CSEF

Centre for Studies in Economics and Finance

WORKING PAPER NO. 591

Testing Policies During an Epidemic

Francesco Flaviano Russo

December 2020

University of Naples Federico II

University of Salerno

Bocconi University, Milan

CSEF - Centre for Studies in Economics and Finance
DEPARTMENT OF ECONOMICS – UNIVERSITY OF NAPLES
80126 NAPLES - ITALY
Tel. and fax +39 081 675372 – e-mail: csef@unina.it
ISSN: 2240-9696

CSEF
Centre for Studies in Economics and Finance

WORKING PAPER NO. 591

Testing Policies During an Epidemic

Francesco Flaviano Russo *

Abstract
I simulate a stochastic SIR-Macro model to study the effects of different testing policies to isolate and quarantine the infectious during an epidemic. I show that contact tracing performs better than random screenings, and than tests on a voluntary basis, only conditionally on having enough capacity to frequently process a large number of tests. Since the testing capacity is difficult to build in the short run through investments, I study the effects of two alternatives: group testing and the use of less sensitive tests. Both strategies, when combined with contact tracing, can significantly smooth the peak of the epidemic, ease the pressure on hospitals, decrease mortality and reduce the overall output loss. I also show that the gains are higher in case of an endogenous reduction of social activities in response to the epidemic.

Keywords: Group-Testing, Tracing, Infection.

JEL Classification: E1, I1, H12.

Acknowledgements: I would like to thank Marco Tarzia for the helpful discussions. The paper includes some of the material that circulated in a working paper titled "Increasing Lab Capacity For Covid-19 Viral Testing". All errors are mine.

* Università di Napoli Federico II and CSEF. E-mail: francescoflaviano.russo@unina.it.

Table of contents

1 Introduction

The ability to identify and quarantine the biggest possible number of infectious individuals is a fundamental asset in the fight against an epidemic, which can avoid the massive disruption of economic and social activities brought by lockdowns. However isolating the infectious could be a daunting task whenever infected, but otherwise asymptomatic, individuals could be pathogen carriers that spread the disease, as in the Covid-19 case. Absent the possibility of frequently screening the entire population, which could perfectly identify all of the infectious, but which is logistically and perhaps also economically unfeasible, the problem is how to optimally design a testing policy targeted to a subset of the population only. In this paper, I propose a stochastic epidemiological model with production to study the comparative performance of three policy alternatives: enforcing periodical random screenings, allowing tests on a voluntary basis only, or implementing contact tracing. This last option, sometimes labeled a TTQ policy (Testing-Tracing-Quarantine), consists in searching and testing all of the exposed individuals who came into contact with the already discovered infected, and which might therefore be infectious. Sophisticated technologies that collect and elaborate cell phone and security camera data through numerical algorithms, although privacy-invading, can be very effective at doing so, as China and South Korea showed when faced with Covid-19.

The model, together with a production function, nests four new features within an otherwise standard Susceptible-Infected-Recovered (SIR) model (Kermack and McKendrick 1954; Allen 2017) of the spread of a contagious disease: voluntary tests, conditional on preferences and on symptoms; an explicit and detailed model of contact tracing, that accounts for the structure of the economy; the possibility of concurrent infections, with similar symptoms; an upper bound to the number of tests that can be processed each period. This test capacity constraint is indeed the crucial determinant of the effects of the alternative testing policies. Simulation results show that contact tracing is the best policy alternative, in terms of smoothing the peak of the epidemic, relieving the pressure on the health system, reducing mortality and reducing the output loss, but only conditionally on having enough capacity to frequently process a large number of tests. In case of limited testing capacity, instead, the performance of all testing policies is not very different from what could be achieved testing only the hos-

pitalized that need a targeted medical treatment, and contact tracing does not significantly improve upon random screenings or upon testing on a voluntary basis.

The question, then, is how to increase the testing capacity. The most obvious way is through investments, either buying more lab material or training more professional. But capacity investments are unlikely to be feasible in the short run, which is exactly when a prompt response to a non-anticipated epidemic is most needed. Training professional requires time, and the very-specialized manufacturers of lab equipment might not be able to meet the demand, especially in case of a pandemic. Moreover, it is unfeasible for many countries to bear the additional debt needed to finance them. Fortunately, capacity investments are not the only possibility to overcome a testing capacity constraint. I study the effects of two alternative strategies, explicitly building them in the model: group testing and the use of less precise tests.

Group testing is a rather old idea. In its simplest version, it consists in collecting two samples from each individual, testing first samples in groups and then processing second samples individually only in case the result is positive for a group. It is well known, at least since the work by Dorfman (1943), that group testing yields a significant reduction in the number of tests needed to screen a population, while keeping the same quality of the original test. Israel and the state of Nebraska are two examples of the use of group testing for Covid-19. My contribution consists in modeling group testing within a SIR-Macro model to simulate its impact on the evolution of the epidemic and on its consequences. It is also well known that group testing works best in case of a small incidence of positive test results, since second samples are seldom processed. Putting it differently, the number of individuals that can be screened with group testing, for fixed testing capacity, is inversely proportional to the ratio of positive tests results to tests performed. In the context of an epidemic, the bigger is the number of tests performed, the smaller is this ratio. There is therefore an important reinforcing effect that makes group testing particularly effective: adopting group testing, coeteris paribus, increases the number of screened individuals, thereby lowering the ratio of positives test results to tests performed, which, in turn, allows for an even bigger testing capacity increase and so on.

The actual number of tests to perform with group testing is contingent on the chosen group

size, which must decrease with the ratio of positives to tests and which must be adjusted frequently to account for the epidemic dynamic. But this ratio is unknown before processing the tests, so it is necessary to continuously forecast it, potentially making group testing difficult to implement. I show however that even a simple adaptive model, with the optimal group size set, each period, according to the positives to tests ratio in the previous period, can allow the identification of a big percentage of the infectious, thereby smoothing the epidemic peak and reducing hospitalizations, deaths and the overall output loss. In the baseline model calibration, when group testing is used within a contact tracing framework, it can reduce mortality by 56% with respect to an alternative scenario with contact tracing but without group testing, and by 63% with respect to the alternative policy with tests to the hospitalized only. The overall output loss, with contact tracing and group testing, turned out to be 38% smaller than in case of contact tracing without group testing and 47% smaller than in case of tests for medical treatment only.

The simulations also show that group testing, within a contact tracing policy, can significantly decreases hospitalizations, mortality and the output loss, even in case of a very limited tracing of the contacts from social activities, which are indeed the most difficult to identify. In the limit, tracing only family members, school peers and colleagues on the workplace, results in a 50% reduction of mortality with respect to benchmark with tests to the hospitalized only. This last result suggests that it is better to first focus the policy efforts on increasing the testing capacity and that it is not necessary to use sophisticated, and privacy-invading, technologies to make contact tracing work.

The second idea to increase the testing capacity consists in using lower quality tests, which are typically more readily available and which do not require specialized professionals. I consider in particular the possibility of using tests with lower sensitivity, that do not always correctly identify the infected. As for the case of group testing, my contribution is evaluating their performance within a SIR-Macro model. I show that using such tests to overcome the capacity constraint significantly eases the pressure of the health system, while also decreasing the mortality and the output loss of the epidemic, especially when used in combination to contact tracing. In the baseline model simulation with contact tracing, if additional tests with 50% sensitivity were available in the same number as the tests with perfect sensitivity,

there would be a 39% mortality reduction and a 28% lower output loss with respect to the benchmark with tests to the hospitalized only. I do not consider the possibility of adopting tests with lower specificity for two reasons: because they will clearly be more effective at reducing infections, since they result in quarantines for some non-infectious, but wrongly classified, individuals, and because the voluntary compliance with such tests will most likely be extremely small, given that they could result in unnecessary individual lockdowns with a high social and economic cost.

The simulation also highlight several additional results. First, the gains from all testing policies are higher in case of an endogenous response to the epidemic, in particular if the agents decreased their social activities proportionately to the observed disease prevalence, in order to reduce the amount of risk to which they are exposed. In this scenario, contact tracing, that allows the identification of a bigger percentage of the infected, yields a steeper social contacts reduction with respect to the other testing policies, with stronger mitigation effects. Similarly, contact tracing is more valuable in case of an endogenous reduction of the infection probability due to the adoption of precautionary measures in response to the diffusion of the disease, such as, for instance, wearing face masks, washing hands more carefully or sanitizing their homes and offices more often. The estimated gains from contact tracing and group testing in the baseline simulation are therefore likely to be lower bounds. Moreover, I show that, when doing contact tracing, it is not worth allowing voluntary tests or implementing additional random screenings: the additional tests will only create further congestion in the system and, since the capacity constraint will be binding more often, the overall percentage o f discovered infected will not change much.

Overall, my simulations show how properly designed testing policies, that focus in particular on increasing the number of screened individuals, can significantly limit the impact of an epidemic on output and mortality.

The rest of the paper is organized as follows. Section 2 discusses the related literature. Section 3 describes the model and explains the calibration and simulation. Section 4 summarizes the simulation results. Section 5 illustrates the gains from increasing the testing capacity through either investments, group testing or less sensitive tests. Section 6 discusses the results robustness to changes in the design of the contact tracing policy. Section 7 discusses the

simulation results in case of an endogenous response to the epidemic. Section 8 summarizes the robustness of the results to several alternative model assumptions and parametrizations. Section 9 concludes.

2 Related Literature

The idea of using group testing to increase lab capacity is originally from Dorfman (1943), who first computed the implied expected reduction in the number of tests needed to screen a population. A huge literature then followed. Alridge, Johnson and Scarlett (2019) survey the most important contributions. In the context of Covid-19, Gollier and Gossner (2020) study group testing applications targeted at estimating the virus prevalence, at allowing negatives to return to work and at screening. The idea of using tests of lower quality was instead put forward by Larremore et al. (2020) and Gans (2020). The former focuses on the lower costs and on the shorter processing time, which allows a more frequent use and, therefore, a more prompt isolation of the infectious. The latter highlights less sensitive tests as superior to test for infectiousness. With respect to those works, my contribution consists in nesting group testing and lower quality tests within a SIR-Macro model where discovered infected individuals are quarantined, in order to study their effect on disease transmission and on effects of the epidemic.

Bergstrom, Bergstrom and Li (2020), also study the effect of alternative testing strategies on the transmission of a disease, focusing on the relationship between testing intervals and infectiousness and on the tests characteristics. I complement their analysis by showing the gains from increasing the testing capacity within a SIR-Macro model. Kasy and Teytelboym (2020) propose testing policies based on the likelihood of being infectious, while Deb, Pai, Vohra and Vohra (2020) propose a methodology to target tests based on occupation and wages in addition to the individual risk of being infectious. I abstract from such allocation mechanisms in my analysis. In a related contribution, Galeotti, Steiner and Surico (2020) develop a method to evaluate alternative tests on the basis of the policy objective. In my analysis, I simply assume the tests characteristics as given.

More generally, my work is related to the epidemiological literature on SIR models (Ker-

mack and McKendrick 1954; Allen 2017), and to the recent, and fast-developing, economic literature that merges an epidemiological structure into small and medium scale economic models, to study the effects of epidemics, and of the policies to to cope with them, on the economy. Examples of these SIR-Macro models include, among others, Eichenbaum, Rebelo and Trabandt (2020a), Atkeson (2020), Farboodi, Jarosch and Shimer (2020), Collard et al. (2020), Garibaldi, Moen and Pissarides (2020), Glover, Heathcote, Krueger and Rios-Rull (2020), Krueger, Uhlig and Xie (2020) and Jones, Philippon and Venkateswaran (2020). Within this SIR-Macro literature, there are few works that study the effects of Testing, Tracing and Quarantine (TTQ) policies. Piguillem and Shi (2020), Berger, Herkenhoff, and Mongey (2020) and Eichenbaum, Rebelo and Trabandt (2020b) find that such policies are more effective and less costly than lockdowns at reducing the overall cost of an epidemic. In a companion paper (Russo 2020), I also use a SIR-Macro model to compare testing policies to lockdowns, finding similar conclusions. I complement all these analysis with a more detailed model of alternative testing policies, which is indeed more stylized in previous contributions. In particular, Piguillem and Shi (2020)and Berger, Herkenhoff, and Mongey (2020) model random tests, while in Eichenbaum, Rebelo and Trabandt (2020b) only a fixed subset of the population is frequently tested. In Russo (2020) the assumption is instead that a fixed fraction of the infected is discovered each period. Alvarez, Argente and Lippi (2020) come to an opposite conclusion, namely that testing policies cannot be considered as a substitute of a lockdown but, rather, as its complement. A similar argument appears in Dewatripont et al. (2020).

3 The Model

I build a stochastic Susceptible-Infected-Recovered (SIR) model (Kermack and McKendrick 1954) with production to study the effect of alternative testing policies on the progression of the epidemic and on its consequences. There are four new features to the model that distinguish it from previous contributions: tests taken by the agents on a voluntary basis, conditional on their preferences and on the onset or not of symptoms; the possibility of developing symptoms independently from the infectious disease; a testing capacity constraint, that limits the number of tests that can be performed each period; a detailed model of the

contact tracing technology based on the structure of the economy. I summarize the model set-up in subsection 3.1 and the details of the testing technology subsection 3.2, while in subsection 3.3 I illustrate the production function and in subsection 3.4 the model dynamics. In subsection 3.5 I explain instead the calibration and the details of the simulations. The model is similar to the SIR-Macro model that I developed in a companion paper (Russo 2020) to study lockdowns, although the production side of the economy is more stylized here, while the testing technology is much more detailed.

3.1 Agents, Firms, Schools and Diseases

A fictional country without population growth is composed by N agents, each represented by a vector $x_{ij,t}$, where i indexes the family to which the agent belongs, j the school or university she attends or her workplace, and where t indexes the time period. At some point, this country is hit by an infectious disease. Each vector x is composed of five binary elements that summarize the individual status with respect to the infectious disease. In particular, $x_{ij,t} = \{f_{ij,t}, a_{ij,t}, s_{ij,t}, u_{ij,t}, d_{ij,t}\}$, where $f_{ij,t} = 1$ in case of infection with symptoms, $a_{ij,t} = 1$ in case of infection without symptoms, $s_{ij,t} = 1$ in case of susceptibility, to infection, $u_{ij,t} = 1$ in case of immunity, upon recovery, and $d_{ij,t} = 1$ in case of death caused by the disease. As in Eichenbaum, Rebelo and Trabandt (2020a and 2020b), I assume that the agents do not know their infection status, which means that tests are valuable. Moreover, I also assume that it is possible to develop the symptoms independently from this infectious disease, so the symptoms are not informative about the infection status, thereby making the tests valuable also for symptomatics. For instance, influenza can induce symptoms which are difficult to distinguish from those induced by Covid-19.

There is a total of I families of different sizes, including the possibility of size one to represent singles. Each family is represented as a set n_h, with $h \in \{1, 2, \dots I\}$. Agents are assigned to families by the function $N(i) = n_h$ if $i \in n_h$. There is also a total of J^f firms and a total of J^s schools, with $J = J^f + J^s$. Each school/university or firm is represented by a set in $\{m_1, m_2 \dots m_J\}$. The cardinality of each set is either the number of workers employed, if j represents a firm, $M_j^f = |m_j|$, $j \in \{1 \dots J^f\}$, or the number of students, if j is a school or university, $M_j^s = |m_j|$, $j \in \{J^f + 1 \dots J^m\}$. Self-employed individuals are represented as

workers in size 1 firms, while I assume that $j = 0$ in case the agent is either unemployed, a pensioner, NEET or if she is simply out of the labor force. For simplicity, I avoid a further division of schools and universities into classes[1]. Workers are randomly assigned to firms and students are randomly assigned to schools, and it is possible, for members of the same family, to be in the same school or firm.

All agents engage in social activities, for instance because they use public transportation, because they join a line outside a movie theater of because they dine with friends at a local restaurant. I model social activities as random matches with agents outside the family, workplace or school. In the baseline simulation, I assume that all agents are matched with a time-invariant fraction of the population regardless of the progression of the epidemic. In other words, they do not voluntarily react to the epidemic by reducing their social activities to decrease the risk of getting the infection, not even close to the peak. This is the case, for instance, if they have a very low degree of risk aversion and if the utility that they derive from social activities is big enough to overcome the expected cost of an infection, including the possibility of dying as a consequence. I make this assumption to propose a starker comparison between the testing policies, independently from the agents' endogenous reaction. In practice, it is very likely for the agents to respond to the epidemic, perhaps also entering in a voluntary lockdown in case of a high observed disease prevalence. I discuss the robustness of the results to this endogenous reactions in section 7.

An infection, for a susceptible agent, is the result of a close contact with one of the infected. The infection probability from a matching with a symptomatic is π, while it is $\bar{\pi} < \pi$ in case of a matching with an asymptomatic. In both cases, the probabilities are independent from the history of matches. I also assume that, upon infection, all agents are asymptomatic. The asymptomatics develop symptoms with fixed probability ρ per period. The probability of developing symptoms independently from the infection is instead κ. Symptoms can be mild or severe, both for the infected individuals and for the non-infected that developed symptoms independently. I assume that symptoms worsen with probability β regardless of the underlying infection. I refer to β as the probability of hospitalization, and I assume, for

[1] Allowing a division of schools and universities into classes results in a smaller probability of getting the infection at school and it would increase the importance of social activities and family relationships as a source of infections. For fixed model parameters, it will also result in a slower progression of the epidemic.

model tractability, that it is always possible for symptoms to worsen independently from the previous history of symptoms severity. Both symptomatics and asymptomatics recover with exogenous probability γ, ignoring that it is, most likely, easier to recover in case of absence of symptoms. Symptomatic agents die with exogenous probability δ. Differently from Favero, Ichino and Rustichini (2020) and Eichenbaum, Rebelo and Trabandt (2020a), among others, I do not model a health system capacity constraint: the death probability does not increase if the number of symptomatics is such that it is impossible to properly treat all of them. As a result, the overall mortality due an uncontrolled epidemic is lower, in the simulation, than its value in case of a health capacity constraint, with the consequence that the gains from mitigation policies that smooth the infection peak are underestimated in my analysis.

3.2 Testing Policies

I build three alternative testing policies in the model. Their goal is to identify and quarantine the infectious, in order to stem the diffusion of the contagious disease. The first consists in allowing voluntary tests only. The second entails frequent random screenings. The third is instead a contact tracing program, whose goal is to identify all recent contacts of each discovered infected agent. In all cases, I assume that hospitalized agents, with severe symptoms, are always tested, because treatment is contingent on the infection status. Differently from Kasy and Teytelboym (2020) and Deb, Pai, Vohra and Vohra (2020), I do not explicitly consider a test allocation mechanism based on individual characteristics other than the presence of severe symptoms, although contact tracing is akin to targeting according to the individual risk of infection.

I denote with $\varepsilon_{i,t} \sim Ber(\kappa)$ a binary variable equal to one in case of the onset of symptoms which are not due to infectious disease, and with $\zeta_{i,t} \sim Ber(\beta)$ a binary variable equal to one in case the symptoms worsen, regardless of their underlying cause. Since $f_{ij,t}$ is equal to one in case of an infection with symptoms, I define[2] with $T^h_{ij,t} = \zeta_{i,t} \max\{f_{ij,t}, \varepsilon_{ij,t}\}$ the binary variable that describes the hospitalization status of each agent and, as a consequence, the need to take a test to target a medical treatment. In this formulation, I assume that the

[2]I assume that it is possible to develop the same symptoms, in a given period, both because of the disease and because of the concurrent infection, and that, in such cases, the severity of the symptoms does not change.

agents that were infected, and that knew about the infection, still have to take a test upon hospitalization, for instance to check if they effectively recovered or not.

Contact tracing consists instead of mandatory tests for all of the identified exposed agents who were matched with each previously discovered infected. It is relatively easy to identify the contacts at home, at work or in school, but it is indeed quite complicated to identify all contacts that result from social activities, at least if privacy-invading tracing technologies such as the ones that use cell phone position and security camera images are either not available or not legally allowed. I realistically assume that only a fraction $0 \leq \phi \leq 1$ of the contacts from social activities is correctly identified. This fraction ϕ, in turn, is a function both of the efficacy of the allowed tracing technologies and of the laws implemented to stem the epidemic, such as, among others, the introduction of mandatory daily registers of bar and restaurant customers. I use the binary variable $z_{ij,t}$ to indicate if agent i is infected and correctly identified as infected in period t (see infra for a complete description). Suppose that $P_{ij,t}$ is the set that includes the individual indexes i of all agents that matched with agent i as a result of social activities in t, and suppose also that the set $P^g_{ij,t} \subseteq P_{ij,t}$ with $|P^g_{ij,t}| = \phi |P_{ij,t}|$ includes instead only the indexes of the correctly identified matches, conditional on the tracing efficacy ϕ. Then an agent i, is required to take a test, as part of contact tracing, if the binary variable $T^g_{ij,t}$ is equal to one, with

$$T^g_{ij,t} = \left(1 - z_{ij,t-1}\right) \mathbb{1}_{[i \in (I_{t-1} \cup P^g_{ij,t-1}) \vee j \in J_{t-1}]} \tag{1}$$

where $I_{t-1} = \{i : z_{ij,t-1} = 1\}$ is the set that includes all family identifiers of the discovered infected agents in $t-1$ and where $J_{t-1} = \{j : z_{ij,t-1} = 1\}$ is the set that includes all firm and school identifiers of the discovered infected agents in $t-1$. This specification implies that immunes are also tested as part of contact tracing. In fact both the asymptomatics and the symptomatics that recover do not actually know that they are immune if they never took a test while infected. Moreover, those who recovered and who took the test while infected, could be called to prove that they recovered fully with an additional test. Alternatively, the health authorities might be uncertain about the possibility of second infections. The main consequence of this assumption is a worse performance of contact tracing, since testing the previously discovered infected contributes to congestion in testing: if they could simply

exhibit a certificate that proves that they were infected, there would be more room to process additional tests each period and a bigger fraction of discovered infected agents. In section 6, I analyze the performance of contact tracing assuming that the previously discovered infected agents are not tested, which turns out to be only mildly better.

A further policy alternative consists in allowing voluntary tests. I assume that tests are costly, both because they cause a physical and mental distress, and as a consequence of their monetary cost. However all agents value the knowledge of their infection status, for instance because they want to protect their family members and co-workers from the risk of infection. I assume that the difference between the benefits and the costs of a test are idiosyncratic, but contingent on the evolution of the epidemic. More specifically, I assume that the higher is the observed disease prevalence, the more the agents are concerned about their infection status and, therefore, the more often they want to take the test. Thus voluntary testing induces an endogenous mitigation mechanism, because, as the observed prevalence increases, it also increases the number of tests and, therefore, the number of quarantined infectious, thereby reducing contagion opportunities, although the testing capacity constraint might impair this mechanism. I also assume that the agents are more willing to take a test in case they have symptoms, independently from the observed prevalence but, since their symptoms might not be the result of the infectious disease, not all symptomatics want to get tested. I denote with $\hat{q}_{ij}$ the threshold value such that the asymptomatic individual i takes a test if the observed prevalence exceeds it. I assume that these thresholds are uniformly distributed between zero and an upper bound, $\hat{q}_{ij} \sim U[0, Q]$, where Q is the observed disease prevalence big enough to persuade all asymptomatics to take a test. I also denote with $\bar{q}_{ij}$ a similar threshold value but for the symptomatics, and I assume that $\bar{q}_{ij} \sim U[0, \hat{q}_{ij}]$. In other words, it takes a smaller observed disease prevalence to persuade an agent to get tested when she has symptoms. This assumption further reinforces the endogenous mitigation of voluntary tests. I denote with $T^v_{ij,t}$ a binary variable equal to one in case agent i voluntarily decides to take a test, with

$$T^v_{ij,t} = \mathbb{1}_{[\hat{I}^D_{t-1} \geq \bar{q}_{ij}]}(1 - \zeta_{i,t}) \max\{f_{ij,t}, \varepsilon_{ij,t}\} + \mathbb{1}_{[\hat{I}^D_{t-1} \geq \hat{q}_{ij}]}(s_{ij,t} + a_{ij,t} + u_{ij,t})(1 - \varepsilon_{ij,t}) \qquad (2)$$

where $\hat{I}^D_{t-1} = I^D_{t-1}/N$ is the observed disease prevalence in $t-1$, equal to the number

of discovered infected agents I^D_{t-1} divided by the population size. As for contact tracing, I assume that the agents behave as second infection were possible, choosing, in particular, to get tested even if they were discovered as infected in some previous period. This actually implies that a subset of the population of individuals who attach a very high value to the knowledge of the infection status (very small $\bar{q}_{ij}$ and $\hat{q}_{ij}$) is almost always tested over the course of the epidemic, which is akin to what Eichenbaum, Rebelo and Trabandt (2020b) assume. In section 8, I test the robustness of the results to an alternative assumption according to which previously discovered infected agents never take tests on a voluntary basis, thereby creating less congestion in testing, showing that the results does not change much.

The last policy alternative consists in implementing random screenings. I assume that the probability to be selected in this random sample ξ is independent from the previous history of tests. One consequence of this assumption is that it possible, albeit unlikely, to be selected several times for random testing while immune upon recovery. This is the case, for instance, if the authorities are not sure about the possibility of second infections or, alternatively, if the agency that selects the individuals for the random screening does not also access the medical records that keep track of the infection status, or if those records either do not exist or are not updated. The alternative assumption would be to select, for random testing, only the agents that have never been discovered as infected, which might be operationally very difficult to do for a centralized agency, especially at the daily or even weekly frequency, which is the reason why I abstract from this possibility. I denote with $T^r_{ij,t} \sim Ber(\xi)$ the binary variable equal to one if the agent is selected for random screenings, which is a draw from a Bernoulli probability distribution with parameter ξ. I assume that such screenings are mandatory, although compliance with such programs could be low, as illustrated by the case of Italy in 2020 in between two Covid-19 waves.

I do not explicitly model a cost for the test, differently from Piguillem and Shi (2020), but I assume that there is a testing capacity constraint. This constraint can be thought either as of a physical constraint, in the sense that it is impossible to do more tests because of the lack of machines and/or professionals to operate them, or as an economic constraint, in the sense that the health system does not have the resources to pay the professionals or to operate the equipment. In case the total number of tests to process Υ_t, at time t, exceeds the

constraint L, I assume that there is random rationing. Rationed tests are moved to the next period, and I assume that there is no mandatory quarantine for the agents waiting to take a test of waiting for a result. With limited testing capacity, this would be in fact equivalent to a confinement of potentially non-infectious individuals as if there was a selective lockdown, which will actually be beneficial in terms of mitigation, but which will move the analysis away from its main goal of evaluating alternative testing policies, and not to compare them to lockdowns. Furthermore, it might be also difficult to legally enforce a quarantine without proof of infectiousness, especially because it would only be the byproduct of a health system inefficiency. I also assume that there is no queue: the laboratories do not process leftover tests before others. This assumption is primary motivated by tractability, as it makes it unnecessary to define, and keep track of, a testing priority order. However such an order might exists, for instance because hospitalized patients are typically granted priority, since treatment, and perhaps also their assignment to a hospital, is crucially contingent on the test results.

The binary variable $\tau_{ij,t}$, which is equal to one in case of a test, either upon hospitalization, as part of contact tracing, as a result of a voluntary choice, as part of random screenings, or because of random rationing in the previous period, is therefore equal to:

$$\tau_{ij,t} = \alpha_{ij,t} \max\{T^v_{ij,t}\,;\, T^g_{ij,t}\,;\, T^l_{ij,t}\,;\, T^r_{ij,t}\,;\, T^h_{ij,t}\} \tag{3}$$

where $\alpha_{ij,t}$ is the idiosyncratic probability to be tested (non-rationed) in t,

$$\alpha_{ij,t} = \begin{cases} 1 & \text{if} \quad \Upsilon_t \leq L \\ \lambda_{ij,t} & \textit{if} \quad \Upsilon_t > L \end{cases} \tag{4}$$

and where $\lambda_{ij,t} \sim Ber(L/\Upsilon_t)$ is a draw from a Bernoulli distribution with probability equal to the ratio of the maximum number of tests that can be processed in a period L (maximum testing capacity) to the total number of tests to perform Υ_t:

$$\Upsilon_t = \sum_{i=1}^{N} \max\{T^v_{ij,t}\,;\, T^g_{ij,t}\,;\, T^l_{ij,t}\,;\, T^r_{ij,t}\,;\, T^h_{ij,t}\} \tag{5}$$

For instance, if the agents require, or are required, to take twice as many tests as the system can process, each one of them has only a 50% chance of getting a test. $T^l_{ij,t}$ is the

binary variable equal to one in case agent i was rationed in the previous period:

$$T^l_{ij,t} = (1 - \alpha_{ij,t-1}) \max\{T^v_{ij,t-1}, T^g_{ij,t-1}, T^l_{ij,t-1}, T^h_{ij,t-1}\} \tag{6}$$

I assume that the available tests deliver a positive result, for an infected, with probability $\sigma \leq 1$. Given this sensitivity, an agent i is identified as infected in t if $z_{ij,t} = \tau_{ij,t}\, S\,(f_{ij,t}+a_{ij,t}) = 1$, where $S \sim Ber(\sigma)$ is a random draw from a Bernoulli distribution with parameter σ. I assume that the tests have perfect specificity: the test result is always negative if there is no infection. Allowing the use of tests with lower specificity results in unnecessary quarantines, that are desirable for a mitigation perspective, albeit at a high individual cost, but which also makes it more difficult, for the agents, to accept to be tested.

Discovered infected agents are quarantined: they are not allowed to go to work or school and they are not allowed to engage in social activities, although they still keep their relationships with the family members[3]. Actually a quarantine could also entail isolation from the family, say in a hotel room or in a dedicated facility (as in Josè Saramago's novel Blindness), but I abstract from this, rather extreme, possibility.

3.3 Production

The production side of the economy is very stylized. I assume that there is no capital, and that the technology does not change over time, so production depends only on employment. I assume that the infected agents are quarantined only in case they tested positive, in which case they cannot work. All others, including the infected with symptoms and the symptomatic, but non-infected, go to work, unless in case of severe symptoms that require hospitalization. I assume that the firms cannot temporarily replace the quarantined workers with the unemployed and, since there is no population growth, it is also impossible for them to hire new workers in case of deaths. The number of workers employed a firm j, at any point in time, is therefore equal to:

[3]The family members of the discovered infected agents are required to take a test, as part of contact tracing, only right after the discovery of the infection in the family. Infections that take place in subsequent periods, as a result of repeated family contacts that are not discontinued, can be therefore overlooked. There is therefore a limit to the percentage of the infected that can be discovered through contact tracing. The alternative would be to frequently test the family members of an infected, but this will also impair the performance of contact tracing by creating congestion.

$$\hat{m}_{j,t} = |m_j| - \sum_{j \in m_j} \{d_{ij,t} + \max[(1 - Z_{ij,t})(f_{ij,t} + a_{ij,t}) \, , \, T^h_{ij,t}]\} \tag{7}$$

where $Z_{ij,t}$ is equal to one if the agent i was never discovered as infected before period t, with $Z_{ij,t} = \prod_{\bar{t}=1}^{t}(1 - z_{ij,\bar{t}})$. Production, for each firm, is equal to $y_{j,t} = \hat{m}^{\alpha}_{j,t}$, while aggregate production is simply obtained summing the individual firms' output: $Y_t = \sum_{j=1}^{J^f} y_{j,t}$. The output gap, in this economy, is the percentage difference between aggregate production and potential production, the latter defined as what would be produced without the epidemic and without any implemented policy, so without deaths, quarantines and hospitalizations. It follows that the testing policies have two contrasting effects on the output gap: the bigger is the number of discovered infected agents, the bigger is the output reduction due to the quarantines; more discovered infected, however, reduce infections and the overall mortality rate, with a lower long-run output reduction.

3.4 Dynamics

The model features five dynamic equations for changes of status with respect to the disease. An agent is infected and symptomatic, in a given period, if she was infected and symptomatic in the previous period and if she did not recover or die, or if it was asymptomatic in the previous period and developed symptoms:

$$f_{ij,t+1} = (1 - \Delta)(1 - \Gamma)f_{ij,t} + P(1 - \Gamma)a_{ij,t} \tag{8}$$

where $\Delta \sim Ber(\delta)$, $\Gamma \sim Ber(\gamma)$ and $P \sim Ber(\rho)$ is a Bernoulli random variable with probability equal to the probability of developing symptoms ρ. An agent is asymptomatic if it was asymptomatic in the previous period and did not recover and did not develop symptoms, or if it was susceptible and infected:

$$a_{ij,t+1} = (1 - \Gamma)(1 - P)a_{ij,t} + H_{ij,t}s_{ij,t} \tag{9}$$

where the binary variable $H_{ij\,t}$ equal to one in case of infection:

$$H_{ij,t} = \begin{cases} 1 & \text{prob} \quad 1-(1-\pi)^{\bar{F}_{i,t}}(1-\bar{\pi})^{\bar{A}_{ij,t}} \\ 0 & \text{otherwise} \end{cases} \tag{10}$$

$\bar{F}_{i,t}$ is the number of symptomatics with whom each susceptible is matched:

$$\bar{F}_{i,t} = \sum_{\{\tilde{i}\in N(i)\,;\,\tilde{i}\neq i\}} f_{\tilde{i}j,t} \quad + \sum_{\{\tilde{j}\in m_j\,;\,\tilde{j}\neq j\}} f_{i\tilde{j},t} Z_{i\tilde{j},t} \quad + \quad \eta^f_{ij,t} \sum_{\{\tilde{i}\notin N(i)\,;\,\tilde{j}\notin m_j\}} f_{\tilde{i}\tilde{j},t} Z_{\tilde{i}\tilde{j},t} \tag{11}$$

where the first term is the number of symptomatics in her family, the second is the number of never-discovered symptomatics in her school or workplace and the third is a fraction $\eta^f_{ij,t}$ of the never-discovered symptomatics met as part of social activities. The number of asymptomatics with whom each susceptible is matched is instead:

$$\bar{A}_{i,t} = \sum_{\{\tilde{i}\in N(i)\,;\,\tilde{i}\neq i\}} a_{\tilde{i}j,t} \quad + \sum_{\{\tilde{j}\in m_j\,;\,\tilde{j}\neq j\}} a_{i\tilde{j},t} Z_{i\tilde{j},t} \quad + \quad \eta^a_{ij,t} \sum_{\{\tilde{j}\notin m_j\,;\,\tilde{i}\notin N(i)\}} a_{\tilde{i}\tilde{j},t} Z_{\tilde{i}\tilde{j},t} \tag{12}$$

where, again, the first term is for asymptomatics in the family, the second for never-discovered asymptomatics on the workplace or in school, and third is a fraction $\eta^a_{ij,t}$ of the never-discovered asymptomatics met in social activities. Since I abstract from the possibility of second infection, I have that an agent is immune if it was immune in the previous period or if she recovered from an infection:

$$u_{ij,t+1} = \Gamma f_{ij,t} + \Gamma a_{ij,t} + u_{ij,t} \tag{13}$$

An agent is instead susceptible if it was susceptible in the previous period and if she was not infected:

$$s_{ij,t+1} = s_{ij,t}(1 - H_{ij\,t}) \tag{14}$$

Finally, infection can result in a death:

$$d_{ij,t+1} = \Delta_t(1-\Gamma) f_{ij,t} + d_{ij,t} \tag{15}$$

3.5 Parameters and Calibration

I calibrate the model to Italy for a generic infectious disease, but that partially resembles what it is currently known about Covid-19. The calibration is also similar to the one that I propose in a companion paper (Russo 2020) that I use to study the effects of lockdowns. Table 1 summarizes the model parameters used in the baseline simulation. I simulate an economy with $I = 500$ families, whose size is set according to the Italian Institute of Statistics (ISTAT) data for 2019: one member for 31% of the families, two members for 27% of them, three for 20%, four for 16% and five or more for 6%. Given the relatively small simulation, I exclude the possibility of more than 5 members. The resulting number of agents N is close to 1200, and the average family size $\bar{n} = N/I \approx 2.5$.

I set the labor force participation such that $W = L/N = 0.6$, consistently with ISTAT. Since, according to ISTAT, 44% of the workers are employed in small firms, I set the number of firms with one employee so that they account for 44% of the stock of all firms. I calibrate the maximum firm size J^{max} and the number of firms J^f, to match the ISTAT average of 3.87 employees per firm, and the 56% fraction of workers employed in non-small firms (of size bigger than one in the simulated economy). The resulting values are $J^{max} = 12$ and $J = 64$. I set the parameter α of the production function to one to have constant returns to scale. I assume that only the agents living in families of three or more can attend schools or university, thereby abstracting from the possibility of single parents. To match ISTAT data, I fix the fraction of those agents to 38.5%. I then calibrate the number of schools J^s so that each school, of equal size, accounts for 2.2% of all school age pupils, once again consistently with ISTAT data. Allowing for schools of different size will only introduce more noise in the simulation results, since the diffusion of the disease will also depend on school size (higher size implies a faster diffusion), but I abstract from this feature.

I assume that the number of random matches from social activities is the same, and time invariant, for all agents and equal to η, and I calibrate this value in order to have an average of 18.5 total matches per individual, before the epidemic, consistently with the estimates by the Istituto Superiore di Sanità (ISS) (2020). These total number of matches TM_{ij} for an individual i, living in family n_i and going to work or school to m_j is equal to $TM_{ij} = |n_i| - 1 + |m_j| - \mathbb{1}_{[|m_j|>0]} + \eta N$, where the indicator function in the middle term

accounts for the possibility of agents not working and not going to school (NEET, pensioners, housewifes etc.). Solving $(1/N)\sum_{i=1}^{N} TM_{ij} = 18.5$, I obtain, for the simulated population size, $\eta = 0.011$. Given this value of η, I draw randomly, each period, the number of symptomatics, asymptomatics, susceptibles, and immunes with whom each agent is matched in social activities. In particular, I extract, for each agent, a random subset of ηN elements from the subset $O_{ij,t}$ of the population of individuals who are not in the agent i family, workplace or school and who are not infected and discovered as such in period t. Then I count the number of symptomatics in this set and, dividing it by the total number of symptomatics in $O_{ij,t}$, I obtain the share $\eta^f_{ij,t}$.I proceed similarly for asymptomatics, immunes and susceptibles.

I set the death probability to the rather high value of $\delta = 0.5\%$ per period, but multiplied by 2, consistently with the assumption in Eichenbaum, Rebelo and Trabandt (2020a), for a subset of the population with higher individual health risk. I set this fraction to 17%, in line with the percentage of the population above 70 in Italy. The probability to develop symptoms in case of infection is $\rho = 0.25$, while the probability to recover, regardless of the onset of symptoms, is $\gamma = 0.1$. Without an infection, and independently from it, there is a probability $\kappa = 0.07$ to develop similar symptoms to the infectious diseases, where 7% is the influenza prevalence registered in Italy in 2019 by the ISS. Thus, in the Covid-19 example, the baseline model simulation refers to the Influenza season, but I will also consider an alternative, summer-like, scenario with a very small incidence of symptoms without infection (see section 8). An Hospitalization is required, on average, for 5% of the agents with symptoms, independently from the infection status, thus $\beta = 0.05$. I set the transmission probability from an asymptomatic to $\bar{\pi} = 1\%$ and I assume that it is twice more likely to get the infection from a symptomatic than from an asymptomatic, so $\pi = 2\%$. Assuming that the period, in the model, is equivalent to one day, the implied reproduction number is $R = 18.5 \cdot 7 \cdot 0.0125 = 1.618$ per week, in line with the estimates by Chowdhury et al. (2020) for Covid-19.

The tests available have perfect sensitivity, $\sigma = 1$, which is nearly true for RT-PCR (Real Time Polymerase Chain Reaction) tests on naso-pharyngeal swabs in the case of Covid-19, at least for non-negligible viral loads. I start from this baseline to then study the effect of the introduction of tests of lower sensitivity in section 5.3. There is not much information available on the preferences over tests. I calibrate the upper bound Q to the uniform distribution of

threshold values for the asymptomatics so that 50% of the symptomatics (recall that the threshold value for a symptomatic is a draw from a uniform distribution between 0 and the threshold value in case she develops symptoms) voluntary decide to take a test when the observed disease prevalence is equal to 10%. The resulting value is $Q = 0.52$. The smaller is Q, the more effective is voluntary testing at fighting the epidemic because of the bigger number of individuals who choose to get a test for given observed prevalence. Since I set this parameter rather arbitrarily, I analyze the robustness of the results in section 8, allowing, in particular, for a much smaller Q. I assume a 50% efficacy of the tracing technology in the baseline simulation, so $\phi = 0.5$, which means that only half of the random matches from social activities are correctly identified for each individual. I study extensively the effects of policies that entail a change of this fraction in section 6.

I set the baseline total testing capacity to $L/N = 0.05$, which is a rather high number. In a country with 50 millions inhabitants, this entails processing 2.5 millions tests per period. The reason is that, in my rather small simulated economy, all agents are closely connected. In a such a dense economy, the number of tests required for contact tracing grows very rapidly with infections, making the capacity constraint binding very often. The main task of the simulation exercises will be to study effects of the epidemic, conditional on the testing policies, for different levels of this maximum capacity. From this analysis, summarized in section 5, it will also be clear that smaller testing capacity constraints will deliver very similar results to my benchmark.

I fix the average fraction of the population tested each period in case of random screenings ξ at half of the maximum testing testing capacity. This fraction is smaller than what it is possible to process in a period for two reasons: to leave room to promptly process tests in hospitals, which might be substantial near the peak of the epidemic, and because compliance with such screening programs is typically low, while the enforcement of mandatory screenings that could be even more complicated than the enforcement of a lockdown. Contact tracing, conversely, is much easier to enforce because the traced agents have a higher risk of being infected, typically resulting in a bigger compliance. That said, I will also analyze the possibility of mandatory random screenings programs that almost exhaust the testing capacity (see section 5.1).

Table 1: **Prameters**

Parameter	Description	Value
π	Infection probability, symptomatics	0.02
$\bar{\pi}$	Infection probability, asymptomatics	0.01
ρ	Probability of symptoms if infected	0.25
κ	Probability of symptoms if not infected	0.07
β	Hospitalization probability if symptoms	0.05
γ	Recovery probability	0.1
δ	Death probability	0.005
L/N	Maximum testing capacity	0.05
σ	Test sensitivity	1
ϕ	Percentage of traced social contacts	0.5
Q	Preference for voluntary tests	0.52
ξ	Random Screenings	0.025
$\bar{n}$	Average family size	2.5
η	Density of the economy	0.011
W	Labor force participation	0.6
α	Technology	1

4 Epidemics and Testing

I simulate 500 epidemics conditional on four different testing policies. The first involves testing the hospitalized agents only, for clinical, rather than screening, purposes. This is the no-policy benchmark against which I will evaluate the performance of all other policies. The second policy consists in allowing voluntary tests alongside tests in hospitals. The third policy consists in the implementation of contact tracing, in addition to tests in hospitals, and in a ban on voluntary tests. The fourth policy consists instead of mandatory random screenings for a fixed fraction of the population in addition to tests in hospitals, coupled again with a ban on voluntary tests. In principle, the optimal testing policy could also consists in a mixture of the last three alternatives. For simplicity of exposition and analysis, I decided to first focus on each alternative in isolation, to then comment on the possibility of mixtures (see section 6). The epidemics start with the random infection of four agents. The testing policies start[4] ten periods after the first infections. Before that, only hospitalized agents are tested. I discard

[4]The starting point of the polices influences their performance. Implementing contact tracing soon enough might effectively stem the epidemic before it fully develops, rapidly bringing contagions to zero. Viceversa, when done too late, it might be ineffective. For the baseline parameters, the epidemics last, on average, for 120 periods, so 10 periods seems a sensible choice, assuming that the authorities need time to organize. I obtained similar results in a small neighborhood of 10, between 8 and 12 periods since first infections.

the simulations without an epidemic, which is a rather unlikely, albeit possible, event, that can be observed in case the initially infected individuals recover before spreading the disease.

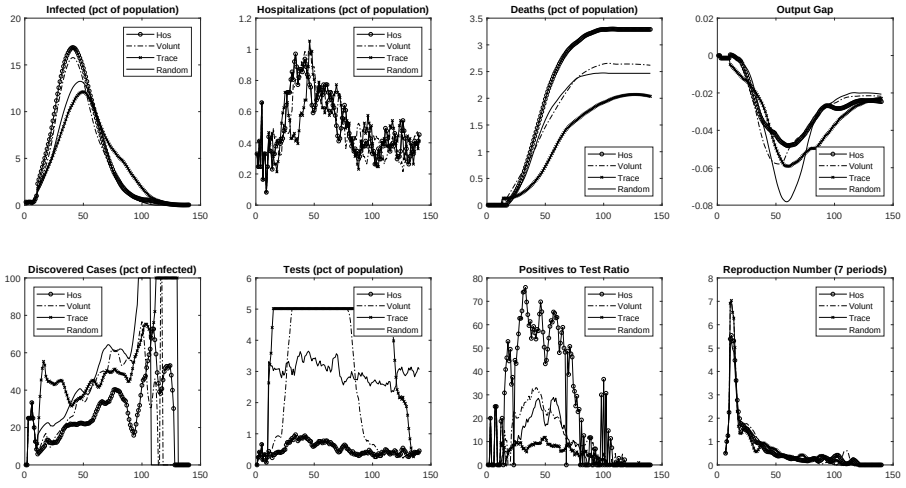
Figure 1 shows one of the possible dynamic paths[5] of the epidemic conditional on the different testing policies. Increasing the number of tests with respect to benchmark, either through voluntary tests, random screenings or contact tracing, increases the number of discovered and quarantined infected, thereby reducing further infection and slowing down the diffusion of the disease. The percentage of discovered infected is small at the beginning of the epidemic, but it increases in time, converging to 100% towards the end, when only few agents are infected. Contact tracing is the policy that allows the identification of the biggest percentage of infected, since it involves testing agents with a higher probability of being infected than the median agent; testing only the hospitalized yields instead worse results, since only a fraction of the symptomatics with severe symptoms is tested. Random tests and voluntary tests yield instead intermediate results. The peak of infections is smoothed for all testing policies with respect to benchmark, and smallest in case of contact tracing. The duration of the epidemic, measured as the time necessary to reach zero infections, is instead roughly the same for the three alternatives. New accesses to hospitals tend to grow with infections, but they are very volatile because of the exogenous health shock independent from the epidemic. They are also bigger, close to the peak of the epidemic, in case of tests to hospitalized agents only. As a consequence of the biggest percentage of identified and quarantined infected, and of the smaller peak of infections, the mortality rate is smallest under contact tracing. The output decreases, as the epidemic progresses, because of the individual quarantines and because of the deaths. When the epidemic ends, the output level is highest for the testing policy that yields the minimum number of deaths. The main problem of contact tracing, but also of voluntary tests, is that the capacity constraint becomes binding almost as soon as the policies are implemented. In the other words, there is not enough testing capacity to make both policies work properly, with the consequence of an insufficient percentage of quarantined infectious. This motivates the main task of the paper, namely to understand how to increase the testing capacity. The positives to test ratio is instead biggest in case of tests to hospitalized agents only, since few tests are administered, and only to individuals with a high probability of infection, while it is

[5]All pictures are smoothed with 5-periods moving averages.

smallest in case of contact tracing, since only a fraction of the traced agents got the infection, due to the relatively small infection probability. Voluntary tests and random tests deliver instead intermediate positives to tests ratios.

Importantly, figure 1 depicts only one of the possible epidemic paths. It is also possible, for contact tracing, to reduce infections at zero soon after its introduction, with a consequent reduction of mortality to almost zero. This is the case, for instance, if a great portion of the initial infections takes place in the family, workplace or at school, rather than as a the result of other social activities, and if there are not many hospitalized agents without the infection, which eases the pressure on the labs. Viceversa, if many of the initial infections resulted from social activities, that are correctly traced only in 50% of the cases, contact tracing might not be as effective at reducing infections. If, in addition, infections grow at a fast pace at early stages, there is also a higher probability of lab congestion, making tracing more difficult, and so on. In this case, contact tracing can even result in a worse performance as compared to voluntary or random tests.

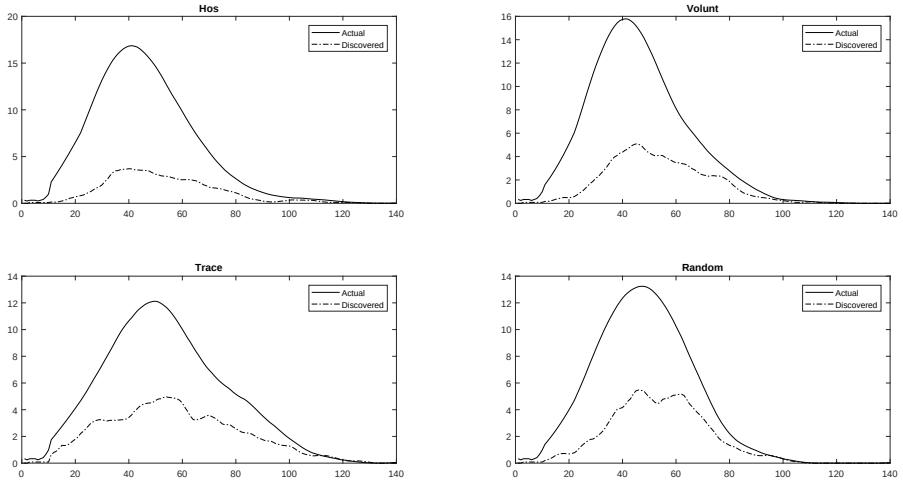
Figure 1: **Epidemic Evolution**



Notes: Clockwise from the upper-left panel: Actual prevalence (infected agents as a percentage of total population), new hospitalizations as a percentage of total population, deaths as a percentage of the population, Output gap, percentage of discovered infected agents (discovered infected over total infected), number of tests performed as a percentage of the population, ratio of positive test results to total tests and 7-periods effective reproduction number. Simulated values under four alternative testing policies: tests to hospitalized agents only (solid line with circles), voluntary tests (solid-dashed line), contact tracing (solid line) and random tests (dashed line). The model parameters used in the simulation are summarized in table 1.

Figure 2 shows instead the relationship, conditional on the testing policies, between the actual disease prevalence, equal to the ratio of infected over population, and the observed disease prevalence, which is instead equal to the ratio of discovered infected agents over population. As the picture suggests, the actual prevalence is, in general, much bigger than what the observed prevalence, and the more so when the epidemic accelerates at early stages. Moreover, the observed prevalence is more informative about the actual prevalence in case of contact tracing or random testing. Thus, in order to correctly interpret the observed epidemiological data, either to compare different countries or to take policy action, it is crucial to account for the implemented testing policy. From these simulations, given the country characteristics, it is also possible to derive multipliers that can be used to transform the official epidemic statistics, into statistics on the actual number of infected, similarly, in spirit, to what Manski and Molinari (2020) and Peracchi and Terlizzese (2020) do. But this is beyond the scope of this paper.

Figure 2: **Actual and Observed Disease Prevalence**



Notes: Actual prevalence (infected agents over population) and observed prevalence (discovered infected agents over population) conditional on four alternative testing policies: tests to hospitalized agents only (upper-left panel), voluntary tests (upper-right panel), contact tracing (lower-left panel) and random tests (lower-right panel). The model parameters used in the simulation are summarized in table 1.

Table 2 summarizes the median performances of the testing policies over the simulation runs with epidemics. I use seven metrics: peak actual prevalence, which is the fraction of the

population infected at peak of the epidemic, to measure the overall stress of the health system; the peak in new hospitalizations, which measures the pressure on hospitals; the mortality rate, whose reduction is the ultimate goal of the policymakers; the percentage of infected individuals who are discovered, which measures the efficacy of the testing policy; the number of tests performed, which measures the cost of the testing policy; the ratio of positives tests to total tests, which measures instead the efficiency of the testing policy. The total output loss as a percentage of potential output, that measures the impact of he epidemic, and of the policies to cope with it, on the economy. There are in principle other metrics that can be used, such as, for instance, the percentage of immunes at the end of the epidemic, which is correlated to the probability of second waves, or the duration of the epidemic, which can proxy for the individual stress caused by the epidemic. I chose to focus on a limited set of indicators for simplicity of exposition

Contact tracing is associated with the best performance, although not by a big margin. In particular, contact tracing reduces the median peak hospitalization rate with respect to the benchmark with tests to the hospitalized only by 17%, while voluntary tests and random tests, respectively, by 5.5% and 10%. The median peak in actual prevalence is also 31% smaller than benchmark with contact tracing, while it is only 11% smaller with voluntary tests and 14% smaller with random tests. The median mortality rate with contact tracing is instead 18% lower than benchmark, while it is 15% lower than benchmark with random tests and 8% lower than benchmark with voluntary tests. The main reason why contact tracing performs better is because it allows the identification of a bigger percentage of the infected: a median of 44%, versus 40% of random testing, 37% of voluntary tests and only 27% of the benchmark. Actually the superior performance of contact tracing could also be a consequence of the calibration, namely of an insufficient number of agents that asks for voluntary tests due to a high Q, or a small number of random screenings. In subsection 5.1, I discuss the consequences of a higher number of randomly screened agents, concluding that random screenings can approach the performance of contact tracing only in case of a small testing capacity. In section 8, I discuss instead the simulation results obtained with a lower value of Q, finding indeed very similar results to the benchmark discussed in this section. Thus it does not seem that the comparison between the testing policies is a by-product of the

calibration.

There are two contrasting effects of the testing policies on output. Identifying more of the infected results in more quarantines and, therefore, in a steeper output reduction during the epidemic. However it also results in less infections and less deaths and, therefore, in a lower output loss in the long run. Contact tracing yields the lowest total output reduction, 6% lower than what testing hospitalized agents only implies. The long run effect of preserving the production potential of the economy prevails. Voluntary tests and random screenings yield instead intermediate results.

One potential drawback of contact tracing is that it is costly: the median number of tests performed over the course of the epidemic is 7.6 times as big as in the baseline simulation with tests to hospitalized agents only, while it is 5.5 times as big in case In case of random tests, and 5.2 times as big in case of voluntary tests. Since contact tracing entails a big number of tests, it is also associated with a lower median positives to tests ratio: a median of 6.4%, versus 12.8% of random tests and 16.5% of voluntary tests, with a baseline median of 50% in case of tests to hospitalized agents only. Therefore contact tracing is also more efficient. A corollary is that comparing the ratio of positives to tests, say across different regions and/or countries, can uncover differences in the capabilities to identify the infected, rather than differences in the progression of the epidemic. Moreover, the median correlation between actual prevalence and the positives to test ratio over the simulated epidemics is around 0.85 in case of either voluntary tests, random tests and tests to hospitalized agents only, but it is only 0.60 in case of contact tracing. Actually in some simulations with contact tracing, the correlation was as low as 0.15, while in some other simulations with random testing, it was above 90%. The conclusion is that the positives to test ratio is less informative about the progression of the epidemic in case of contact tracing and more informative in case of random or voluntary tests.

All in all, the simulation results show that contact tracing can ease the pressure on the health system and can reduce mortality, while also containing the output loss associated with the epidemic, but its performance is not much different from what could be achieved by simple random testing or by allowing tests on a voluntary basis only. The problem is that the median percentage of discovered infected is rather low, in absolute terms, even in case of contact tracing: a median of 44%, with 90% of the simulation displaying a average discovered

rates between 20% and 52%. There are two forces behind this result. First, not all social contacts can be traced, either because the technology to do so is not available or not usable. Second, contact tracing puts a lot of pressure on the laboratories, creating congestion: in many cases, the required number of tests exceeds the capacity constraint and the labs are not able to promptly process all tests, with a result of an insufficient number of identified infectious[6]. So the question is if infections, hospitalizations, mortality and output loss can be further reduced with an increase in testing capacity. I study the effect of three alternatives to increase the testing capacity in the next section.

5 Increasing Testing Capacity

In this section I study the effects of three alternative strategies to increase the testing capacity. The first consists in new investments, which can entail, either buying new equipment, hiring more professional, or both. I summarize the performances of the testing policies in case of capacity investments in subsection 5.1. The problem with those investments is that they might not be feasible in the short run, which is exactly when testing capacity is most needed to promptly respond to an epidemic. For instance, there might be not enough trained professional to hire, and training new ones does require time. New equipment, in turn, needs time to be manufactured, and the few specialized firms that produce them might not be able to quickly evade all orders, especially if many countries are hit simultaneously. Indeed, the shortage of reagents for RT-PCR Covid-19 tests at the beginning of 2020 was perhaps the main challenge faced by most governments and that, most likely, paved the way to generalized lockdowns. There are however alternatives to investments to increase the testing capacity. A first possibility is group testing, which I describe in detail in subsection 5.2. Israel and the state of Minnesota are two examples of the use of group testing to cope with Covid-19 (Gollier and Gossner 2020). The second consists instead in using, if available, alternative, less sensitive tests, that identify the infected with a probability, rather than with certainty. Larremore et al. (2020) and Gans (2020) advocated their use to cope with Covid-19 because of their superior performance to test for infectiousness. I study the effect of their introduction

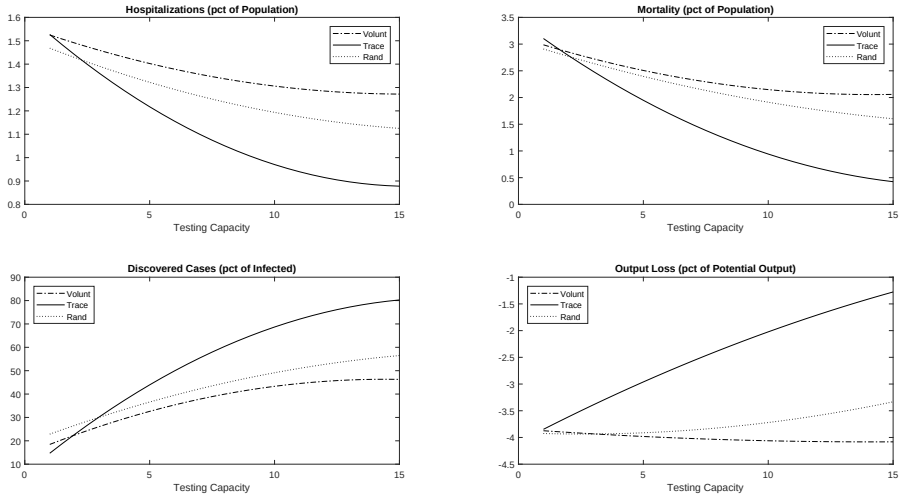
[6]In section 6, I also show that this congestion is not the by-product of the assumption that previously discovered infected agents are required to take tests for contact tracing.

in subsection 5.3.

5.1 Testing Capacity Investments

Figure 3 plots the median mortality rate, the median peak in new hospitalizations, the median percentages of identified and quarantined infected and the median cumulative output gap, or total output loss of the epidemic, as functions of the testing capacity for three alternative testing policies: voluntary tests, random tests and contact tracing. I do not plot the results obtained with tests to the hospitalized only because the capacity constraint is never binding, with unchanged results. All functions are smoothed with a quadratic regression. Since capacity investments are unlikely to be effective in the short run, these results show either what could be achieved in the medium-long run or, alternatively, they could be used to compare the performances of different country/regions with similar characteristics except for the testing capacity.

Figure 3: **Testing Capacity**



Notes: Upper-left panel: Median peak in new hospitalizations as a percentage of the population. Upper-right panel: median death rate. Lower-left panel: Median ratio of discovered infected to total infected. Lower-right panel: median total output loss as a percentage of potential output (cumulative output gap). Maximum testing capacity on the x-axis. Solid line: contact tracing; Solid-dashed line: tests on voluntary basis; Dashed line: random screenings. All pictures are smoothed with a quadratic regression. The parameters used in the simulations are summarized in table 1.

The upshot of the figure is that contact tracing benefits the most from an increased testing

capacity. The reason is that, as already stressed in section 4, the number of required tests under contact tracing grows very rapidly with infections, especially when the epidemic accelerates, so it requires a non-trivial amount of testing capacity in order to work properly, that is to identify a sufficient number of infected agents. In particular, A median identification of 60% of the infected, with contact tracing, requires a testing capacity increase, everything else equal, from 5% to 8.5%. To identify a median of three quarters of the infected, contact tracing needs instead at least a 13.5% capacity. If it was possible to increase the capacity up to this point, the median peak hospitalization rate would decrease by more than 25%, while the median mortality rate by more than 74% and the median output loss by 72%. The picture also shows decreasing returns in the gains from capacity increases, which is a consequence of the baseline calibration according to which only half of the social contacts are correctly identified by tracing. In other words, there is a limit to what contact tracing can achieve even in case of a very big testing capacity.

The voluntary tests policy also benefits from more capacity, especially close to the peak of the epidemic when, as a consequence of the big observed prevalence, many more individuals ask for a test. However the percentage of discovered infected grows less rapidly than in case of contact tracing, because the additional tests allowed by the bigger capacity are administered to agents with a lower probability of being infected (not necessarily exposed). The consequence is a modest decrease of both the median mortality and of the median peak hospitalization rate. Moreover, voluntary tests deliver a slightly bigger output loss when the testing capacity increases, since there are more discovered infected and, therefore, more quarantines, but not enough to significantly decrease mortality rate.

Random testing yields intermediate results: since the number of tests grows linearly with the capacity, there is a bigger number of discovered infected, but since the median ratio of randomly screened infected to randomly screened non-infected does not change with the capacity increase, the percentage of discovered infected agents does not grow much. As stressed in section 4, it is possible that contact tracing performs better than random screenings, in the baseline simulations, because of an insufficient scope of this last program, which does not exhaust the testing capacity. From this last simulation exercise, I can actually assess the extent to which the assumption about the number of the randomly tested agents influences

the comparison. Figure 3 shows that the median mortality rate achieved by contact tracing in the baseline model with 5% maximum capacity can be reached, in case of random screenings, with slightly more than 10% maximum capacity. In other words, contact tracing with limited testing capacity is not very different from implementing random screenings. However the performances of the two policies diverge as the testing capacity increases. For instance, the mortality rate that can be achieved with contact tracing and 10% maximum capacity does not seem to be reachable by random screenings for reasonable numbers of screened individuals per period.

The conclusion from this analysis is that it is possible to significantly reduce the mortality and output loss of an epidemic by increasing the testing capacity, and the more so in case of contact tracing. An alternative interpretation is that contact tracing works best in the countries or territories with bigger testing capacity. The question, once again, is how to increase this capacity in the short run, since capacity investments are hardly feasible in the midst of an epidemic and since they require time to be effective. I propose to alternatives, studying their performance: group testing and the use of less sensitive tests. Since contact tracing benefits the most from a bigger testing capacity, I will focus exclusively on it.

5.2 Group Testing

Dorfman (1943) was the first to propose group testing to increase the testing capacity while keeping the same test sensitivity[7] and specificity. The idea consists in collecting two samples per individual. First samples are bunched into groups of fixed size. If the result for the the group is negative, then the result is a negative for all of the group's members, and second samples are discarded. In case the group result is positive, second samples are processed individually. Suppose that p is the probability that a test turns positive and that first samples are bunched in groups of size g. The expected number of tests needed for n individuals, assuming that the number of groups n/g is integer and that the test results are independent[8],

[7]This statement assumes that bunching samples together does not result in a "dilution" that makes it more complicated to detect the pathogen responsible for the contagious disease. In such cases, the group size must be reduced in order to restore the test sensitivity.

[8]This excludes the possibility that clusters of individuals who were in contact with each other are tested in the same session or labs. So I assume, unrealistically, that the samples in those clusters are processed in different labs.

is:

$$nG(p,g) = \frac{n}{g}\left\{(1-p)^g + (1+g)[1-(1-p)^g]\right\} \tag{16}$$

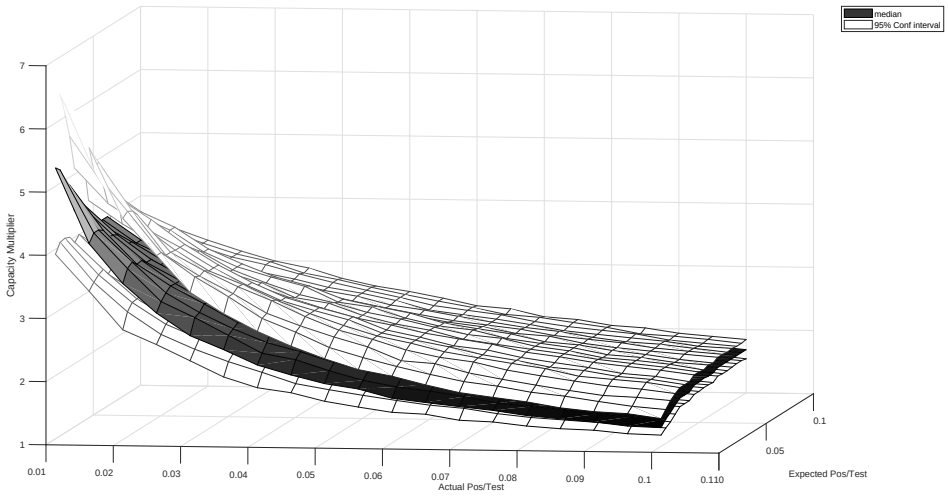
Basically with probability $(1-p)^g$, only one test is needed for g individuals, while $g+1$ tests are needed with probability $1-(1-p)^g$, one for thee group and g for each member. Minimizing with respect to g, for fixed n and p, gives the optimal group size $g^*(p)$ and the resulting number of tests required $nG^*(p)$. The expected testing capacity multiplier upon adoption of the optimal group testing strategy is $S(p) = 1/G^*(p)$.

In the context of my model, the adoption of group testing makes it possible to overcome the upper bound L to the number of tests that can be performed each period. The actual capacity multiplier is a function of the probability p, equal to the ratio of positives results to total tests performed, on which the optimal group size depends: the lower this ratio, the lower the probability that group tests will yield a positive and, therefore, the bigger the optimal group size and the lower the number of tests needed to screen the population (bigger capacity multiplier). The problem is that the actual fraction of positives to tests is unknown before the tests are processed, so the labs need to forecast it when setting the group size, and the final capacity multiplier will be a function of how accurate the forecast is. Formally, the actual capacity multiplier $S_t(p_t, g(p^e_{t-1,t}))$, in a given period t, is a function of the actual positives to tests ratio p_t and of the group size g chosen, according to Dorfman's solution, based on the forecast $p^e_{t-1,t}$ of the positives to tests ratio in t formulated in $t-1$, If the actual ratio of positives tests is bigger than the forecast, the group size is too big, resulting in more tests needed; Viceversa, if the ratio is smaller than the forecasts, less tests are needed.

The capacity multipliers in case there are differences between the actual and the forecasted ratio of positives to tests can be quantified by simulation. Figure 4 shows the results in case of an actual fraction of positives between 0.1% and 10% and of a group size chosen according to forecasts of the ratio of positives between 0.1% and 10%. I plot the the median capacity multipliers, as well the lower and upper bounds of their 95% confidence interval, over 1000 simulations of societies with 1000 individuals. The pictures shows that gains from group testing could be substantial, although they decrease with the actual fraction of positives.

The evidence summarized in figure 4, together with the results discussed in section 5.1,

Figure 4: **Testing Capacity Multipliers with Group Testing**



Notes: Median and 95% confidence interval of the testing capacity multiplier obtained with group testing in case the actual (ex-post) fraction of positives to tests is different from the forecasted fraction of positives to tests used to set the group size according to Dorfman (1943).

suggests that there could be gains from the introduction of group testing. In the context of an epidemic, those gains could be substantial because of an endogenous reinforcement effect. If there is a binding capacity constraint, only few tests are performed and, with many infections in the economy, this results in big ratios of positives to tests, which make group testing useless. However implementing group testing, by increasing the number of tests, reduces the ratio of positives to tests for fixed number of infected individuals, allowing an even bigger capacity increase that further decreases the positives to tests ratio and so on. In other words, group testing introduces a complementarity that makes it possible to identify and quarantine a big percentage of infectious and, therefore, to reduce infections, hospitalizations and deaths.

A complete group testing strategy entails also a forecast model of the ratio of positives to tests. Actually complicated forecast models are potentially difficult to implement for two reasons. First, and foremost, because there are not enough time series data to estimate or train them, especially at early stages of the epidemic and in case of a new disease such as the Covid-19. Second, the forecasts must be updated very frequently, and individual labs do not necessarily have the know-how to do so, while in case of an external, centralized agency that

pools and processes the data, there would coordination and communication problems that are not easy to solve. Given this concerns, I will focus, for simplicity, on a simple adaptive algorithm to set the optimal group size: a rule-of-thumb according to which the forecasts of the positives to test ratio is equal to the positives to test ratio in the previous period: $p^e_{t-1,t} = p_{t-1}$. The upper bound to the number of tests that can be performed each period with group testing is therefore:

$$L^g = \lfloor L\, S_t(p_t, g(p_{t-1})) \rfloor \tag{17}$$

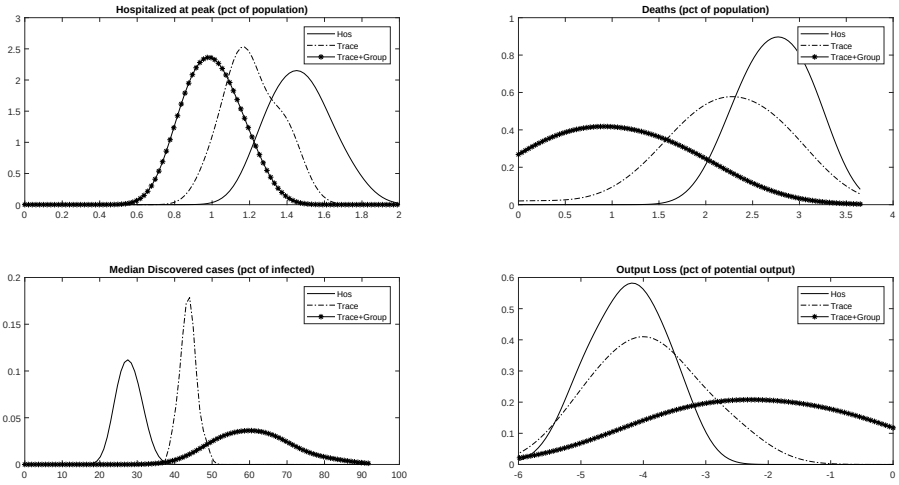
I simulate the epidemics building explicitly group testing in the model. I assume that there is a unique lab that sets the group size according to the previous period positives to tests ratio. The tests are randomly bunched in groups and processed according to the group testing protocol. I then compute the total number of tests processed each period and the total number of tested agents each period. Table 3 compares, for the baseline 5% maximum test capacity, the performance of contact tracing with group testing and rule-of-thumb forecasts to contact tracing without group testing and to the benchmark with tests to hospitalized agents only. In addition to the metrics used in subsection 4, I also consider the number of screened individuals, which is not equal to the number of processed tests, as in the previous simulations, but bigger than that because of group testing.

Group testing increases the median percentage of screened individuals by 36% with respect to contact tracing only, resulting in an increase in the median percentage of identified infected up to 60%, or 35% more with respect to contact tracing alone. Having more quarantined infectious smooths the peak of the epidemic, with a 54% reduction of median actual prevalence peak and a 19% reduction of the median peak of new hospitalizations. Median mortality, as a result, drops by 56%. As compared to the alternative of testing only the hospitalized, group testing reduces mortality by 63%. The median overall output loss is also 38% smaller with contact tracing and group testing with respect to contact tracing alone, and 47% smaller than benchmark. Moreover, the total number of tests performed with group testing decreases by 26%, so group testing will also ease the pressure on labs.

Figure 5 show the empirical distributions over simulation runs of the mortality rate, of the peak hospitalization rate, of the median percentage of discovered infected and of the

total output loss for the baseline model simulation with tests to hospitalized agents only, for contact tracing and for contact tracing coupled with group testing. The picture shows that the gains can be actually achieved in most simulations. Moreover, there is also a considerable probability mass at zero mortality in case of contact tracing with group testing, which is a goal seldom reached by contact tracing alone. The cumulative output loss is also typically lower in case of group testing.

Figure 5: **Contact Tracing and Group Testing During an Epidemic**



Notes: Empirical distributions of the peak in new hospitalizations (upper-left panel), mortality (upper-right panel), median percentage of discovered infected (lower-left panel) and output gap (lower-right panel) over simulation runs. Solid line: tests to hospitalized agents only; Solid-dashed line: contact tracing; Solid line with asterisks: contact tracing with group testing. 500 simulations conditional on each testing policy. The parameters used for the simulation are summarized in table 1.

In addition to the two-rounds group testing by Dorfman (1943), I also consider an alternative with three samples per individuals and three rounds of testing. First samples are bunched in groups of $\hat{g}$. In case of positive results, second samples are then bunched in subgroups of g, with $1 \leq g \leq \hat{g}$ and such that $\hat{g}/g$ is integer. In case of a positive result for a subgroups, third samples are processed individually. The expected number of tests, in this case, is equal to:

$$nG_3(p, g, \hat{g}) = \frac{n}{\hat{g}} \left\{ (1-p)^{\hat{g}} + \left[1 - (1-p)^{\hat{g}}\right] \left[1 + \frac{\hat{g}}{g} G(p, g)\right] \right\} \tag{18}$$

where $G(p, g)$ is defined in equation (16). The optimal group sizes $\hat{g}$ and g are the result of

the minimization of the expected number of tests. The testing capacity multiplier associated with this three-samples strategy is $1/G_3^*(p)$, where $nG_3^*(p)$ is the number of required tests at the optimal choices of $\hat{g}$ and g. As in the previous group testing simulation, I consider a simple rule-of-thumb forecast model for the positives to tests ratio. The results are summarized again in table 3. The gains from this strategy are actually very similar to the baseline, two-rounds, alternative. Most likely, the reason is that the number of tests grows much more rapidly, with respect to the two-samples case, when the ex-post ratio of positives to tests is bigger than the forecast, which often happens before the peak, when infections accelerate. Since the costs of collecting three samples is arguably higher, both for the laboratories and for the tested individuals, the conclusion is that two samples are sufficient.

5.3 Less Sensitive Tests

The main reason why capacity investments might not be feasible in the short run is the complexity of the tests themselves, that require both specialized professionals and dedicated equipment. However less precise tests, that require less expertise and less time, and little to no extra equipment, are typically available. For instance, antigen tests, in the covid-19 case, are a less precise, but much quicker and cheaper alternative to RT-PCR tests. These lower-precision tests are inadequate in case of severe symptoms, because medical treatment is typically contingent on the infection status. The question is if they are useful for mitigation. Tests with lower specificity result in unnecessary quarantines of non-infectious agents, that, by reducing the density of the society, reduce infections. Such tests are clearly more effective at stemming the diffusion of the disease, albeit at a high individual cost. Their main problem, however, is that not many agents will agree to take them, since they might yield an individual, selective, lockdown, not motivated by the infectiousness. Moreover, enforcing a quarantine can be legally difficult without an almost-sure proof of infectiousness. In short, less specific tests are unlikely to be usable, so I abstract from them. I will focus instead on the adoption of tests with lower sensitivity, which do not always identify the infected. As such, they result in an insufficient number of quarantined infectious agents at each testing round, but, since they can be used more frequently, they might also imply a bigger percentage of identified infected (Larremore et al.2020 and Gans 2020). My contribution is to study the effects of less sensitive

tests within the SIR-Macro model.

I assume that $\hat{L}$ alternative test with sensitivity $\hat{\sigma} = 0.5$ are available. The labs perform the tests with perfect sensitivity first, until exhaustion, and then the tests with lower sensitivity. The assignment of agents to tests is random. If the number of tests to perform exceeds the total capacity $L + \hat{L}$ there is random rationing, with an additional random assignment of individuals to tests. This formulation assumes the existence of a central authority that allocates tests to labs or, alternatively, the uniqueness of the lab. In reality, it is possible for some labs to exceed their individual capacity even if the capacity constraint is not binding at the aggregate level, with some other labs registering an excess capacity. I decided to abstract from these complications to avoid an explicit model of the lab industry and of the distribution process of tests to labs.

In case $\hat{L} = L$, meaning that there are as many low sensitive tests available as perfect tests, the median mortality rate with contact tracing turned out to 39% lower than in case only the perfect tests were used, and 50% lower than its value with tests to the hospitalized only for the baseline 5% maximum capacity. The reduction in the median peak of new hospitalization with respect to contact tracing with perfect tests only is instead 19%, while the median output loss is 28% lower. The gains are lower in case of voluntary tests: a 7% reduction in median mortality as compared to the use of the perfect tests only, with a 3% lower median peak hospitalization rate and a 2% lower median output loss. In case of random screenings, the performance is actually slightly worse than benchmark, although the difference is most likely due to random variation. Adding twice as much tests of lower quality, $\hat{L} = 2L$, in case of contact tracing, results in a 61% lower median mortality as compared to the use of the perfect tests only, a 19% lower median peak hospitalization rate and a 56% lower median output loss. The conclusion is that using tests with lower sensitivity, when doing contact tracing, can also reduce mortality and the overall cost of the epidemic in terms of output lost.

I also performed a different exercise in line with the analysis by Larremore et al. (2020) and Gans (2020). In particular, I simulated the economy assuming that only less precise tests were available, but in greater quantity than perfect tests, and compared the performance to the benchmark simulation. The idea is to perform less precise tests with greater frequency in order to promptly quarantine the infectious. In my model, however, I do not account

specifically for a latency period between the infection itself and the time from which the infected becomes infectious, nor I explicitly model infectiousness as a function of symptoms or pathogen characteristics (for instance the viral load in case of viral diseases), so my results must be considered as suggestive. In particular, they are indicative of a lower bound to the actual advantage of using such tests because of the delay in identifying the infected of more precise tests. In the simulations with $0.1 \cdot N$ tests with 50% sensitivity and contact tracing, I had the same median mortality rate that I obtained, with contact tracing, in case of $L = 0.05 \cdot N$ tests with 100% sensitivity. The median peak in new hospitalizations is also very close, while the output loss 2% lower. In other words, doubling the capacity ad halving the sensitivity does not yield significant change in the results. In the simulations with $0.2 \cdot N$, for the same 50% sensitivity, I had instead a 40% lower median mortality rate with respect to what contact tracing yields in the benchmark simulation, a 12% lower median peak of new hospitalizations and a 28% lower median output loss. The conclusion is that using tests with lower sensitivity yields better results only in case they are feasible to frequently process in very large numbers.

6 Improving Upon Contact Tracing

In this section I study how to improve contact tracing along four dimensions: adopting more effective tracing technologies, allowing individuals to voluntary take tests even if they are not required to do so for contact tracing, implementing additional random screenings and avoiding tests to the previously discovered as infected.

The choice of the efficacy at tracing social contacts $\phi = 0.5$ in the baseline simulation is rather arbitrary. As already stressed, this parameter is a function of both the tracing technologies, which range from customers registries filled by restaurants and shops to cell phone localization and security cameras image recognition, and of the legal framework which might or might not allow intrusive and privacy-violating practices without explicit consent. Since there might be significant cross country differences both in the ability to manage those technologies and in their legal feasibility, it is interesting to assess the performance of contact tracing for different tracing efficacy. I simulated the model with two alternative values of ϕ,

respectively 25%, meaning that only one out of four contacts from social activities is correctly identified and tested, and 75%, with three out of four contacts identified. The resulting decreases in the median peak of infections, hospitalizations mortality and output loss obtained with contact tracing turned out to be not very different from the baseline simulations with $\phi = 0.5$. This means that contact tracing collapses, in the baseline simulation, because of a binding test capacity constraint, not because of an insufficient number of identified social contacts. Putting it differently, sophisticated tracing technologies are not very useful if there is not enough testing capacity: for small capacity, tracing within the family, school and workplace is enough to create congestion, and the extra gains from a correct identification of a bigger percentage of the social contacts are minimal. Moreover, I also find that using group testing to increase the testing capacity delivers good results even in case of an impossibility to trace social contacts. Model simulations with two-rounds group testing, rule-of-thumb forecasts and $\phi = 0$ result in a 31% lower median mortality rate with respect to a scenario with contact tracing, no group testing and $\phi = 0.5$, or 43% lower than in the benchmark simulation with tests to the hospitalized. The conclusion is that it is best to first increase the testing capacity, rather than adopting more sophisticated, and potentially privacy-invading, tracing technologies.

I also simulated the model allowing voluntary tests in addition to contact tracing. The idea is that, since symptomatic agents take the test frequently, this can allow the identification of more of the infectious, therefore improving upon the performance of contact tracing alone. The problem, however, is that the capacity constraint becomes binding more often, and, since the extra tested agents have a lower probability of being infected, the resulting congestion might reduce the overall percentage of identified infected. The simulation results show that the two effects balance, and the overall performance of contact tracing is almost unchanged, even in case of a 10% maximum capacity. The conclusion is that allowing voluntary tests, when doing contact tracing, is not useful, putting only more pressure on labs.

For similar considerations, it is not useful to implement random screenings in addition to contact tracing, given also that the probability of finding an additional infected individual with a random screening is even smaller than in case of voluntary tests: infected individuals can develop symptoms and, therefore, ask for voluntary tests more often than non-infected.

I simulated the model with contact tracing and with additional random tests to 0.5% of the population, equivalent to one tenth of the baseline testing capacity. The results were similar to the benchmark simulation, also in case of a 10% maximum test capacity. The conclusion is that implementing random tests when doing contact tracing only increases the pressure on labs without tangible gains.

One of the strong assumptions of my simulation exercises is that all testing policies, including contact tracing, are designed as if second infections were possible. In particular, agents are required to take tests, as part of contact tracing, even if they took a test while infected in some previous period, so even if they are immune and if they know about their immunity. The effect is an avoidable congestion in the system that dwarfs the performance of contact tracing. To evaluate the importance of this assumption, I simulated the model again assuming that identified infected agents are not tested when contact tracing is implemented. Thus an agent is required to take a test, under contact tracing, if the binary variable $\bar{T}^g_{ij,t}$ is equal to one, with:

$$\bar{T}^g_{ij,t} = Z_{ij,t-1} \, \mathbb{1}_{[i \in (I_{t-1} \cup P^g_{ij,t-1}) \text{ or } j \in J_{t-1}]} \tag{19}$$

For the baseline model parameters, the testing capacity constraint becomes binding less often, with however only a slightly better overall performance. Median mortality, in particular, is 6% lower than its value in case of contact tracing without memory of the discovered infected, while the median infection peak roughly 5% smaller. The median peak in new hospitalizations is actually unchanged, while the output loss 2.5% smaller. These results suggest that the tests to previously discovered infected agents are not the main force behind the frequently binding testing capacity constraint under contact tracing.

7 Endogenous Response to the Epidemic

In the model, there is no endogenous agents' response to the epidemic. This is actually unlikely to be the case, since some agents might voluntary decide to change their daily routines to decrease the infection risk to which they are exposed, for instance avoiding eating out or wearing a protecting device. In this section, I study the effect, on the testing policy

performances, of an endogenous density reduction proportionately to the observed disease prevalence. In particular, I assume that the effective density $\hat{\eta}_t$ at time t is:

$$\hat{\eta}_t = \max\left\{\eta - \theta \hat{I}^D_{t-1}; 0\right\} \tag{20}$$

where $\hat{I}^D_{t-1}$ is the observed disease prevalence and where the parameter θ governs how fast the agents reduce the social activities. I further set $\theta = \eta/10$, meaning that all agents undertake a voluntary social lockdown if the observed prevalence becomes bigger than 10%. This reduced form modeling can be considered as the outcome of the both the agents' choice problem over η and of the potential coordination problems that can arise, given that everyone could simply free-ride on the social activities reduction of others, with the consequence of an insufficient aggregate reduction[9].

This endogenous response of the agents slows down the epidemic independently from the testing policy. It smooths the peak of infections and the peak number of new hospitalizations, with a resulting reduction of mortality. There are however two contrasting effects on the performance of testing policies: the policies that identify more of the infected increase the ratio between the actual and the observed prevalence, thereby increasing the perceived individual risk of infection for all levels of the actual disease prevalence. However, since more identified infected means also a smaller number of infections, they also reduce the actual prevalence, which limits the endogenous response. The simulation results show that the gains from contact tracing, which is the best policy alternative in terms of identified infected, are higher in case of an endogenous response. In particular, in case of a maximum 5% capacity, the median mortality rate under contact tracing is 33% lower than in case of tests to hospitalized agents only. Conversely, an endogenous density reduction within a voluntary testing or random screenings policy do not result in significantly bigger gains. In case of a 10% maximum capacity, contact tracing with the endogenous density reduction reduces mortality by 77% with respect to the benchmark with tests to hospitalized agents only, while voluntary tests by 24% and random screenings by 45%. Thus the gains are higher, with more testing capacity, because a bigger fraction of the infectious is identified at earlier stages, and the more so in

[9]this is also the reason why a lockdown might be necessary even in case of a voluntary reduction of social activities by the agents, as in Eichenbaum, Rebelo and Trabandt (2020a).

case of contact tracing. Implementing group testing in two rounds, in addition to contact tracing, also results in bigger gains in case of an endogenous response, with, in particular, an extra 45% reduction of mortality with respect to what could be achieved with contact tracing alone.

Obviously all of these extra gains determined by the endogenous density reduction are a function of θ, which governs the speed at which social activities shrink with the observed prevalence: the bigger is θ, the bigger the gains from adopting contact tracing with group testing. In terms of primitives, θ is a function both of the preferences towards risk and of the utility assigned to social activities, but it could be also be interpreted as an effect of additional implemented policies to cope with the epidemic. For instance, closing museums, libraries and bars, reducing the maximum capacity of public transportation or, in general, all lockdown policies, will decrease θ. According to this interpretation, the simulation results show that contact tracing, when implemented alongside lockdowns, is particularly effective at stemming the diffusion of a contagious disease.

The simulation results turned out to be very similar in case of an endogenous reduction of the infection probabilities π and $\bar{\pi}$, in case, for instance, the agents took precautionary measures to protect themselves from infections as the observed prevalence increases. Thus, in case both the density and the contagion probability decreased with the observed prevalence, the performance of contact tracing would be further enhanced.

In conclusion, the simulation discussed in this section suggest that the baseline results in the paper are likely to be lower bounds to what could be gained with contact tracing and group testing. In other words, the baseline model simulation is likely to propose a worse-case scenario evaluation, which is also the reason why I did not explicitly built an endogenous agents' response in the model in the first place.

8 Robustness

The first robustness exercise consists in simulating the model with a lower incidence of symptoms due to concurrent infections, setting $\kappa = 0.5\%$. In the Covid-19 example, these scenario can be thought as the summer, when respiratory pathologies are infrequent, while the baseline

simulation refers to winter months, when Influenza typically strikes. With less symptomatics that worsen, the pressure on hospitals is generally lower. Moreover, voluntary tests are associated with a worse performance, because fewer agents take tests, resulting in a lower overall percentage of discovered infected agents. Contact tracing, instead, reduces median mortality by 17% with respect to benchmark in case of a 5% maximum capacity and, when combined with group testing in two rounds, by 50%. The other results turned out to be similar. So the main paper result stands even in case of the presence of relatively unimportant concurrent infections.

For similar considerations, the main results are also robust in case of a lower probability of requiring hospitalizations. In this case, there are bigger gains from all testing policies because testing the hospitalized only is associated with a worse performance. A higher death probability results instead in bigger gains, highlighting, once again, contact tracing as the best alternative. Similar considerations would apply in case the model featured a health system capacity constraint, according to which a proper treatment all patients is impossible if their number grows too much, with a resulting increase in mortality. Reducing infections and hospitalizations, in this case, is particularly valuable and, therefore, increasing testing capacity will yield even bigger gains. In case of a higher density η, contact tracing results in a testing capacity constraint that becomes binding more often, thereby worsening its performance. In such cases, there are even bigger gains from the adoption of strategies that increase the testing capacity, and especially from adopting group testing. A similar reasoning applies to countries with bigger average workplace and school size. In case of higher average family size, since it is not possible to isolate the discovered infected from their family members, there is, in general, a worse performance of all testing policies, which are not as able to stem the disease diffusion as in the benchmark simulation. Contact tracing is still however the best policy alternative.

Perhaps the reason why voluntary tests yield sub-optimal results is because of a high calibrated value of Q, which translates into an insufficient number of agents that want to take tests even in case of a high observed disease prevalence and, therefore, into a mild endogenous mitigation. To check the robustness of the results, I simulated the model under a rather extreme alternative parametrization, in which 90% of the symptomatics want to get tested in case of a 10% observed prevalence. The resulting calibrated value of Q is 0.17. The results

in case of a 5% maximum test capacity are not very different from the baseline simulation. Even in case of a 10% maximum capacity, the mortality rate in case of voluntary tests is 30% smaller than in case of tests to hospitalized agents only, but 4% bigger than in case of random screenings and 1.5 times bigger than in case of contact tracing. Voluntary tests are still a dominated choice.

I also simulated the model assuming that all agents choose to get tested on a voluntary basis only in case they were never identified as infected in some previous period, or, in other words, assuming perfect knowledge of the impossibility to get second infections. The indicator variable for the individual decision to take a test becomes, in this case:

$$\bar{T}^{v}_{ij,t} = Z_{ij,t-1}\Big[\mathbb{1}_{[I^{D}_{t-1}\geq\bar{q}_{ij}]}(1-\zeta_{i,t})\max\{f_{ij,t},\varepsilon_{ij,t}\} + \mathbb{1}_{[I^{D}_{t-1}\geq\hat{q}_{ij}]}(s_{ij,t}+a_{ij,t}+u_{ij,t})(1-\varepsilon_{ij,t})\Big] \quad (21)$$

This new setting implies that fewer tests are taken on a voluntary basis, thereby inducing less congestion. However this turned out to be not enough to significantly improve upon the performance of voluntary tests for the baseline simulation. The median percentage of discovered infected is in fact 5% bigger with respect to the case of voluntary tests of previously discovered infected agents, while the median infection peak 7% smaller, with a 3.5% lower median mortality and an almost equal output loss. These results, when compared to what could be obtained with contact tracing and no tests to the previously discovered infected agents, stress that the comparison between voluntary tests and contact tracing is not influenced by the assumption of memoryless testing.

As already stressed in section 5.2, the rule-of-thumb forecast used to set the group size in case of group is not optimal and it is, in principle, possible for group testing to yield even bigger gains in case of a more sophisticated forecast model, although, at early stages of the epidemic, the lack of a sufficient time series of data impairs the forecast ability. I considered a further adaptive forecast model based on a simple rolling regression, according to which the forecast in t is based on information on the evolution of the epidemic up to date $t-1$. I included four variables in the regression: the lag of the positives to test ratio, the number of screened individuals, the stock of discovered infected and the fraction of the population which is hospitalized. The results turned out to be very similar to the rule-of-thumb alternative,

with a slightly lower median percentage of discovered infected and a slightly smaller death rate. Actually this forecast model is also rudimentary, and this simulation results do not exclude the possibility that more sophisticated forecasts such as, for instance, the ones based on deep learning algorithms, might deliver better results. Once again, given the substantial implementation difficulties of complicated forecasts, it seems to be enough to focus on a simple rule-of-thumb.

9 Conclusion

I showed that increasing the testing capacity, during an epidemic, can smooth the peak of infections, ease the pressure on hospitals, reduce mortality and curb the output loss due to the epidemic, especially in case contact tracing is used to target the individuals to test. Investments in new lab equipment, or in the training of qualified professionals, are however typically unfeasible in the short run, as a quickly unfolding epidemic requires prompt actions. I identified two viable and effective alternatives: group testing and the use of less sensitive tests. Both strategies are associated with significant gains, especially if combined with contact tracing. Considering that group testing does not require any additional investment or equipment, and that it is easy to implement with a rule-of-thumb forecast for the positives to tests ratio, it qualifies as a leading choice. There are however costs associated with it. First, there is an additional, albeit minimal, discomfort for the tested individuals due to the collection of two samples which, in the case of Covid-19, means two naso-pharingeal swabs rather than one. Second, testing a large fraction of the population could be logistically complicated: group testing, according to my simulation results, implies, in a country of 50 millions inhabitants, the screening of a median of roughly 2.5 millions individuals per period, which might be not easy to achieve without, say, long lines and lengthy waits in front of laboratories and hospitals. I also showed that adopting complicated, costly, and privacy-invading tracing technologies is not worthwhile in case of a limited testing capacity. Moreover, I showed that, when doing contact tracing, it is not useful to also allow voluntary tests or to implement additional random screenings. All in all, the simulation results suggest that lockdowns might not be necessary to stem an epidemic even in case of a limited ability to trace social contacts

and even in case of a relatively small testing capacity, which is easy to overcome with group testing.

References

Aldridge, Matthew, Oliver Johnson and Johnathan Scarlett (2019). "Group Testing: An Information Theory Perspective.", *Foundations and Trends in Communications and Information Theory* 15(3-4): 196-392.

Allen, Linda J.. 2017. "A Primer on Stochastic Epidemic Models: Formulation, Numerical Simulation, and Analysis." *Infectious Disease Modelling* 2: 128-142.

Alvarez, Fernando, David Argente, and Francesco Lippi. 2020. "A Simple Planning Problem for COVID-19 Lockdown." *Covid Economics* 14: 1-32.

Collard, Fabrice, Christian Hellwig, Tiziana Assenza, Sumudu Kankanamge, Martial Dupaigne, Nicolas Werquin and Patrick Feve. 2020. "The Hammer and the Dance: Equilibrium and Optimal Policy During a Pandemic Crisis." Center for European Policy Research Discussion Papers no. 14731.

Atkeson, Andrew. 2020, "What Will be the Economic Impact of Covid-19 in the US? Rough Estimates of Disease Scenarios." National Bureau of Economic Research Working Paper no. 26867, Cambridge (MA).

Berger, David W, Kyle F. Herkenhoff, and Simon Mongey. 2020, March. "An SEIR Infectious Disease Model with Testing and Conditional Quarantine." National Bureau of Economic Research Working paper no. 26901, Cambridge (Ma).

Bergstrom, Ted, Carl T. Bergstrom and Haoran Li. 2020. "Frequency and Accuracy of Proactive Testing for COVID-19." Mimeo.

Chowdhury, Rajiv, Kevin Heng, Shajedur Rahman Shawon, Gabriel Goh, Daisy Okonofua, Carolina OchoaRosales, Valentina GonzalezJaramillo, Abbas Bhuiya, Daniel Reidpath, Shamini Prathapan, Sara Shahzad, Christian L. Althaus, Nathalia GonzalezJaramillo, Oscar H. Franco. 2020. "Dynamic Interventions to Control COVID-19 Pandemic: a Multivariate Prediction Modelling Study Comparing 16 Worldwide Countries." *European Journal of Epidemiology* 35: 389-399.

Deb, Rahul, Mallesh Pai, Akhil Vohra and Rakesh Vohra. 2020. "Testing Alone is Insufficient." Mimeo.

Dewatripont, Mathias, Michel Goldman, Eric Muraille, and Jean-Philippe Platteau. 2020. "Rapid Identification of Workers Immune to COVID-19 and Virus-free: A Priority to Restart the Economy." Discussion Paper, Université Libre de Bruxelles.

Dorfman, Robert. 1943. "The Detection of Defective Members of Large Populations." *The Annals of Mathematical Statistics*: 14(4): 436-440.

Eichenbaum, Martin S., Sergio Rebelo and Mahias Trabandt. 2020a. "The Macroeconomics of Epidemics." Working Paper.

Eichenbaum, Martin S., Sergio Rebelo and Mahias Trabandt. 2020b. "The Macroeconomics of testing and Quarantining." Working Paper.

Farboodi, Maryam, Gregor Jarosch and Robert Shimer. "Internal and External Effects of Social Distancing in a Pandemic." Becker-Friedman Institute Working Paper no. 2020-47.

Favero, Carlo, Andrea Ichino and Aldo Rustichini. 2020. "Restarting the Economy While Saving Lives Under Covid-19." CEPR Discussion Paper no. DP14664.

Galeotti, Andrea, Jakub Steiner and Paolo Surico. 2020. "Merit of Test: Perspective of Information Economics." *Health Policy and Technology*, forthcoming.

Garibaldi, Pietro, Espen R. Moen and Christopher Pissarides. 2020. "Modelling Contacts and Transitions in the SIR epidemics Models." *Covid Economics* 5: 1-20.

Glover, Andrew, Jonathan Heathcote, Dirk Krueger and Jose-Victor Rios-Rull. 2020. "Health Versus Wealth: On the Distributional Effects of Controlling a Pandemic." Working Paper.

Gollier, Christian and Olivier Gossner. 2020. "Group Testing Against Covid-19." *Covid Economics* 2: 32-42.

Jones, Callum J., Thomas Philippon and Venky Venkateswaran. 2020. "Optimal Mitigation Policies in a Pandemic: Social Distancing and Working from Home." National Bureau of Economic Research Working Paper no. 26984, Cambridge (Ma).

Kasy, Maximilian and Alexander Teytelboym. 2020. "Adaptive Targeted Infectious Disease Testing." *Oxford Review of Economic Policy* 36(1): 77-93.

Kermack, William. O., and Anderson. G. 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.

Krueger, Dirk, Harald Uhlig and Taojun Xie. 2020. "Macroeconomic Dynamics and Reallocation in an Epidemic." National Bureau of Economic Research Working Paper no. 27047, Cambridge (Ma).

Farboodi, Maryam, Gregor Jarosch, and Robert Shimer. 2020. Internal and External Effects of Social Distancing in a Pandemic." Becker Friedman Institute Working paper no 2020-47.

Gans, Joshua S. 2020. "Test Sensitivity for Infection versus Infectiousness of SARS-Cov-2." *Covid Economics* 47: 1-16.

Istituto Superiore di Sanità. 2020."Valutazione di Politiche di Riapertura Utilizzando Contatti Sociali e Rischio di Esposizione Professionale." Rome, Italy.

Larremore, Daniel B., Bryan Wilder, Evan Lester, Soraya Shehata, James M. Burke, James A. Hay, Milind Tambe, Michael J. Mina and Roy Parker. 2020. "Test Sensitivity is Secondary to Frequency and Turnaround Time for COVID-19 Surveillance." MedRXiv.

Manski, Charles F. and Francesca Molinari. 2020. "Estimating the COVID-19 Infection Rate: Anatomy of an Inference Problem." *Journal of Econometrics*, forthcoming.

Peracchi, Franco e Daniele Terlizzese. 2020. "Estimating the Prevalence of the Covid-19 Infection, With an Application to Italy." *Covid Economics* 43:19-41.

Piguillem, Facundo and Liyan Shi. 2020. "The Optimal COVID-19 Quarantine and Testing Policies." EIEF working paper.

Russo, Francesco Flaviano. 2020. "Epidemics and Policy: the Dismal Trade-off." CSEF working paper no. 570.

Table 2: **Testing Policies**

	Test	Mean	Median	Std	95 Conf L	95 Conf U
Infpeak	Hos	17.64	18.27	4.52	0.48	22.06
	Volunt	15.31	16.25	4.25	0.48	19.18
	Trace	11.63	12.57	4.68	0.40	17.07
	Rand	14.31	15.66	5.05	0.32	19.10
Hospeak	Hos	1.42	1.47	0.21	0.90	1.72
	Volunt	1.33	1.38	0.19	0.90	1.57
	Trace	1.19	1.22	0.19	0.82	1.47
	Rand	1.29	1.31	0.20	0.82	1.56
Deaths	Hos	2.72	2.78	0.81	0.08	3.73
	Volunt	2.43	2.57	0.74	0	3.26
	Trace	2.06	2.29	0.86	0	3.03
	Rand	2.25	2.38	0.87	0.08	3.18
Yloss	Hos	-2.93	-2.99	0.86	-3.75	-1.88
	Volunt	-2.87	-3.09	1.05	-4.22	-0.01
	Trace	-2.48	-2.81	1.14	-3.74	-0.02
	Rand	-2.85	-3.17	1.14	-4.01	-0.01
Tests	Hos	0.72	0.73	0.06	0.51	0.78
	Volunt	3.63	3.77	0.85	0.52	4.35
	Trace	5.13	5.50	1.45	0.62	6.42
	Rand	3.95	3.96	0.08	3.77	4.06
Discovered	Hos	28.66	27.87	4.32	21.43	33.91
	Volunt	37.66	37.90	4.09	30.34	44.40
	Trace	45.29	44.29	6.86	20.00	51.89
	Rand	40.09	39.76	6.09	20.00	48.90
Postest	Hos	47.17	50.00	6.30	33.33	53.33
	Volunt	16.95	16.53	4.01	12.20	20.00
	Trace	6.17	6.45	3.01	1.96	8.20
	Rand	12.93	12.82	4.73	2.63	16.44

Notes: Infpeak is the percentage of the population infected at the peak of the epidemic (peak actual prevalence). Hospeak if the maximum fraction of the population to access an hospital. Deaths is the mortality rate as a percentage of the population, computed at the end of the epidemic. Yloss is the cumulative output loss (total output gap). Tests is the total number of tests performed over the course of the epidemic expressed in multiples of the population. Discovered is median of the ratio of discovered infected to total infected during the epidemic. Postest is the median ratio of new discovered infections to total performed tests during the epidemic. Hos refers to tests to hospitalized agents only. Volunt refers to voluntary tests in addition to tests to hospitalized agents. Trace refers to contact tracing in addition to tests to hospitalized agents. Rand refers to random tests in addition to tests to hospitalized agents. Test refers to the testing policy. Mean, median, standard deviation (Std), lower bound of the 95% confidence interval (95 Conf L) and upper bound of the 95% confidence interval (95 conf U) refer to the distribution of the results over all simulation runs with an epidemic. The parameters used in the simulation are summarized in table 1.

Table 3: **Contact Tracing with Group Testing**

	Test	Mean	Median	Std	95 Conf L	95 Conf U
Infpeak	Trace	11.63	12.57	4.68	0.40	17.07
	Group	5.89	5.80	4.16	0.32	12.67
	Group 3	6.64	6.65	4.70	0.32	14.73
Hospeak	Trace	1.19	1.22	0.19	0.82	1.47
	Group	1.02	0.98	0.20	0.74	1.31
	Group 3	1.06	1.05	0.19	0.81	1.39
Deaths	Trace	2.06	2.29	0.86	0	3.03
	Group	0.98	1.02	0.76	0	2.21
	Group 3	1.08	1.18	0.84	0	2.20
Yloss	Trace	-2.48	-2.81	1.14	-3.75	-0.03
	Group	-1.54	-1.60	1.04	-3.08	-0.11
	Group3	-1.71	-1.65	1.12	-3.42	-0.11
Tests	Trace	5.13	5.50	1.45	0.62	6.42
	Group	3.45	4.07	1.65	0.24	5.49
	Group 3	3.40	3.94	1.58	0.34	5.08
Discovered	Trace	45.29	44.29	6.86	20.00	51.89
	Group	62.95	60.47	12.59	16.23	81.67
	Group 3	63.00	60.13	14.82	33.33	87.50
Postest	Trace	6.17	6.45	3.01	1.96	8.20
	Group	7.58	7.49	1.86	4.00	10.00
	Group 3	8.62	7.69	4.47	5.88	10.10
Screened	Trace	5.13	5.50	1.45	0.62	6.42
	Group	6.71	7.49	3.03	0.60	10.66
	Group 3	6.69	7.48	2.94	0.82	10.04

Notes: Infpeak is the percentage of the population infected at the peak of the epidemic (peak actual prevalence). Hospeak if the maximum fraction of the population to access an hospital. Deaths is the mortality rate as a percentage of the population, computed at the end of the epidemic. Yloss is the cumulative output loss (total output gap). Tests is the total number of tests performed over the course of the epidemic expressed in multiples of the population. Discovered is median of the ratio of discovered infected to total infected during the epidemic. Postest is the median ratio of new discovered infections to total performed tests during the epidemic. Screened is the effective number of tests (bigger than the number of tests performed because of group testing. Test refers to the testing policy. Trace refers to contact tracing and tests to hospitalized agents. Group refers to group testing in two rounds (1 grouping) in addition to contact tracing and tests to hospitalized agents only. Group 3 refers to group testing in three rounds (2 groupings) in addition to contact tracing and tests to hospitalized agents only. Mean, median, standard deviation (Std), lower bound of the 95% confidence interval (95 Conf L) and upper bound of the 95% confidence interval (95 conf U) refer to the distribution of the results over all simulation runs with an epidemic. The parameters used in the simulation are summarized in table 1.