COVID-19.

(iv) Data on confinement decisions. This information is also based on phone mobility data. At the postal code and daily level, we have the proportion of phones with home in postal code m which did not go outside to work or to consume at day t. We denote these observed variables as $\overline{Q}^{\mathcal{W}}_{mt}$ and $\overline{Q}^{\mathcal{C}}_{mt}$, respectively.

(v) Data on electricity consumption. We have data from the Ontario's *Independent Electricity System Operator* (IESO) *Smart-metering program* on electricity consumption at the hourly and user level. Based on information on the capacity installed and the peak hours of consumption before COVID-19, we distinguish two types of clients: residential (home) and commercial (workplace). We construct the variables $E^{\mathcal{H}}_{mt}$ and $E^{\mathcal{W}}_{mt}$ that represent electricity consumption in postal code m, day t, by households and workplaces, respectively.

Using data from the network structure before COVID-19 in $\{G_{m_{\mathcal{H}},m_{\mathcal{W}},m_{\mathcal{C}}}\}$, we can construct the variables $g^{\mathcal{H}}_m$ and $g^{\mathcal{W}}_m$ that represent the share of households with home in postal code m, and the share of workers with workplace in m, respectively. Combining these variables with the aggregate consumptions $E^{\mathcal{H}}_{mt}$ and $E^{\mathcal{W}}_{mt}$, we can construct the consumption per household ($e^{\mathcal{H}}_{mt}$) and the consumption per workplace ($e^{\mathcal{W}}_{mt}$):

$$e^{\mathcal{H}}_{mt} \equiv \frac{E^{\mathcal{H}}_{mt}/N}{g^{\mathcal{H}}_m} \quad \text{and} \quad e^{\mathcal{W}}_{mt} \equiv \frac{E^{\mathcal{W}}_{mt}/N}{g^{\mathcal{W}}_m} \tag{20}$$

[15] See Chen and Pope (2020) on the use of cell phone mobility data for US.

[16] This definition is based on the condition that each individual has one home, workplace, and consumption place, but it can be trivially extended to more than one.

We use these variables to estimate the parameters in the production function.

Using electricity consumption as a measure of output needs further explanation. Let $Y_{mt}^{\mathcal{H}}$ and $Y_{mt}^{\mathcal{W}}$ be the output per individual producing home goods and tradable goods, respectively. We assume that the production technology is Leontief (constant coefficients) with respect to electricity. That is,

$$Y_{mt}^{\mathcal{H}} = \min\left\{K_{mt}^{\mathcal{H}}\ ,\ \frac{1}{\Psi_m^{\mathcal{H}}} e_{mt}^{\mathcal{H}}\right\} \quad \text{and} \quad Y_{mt}^{\mathcal{W}} = \min\left\{K_{mt}^{\mathcal{W}}\ ,\ \frac{1}{\Psi_m^{\mathcal{W}}}\ e_{mt}^{\mathcal{W}}\right\} \tag{21}$$

where $K_{mt}^{\mathcal{H}}$ and $K_{mt}^{\mathcal{W}}$ represent the contribution of inputs other than electricity, and $\Psi_m^{\mathcal{H}}$ and $\Psi_m^{\mathcal{W}}$ are technological coefficients that measure the amount of electricity required to produce a unit of output. Note that these technological coefficients may vary across geographic locations because variation in the industrial composition and in the size and type of households.

Given this production function and the assumption that inputs are used efficiently, we have the following relationship between electricity consumption and output:

$$\ln e_{mt}^{\mathcal{H}} = \ln Y_{mt}^{\mathcal{H}} + \ln \Psi_m^{\mathcal{H}} \quad \text{and} \quad \ln e_{mt}^{\mathcal{H}} = \ln Y_{mt}^{\mathcal{H}} + \ln \Psi_{mt}^{\mathcal{H}}$$

Based on this expression, we have that the time difference in log electricity consumption is equal to the time difference in log output: i.e., $\Delta \ln e_{mt}^{\mathcal{W}} = \Delta \ln Y_{mt}^{\mathcal{W}}$.

(vi) Data on public policies. Dates of "State at Home" orders, and industries/sectors affected.

3.2 Identification and Estimation

We can distinguish three different groups of structural parameters in the model: the epidemiological parameters, θ_π; the testing parameters, θ_λ; the production function parameters, θ_Y; and the preferences parameters, θ_U.

$$\begin{cases} \theta_\pi & \equiv & (\rho_I,\ \pi_{-A},\ \pi_{+A},\ \pi_{-SU},\ \pi_{+SU},\ \pi_{-SD},\ \pi_{+SD})' \\ \theta_\lambda & \equiv & (\lambda_A,\ \lambda_S)' \\ \theta_Y & \equiv & (\alpha(0),\ \alpha(1),\ \beta(0,0),\ \beta(0,1),\ \beta(1,0),\ \beta(1,1), \gamma(0),\ \gamma(1))' \\ \theta_U & \equiv & (\delta,\ \phi_{alive},\ \phi_{health},\ \phi_{immu},\ \omega(0,1),\ \omega(1,0))' \end{cases} \tag{22}$$

We consider the following sequential approach for the estimation of these structural parameters.

STEP 1. Parameters θ_π and θ_λ are estimated from the epidemiological data and the testing data.

STEP 2. Production function parameters are estimated using data of log electricity per household and per workplace, $\ln e_{mt}^{\mathcal{H}}$ and $\ln e_{mt}^{\mathcal{W}}$, average confinement, $\overline{Q}_{mt}^{\mathcal{W}}$ and $\overline{Q}_{mt}^{\mathcal{C}}$, and shares of health states, $S_{mt}(\widetilde{x})$.

STEP 3. Given estimates of θ_π, θ_λ, and θ_Y, we estimate the preference parameters, θ_U, using a Simulated Method of Moments (SMM) estimator. This estimator minimizes the distance between actual moments from the data $\{S_{mt}(\widetilde{x}),\ \overline{Q}^{\mathcal{W}}_{mt},\ \overline{Q}^{\mathcal{C}}_{mt},\ \ln e^{\mathcal{H}}_{mt},\ \ln e^{\mathcal{W}}_{mt}\}$ and moments using simulated data from solving the model.

4 Numerical experiments

In this section, we present several numerical experiments to illustrate the properties and predictions of the model. These results are preliminary as they are based on a simplified version of the model and a rudimentary calibration of the model parameters. Based on our calibration, we solve the model and simulate the path of the endogenous variables under different experiments. Experiment 1 is our benchmark scenario and is characterized by a ring network structure, an initial herd immunity of 0%, and no public interventions — no testing and no subsidies. Each of the other experiments incorporates a specific modification with respect to this benchmark. In experiment 2, the initial level of herd immunity is 67%. In experiment 3, we incorporate a government subsidy to work from home. We present results for three levels of this subsidy: 20%, 30%, and 40% of an individual's earnings if her workplace works at full capacity, i.e., when all the workers are active and working in site. In experiment 4 we introduce testing. We present results for three different levels of the testing rate: 2%, 10%, and 20% for asymptomatic individuals, and 80% for symptomatic. Finally, in experiment 5 we modify the structure of the network of social connections.

We start by describing the simplified version of the model and our calibration of the parameters.

Table 1
Parameters in Benchmark Scenario (Experiment 1)

• Population: $N = 100,000$	**Production function**
• # individuals infected at period 1: 10	$\alpha(0) + \beta(0,0)\ \lvert\mathcal{W}\rvert = F$
• Number of team members: $\lvert\mathcal{W}\rvert = 10$	$\alpha(0) + \beta(0,1)\ \lvert\mathcal{W}\rvert = 0.40\ F$
	$\alpha(1) + \beta(1,0)\ \lvert\mathcal{W}\rvert = 0.35\ F$
Epidemiological parameters	$\alpha(1) + \beta(1,1)\ \lvert\mathcal{W}\rvert = 0.20\ F$
• $\pi_{-A} = 1/6$; $\pi_{+A} = 1/7$	$\alpha(0) = 0.20\ F$
• $\pi_{+SU} = 1/14$; $\pi_{-SU} = (10/90)(1/14)$,	$\alpha(1) = 0.05\ F$
implying mortality rate at $SU = 10\%$.	$\Lambda\left(\alpha(1) - \alpha(0) + [\beta(1,0) - \beta(0,0)]\lvert\mathcal{W}\rvert\right) = 0.005$
• $\pi_{+SD} = 1/10$; $\pi_{-SD} = (5/95)(1/10)$	
implying mortality rate at $SD = 5\%$	**Preferences**
• Infection rate: $\rho_I = 0.108$	$\Lambda(\alpha(1) - \alpha(0) + [\beta(1,1) - \beta(0,1)]\lvert\mathcal{W}\rvert + \delta\phi_{health})$=0.99
implying basic rep.number $R_0 = 3.5$	$\delta = 1.0$

4.1 Simplified model

The results that we present here are based on a version of the model that incorporates the simplifying assumptions $A1$ and $A2$ below. The only reason why we include these conditions is computational. Under these assumptions, the equilibrium objects Q_t and P_{it} are scalars and not vectors that depend on all the state variables. Computing an equilibrium in this simplified model takes only 5 minutes – approximately – while for the general model it requires a few hours. By starting with the simplified model, we have been able to try many different specifications and parameterizations at a very low cost of time.

A1. Individuals in the recovered states RSU, RAD, and RSD always choose to work outside. Individuals in state SU always choose to work at home. This behavior is common knowledge. According to this assumption, the only individuals free to choose are those in state $\widetilde{H} = \{H \cup AU \cup RAU\}$.

A2. Individuals are quasi myopic. They are forward looking only in terms of how today's decision affects their own risk of being infected next period.

4.2 Parameterization / Calibration

Table 1 presents then parameters used for the benchmark scenario (Experiment 1).

Population. We consider a relatively large population with $100,000$ individuals.

Number of individuals infected at the initial period. At day 1, there are 10 individuals (i.e., 0.01% of the population), randomly selected, who are infected and undiagnosed (state AU). The rest of the individuals are in state H.

Network structure. In these experiments we consider four different types of networks: a ring lattice; a small world network; a caveman graph; and a randomly rewired caveman graph. Figure 5 presents examples of these network structures with 25 individuals.

A *ring lattice* is a regular graph – each node has the same degree – where nodes are arranged in a circle with each node connected to $|\mathcal{W}|$ nearest neighbors. A ring network with $N >> |\mathcal{W}|$ has a high clustering coefficient and high average path length. A ring lattice does not capture the observed degree heterogeneity that we find in actual social networks. The network at the bottom left of figure 5 is a ring lattice with 25 individuals each having 5 edges.

The *small world network* is a variation of a ring lattice where some nodes are randomly rewired. This random rewiring has the effect of reducing the average path length. An example of a small

world network is shown at the lower right of figure 5.

A *caveman graph* consists of several local clusters where nodes within a cluster are highly connected but there is very little connection between clusters. In our model, this graph may represent an economy where there is very small overlapping between production teams. Compared to a ring lattice, a caveman network shares the feature of a large average path length, but their local structures are very different.

Finally, we consider a variation of the caveman graph where some nodes are randomly rewired. Figure 5 (upper right) shows an example, with 25 individuals, 5 local clusters, and rewiring probability of 0.5. Rewiring decreases average path length making the network more connected.

In our benchmark economy (experiment 1), the network consists of ring lattice with $N = 100,000$ and $|\mathcal{W}| = 10$. We have chosen $|\mathcal{W}| = 10$ workers per team because it is close to median plant size in many industries. In experiment 5, we modify this network structure by changing the value of the parameter $|\mathcal{W}|$ and by considering the other three types of networks.

Epidemiological parameters for COVID-19. We let $\pi_{-A} = 1/6$ (i.e., average incubation period of 6 days), and $\pi_{+A} = 1/7$ (i.e., 7 days of average waiting time to recovery if no symptoms).

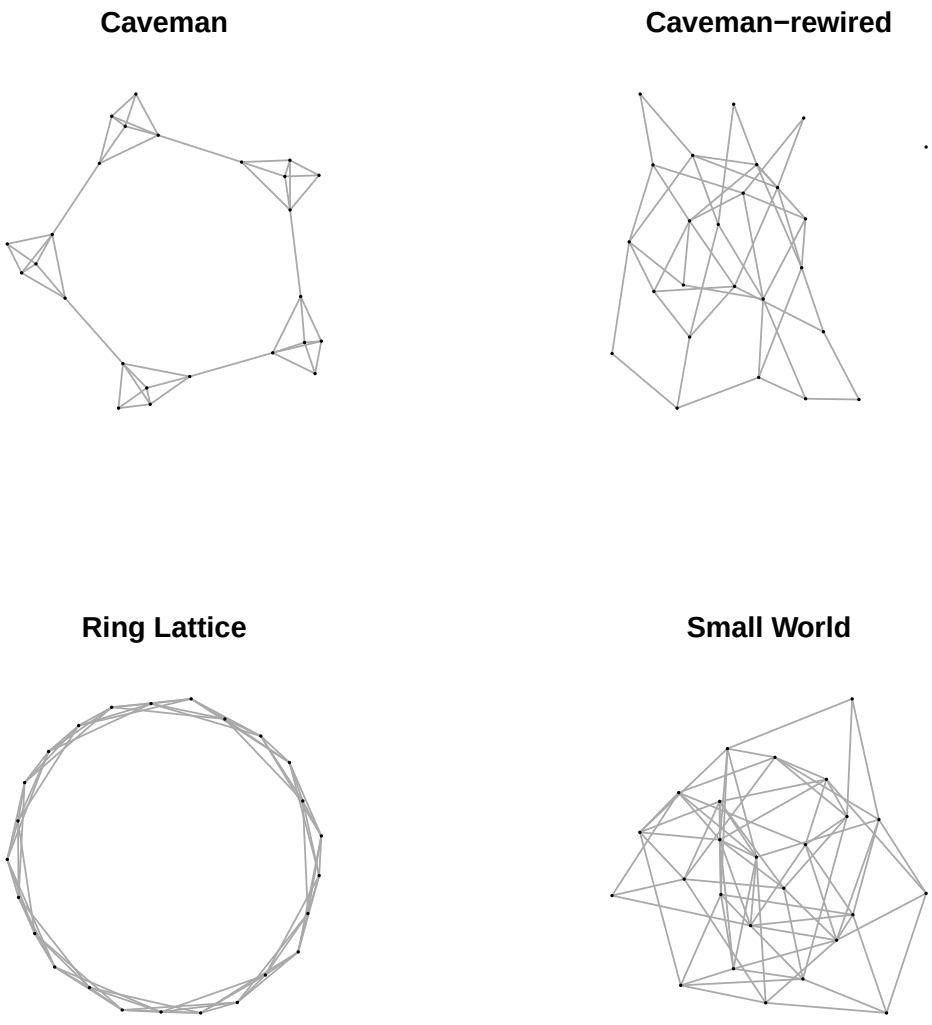
For the recovery of symptomatic-undiagnosed individuals (state SU), we set $\pi_{+SU} = 1/14$, i.e., 14 days of average waiting time to recovery after developing symptoms. Parameter π_{-SU} is set to match a mortality rate of 10% for these undiagnosed (untreated) individuals. That is, $\pi_{-SU}/\pi_{+SU} = 10/90$, and this implies $\pi_{-SU} = 0.0079$.

For the recovery of symptomatic-diagnosed individuals (state SD), we set $\pi_{+SD} = 1/10$, i.e., 10 days of average waiting time to recovery with diagnosis and treatment. Parameter π_{-SD} is set to match a mortality rate of 5% for these diagnosed (treated) individuals. That is, $\pi_{-SD}/\pi_{+SD} = 5/95$, and this implies $\pi_{-SD} = 0.0053$.

To calibrate the parameter ρ_I in the probability of infection, we match the value of the basic reproduction number R_0. For COVID-19, different studies provide values of R_0 between 3 and 4. We choose a value $R_0 = 3.5$. Using the standard representation of R_0 in the epidemiological literature, we have that $R_0 = |\mathcal{W}|\ \rho_I/(\pi_{+A} + \pi_{-A})$. This condition – together with the choice of $|\mathcal{W}| = 10$ – implies a value of $\rho_I = 0.108$.

Production function. For the moment we have chosen $\gamma(0) = \gamma(1) = 0$. Under this condition, the production function becomes $Y_i = \alpha(a_i) + \beta(a_i, 0)\ n_i^{(a=0)} + \beta(a_i, 1)\ n_i^{(a=1)}$, where $n_i^{(a=0)}$ is the number of other members physically present in the workplace, and $n_i^{(a=1)}$ is the number of those

Figure 5: Examples of Network Structures (25 individuals)



working from home. The selection of the parameters $\alpha(0)$, $\alpha(1)$, $\beta(0,0)$, $\beta(0,1)$, $\beta(1,0)$, and $\beta(1,1)$ is based on the following conditions that relate these parameters with the amount of output at full capacity, F.

a. Full capacity output: everybody works outside: $\alpha(0) + \beta(0,0)\ |\mathcal{W}| \equiv F$.

b. Own individual works outside, team members work at home: $\alpha(0) + \beta(0,1)|\mathcal{W}| = 0.40\ F$.

c. Own individual works at home, team members work outside: $\alpha(1) + \beta(1,0)|\mathcal{W}| = 0.35\ F$.

d. Everyone works from home: $\alpha(1) + \beta(1,1)|\mathcal{W}| = 0.20\ F$.

e. Own individual works outside, all members inactive: $\alpha(0) = 0.20\ F$.

f. Own individual works at home, all members inactive: $\alpha(1) = 0.05\ F$.

It is straightforward to verify that conditions (a) to (f) imply that function $\beta(a, a')$ is: monotonic in its two arguments, i.e., $\beta(0,0) > \beta(1,0) > \beta(1,1)$, and $\beta(0,0) > \beta(0,1) > \beta(1,1)$; and supermodular, i.e., $\beta(0,0) - \beta(0,1) - \beta(1,0) + \beta(1,1) > 0$.

The value for full capacity output F is chosen such that the probability of working at home when there is zero risk of infection and all the team members are working outside is 0.5%. That is, $\Lambda(\alpha(1) + \beta(1,0)|\mathcal{W}| - \alpha(1) - \beta(0,0)|\mathcal{W}|) = 0.005$, and given conditions (a) to (f), this implies that $\Lambda(-0.65F) = 0.005$.

Utility of health status. The parameter ϕ_{health} that captures the utility of being uninfected is chosen such that an individual with no active team members and with probability one of getting infected chooses working at home with probability 99%. That is, $\Lambda(\alpha(1) - \alpha(0) + [\beta(1,1) - \beta(0,1)]|\mathcal{W}| + \delta\phi_{health}) = 0.99$. The daily discount factor δ is equal to one.

4.3 Results

4.3.1 Experiment 1: No interventions

Figure 6 presents the simulated paths of endogenous variables in our benchmark scenario with a ring network without testing and without subsidies.

Number of new infections per day (row 2, column 1) and cumulative share of infected individuals (row 1, column 4). During the first 10 days, the rate of infection is quite small. Then, it increases very rapidly and by day 30 practically the whole population has got infected. At its peak, the number of new infections per day reaches $14,000$.

Probability of confinement (row 1, column 1) . This probability responds endogeneously – with some lag – to the rapid expansion of the virus. However, at its peak only 2% of the workers decide (voluntarily) to be confined.

Share of deceased (row 1, column 2) and share of immune (row 1, column 3). The rapid expansion of the virus generates a fast convergence to the new steady-state. This steady-state is practically achieved after only 80 days. The share of deaths is 5%.

Aggregate output (row 2, column 2). The long-run (permanent) effect of the virus on aggregate output is a reduction by 7.5%. This long-run effect is due to 5% reduction in the labor force (deaths) together with the complementarity in the production function. The steady-state amount of output is reached after 80 days periods, and it follows after a very deep crisis that lasts approximately 30 days. During this deep recession, output becomes as low as 58% of its full capacity.

4.3.2 Experiment 2: Herd immunity

Figures 7 presents the results for experiment 2 where everything is the same as in our benchmark except that at period $t = 1$ the share of immune individuals is 67%.[17] The goal of this experiment to show the value of immunity. We can interpret this immunity as the result of vaccination. We can also interpret it as a second wave of the virus that arrives once a substantial proportion of the population is already immune because of previous recovery after infection.

The evolution of all the endogenous variables is dramatically different than under the benchmark scenario. The diffusion of the virus is very slow. The number of infected individuals per day is always lower than 30 and the effect on output is negligible. Herd immunity has a nonlinear effect on the diffusion of the virus.

[17] Based on standard SIR models, the simplest formula for the *herd immunity threshold* is $1-1/R_0$ where R_0 is the basic reproduction number. According to this formula, and our choice of $R_0 = 3.5$, the herd immunity threshold in our benchmark scenario is 71%. An initial share of immune individuals of 67% is slightly lower than this threshold. Of course, we expect our model to deliver threshold values different than $1-1/R_0$. In fact, we find in this experiment that a 67% of immune individuals is enough to generate a very slow diffusion of the virus.

Figure 6: Experiment 1: No Public Interventions

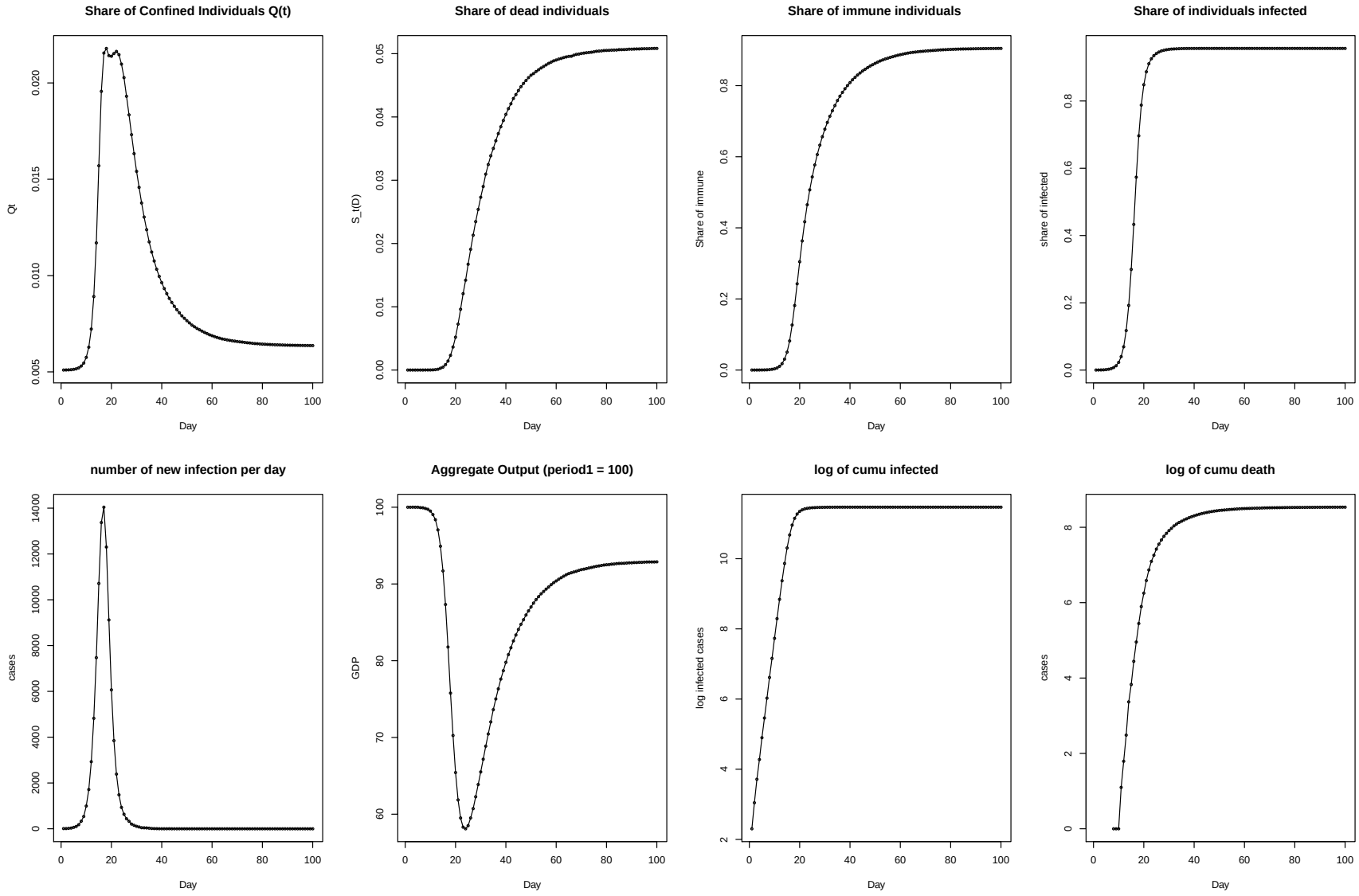
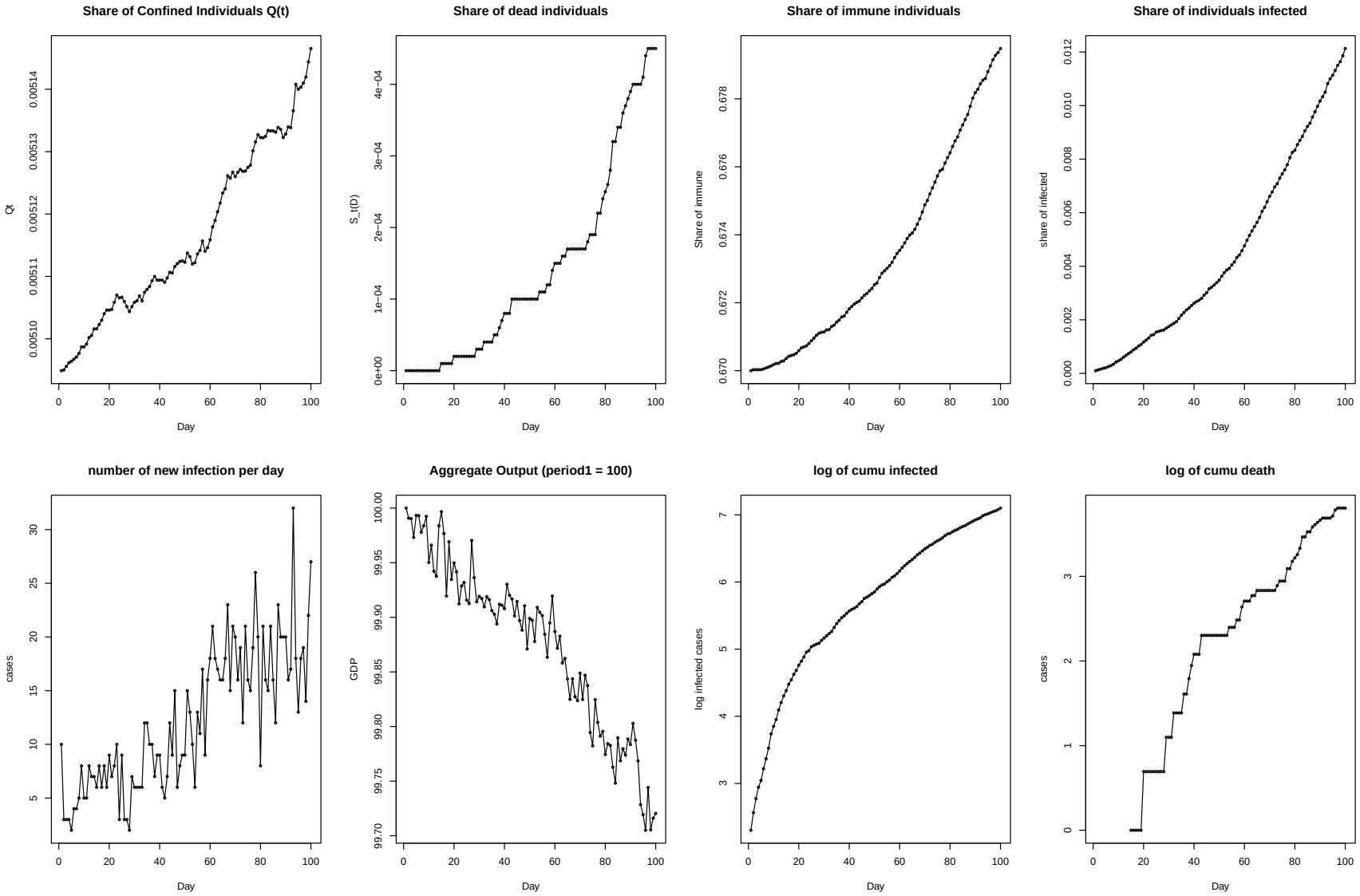


Figure 7: Experiment 2: Initial Herd Immunity of 67 percent



4.3.3 Experiment 3: Subsidy to work from home

In experiment 3, we modify the benchmark scenario by including a subsidy to work from home. We implement experiments for three different values of the subsidy: 20%, 30%, and 40% of the full capacity output. That is, $\tau(0) = 0$ (no subsidy or tax for working outside), and $\tau(1) = 0.2F$, or $0.3F$, or $0.4F$, respectively.

Figure 8 shows the effects of each subsidy rate with respect to the benchmark. The subsidy – even at 20% which is not large as a percentage of the full salary – has a very strong positive effect on confinement decisions, especially during the peak of the virus expansion. At this peak, the probability of working at home increases from 2% without the subsidy to more than 12%. It has also an important effect on the number of infections per day during the peak: it goes from $14,000$ individuals to less than $11,000$. As expected, all these effects become larger when the subsidy rate increases.

4.3.4 Experiment 4: testing

In experiment 4, we modify the benchmark scenario by introducing testing, both to symptomatic individuals – with a probability $\lambda_S = 80\%$ – and to asymptomatic individuals – with three experiments where the probability λ_A takes the values 2%, 10%, and 20%, respectively.

Figure 9 presents the effects with respect to the benchmark. Very interestingly, we find that the effects of introducing testing are basically the complements of the subsidy to confinement. Testing has practically zero effect on the number of infections per day. However, it has an important effect on the timing of confinement and on output, and in the number of deaths. Testing identifies infected individuals and removes them from the labor force. Due to complementarity in the production process, this has a positive incentive on confinement. Now, the peak of confinement occurs earlier than in the benchmark. This has also an effect on output: the recession is slightly not as deep and it is substantially shorter. The long run effects are also smaller, due to the savings of lives.

As the rate of testing asymptomatic individuals increase – to 10% and 20% – the effects on infection start to show, with a lower peak of infectious cases, further reduction in death toll, and even stronger reaction for the dynamics of decision to work from home.

4.3.5 Experiment 5: Modifying the network structure

Here we include a set of experiments to illustrate how the structure of the social/production network influences virus diffusion and its economic impact. Very interestingly, we show that the network

Figure 8: Experiment 3: Subsidy to Work from Home. Differences with Benchmark

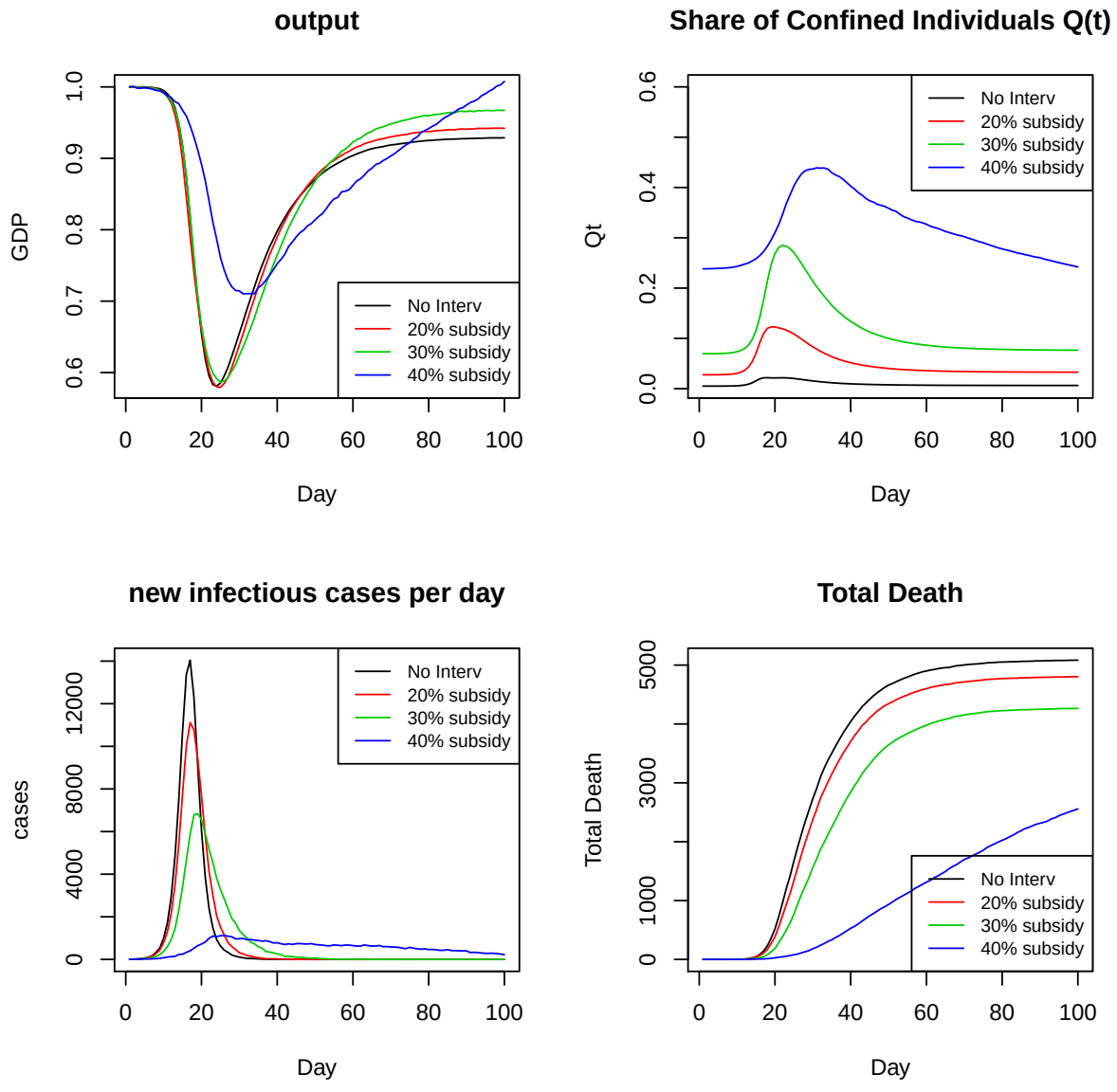
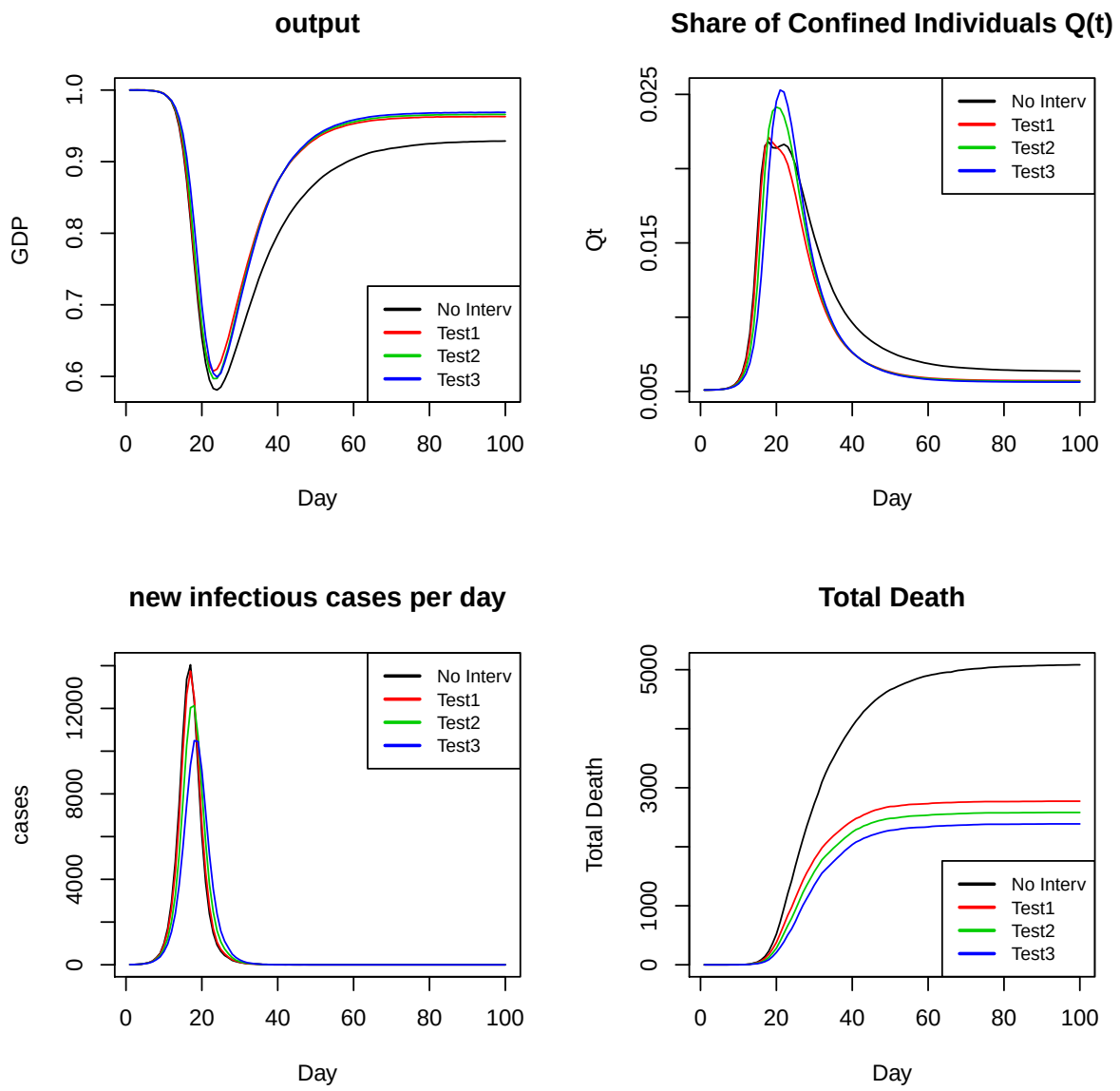


Figure 9: Experiment 4: Testing. Differences with Benchmark



structure also affects individuals' confinement decisions.

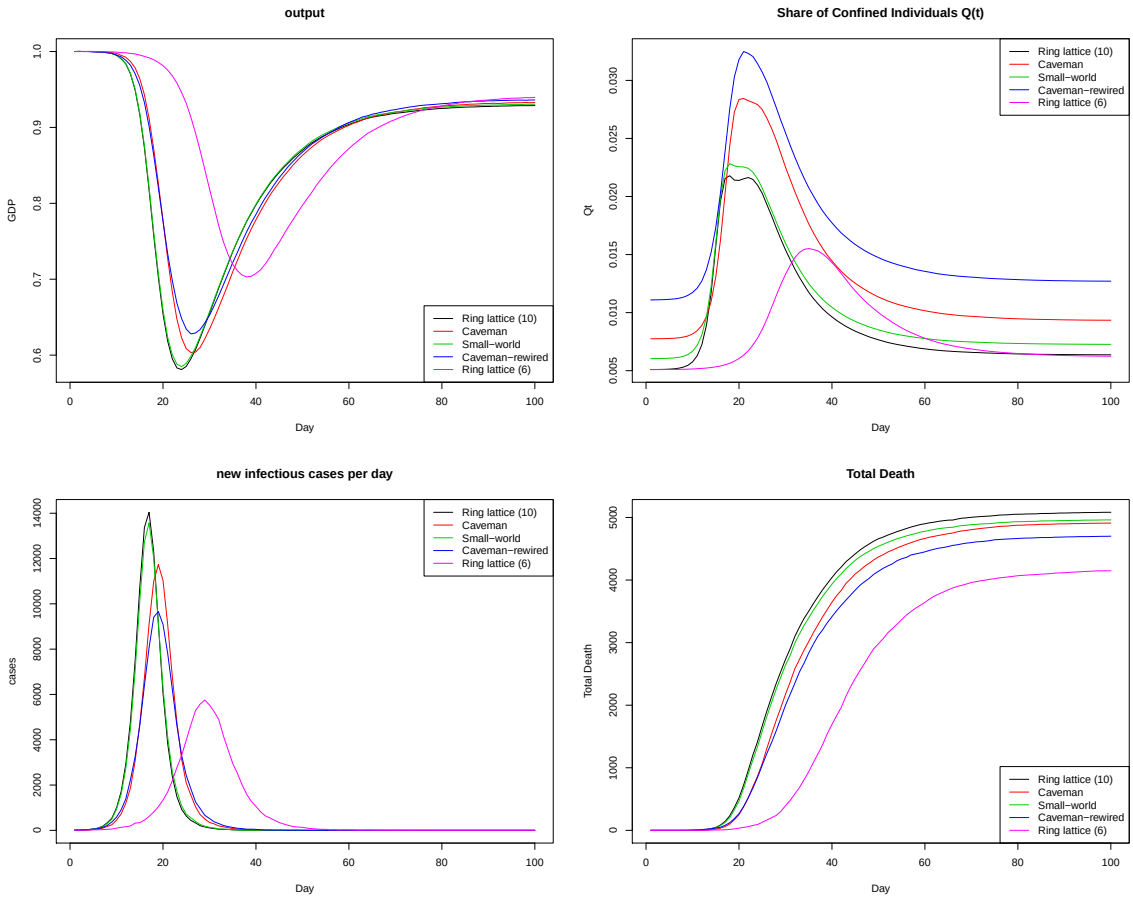
Figure 10 presents the simulated paths of the endogenous variables compared to the benchmark model. The small world network (green curve) is generated from the same ring lattice as the benchmark but with a re-wiring probability of 0.5 using the Watts-Strogatz algorithm. It shows a very similar virus diffusion pattern with similar peak and number of new cases per day. Although the small world network is more connected than the benchmark, the virus does not spread much faster because individuals respond with a higher probability of confinement.

The ring lattice with a smaller degree – $|\mathcal{W}| = 6$ instead of $|\mathcal{W}| = 10$ – delivers the largest differences compared to the benchmark. In this network, with larger social distance, the virus diffuses more slowly. As a result, the share of individuals who choose to work from home is smaller and it peaks much later. Output also gets a much smaller hit.

The caveman network (red curve) shows also different diffusion dynamics than the benchmark. Since individuals are strongly connected within the local clusters, there is a much stronger response to work from home, which then leads to a slower spread of the virus. The economy is hit not as hard as in the benchmark and the total death toll is also smaller. Its variation with a rewiring probability of 0.5 (blue curve) depicts an even more optimistic picture compared to the benchmark, thanks to a stronger propensity to work from home.

Figure 11 shows the logarithm of cumulative cases for the different networks. Though the differences between these four networks are relatively small, they generate substantial differences in the growth rate of the total number of infections. The ring lattice with $|\mathcal{W}| = 10$ and the small world network have the largest growth rate of 57%. The caveman network and its variation have about 47% growth rate. The ring lattice with $|\mathcal{W}| = 6$ generates a growth rate of 30%.

Figure 10: Experiment 5: Network structures. Differences with Benchmark



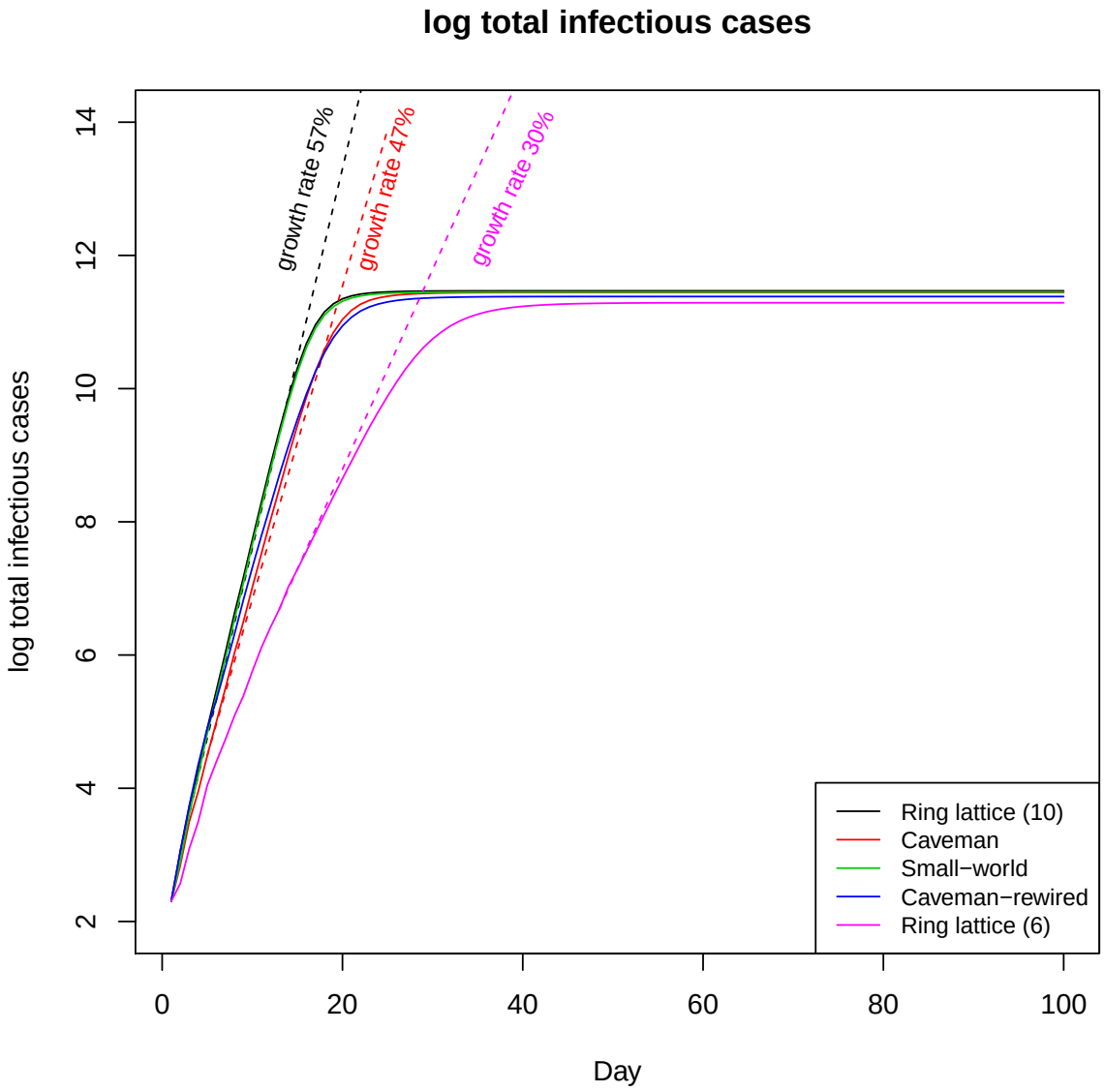
5 Conclusions

The COVID-19 pandemic has generated important challenges in our societies. The academic economics profession has responded quickly and enthusiastically to these challenges, with notorious engagement. We, as many other academics, believe that COVID-19 is a unique opportunity to make substantial research progress on some issues which are important for the economics of pandemics but which are not only constrained to this problem. This paper is a first report of an ongoing research project motivated by some of these problems.

In this paper, we propose dynamic structural network game to study individuals' working and consumption decisions during the COVID-19 pandemic. The model can be estimated using daily data on the spread of the virus, individuals' mobility choices, and electricity consumption for households and establishments. This type of data is now available for multiple countries or regions.

The topics that we plan to investigate in research program are especially connected to COVID-19 but they are also of broader interest in economics and especially in empirical IO and structural

Figure 11: Experiment 5: Log total infection cases for different network structures



econometrics. (1) Incorporate, in structural models, high-frequency granular data on individuals' mobility choices. (2) Estimation of production functions that account for different productivity of working in-site or at home and for complementarity in these decisions. (3) Estimation of games allowing for flexible information structures, biased beliefs, and learning. (4) In dynamic social network games, definition of equilibrium concepts which are realistic and can be computed at a reasonable cost when applied to the social network structures that we observe in reality.

References

[1] Acemoglu, D., K. Bimpikis, and A. Ozdaglar (2014): "Dynamics of information exchange in endogenous social networks," *Theoretical Economics*, 9(1), 41-97.

[2] Acemoglu, D., V. Chernozhukov, I. Werning, and M. Whinston (2020): "A Multi-risk SIR model with optimally targeted lockdown," NBER Working Paper No. 27102.

[3] Acemoglu, D., M. Dahleh, I. Lobel, and A. Ozdaglar (2011): "Bayesian learning in social networks," *The Review of Economic Studies*, 78(4), 1201-1236.

[4] Adda, J. (2016): "Economic activity and the spread of viral diseases: Evidence from high frequency data. The *Quarterly Journal of Economics*, 131(2):891-941.

[5] Aguirregabiria, V., and P. Mira (2007): "Sequential estimation of dynamic discrete games," *Econometrica*, 75, 1-53.

[6] Alvarez, F., D. Argente, and F. Lippi (2020): "A Simple Planning Problem for COVID-19 Lockdown," No. 20/05. Einaudi Institute for Economics and Finance (EIEF).

[7] Atkeson, A. (2020): "What will be the economic impact of COVID-19 in the US? Rough estimates of disease scenarios," Staff Report 595, Federal Reserve bank of Minneapolis.

[8] Auld, M. (2003): "Choices, beliefs, and infectious disease dynamics," *Journal of Health Economics*, 22(3), 361-377.

[9] Bala, V., and S. Goyal (1998): "Learning from neighbours," *The Review of Economic Studies*, 65(3), 595-621.

[10] Barrot, J., B. Grassi, and J. Sauvagnat (2020): "Sectoral Effects," *COVID Economics.* Issue 3.

[11] Berger, D., K. Herkenhoff, and S. Mongey (2020): "An SEIR infectious disease model with testing and conditional quarantine," No. w26901. National Bureau of Economic Research.

[12] Boeri, T., A. Caiumi, and M. Paccagnella (2020): "Work versus Safety," *COVID Economics.* Issue 2.

[13] Borjas, G. (2020): "Demographic Determinants of Testing Incidence and COVID-19 Infections in New York City Neighborhoods," NBER WP #26952.

[14] Chan, T., B. Hamilton, and N. Papageorge (2016): "Health, risky behaviour and the value of medical innovation for infectious disease," *The Review of Economic Studies*, 83(4), 1465-1510.

[15] Chen, K., and D. Pope (2020): "Geographic Mobility in America: Evidence from Cell Phone Data," NBER Working Paper No. w27072.

[16] Eichenbaum, M., S. Rebelo, and M. Trabandt (2020): "The macroeconomics of epidemics," No. w26882. National Bureau of Economic Research.

[17] Ericson, R., & Pakes, A. (1995): "Markov-perfect industry dynamics: A framework for empirical work," *Review of Economic Studies*, 62, 53-82.

[18] Engle, S., J. Stromme, and A. Zhou (2020): "Staying at Home: Mobility Effects of COVID-19," *COVID Economics.* Issue 4.

[19] Fang, H., L. Wang, and Y. Yang (2020): "Human Mobility Restrictions and the Spread of the Novel Coronavirus (2019-nCoV) in China," NBER Working Paper No. 26906.

[20] Fershtman, C., and Pakes, A. (2012): "Dynamic games with asymmetric information: A framework for empirical work," *Quarterly Journal of Economics*, 127, 1611-1661.

[21] Gale, D., and S. Kariv (2003): "Bayesian learning in social networks," *Games and Economic Behavior* 45(2), 329-346.

[22] Geoffard, P, and T. Philipson (1996): "Rational Epidemics and Their Public Control," *International Economic Review*, 37(3), 603-624.

[23] Golub, B., and M. Jackson (2010): "Naive learning in social networks and the wisdom of crowds," *American Economic Journal: Microeconomics,* 2(1), 112-149.

[24] Greenwood, J., P. Kircher, C. Santos, and M. Tertilt (2019): "An equilibrium model of the African HIV/AIDS epidemic," *Econometrica*, 87(4), 1081-1113.

[25] Hall, R., C. Jones, and P. Klenow (2020): "Trading Off Consumption and COVID-19 Deaths." Mimeo, Stanford University.

[26] Heidhues, P., & Köszegi, B. (2018): "Behavioral industrial organization". In B. Bernheim, S. DellaVigna, and D. Laibson (eds.) *Handbook of Behavioral Economics*, Elsevier.

[27] Inoue, H., and Y. Todo (2020): "Lockdowns and Supply Chains," *COVID Economics.* Issue 2.

[28] Jones, J., T. Philippon, and V. Venkateswaran (2020): "A Note on Efficient Mitigation Policies," *COVID Economics.* Issue 4.

[29] Kermack, W., and A. McKendrick (1927): "A contribution to the mathematical theory of epidemics," *Proceedings of the Royal Society of London. Series A, Containing papers of a mathematical and physical character,* 115(772), 700-721.

[30] Kremer, M. (1996): "Integrating behavioral choice into epidemiological models of AIDS," *The Quarterly Journal of Economics*, 111(2), 549-573.

[31] Kuchler, T., D. Russel, and J. Stroebel (2020): "The Geographic Spread of COVID-19 Correlates with Structure of Social Networks as Measured by Facebook," NBER WP #26990.

[32] Maskin, E. and J. Tirole. (1988): "A theory of dynamic oligopoly, II: Price competition, kinked demand curves, and Edgeworth cycles," *Econometrica*, 3, 571-599.

[33] Mossel, E., M. Mueller-Frank, A. Sly, and O.Tamuz (2020): "Social learning equilibria," *Econometrica*, 88(3), 1235-1267

[34] Pichler, S. (2015): "Sickness absence, moral hazard, and the business cycle," *Health Economics*, 24(6), 692-710.

[35] Piguillem, F., and L. Shi (2020): "The Optimal COVID-19 Quarantine and Testing Policies," No. 20/04. Einaudi Institute for Economics and Finance (EIEF).