Analytics Saves Lives during the Covid Crisis in Chile

Leonardo J. Basso, Marcel Goic, Marcelo Olivares, Denis Sauré, Charles Thraves
Facultad de Ciencias Fisicas y Matematicas, Universidad de Chile and Instituto Sistemas Complejos de Ingeniería, email: lbasso@ing.uchile.cl, mgoic@uchile.cl, molivares@uchile.cl, dsaure@dii.uchile.cl,chtraves@dii.uchile.cl

Aldo Carranza, Gabriel Y. Weintraub
Graduate School of Business, Stanford University, email: aldogael@stanford.edu, gweintra@stanford.edu

Julio Covarrubias, Cristian Escobedo, Natalia Jara, Antonio Moreno
ENTEL Ocean, Empresa Nacional de Telecomunicaciones, email: jcovarrubia@entel.cl,cescobed@entel.cl, njaras@entel.cl,acmoreno@entel.cl

Demian Arancibia, Manuel Fuenzalida, Juan Pablo Uribe, Felipe Zuñiga, Marcela Zuñiga
Gobierno de Chile, email: darancibia@minciencia.gob.cl, manuel.fuenzalida@minsal.cl, juan.uribe@minsal.cl, felipe.zuniga@minsal.cl,marcela.zuniga@minsal.cl

Miguel O'Ryan, Emilio Santelices, Juan Pablo Torres
Facultad de Medicina, Universidad de Chile, email: moryan@uchile.cl,emilio.santelices@uchile.cl,jptorres@minsal.cl

Magdalena Badal, Mirko Bozanic, Sebastian Cancino, Eduardo Lara, Ignasi Neira
Instituto Sistemas Complejos de Ingeniería, email: m.badal@isci.cl, m.bozanic@isci.cl, s.cancino@isci.cl, e.lara@isci.cl, i.neira@isci.cl

During the Covid-19 crisis, the Chilean Ministry of Health and the Ministry of Sciences, Technology, Knowledge and Innovation partnered with the Instituto Sistemas Complejos de Ingeniería (ISCI) and the telecommunications company ENTEL, to develop innovative methodologies and tools that placed operations research and analytics at the forefront of the battle against the pandemic. These innovations have been used in key decision aspects that helped shape a comprehensive strategy against the virus, including tools that: (i) shed light on the actual effects of lockdowns in different municipalities and over time; (ii) helped allocate limited intensive care capacity; (iii) significantly increased the testing capacity and provided on-the-ground strategies for active screening of asymptomatic cases; and (iv) implemented a nationwide serology surveillance program that significantly influenced Chile's decision regarding vaccine booster doses and that also provided information of global relevance. Important challenges during the execution of the project included the coordination of large teams of engineers, data scientists, and health care professionals in the field; how to effectively communicate information to the population; and the handling and use of sensitive data. The initiatives enjoyed ample press coverage and, by providing scientific evidence supporting the decision-making behind the Chilean strategy against the pandemic, they helped provide transparency and objectivity to decision-makers and the general population. According to conservative estimates, the number of lives saved by all of the initiatives together is close to 3,000, equivalent to more than 5% of the total death toll in Chile during the pandemic. The saved resources associated with testing, ICU beds, and working days amount to more than 300 million USD.

Key words: Covid Pandemic, Data Science, Decision making, Public Policy

1. Introduction

Chile, a country of roughly 19.7 million inhabitants (INE 2019), had its first recorded Covid-19 case in March 3, 2020 with a traveler from Singapore. After almost two years, Chile has reported about 2.1 million cases and 39.7 thousand deaths due to the Covid pandemic (Ministry of Health 2022b,c). According to the Bloomberg's Covid Resilience Ranking, Chile was the best country in the world to live in during the pandemic as of the end of 2021 (Bloomberg 2022). In this paper, we describe how applied analytics played and continues to play a key role in Chile's favorable situation regarding overall pandemic management.

Chile's strategy to contain the pandemic consisted in a multilayered approach based on three strategic pillars:

1. Contagion prevention: lockdowns, testing and tracing strategies aimed at preventing infections, and identifying early on infected individuals so as to minimize spread.
2. Nationwide centralized management of critical beds: balancing the capacity of intensive care beds with growing and variable demand throughout the country.
3. Vaccine roll-out: designing and implementing a vaccination strategy to optimize its effectiveness based on a limited supply of a diversified pool of vaccines, using different technologies with uncertainty as to their effectiveness.

The collaboration between the Ministry of Health and the Ministry of Sciences, Technology, Knowledge and Innovation with the Instituto Sistemas Complejos de Ingeniería (ISCI, an interdisciplinary research center of engineering and economics) and ENTEL (the largest telecommunications company in Chile) was formed to develop a scientific approach that could support key decision-making in each of the three pillars of the Chilean strategy with the key goal of saving lives. In this respect, analytics proved to be fundamental to solving the different complex problems of this crisis.

At its core, the collaboration was supported by the collection, processing, and analysis of massive amounts of critical data, which were used to feed various models and then to develop decision-support tools using different dashboards, allowing decision-makers to have ready access to them. These data science, analytics, and operations research efforts resulted in an end-to-end cutting-edge technological pipeline, which was developed in a record amount of time given the urgency, and according to the exigencies of the pandemic dynamics. The models use advanced methods from statistics, machine learning, and operations research to support a proactive decision-making process and resulted in important scientific contributions reported in no less than four publications in prestigious journals (Carranza et al. 2022, Goic et al. 2021, Basso et al. 2021, Sauré et al. 2022).

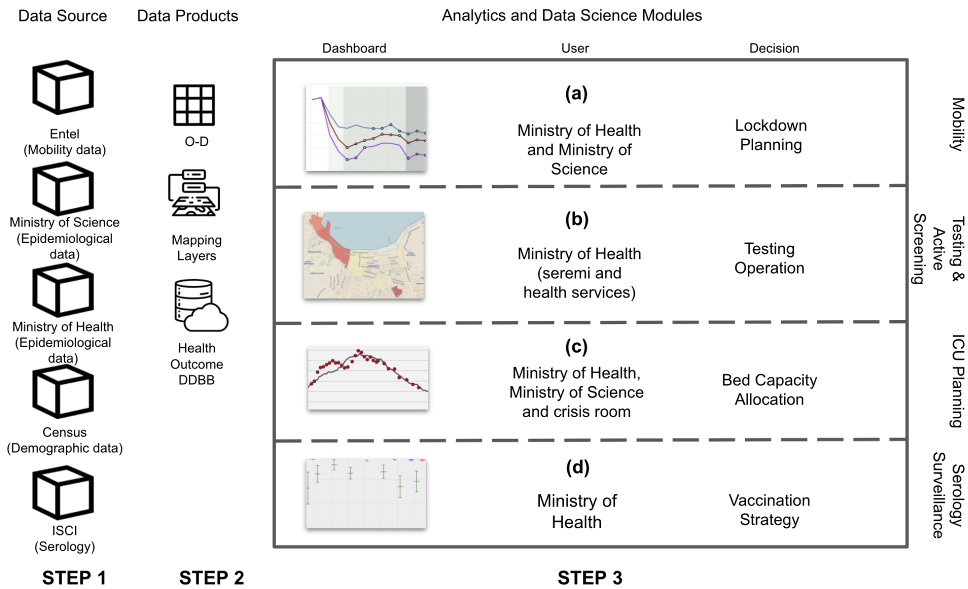


Figure 1 **The data and analytics pipeline of the project.**

As described in Figure 1, at the core of the project is the production of data on mobility patterns, using granular anonymous cellphone data provided by ENTEL. Another key data source was the Covid-19 data repository, a fundamental and strategic initiative of the Ministry of Science. The repository contains open aggregated data on confirmed cases, ICU occupancy, deaths, and the vaccination campaign, among other relevant information (and mobility data was added to this data hub as well). The mobility data was combined, on the one hand, with granular socioeconomic data and, on the other, with granular epidemiological data and serology data (Step 1 "Data Source" in Figure 1). All these data were cleaned and processed to produce several data products (Step 2 "Data Products" in Figure 1), including (i) origin destination matrices capturing mobility patterns; (ii) graphical mapping layers describing these and other population characteristics, including infections and demographics; and (iii) several health variables and outcomes. These data products then fed different analytics and data science modules (including dashboards and models) that allowed a series of decision-makers/users to take action regarding contagion prevention, management of hospital resources, and vaccination roll-out.

Complementing Figure 1, Figure 2 presents an alternative perspective on the project in the form of a timeline of the various milestones associated with the different modules in the project. We present such a timeline as a reference to have at hand while reading this manuscript, so as to help the reader envision of the unique situation the team and Chile was facing at any given point. Next, we describe the components of the project in more detail.

Regarding contagion prevention, our team transformed aggregated and anonymized telecom data into granular mobility data, which was then used to create public dashboards with this mobility

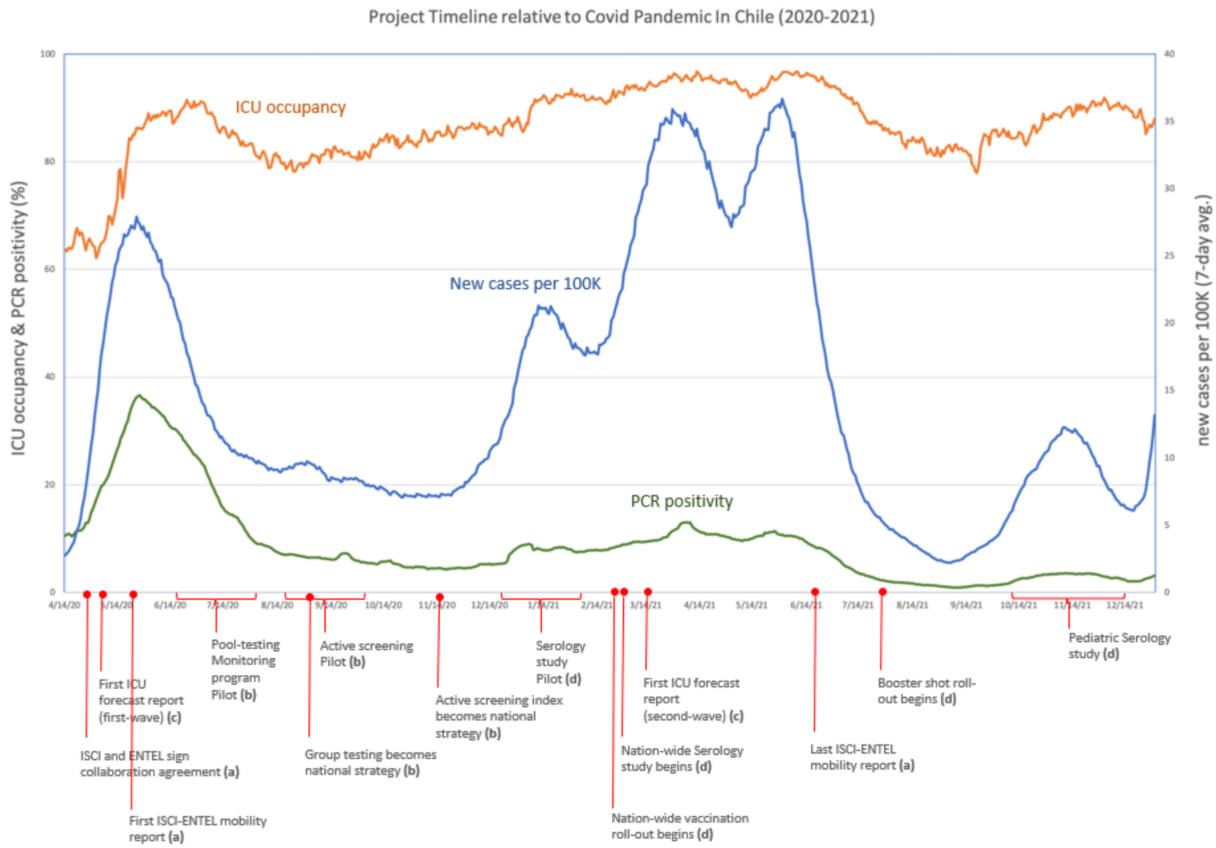


Figure 2 **Key pandemic indicators and project timeline during 2020 and 2021. Each milestone on the time-axis includes the module in Figure 1 with which it is associated with.**

data showing the impact of lockdowns and voluntary shelter-at-home decisions across the country. The graphs and data are downloadable and are routinely added to the Ministry of Science's GitHub. We also developed econometric models showing that the impact of social-distancing measures and lockdowns is highly heterogeneous and dependent on socioeconomic levels. The econometric analysis also showed that reducing mobility correlates significantly with the reduction of infections. The team produced 36 publicly available mobility reports describing these data and results, which provided key input to (i) to the government's lockdown strategy throughout the country, and ii) its plan to support lower income populations with complementary measures to increase compliance (Step 3a "Mobility" in Figure 1). An estimate of impact suggests that these efforts prevented between 6,500 and 12,800 infections, between 280 and 551 hospitalizations in intensive care units (ICU), and between 189 and 371 deaths in the first wave of the pandemic; see Section 2.2 for details.

With respect to contagion prevention through testing and contact tracing (Step 3b "Testing and Active Screening" in Figure 1), the team developed and implemented pilots on group-testing techniques, which shaped key aspects of the national testing strategy, significantly increasing testing capacity by more than 28% leading to savings of more than 236 million USD; see Section 3.3 for details.

In addition, using the mobility and granular geo-referenced epidemiological data mentioned above, our team developed index-based nationwide heat maps to guide surveillance testing efforts to detect asymptomatic cases in public spaces (Active Screening of Cases, which we refer to as "BAC" for its Spanish acronym), which were adopted as integral components of the national testing strategy. The BAC heat maps were adopted by all 16 regional Undersecretaries of Public Health, and all 29 Health Services (HS) that form the front line of the Chilean public health-care system. As of January 2022, the index continues to be used, and has proven instrumental in detecting tens of thousands of asymptomatic cases. An estimate of impact suggests that BAC prevented between 16,205 and 23,586 infections, between 599 and 886 hospitalazaitons in ICU, and bewteen 296 and 448 less deaths from November 2020 to January 2022; see Section 3.3 for details.

Despite the prevention efforts described above, many people were nevertheless infected, and a fraction of them were severely ill, thus requiring hospitalization. To support the centralized management of critical beds (Step 3c "ICU Planning" in Figure 1), the team used predictive machine-learning models to produce short-term demand forecasts of intensive care beds at the regional level throughout the country. These forecasts were key inputs to inform the coordinated efforts to adapt and augment the supply of this scarce resource, and reallocate patients across regions. Notably, the ICUs were never overdemanded in Chile (iCovid Chile 2022). An estimate of impact suggests that this initiative prevented between 467 and 1,1017 deaths from May 2020 to August 2021; see Section 4.3.

As the pandemic advanced, vaccines became a key instrument to prevent infections, severe illnesses, and deaths. Chile decided very early that it would follow a multi-platform approach, favoring availability over the choice of a specific vaccine technology. Chile thus signed contracts with vaccine manufacturers with different technologies, such as Sinovac (inactivated virus), Pfizer BioNTech (mRNA), and AstraZenaca (viral vector). This raised the challenge of generating information about the results of different vaccines that would not be easily obtained elsewhere. In the context of vaccine roll-out, the team designed and implemented a centralized surveillance system that monitored the presence of Immunoglobulin G antibodies (IgG) in adults and children, inoculated with different types of vaccines (Step 3d "Serology Surveillance" in Figure 1). The information provided by this system, which uses the mobility data to design the sampling mechanism in the general population and fed a statistical model of IgG waning dynamics, was instrumental to the government's decision to implement heterologous booster shots. Chile became one of the world pioneers in booster shots and, arguably, such boosters avoided a significant third wave in 2021, preventing between 6,600 and 29,000 infections, between 290 and 1,050 hospitalizations in ICU, and between 300 and 1,100 deaths; see Section 5.4 for details.

Table 1 summarizes the overall impact of the project and its four initiatives: an estimated total prevention of between 29,305 and 65,386 infections, between 1,169 and 2,486 hospitalizations in ICU, and between 1,252 and 2,936 deaths from May 2020 to January 2022. In Table 1 we also quantify the monetary savings associated with a decrease in workdays due to each infected individual and to costs associated with ICU hospitalizations. These, in addition to the savings associated with implementation of group-testing for SARS-CoV-2 massive diagnosis, amount to between 290.6 and 416.2 million USD. These calculations are explained in Appendix B.

Table 1 Summary of impact and cost savings

Module	Infections	Lost workdays (M$)	ICU hospitalization	ICU costs (M$)	Deaths
Mobility platform					
Conservative	6,500	2.4	280	10.48	189
Less Conservative	12,800	4.7	550	20.64	371
BAC					
Conservative	16,205	5.9	599	22.46	296
Less Conservative	23,586	8.6	886	33.23	448
ICU forecast					
Conservative	-	-	-	-	467
Less Conservative	-	-	-	-	1,017
IgG surveillance					
Conservative	6,600	2.4	290	10.88	300
Less Conservative	29,000	10.6	1,050	39.38	1,100
Total					
Conservative	29,305	10.7	1,169	43.8	1,252
Less Conservative	65,386	23.9	2,486	93.2	2,936
Total $ Savings			**plus Test Savings**		
Conservative	**M$ 54.6**		**M$290.6**		
Less Conservative	**M$117.2**		**M$416.2**		

The initiatives in the project received a tremendous amount of attention in the press and the general population. A sample of the coverage can be found here, and all media coverage of ISCI activities, including Covid, may be found here (in Spanish). As an example, only the mobility platform alone was covered by more than 25 TV and 35 radio interviews of the researchers and more than 120 newspaper articles commenting on the reports generated by ISCI researchers.

To summarize, the collaboration between the Chilean government, ISCI, and ENTEL resulted in a data-driven and science-based decision-making approach that helped the people of Chile endure the pandemic. In the words of the Undersecretary of Public Health (2018-2022) Paula Daza, "This collaborative work had a strong impact on the response capacity that we have developed as a country to face the pandemic", to which the Minister of Science (2018-2022) Andres Couve adds "We achieved the remarkable and challenging goal of coordinating between the government, the scientific community, and the private sector. And we showed that it can be done in a small and faraway country that is preparing with the scientific community to tackle future challenges. And we also showed that Chile can be an example of how science can be used in public policy."

Organization of the paper. In Sections 2 to 5 we provide details on each of the four initiatives of the project, in terms of the challenges faced, the analytics behind the solutions implemented, and the estimation of impact associated with them. In Section 6 we elaborate on the various challenges faced during the two years of execution of the project, and on the opportunities for transportability associated with the project. Section 7 presents our conclusions. The Appendix includes further technical details about the methods and models implemented, and their analysis.

2. Mobility and Contagion Prevention

2.1. Background

During the early stages of the pandemic, lockdowns aimed at inducing social distancing were fundamental to slowing down the spread of the virus. Following the first cases in Chile in early March of 2020, the rapid spread of infection – with a doubling time of 3 days on average – led to school closures and lockdowns in the city of Santiago during March, where most of the cases were concentrated. As shown in the top panel of Figure 3, most of the cases at that point were located in high- and middle-income municipalities, which led health authorities to adopt a *localized lockdown* strategy, enforcing mobility restrictions in these areas with a higher case incidence.

The top panel of Figure 3 shows the evolution of cases in three groups of municipalities: the first two groups include the high- and middle-income localities where the first lockdowns were located, and the third group includes the municipalities in the lockdowns that followed (starting in the third week of April), the lower-income population lives. The dots on each curve show the periods in which each group was in lockdown, revealing a significant disparity of these lockdowns in their role in better controlling the infection: lockdowns and school closures were effective at reducing cases in high-income areas, while in middle- and low-income areas cases continued to grow at a faster rate than high income areas that were not under lockdown.

Unfortunately, the infection data revealing this heterogeneous effect of lockdowns across regions came too late. For one, interventions intended to promote social distancing are aimed at reducing transmission, and their effect on secondary infections are observed with a week of delay, which

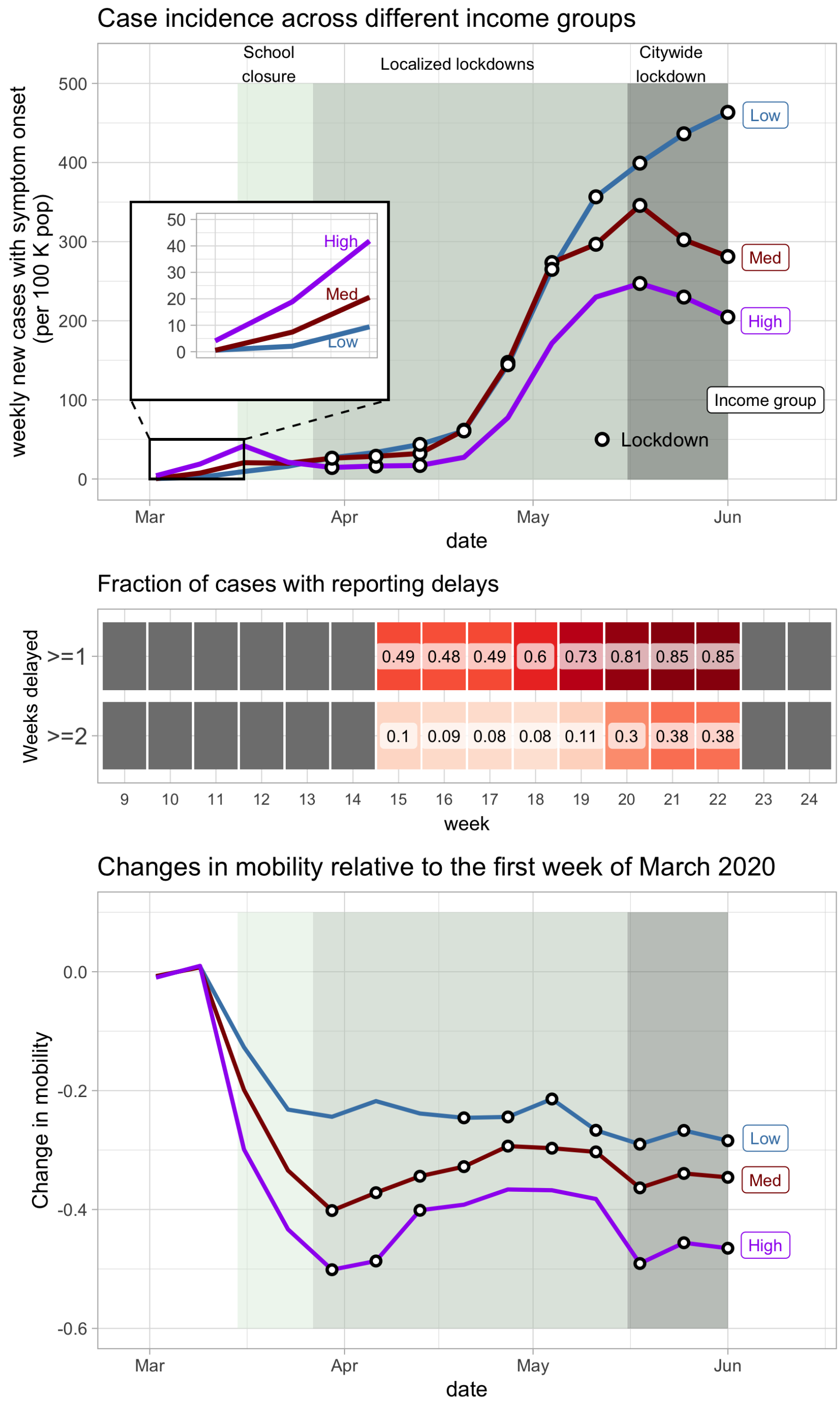


Figure 3 **The top panel shows weekly new cases / 100 K population for three groups of municipalities. White dots indicate weeks under lockdown for each group. The middle panel shows reporting delays, measured from symptom onset to case confirmation. The bottom panel shows the changes in mobility for these three groups of municipalities, relative to the mobility observed in the first week of March 2020.**

is approximately the average serial interval for the original strain of Covid-19 infections (Lauer et al. 2020, Rai et al. 2021). In addition, the dramatic escalation of cases coupled with insufficient testing capacity – positivity rates reaching as high as 38% at the end of May (see Figure 2) – led to significant delays in the reporting of infections. The middle panel of Figure 3 shows the delay in the reporting of new cases, indicating the proportion of cases reported with more than a week of delay (≥ 1) and more than two weeks (≥ 2) after symptom onset (Arroyo et al. 2020). Note that by mid-April, about half of the cases took more than one week to be confirmed, and by the end of May this proportion reached 85%. At that time, about 40% of cases took more than two weeks to be confirmed following the start of symptoms, with an average delay of 9.5 days. Because of these reporting delays, there was little visibility into how the infections were evolving and therefore no prompt feedback on the effectiveness of lockdowns in controlling outbreaks: by the time the cases were reported, the situation was already out of control. Hence, it became evident that alternative indicators were needed to provide timely feedback on the effectiveness of lockdowns in order to plan them efficiently.

2.2. Analysis of Mobility Patterns

Tracking the mobility patterns of the population presented an opportunity to monitor whether lockdowns were indeed inducing social distancing in the population, thereby providing valuable information *before* outbreaks occured. Researchers around the world were using Google mobility reports and other region-specific data from cellphones to track mobility. However, these data sources were not well suited for the health authorities to assess the localized lockdown strategies used in Chile because (i) the data sources were not sufficiently granular to monitor mobility at the level at which the lockdowns were implemented, and (ii) they provided information about movements within a given area, but did not provide information about travel patterns between households and destinations, which was useful to understand the spread of outbreaks in order to plan localized lockdowns.

In collaboration with ENTEL Ocean, the digital solutions branch of ENTEL, we constructed mobility indicators at a more granular level, based on the usage of telecommunications infrastructure: devices are geo-referenced during each connection using antenna triangulation, which provides a more representative sample of devices relative to the alternative of using GPS-based location. Connections during night hours (9—11pm) were used to assign a *household* to each device, which was aggregated at the census zone level to maintain anonymity. (Each census zone is composed of several census blocks and has an average population of 3,661.) To identify mobility patterns, we tracked the location of each anonymous device during two time blocks during working hours (10am—1pm and 2—5pm), aimed at capturing work-commuting patterns during weekdays, excluding weekends and holidays. We then assigned a location to each time block, using the most frequent census zone where the device was positioned. The data was then aggregated using origin—destination matrices at the census zone level, thus maintaining anonymity of travel patterns; see Carranza et al. (2022) for details on how these mobility indicators are calculated.

The bottom panel of Figure 3 shows changes in mobility relative to the first week of March, considered as the pre-pandemic baseline. The figure reveals a huge disparity in mobility reductions across municipalities: higher-income areas reduced mobility by 40%—50%; middle income areas reduced mobility by 30%—40%, and lower-income areas reduced mobility only by 20%—25%. This heterogeneous compliance with shelter-in-place mandates is an important factor in explaining why lockdowns were not as effective at controlling infections during the first wave of cases. A key advantage of having these mobility measures promptly available is that health authorities can respond faster to this information (compared to a response based on infection rates). Further, the heterogeneous socioeconomic response can suggest other measures to complement lockdowns and reduce social distancing, such as food and financial support.

Beyond these descriptive statistics, we conducted a rigorous econometric analysis of these mobility patterns with two objectives in mind. First, we sought to measure the impact of lockdowns in reducing mobility, and to better understand why this effect varied across locations. Second, we studied if there was a causal relationship between reductions in mobility and controlling infection rates, which was critical to validate that these new metrics were useful for anticipating infection outbreaks, thereby providing faster feedback on the effectiveness of lockdowns. For such purposes, we developed panel data linear regression models that tackled these two objectives. The models exploit the staggered implementation of localized lockdowns across municipalities to separate seasonal effects from the causal impact of the lockdowns on mobility. See Carranza et al. (2022) for more details on the model and results.

During the implementation of localized lockdowns, the high-income areas reduced their mobility twice as much as low- and middle-income areas; similar differences were observed during the citywide lockdown period. Furthermore, by comparing infection rates across regions that were subject to similar lockdown conditions, we measured how differences in mobility are associated with virus transmission. We constructed a specific measure of mobility for this purpose, which measures the risk of infection on a focal location based on travel patterns from other locations, accounting for asymptomatic cases from these origins. We showed that this measure of risk-adjusted mobility accounts for 37% of infection rates in low-income populations due to their higher mobility. This risk-adjusted mobility metric also became useful for better targeting testing efforts to detect asymptomatic cases circulating in the community, which we explain in the next section.

Many of the mobility indicators constructed fed an online visualization platform that was made available to the general public and health authorities. The platform visualization interface is shown in Figure 4; figures are customizable and downloadable, as is the generating data, which has also been made available at the GitHub of the Ministry of Science.

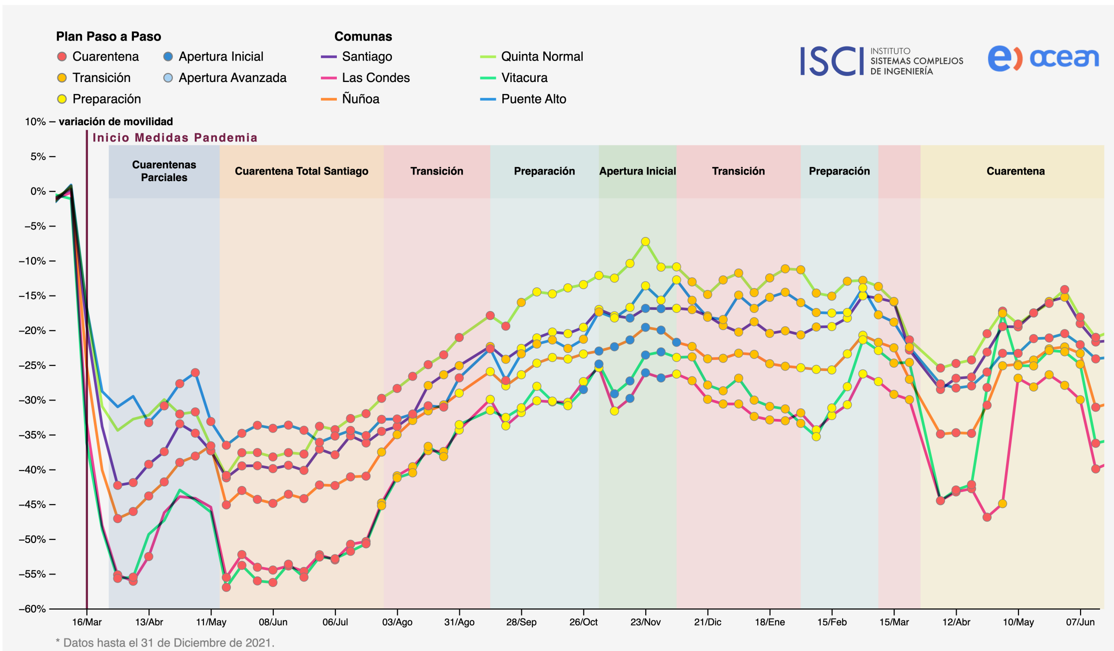


Figure 4 **Mobility visualization platform. The graph shows a weekly time series of mobility during 18 months since the start of the pandemic, measured as the relative change in the fraction of the population leaving their census zone compared to pre-pandemic baseline. Dots indicate the level of mobility restrictions of the municipality on each week. Background panels indicate the predominant level of restrictions for the city. This public-access platform is available at ISCI's website.**

The mobility platform received huge coverage by the media, with more than 25 TV and 35 radio interviews of the researchers and more than 120 newspaper articles on the reports generated by ISCI researchers. Providing public access to this information and its broad dissemination through the media helped to build confidence among the population that the huge efforts that were made to comply with social distancing were indeed having an impact on mitigating the pandemic. The mobility platform was expanded to include other major cities in Chile and trips between cities, which was useful for monitoring travel patterns during the summer vacation period (January—February 2021).

2.3. Impact: Increasing Shelter-in-Place Adherence

The information provided by this nationwide mobility platform became a useful tool for the health authorities to assess the localized lockdown strategy used in Chile. The main impact of these mobility indicators was: (i) to increase confidence among the population that social distancing was effective at reducing virus transmission; (ii) to provide information to the health authorities that would enable them decide where and when to impose social distancing mandates and provide resources to increase the public's compliance with them (e.g., distributing food and financial support in low-income neighborhoods); and (iii) to identify areas where lockdowns were simply not working and other mitigation measures were needed to mitigate transmission (personal communications with the Minister of Science Andres Couve).

Altogether, these efforts contributed to reducing the mobility of the population and thereby lowering infection rates. It is difficult to isolate the specific contribution that the mobility platform had in the implementation of these multiple initiatives. Hence, we used a conservative approach to quantify this potential contribution, evaluating the impact of a small change in mobility due to an increase in shelter-in-place compliance. We used one standard deviation (σ) of the risk-adjusted mobility observed during the March—July period as a reference to measure changes in mobility. A 5-percentage-point improvement in the compliance with shelter-at-home mandates translates into a $\sigma/10$ reduction in mobility, equivalent to one-tenth of the observed standard deviation, a relatively small change. Let's assume that the implementation of the platform generated this small mobility reduction during a limited time period, from the beginning of June 2020 (the first press releases of the platform) to the beginning of August 2020 (the lifting of the citywide lockdown), a period of 10 weeks in total. Using the econometric model that we developed in Carranza et al. (2022) to link mobility with infections, this small change in mobility translates into a 1.2% reduction in the weekly infection rate, equivalent to 6,500 fewer cases during this 10-week period (4.5% of the infections registered in the period). In turn, this implied 279.5 fewer ICU hospitalizations and 189 fewer deaths (calculated based on the hospitalization and mortality rates from Table 2 in Appendix B). If we assume a less conservative estimate of a 10% increase in shelter-at-home compliance due to the platform implementation (a $\sigma/5$ reduction in mobility), the calculation yields 12,800 fewer infections, 550 fewer ICU hospitalizations, and 371 fewer deaths.

As shown in Figure 4, the mobility indicators suggest a *lockdown fatigue* effect, where mobility became less responsive to shelter-in-place mandates over time, a phenomenon that has been documented in related work (Joshi and Musalem (2021), Li et al. (2022)). Although lockdowns were an effective measure for mitigating the spread of infections in the initial phase of the pandemic in some regions, their effect became weaker over time. Hence, alternative measures were needed to contain the pandemic. In the next section, we show how the mobility indicators constructed from telecom data improved the efficiency of one important such strategy: the testing and screening of asymptomatic cases that followed the first wave of the pandemic in Chile.

3. Testing and Active Screening of Cases

Given the large proportion of infections transmitted by pre-symptomatic, asymptomatic, and mildly symptomatic cases of Covid-19, massive testing in the general population became a useful mitigation measure for reducing community transmission (Mercer and Salit 2021). However, capacity was relatively low; in fact, positivity rates reached up to 40% at the peak of the first wave of the pandemic in Chile. Hence, it became critical to expand testing capacity and improve the efficiency of testing strategies. This section describes the efforts developed to achieve these objectives.

3.1. Group Testing Strategy

The first initiative was to implement a large-scale *group-testing* approach to expand testing capacity at lower costs. The group-testing technique consists of producing a combined sample out of individual samples, which is tested using a single PCR reaction in order to detect if *any* of the individual samples is infectious, in which case each sample is tested individually. This technique is effective at reducing laboratory costs and reporting delays when positivity rates are low (Dorfman 1943). The technique was first validated for Covid-19 by Yelin et al. (2020), but required a local validation for its use in Chile (Farfan et al. 2020).

In order to validate the method before using it in a large-scale nationwide implementation, we first conducted a pilot study through a group-testing monitoring program in long-term care facilities (LTCF), while accounting for the particularities of the setting (Basso et al. 2021). The success of the pilot showed that it was possible to diagnose patients at a much lower cost (in terms of PCR reactions), and therefore was adopted as a national testing strategy and thus followed by several laboratories. In particular, the number of tests performed via group-testing accounts for approximately 10 to 20% of the total testing conducted (Ministry of Health 2022a). We elaborate on the impact of this effort at the end of the section.

The use of the group-testing technique coupled with a decrease in positivity rates after the first wave (see Figure 2) allowed testing capacity optimization, which enabled the health authorities to initiate a testing plan to screen asymptomatic cases in the community.

3.2. Active Screening of Asymptomatic Cases

Unlike symptomatic individuals, who are tested once symptoms appear, asymptomatic individuals are generally unaware of their condition and therefore are likely to continue with their daily activities without adopting additional care or restrictions, thus spreading the virus in the community. Health authorities designed a testing strategy to identify these cases by implementing mobile PCR testing stations in public locations. We developed a system to efficiently locate these stations that combines mobility, granular georeferenced epidemiological data, and demographic information.

Specifically, we developed an active screening index (also referred to as the “BAC index”, after the Spanish acronym for *Búsqueda Activa de Casos*) that estimates the likelihood of finding asymptomatic cases in localized public-access areas. The index – which is computed for each urban census zone and each day – measures the risk of finding an asymptomatic case by weighting the positivity of the incoming people in a given census zone by the density of the inflow of asymptomatic cases. The latter term is incorporated to account for the volume and spread of asymptomatic cases in the area, given that stations have a limited testing radius. In this sense, the index combines epidemiological and mobility data. Further details of the BAC index computation are provided in Appendix A.

The BAC index was integrated into the Ministry of Health’s visualization platform (in the form of heat maps) in order to guide testing efforts throughout the country. The pilot was launched in October 2020, and the nationwide program in November 2020, covering a population of 15 million residents in 256 towns and cities (see Figure 2). The index, which is computed at the beginning of each week, is normalized so as to emphasize the relative comparison across geographic units, thus facilitating the visualization for the decision-makers that choose which locations to prioritize testing.

The independent dashboard developed by ISCI, also available to the health authorities, is shown in Figure 5. There, the census zones shown in dark red are the ones with the highest index value, whereas those in yellow have a lower index value. A higher index value implies a higher estimated chance of finding asymptomatic cases and, therefore, a better candidate locate for a testing station. Users were able to filter the index, e.g., by day of the week, which is an important feature, as mobility patterns vary depending on the day of the week and so does the BAC index, accordingly. This heat-map platform is updated every Sunday with the index values for the upcoming week.

3.3. Impact: Increasing Detection of Asymptomatic Cases

The success of the group-testing pilot led to the adoption of this technique as a nationwide strategy in September 2020, which in turn led to a dramatic increase in testing capacity: considering that between 10% and 20% of each PCR reaction used this technique, with a group size of 5 samples, this translated into an overall increase in testing capacity of 28% to 35%, that is, in the order of 7.5 to 9.5 million tests between the beginning of the pandemic and January 2022. The resulting savings were in the order of 236 to 299 million USD, assuming a cost of 31.25 USD for each PCR test.

The active screening efforts supported with the BAC index visualization platform detected 46,000 asymptomatic cases from November 2020 to January 2022, with a total number

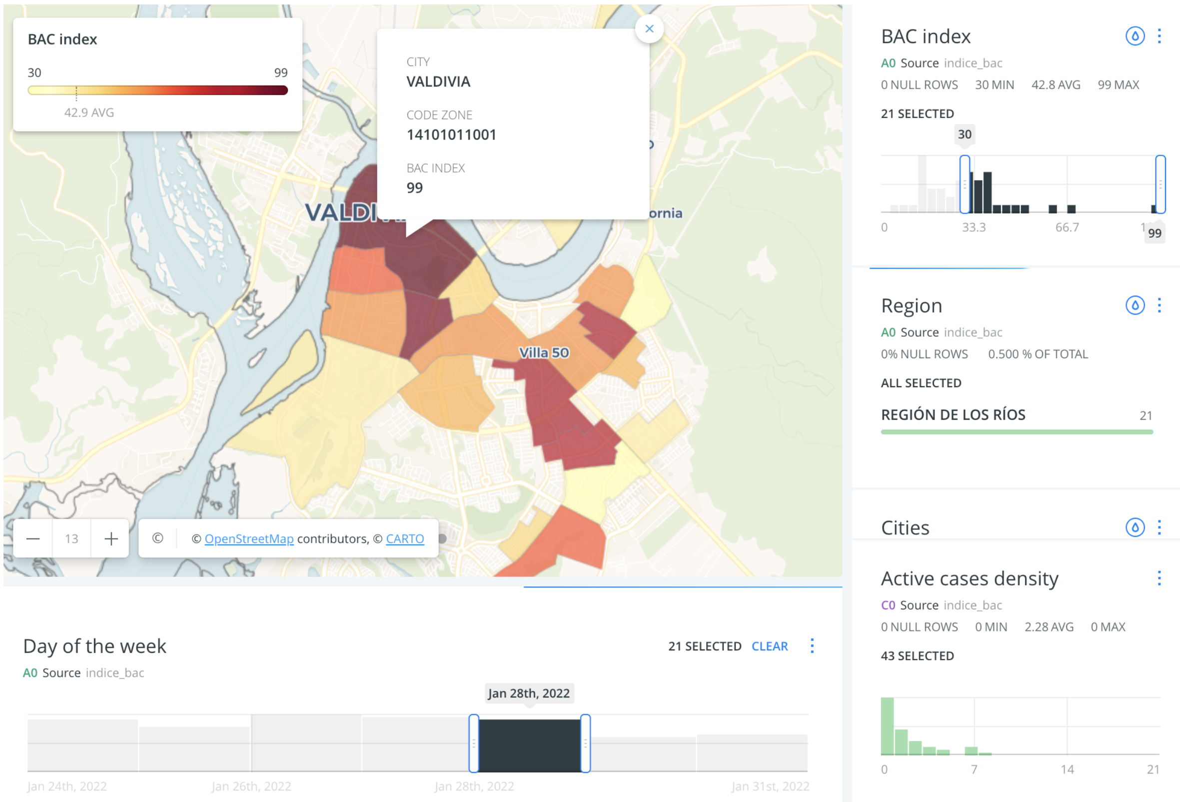


Figure 5 **Visualization heat-map tool of the BAC index. (The text of the visualization has been translated from Spanish to English.)**

of (active screening) tested individuals of 2,173,126 (see products 63, 64, and 65 at the GitHub of the Ministry of Science).

To estimate the impact of the Testing and Active Screening module, we considered a counterfactual scenario where the same number of tests were distributed at random in the population within each town and city. The test positivity under this counterfactual was 0.68%, compared to 2.12% using the BAC index, which yields 31,324 additional asymptomatic cases detected by the system.

Using the effective reproductive number (R_e) – which represents the expected number of secondary infections of an infected case (Grassly and Fraser 2008) – we estimated the averted infections from the detection of these incremental asymptomatic cases. We used only half of the estimated R_e since some of the infections may have occurred before detection (see Appendix B for details on the parameters used). The calculations yields 16,205 fewer infections, 599 fewer ICU hospitalizations, and 296 fewer deaths. A less conservative scenario, in which we compare to the alternative without any testing to detect asymptomatic cases, results in 23,586 fewer infections, 886 fewer ICU hospitalizations, and fewer 448 deaths.

4. ICU Capacity Planning

4.1. Background

In early May 2020, infection rates started to increase rapidly (see Figure 2), threatening the ability of the health-care network to accommodate all incoming Covid-19 cases. By that time, the Ministry of Health decided to assume centralized control of all ICU beds in public and private hospitals in order to provide enough critical care to all those who might need it. This centralized design considered expanding bed capacity at the regional level and moving patients across regions to balance bed utilization. These decisions required a detailed forecast of how many beds were going to be needed in the near future for each region in the country.

By mid-May, the Chilean Society of Intensive Medicine (SOCHIMI) reported a worrisome occupation rate of ICU beds of more than 95% in the capital city of Santiago, where most of the cases were concentrated; hence, ICU capacity planning became a first-order concern. On May 12, we were urged to prepare short-term forecasts of ICU occupancy rates for those regions with the highest utilization rates. Within 24 hours, we submitted our first report. From then on, we prepared forecasts every two days for several weeks during the whole duration of the hospital crisis.

4.2. Short-term ICU Occupancy Forecasts

The forecasts were based on compartment a model where patients stochastically evolve through different states (details of the model are described in Goic et al. (2021)). For each region, we replicated the behavior of the ICU process, balancing inbound and outbound flows of patients in three different state variables: first, the number of infectious individuals who show symptoms of Covid-19; second, the number of critically ill people who need an ICU bed; and finally the number of individuals who are discharged from an ICU. To describe transitions between states, we estimated how likely it was that a given patient would evolve to another state, and a probability distribution for the duration of that transition. Our model considers that these events depend on the characteristics of the patients, and that the duration of stay can be highly heterogeneous. For instance, while some patients require ventilation for only a few days, other stay in an ICU for several weeks (Vekaria et al. 2021). To accommodate these variations, we used flexible distributions to characterize the length of stay, even allowing for bi-modal distributions.

The calibration of the compartment model required precise estimates of the relevant epidemiological and clinical parameters, which were expected to change as the pandemic evolved. For example, the proportion of patients requiring mechanical ventilation can change over time, and so can the clinical criteria for releasing them from the ICU: because Covid-19 was a new virus it involved continuous learning by medical teams. In this regard, the Chilean Society of Intensive Medicine indicated that compared to early cases, patients were reducing their length of stay due

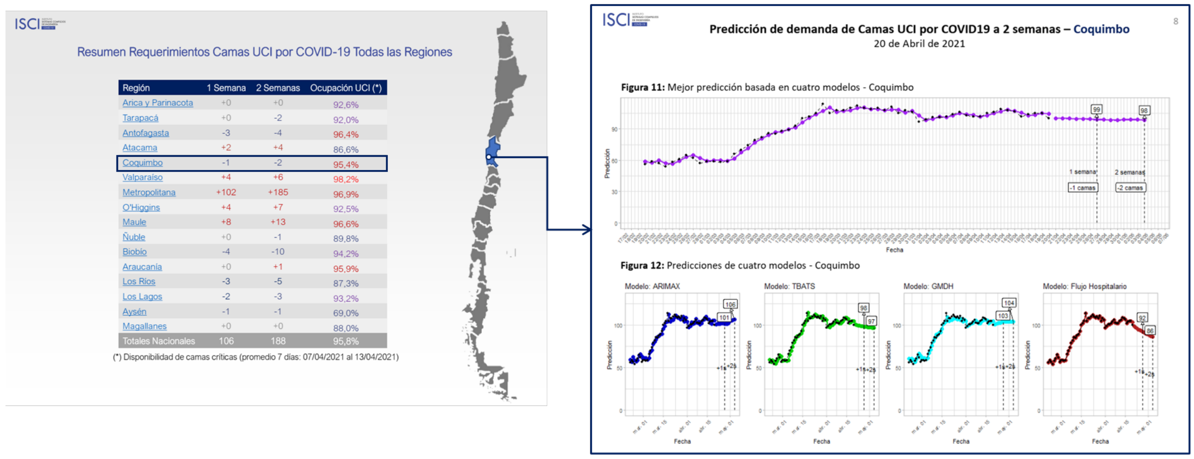


Figure 6 **Schematic representation of ICU forecasts dashboard.**

to less frequent use of mechanical ventilators and a more effective selection of the more serious cases. Finally, the systems for generating the data were under constant stress and, therefore, the information that we had available could be lagged.

To accommodate all these short-term variations in the process we combined the compartment model with a variety of autoregressive and machine learning models that have been shown to provide flexible estimation of complex dynamic interactions (Zhang 2003, Zhang and Qi 2005). To produce an integrated forecast for each region, we used ensembles that have been shown to provide more accurate predictions in different domains (Montgomery et al. 2015, Wu and Levinson 2021). Considering that the medical staff in charge of ICU capacity decisions had a very intuitive interpretation of the compartment model, we always included it in the ensemble.

These forecasting models were built into a fully automated dashboard system that generated detailed reports as needed (we generated 30 in 2020 and 26 in 2021). Figure 6 provides a schematic illustration of these dashboard-style reports, which were designed to facilitate a quick evaluation on the part of health officials and SOCHIMI. For each region, we provided a graphical summary of the actual requirements and our predictions for the next 14 days and, we highlighted the number of additional ICU beds that would be required in exactly one and two weeks ahead. The report included a summary table with one- and two-week predictions for all country regions, allowing the health authorities to visualize those regions in most critical need of new beds.

4.3. Impact: Benefits of Accurate Forecasting

Each short-term ICU forecast was reported by the Minister of Science to the President and his Crisis Advisory Committee, who used them as input for focusing efforts to adjust the supply of intensive care beds. In particular, this interaction resulted in a progressive and directed increase in the number of ICU beds, leading to more than doubling the national capacity of ICU beds from

1,563 at the beginning of the pandemic to 3,102 beds at the peak of the first wave, and then to 4,538 beds at the peak of the second wave. Most of this additional capacity resulted from converting traditional beds into ICU beds. To put this effort into context, this increment allowed hospitals to provide critical care to 7,870 patients, saving a large fraction of them.

Our forecasts were instrumental to assisting health officials to increment capacity at the right pace. Certainly, larger forecasting errors would lead to different costs depending on the direction of the error. While overpredictions are associated with larger investments and longer delays of other medical procedures (Sud et al. 2020), underpredictions would have implied a potentially high number of Covid-19 patients without critical care. To assess the impact of having more accurate predictions, we assumed that the ICU bed capacity decisions closely follow the forecasts, which is reasonable because conversion of ICU beds is expensive and has a large opportunity cost in terms of other procedures; hence, large safety stocks are rarely allocated. Monte Carlo simulations indicate that having worse forecasting accuracy would have caused between 575 and 1,151 individuals not to receive the mechanical ventilation they needed. Considering the ICU survival rate derived from Table 2, we estimate that our system directly assisted in saving between 467 and 1017 lives. See Appendix C for more details.

Our forecasts also helped to adjust the lockdown plans and other policies in order to prevent ICU capacity overflow. This proactive planning of ICU capacity was successful in preventing the overflow of patients into ICUs that was observed in other countries and, later in the process, in allowing hospitals to readily re-schedule a significant number of surgeries that were delayed during the most critical phases of the pandemic.

5. Serological Surveillance

5.1. Background

By the second half of 2020 it became clear that vaccine roll-out would become central in most countries' strategies to contain the pandemic, once vaccines became available. It was also clear that vaccine supply was uncertain, and that prioritization of high-risk groups would likely be required. At the time, Chile had acquired a significant amount of the inactivated virus vaccine CoronaVac from the Chinese manufacturer Sinovac, but did not have priority access to large volumes of mRNA (e.g., Pfizer) or viral vector (e.g., AstraZeneca) vaccines. At this juncture, the government opted for pursuing a vaccine roll-out strategy that would combine vaccines with different technologies. This strategy quickly placed Chile as a leader in the vaccination campaign worldwide, with about 75% of its population vaccinated with CoronoVac. However, it was unclear whether this inactivated virus vaccine would have a lasting effect on the population. Since the Chilean vaccination strategy was unique, health authorities could not rely on international studies on vaccine protection, and therefore needed to develop their own system to monitor vaccine effectiveness. Further, a key aspect was that the system would need to have the capacity to evaluate changes in potential vaccine protection over time and for the different vaccines.

Experience with the implementation of lockdowns revealed that waiting to analyze large-scale infection data in order to monitor the effectiveness of the vaccine would lead to a long delay to take action. An alternative was to implement a nationwide serology study to monitor the immune response of the inoculated population over time and evaluate how this was affected by prior infection, demographic variables, comorbidities, and the vaccine used, among other factors. A key design aspect of this serological monitoring system was to decide the location of the antibody testing stations in order to maximize the representativeness of the sample. An efficient design would increase the statistical power of the study, which in would turn reduce the time needed to reach conclusions and take action.

A first pilot program was implemented in March 2021, in parallel with the start of the nationwide vaccination campaign, in which antibody testing stations were installed at high-traffic locations in Santiago. The locations were chosen based on the mobility patterns that were used to construct the mobility platform and the BAC system (see Figure 1).

The tests were free of charge for voluntary patients and individuals who volunteered, signed an informed consent, and were tested for IgG antibodies using a lateral flow test (LFT) that produced results within 15 minutes; while waiting for results, subjects responded to a web-based questionnaire (under a health-care worker's supervision) asking for clinical and geo-demographic data, which populated our database in real time.

The pilot concluded successfully with over a thousand people tested, and the results were made publicly available through a web-based dashboard. Among lessons learned during the pilot were the need to account for evolving vaccination status information (e.g., number of doses), and to carefully communicate the non-diagnostic nature of the test and that a negative result did not necessarily translate into lack of protection.

5.2. Nationwide Implementation

Starting in March 2020, 29 testing stations were deployed throughout the most populated cities in Chile (see Figure 2). The operation was carried out by personnel from the 29 health services that form the front line of the Chilean public health-care system. The testing stations were located strategically so as to obtain a geographically representative sample of the general population. For this, we used our data on mobility patterns and demographics to formulate an integer program (IP) that selected the census zones in which to install testing stations so that the geographical distribution of the overall sample collected would replicate that of the general population, subject to some operational constraints. Because mobility patterns change from day to day, sometimes from morning to afternoon, the model considered the decisions of where to locate each station, on what day, and in which time block (morning/afternoon). The model considered a planning horizon of one month, which allowed us to maintain representativeness over time during the duration of the study. The overall problem is separable across Chile's 16 regions (first-level administrative division); thus, we solved 16 separate formulations to find an optimal nationwide allocation. The formulation of the IP can be found in Appendix D

5.3. Operation and Results

The serological study described above operated for about 9 months, during which over 70 thousand samples were collected. Throughout the study, we constantly monitored the sample, in terms of its representativeness with regard to various sociodemographic variables (most notably the geographic one, which was the one that we were actively managing) so as to detect any anomaly. In this regard, because of idiosyncratic discrepancies between mobility data and realized mobility, biases associated with testing volunteers in public spaces, and unplanned modifications to station allocation (e.g., at times some stations could not change locations at the required frequency, and some locations were either permanently or sporadically unavailable), the geographic representativeness of partial samples did not always match that prescribed by the solution to the IP model. We dynamically corrected partial deviations from the planned sample by implementing the IP solution on a rolling horizon basis, fixing the composition of the collected sample, and planning data collection for the rest of the horizon. Figure 7 shows the final composition of the sample for the most populated city in Chile.

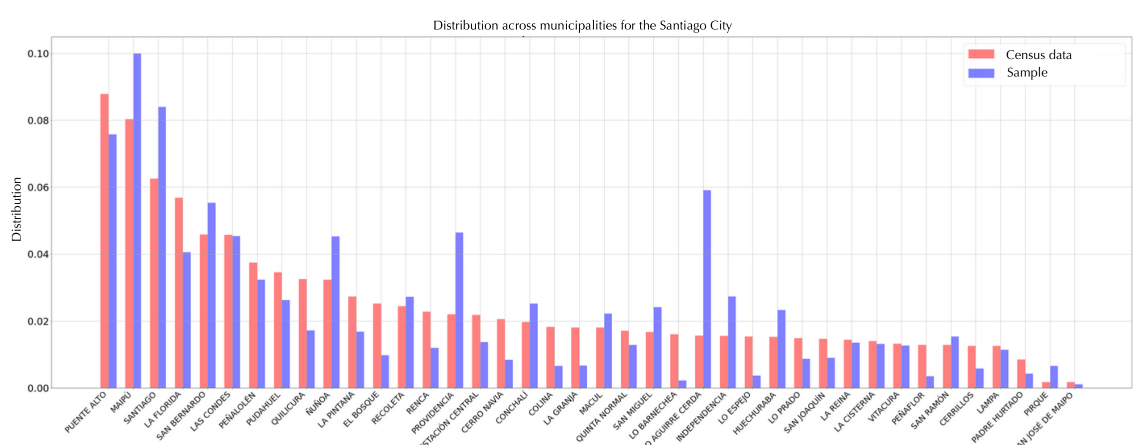


Figure 7 **Sample vs. census population distribution across the most populated municipalities of Santiago.**

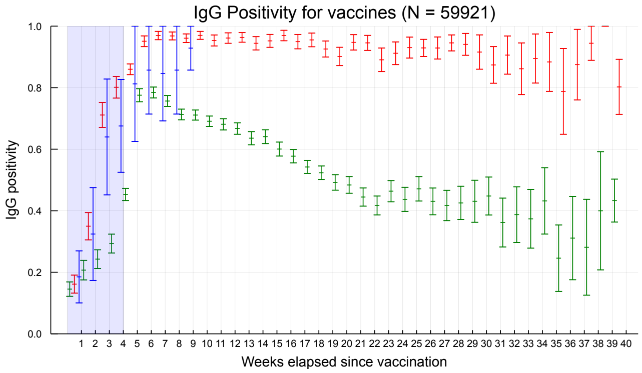
By September 2021, most of the samples being collected corresponded to individuals with either partial or full vaccination schemes. A first analysis showed that the collected data allowed us to monitor the dynamic IgG response to vaccination in the general population. In light of the result above, the study was extended for six more months, in the form of a monitoring program of IgG dynamics in a vaccinated population (where geographical representativity was less relevant) that continued operating with a reduced number of stations, collecting data mostly on IgG response to primary and booster doses. By the time of this writing, the aforementioned monitoring program had tested over one hundred thousand volunteers from the largest cities in Chile.

Because of the profound public policy implications of our results (see below), we extended our study along several dimensions. On the one hand, with the support of the Ministry of Education, we extended our study to include the pediatric population, which was incorporated into the vaccine roll-out program late in the second half of 2021, see Torres et al. (2022b). Because inoculations were carried out at public schools and during school days, our study formulated an IP model to select about two dozen schools, across three regions, so as to obtain a representative sample of the pediatric population of those regions. In said IP, school enrollment data took the role that mobility data had in the nationwide serological study. On the other hand, we are currently studying the relationship between IgG positivity and quantitative levels of neutralizing antibodies, which will help connect IgG positivity to more robust measures of protection.

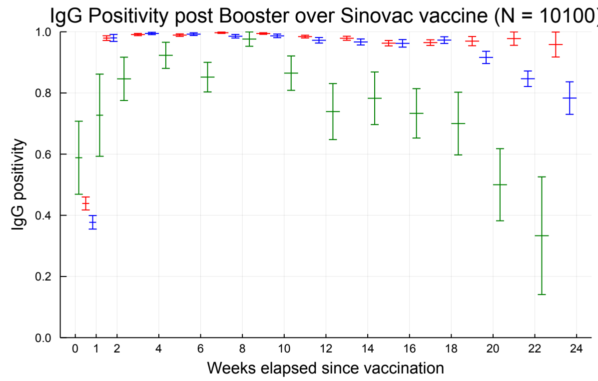
5.4. Impact: Advancing Booster Roll-out

The centralized IgG surveillance system demonstrated the waning of IgG positivity among recipients of a primary vaccination scheme using Sinovac (about 75% of the eligible population) after the second week since receiving a second dose, specially among those 60 years of age and older; see the left panel in Figure 8, and Sauré et al. (2022) for more details. This information was crucial to supporting the decision to begin rolling out booster doses in early August (see Figure 2): heterologous regimes using either Pfizer or AstraZeneca as boosters (but not Sinovac) were prioritized, specially among older segments of the population. More recently, information collected by the surveillance system guided the decision regarding a second booster dose (right panel in Figure 8; see Torres et al. 2022a for more details), which began roll-out in January 2022.

The IgG surveillance system can be credited in part for the early roll-out of booster shots in Chile. In order to quantify the impact of this decision, we estimate: i) the overall impact of booster doses using data available reporting the incidence of infections, ICU hospitalizations, and deaths during September–December 2021 for populations without booster shots, and ii) the impact of



(a) IgG dynamics following primary vaccination scheme



(b) IgG dynamics following booster dose

Figure 8 IgG positivity among recipients of primary vaccination scheme (left panel) and booster dose over two-dose Sinovac primary vaccination (right panel); error marks depict 95% confidence intervals; the series correspond to the Pfizer BioNTech (red), Oxford/AstraZeneca (blue), and Sinovac (green) vaccines.

earlier implementation by considering the alternative scenario where the decision is delayed until booster doses are approved by the US Food and Drug Administration (FDA).

Regarding i) above, Appendix E provides further details about the data and the model used for estimating the impact of the booster shot on health outcomes. Using these predictions, we estimated for each week and each health outcome the number of occurrences that were prevented, considering the actual population of different age groups that received the booster up to that week. The weekly averted occurrences for each outcome are illustrated in the top panel of Figure 9, where the colors indicate different age groups. Averted infections were predominantly in the younger group, where transmission is higher. Averted hospitalizations and deaths were concentrated in the older age groups, which is the main reason why these age groups were prioritized during the booster vaccination campaign. The bottom panel of the figure shows the cumulative averted cases at the end of each month.

Regarding ii), the FDA approved Pfizer boosters during November 2021; thus, the alternative roll-out would have been delayed by 2 to 3 months, which we used to calculate the lower and upper bounds of the impact of the serological surveillance system. Delaying the campaign by 2 months would have caused 6,600 additional infections, 290 ICU hospitalizations, and 300 deaths, which correspond to the conservative estimates of the impact. An additional month of delay (3 months total) would have causes infections to increase by 29,000, leading to 1,050 additional ICU hospitalizations and 1,100 additional deaths.

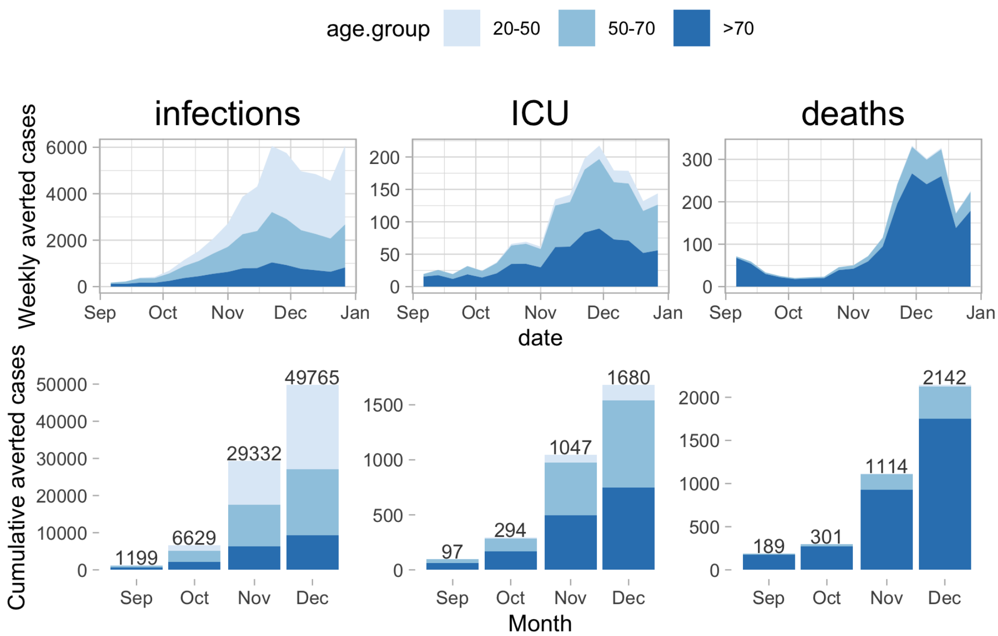


Figure 9 **Averted infections, ICU hospitalizations, and deaths associated with the booster shot campaign through September–December 2021. The top panel shows the weekly averted cases for different age groups. The bottom panel shows the cumulative averted cases at the end of each month.**

6. Implementation Challenges and Transportability

6.1. Implementation Challenges

In addition to the technical difficulties that were encountered in developing the different modules of the project, there were also significant practical challenges at different stages of the implementation that set the project apart from more traditional implementations of analytics and operations research. These challenges needed to be promptly addressed in order to enable a timely adoption of the system and maintain engagement of the different stakeholders.

A first major challenge was urgency: as opposed to other applications of analytics that can be planned carefully, here we were dealing with an always changing situation, where needs arose suddenly and solutions had to come swiftly. Responding in such an environment required a significant amount of spirit, and continuous improvement.

A second challenge was the coordination of a large number of professionals with different backgrounds and working in different institutions, each of which were fundamental in the development of the system. The implementation required building collaborations among: (i) engineers from different institutions with training in data science and modeling, who were in charge of developing the data pipeline and analytics backbone of the system; (ii) public officials in charge of adapting the regulatory environment that was needed for the project to be implemented in practice; (iii) health-care professionals that formed the front line workforce in charge of making the system work in the field; and (iv) public relations teams from many institutions, in charge of communicating the efforts and results obtained in the context of the project. Many of these professionals were also geographically dispersed, which further complicated coordination in the nationwide implementation of some of the modules.

A third challenge was to encourage participation and adoption by the different health authorities that form Chile's decentralized health-care system. Although there was full alignment among the high-level leadership of the government (i.e., at the ministry level), a full-fledged nationwide implementation required the sponsorship of the local authorities in each of the regions in Chile. For example, the launch of the Testing and Active Screening module was implemented simultaneously in four cities in the presence of health authorities and ISCI researchers at each location, which included media coverage on national TV, and then extended nationwide with similar media coverage.

A fourth important challenge was to define the data that could be actually used as an input to the system. Early in the project, it was clear that mobility data would provide key information to guide decision-making, but it was fundamental to mitigate any concerns related to consumer privacy that could hinder the use of these data. Similar concerns were raised in using granular epidemiological data that was not publicly available. Models needed to be adapted in order to use anonymous data that was properly aggregated in order to protect the confidentiality of users and patients; research teams had to agree on nondisclosure agreements and work through an iterative process among the institutions involved in order to reach consensus on the precise level of granularity beyond which the data could not be used, balancing model precision with privacy concerns.

Fifth, we faced a critical challenge on how the implementation of the system and the information that it was providing was communicated to the general public. During the implementation of the Mobility module, there was strong evidence that certain municipalities were exhibiting low levels of adoption of the quarantine mandates. Revealing this information without an adequate level of communication could generate resistance in the population if it was viewed as too invasive or controlling. To mitigate this risk, ISCI researchers were directly involved in the communication campaign through the media, highlighting the scientific evidence provided by the mobility platform, mitigating concerns related to user privacy, and suggesting the need to provide aid to those areas that were facing difficulties in complying with lockdown mandates.

Similar communication interventions were needed to disseminate the findings of the Serology Surveillance module, which was revealing a decay of the (IgG) immunity response of the CoronaVac vaccine in the elderly population after 3 months of inoculation. It was important to highlight the value that this early and fast vaccination campaign, focused initially on this high-risk population group, had in reducing mortality and hospitalization rates during the second wave of cases in

Chile, in which the Gamma variant was prevalent in South America. Researchers in Chile from different disciplines worked in several studies that provided scientific evidence of the effectiveness of the vaccination campaign during this second wave, thereby demonstrating that it had fulfilled its purpose within the overall vaccination plan that Chile pursued (Jara et al. 2021, Sauré et al. 2021). These studies complemented the evidence provided by the Serology Surveillance module, showing that there was indeed an immune response from the vaccine during those critical months, but that it was necessary to reinforce it through the booster shots in anticipation of the Delta wave. These conclusions – which were derived from the results published in Sauré et al. (2022) – were broadcasted on national TV, in a joint presentation by the Ministry of Health, the Ministry of Science, and ISCI researchers. Reinforcing the positive aspects of Chile's unique vaccination strategy based on scientific evidence is one of the factors that contributed to the high level of vaccine acceptance in the Chilean population.

Finally, many of the solutions proposed were quite delicate in nature: they revealed confidential data, anticipated potential scenarios with high morbidity/mortality levels, bad outcomes for the community, or spoke to the efficient use of public resources. In this regard, each proposed solution had to be thoroughly discussed with the health authority before their dissemination/implementation, so as to consider all possible political ramifications.

6.2. Transportability

The initiatives in the project share the common trait of using evidence-based approaches to guide decision-making for efficient use of scarce resources. A key role in said approaches is that played by the mobility data, which comes from the collaboration agreement and work by ENTEL and ISCI.

In this regard, there are many future application areas that will greatly benefit from the availability of the mobility data produced in the context of this project; moreover, it is possible to process the raw data so as to produce different data products that suit the specific needs of other application areas.

Regarding the analysis of mobility patterns as used in this project, it is possible to directly adapt the Serology Surveillance and the Testing and Active Screening modules using epidemiological data associated with different epidemiological settings, or with other endemic diseases, such as influenza. Indeed, beyond the context of the pandemic, the IP-based methodology developed for serology surveillance can be applied to conducting population studies in other context so as to overcome the logistical challenges associated with obtaining a geographically representative sample of the population.

More broadly, the analysis of the mobility data, as in this project, could provide key input to help improve a number of public services. For example, IP methods could help decide where to allocate temporary resources, such as the mobile vehicle-registration stations that are used at the beginning of every year close to the annual registration deadline, so as to maximize coverage or minimize travel time. Similarly, a different processing of the raw mobility data could help to assign each registered citizen to a voting site, so as to minimize travel time and maximize turnout.

The potential for developing useful tools alluded to above led, in early 2022, to a new collaboration agreement for research and innovation between ISCI and ENTEL, going beyond the pandemic response. This new alliance aims at leveraging the use of mobility data in public and commercial projects. An early example of this new collaboration effort is a grant funded by a state-run agency, already in progress, to produce origin–destination matrices to help improve planning of freight transport within cities.

Another situation where we foresee future applications of the efforts we have carried out is in the forecast of ICU occupancy during the winter season, when influenza and other viruses circulate. It is almost always the case that during winter in Chile, the ICU system is stressed; we deem it possible, and will indeed do research on it, that with some knowledge of the circulation of the different viruses, we could predict peaks of ICU demand.

7. Conclusions

In this work, we describe the collaboration between the Chilean government, Instituto Sistemas Complejos de Ingeniería (ISCI), ENTEL, and other partners in specific projects, such as the faculty of medicine at University of Chile, in developing a data-driven and science-based proactive approach to decision-making that helped the people of Chile better endure the Covid-19 pandemic.

There are many aspects of this collaboration that go beyond traditional practical applications of analytics and operations research and that, we believe, set this project apart.

This project had real-life, nationwide **implementation**: all modules of the project had a national reach, including the fact that health-care workers were on the ground, working at places that were specifically identified by our advanced analytics models.

This project had a profound **impact**, by providing scientific evidence supporting the decision-making behind the three pillars of the Chilean strategy. This helped to provide transparency and objectivity to decision makers and the general population. But, of course, the most important measure of impact is that thousands of lives were saved: without the project, the death toll in Chile could have been at least 5% higher. Also, millions of dollars were saved. All this was achieved because the project offer technical solutions that contributed to each of the three main strategic pillars in the fight against Covid.

The **technical solutions** provided were innovative, crafted specifically for each challenge, mixing all the tools we had at hand in OR, and interacting with other disciplines such as with geography and epidemiology. The fact that many of our solutions were eventually published in top journals is a manifestation of their novelty.

The project faced a large number of **difficulties**. First, the project was carried out in, possibly, one of the most complicated medical emergencies in living memory, where solutions were needed urgently. Second, the only way the project and its four initiatives had any chance to succeed was through the joint work of a quite heterogeneous group of people, ranging from engineers and researchers, to political authorities, to health care workers on the ground, who were facing huge amounts of stress. Building trust among everyone was fundamental, so that the efforts of one group of people were followed up by the next group in the line of work. It is actually fairly easy to envision many ways in which the project could have failed. Third, it required a huge amount of coordination between institutions and its officials. And fourth, as it dealt with very delicate information, the results had to be communicated with care, a task that required further interaction between authorities and scientists, and had them both explaining on national media. In fact, the media coverage of the project was vast.

The project and the solutions have a high potential for **transportability**. Indeed, a new agreement between Entel and ISCI has been signed in order to explore further uses of the mobility data, while large scale population studies following the framework used in the IgG surveillance program can be attempted. We may also adapt our ICU demand prediction model for when influenza and other winter viruses hit the population and the health care system.

A key aspect for transportability is the fact that, during the project, hundreds of people were involved and trained, including the support of over a dozen research assistants, most of them graduate students. These human resources now have the knowledge and experience to push further this successful interaction between engineering and public health. In fact, we are already collaborating with the new government that took office in March 2022 to extend under their administration the Covid initiatives presented in this paper and also explore further opportunities for collaboration. To conclude we quote the current Minister of Health Begoña Yarza: "Our government is just getting started. We are very excited to continue these collaborations using science and analytics to support decision-making in the pandemic and, why not, in other public health challenges we need to address."

Acknowledgments

We first and foremost thank the health authorities for supporting evidence-based decision-making in times when following science is not always the easiest or safest path. We thank the Servicio de Salud Metropolitano Centro for their support in setting up the pilots for the active screening of asymptomatic cases and the serology surveillance initiatives, and the Asociación Chilena de Seguridad (ACHS) and the Servicio Nacional del Adulto Mayor (SENAMA) for their participation in the pilot of the group-testing monitoring program. We thank the various research assistants and faculty that contributed with their hard work in setting up the tools developed in this project, and in communicating its results. We also thank the hundreds of individuals involved in the fieldwork throughout Chile. This work was supported by an ANID PIA/APOYO grant (No AFB180003), used to hire research assistants for developing solutions, information systems, data management, and training of field staff. Fieldwork was funded by the Ministry of Health of the Government of Chile. M. Goic acknowledges the financial support of Fondecyt (Project No 1221711) and the Institute for Research in Market Imperfections and Public Policy (IS130002 ANID). Finally, we thank the Stanford RISE COVID-19 Crisis Response Faculty Seed Grant Program.

References

Arroyo C, Engel E, Pardow D, Simonetti P (2020) Informe sobre la evolución de la epidemia de covid-19 en Chile, Espacio Público.

Basso L, Salinas V, Sauré D, Thraves C, Yankovic N (2021) The Effect of Correlation and False Negatives in Pool Testing Strategies for Covid-19. *Health Care Management Science* 25(1):146–165.

Bloomberg (2022) The Covid Resilience Ranking, URL `https://www.bloomberg.com/graphics/covid-resilience-ranking/`, accessed: xx/xx/2022.

Carranza A, Goic M, Lara E, Olivares M, Weintraub G, Covarrubia J, Escobedo C, Jara N, Basso L (2022) The Social Divide of Social Distancing: Shelter-in-Place Behavior in Santiago during the Covid-19 Pandemic. *Management Science* 68(3):2016–2027.

Dorfman R (1943) The detection of defective members of large populations. *The Annals of Mathematical Statistics* 14(4):436–440.

Farfan MJ, Torres JP, ORyan M, Olivares M, Gallardo P, Salas C (2020) Optimizing RT-PCR detection of SARS-CoV-2 for developing countries using pool testing. *Revista Chilena de Infectología* 37(3):276–280.

Goic M, Bozanic-Leal M, Badal M, Basso LJ (2021) COVID-19: Short-term forecast of ICU beds in times of crisis. *PloS one* 16.1 e0245272, URL `https://doi.org/10.1371/journal.pone.0245272`.

Grassly NC, Fraser C (2008) Mathematical models of infectious disease transmission. *Nature Reviews Microbiology* 6(6):477–487.

iCovid Chile (2022) Hospital capacity, URL `https://www.icovidchile.cl/capacidad-hospitalaria`, accessed: 18/01/2022.

INE (2019) Proyecciones de poblacion, URL `https://www.ine.cl/estadisticas/sociales/demografia-y-vitales/proyecciones-de-poblacion`.

Jara A, Undurraga EA, González C, Paredes F, Fontecilla T, Jara G, Pizarro A, Acevedo J, Leo K, Leon F, et al. (2021) Effectiveness of an inactivated SARS-CoV-2 vaccine in Chile. *New England Journal of Medicine* 385(10):875–884.

Joshi YV, Musalem A (2021) Lockdowns lose one third of their impact on mobility in a month. *Scientific Reports* 11(1):1–10.

Lauer SA, Grantz KH, Bi Q, Jones FK, Zheng Q, Meredith HR, Azman AS, Reich NG, Lessler J (2020) The incubation period of coronavirus disease 2019 (COVID-19) from publicly reported confirmed cases: Estimation and application. *Annals of Internal Medicine* 172(9):577–582.

Li Y, Undurraga EA, Zubizarreta JR (2022) Effectiveness of localized lockdowns in the SARS-CoV-2 pandemic. *American Journal of Epidemiology* 191(5):812–82.

Mercer TR, Salit M (2021) Testing at scale during the COVID-19 pandemic. *Nature Reviews Genetics* 22(7):415–426.

Ministry of Health (2022a) , private communication.

Ministry of Health (2022b) Accumulated confirmed cases of coronavirus nationwide, URL `https://www.minsal.cl/nuevo-coronavirus-2019-ncov/casos-confirmados-en-chile-covid-19/`, accessed: 10/01/2022.

Ministry of Health (2022c) Plataforma Territorial Ministerio de Salud. Situación COVID-19, URL `https://storymaps.arcgis.com/stories/fbf237db339b4422b51cfbd707db2810`, accessed: 10/01/2022.

Montgomery JM, Hollenbach FM, Ward MD (2015) Calibrating ensemble forecasting models with sparse data in the social sciences. *International Journal of Forecasting* 31(3):930–942.

Rai B, Shukla A, Dwivedi LK (2021) Estimates of serial interval for COVID-19: A systematic review and meta-analysis. *Clinical Epidemiology and Global Health* 9:157–161.

Sauré D, Neira I, Goic M, O'Ryan M, Torres JP, Bruhn A, Ferres M, Angulo J, Vera M, Basso LJ (2021) Infection with SARS-CoV-2 Variant Gamma (P.1) in Chile Increased ICU Admission Risk Three to Five-Fold, Instituto Sistemas Complejos de Ingenieria Working Paper Series.

Sauré D, O'Ryan M, Torres J, Zuniga M, Santelices E, Basso L (2022) Dynamic IgG Seropositivity in Chile following rollout of an inactivated SARS-CoV-2 vaccine (CoronaVac) and an mRNA vaccine (BNT162b2): A sentinel surveillance study. *The Lancet Infectious Diseases* 22(1):56–63.

Sud A, Jones ME, Broggio J, Loveday C, Torr B, Garrett A, Nicol DL, Jhanji S, Boyce SA, Gronthoud F, et al. (2020) Collateral damage: The impact on outcomes from cancer surgery of the COVID-19 pandemic. *Annals of Oncology* 31(8):1065–1074.

Torres J, Basso L, O'Ryan M, Zuñiga M, Sauré D (2022a) Population-based COVID-19 IgG Seropositivity up to 16 weeks after a heterologous or homologous booster dose among Sinovac recipients in Chile, Instituto Sistemas Complejos de Ingenieria Working Paper Series.

Torres J, Sauré D, Basso L, Zuniga M, Cazor A, O'Ryan M (2022b) SARS-COV-2 IgG positivity in vaccinated and non-vaccinated Chilean children: A national cross-sectional study in schools. *International Journal of Infectious Diseases* URL `https://doi.org/10.1016/j.ijid.2022.04.039`.

Vekaria B, Overton C, Wiśniowski A, Ahmad S, Aparicio-Castro A, Curran-Sebastian J, Eddleston J, Hanley NA, House T, Kim J, et al. (2021) Hospital length of stay for COVID-19 patients: Data-driven methods for forward planning. *BMC Infectious Diseases* 21(1):1–15.

Wu H, Levinson D (2021) The ensemble approach to forecasting: A review and synthesis. *Transportation Research Part C: Emerging Technologies* 132:103357.

Yelin I, Aharony N, Tamar ES, Argoetti A, Messer E, Berenbaum D, Shafran E, Kuzli A, Gandali N, Shkedi O, et al. (2020) Evaluation of COVID-19 RT-qPCR test in multi-sample pools. *Clinical Infectious Diseases* 71(16):2073–2078.

Zhang GP (2003) Time series forecasting using a hybrid ARIMA and neural network model. *Neurocomputing* 50:159–175.

Zhang GP, Qi M (2005) Neural network forecasting for seasonal and trend time series. *European Journal of Operational Research* 160(2):501–514.

Appendix A: BAC Index

Let us denote by I the set of urban census zones, and by D a set of days. For each urban census zone $i \in I$, we denote its inhabitants by n_i and its area by s_i. In addition, the symptomatic cases in urban census zone i on day $d \in D$ are denoted by a_{id}. In any census zone, the BAC index is computed as the geometric mean of the two following elements: (i) the positivity of inflow, and (ii) the inflow density of symptomatic cases. The intuition for why the index has these two elements is that, on the one hand, we would like to test in those census zones that have the highest positivity in their floating (i.e., inflow) population. On the other hand, those symptomatic cases among the inflow population should be normalized by the total area of the census zone we are looking into. For each census zone $j \in I$ on day $d \in D$, we define the inflow of symptomatic cases as F^a_{jd}, and the inflow as F_{jd}; then the BAC index in census zone $j \in I$ on day $d \in D$ is defined as

$$\begin{aligned} BAC_{jd} &= \sqrt{\frac{F^a_{jd}}{F_{jd}} \cdot \frac{F^a_{jd}}{s_j}} \\ &= \sqrt{\frac{\sum_{i\in I} \frac{a_{id}}{n_i} f_{ij}}{\sum_{i\in I} f_{ij}} \cdot \frac{\sum_{i\in I} \frac{a_{id}}{n_i} f_{ij}}{s_j}}. \end{aligned}$$

The reason why (i) and (ii), i.e., the positivity of inflow and the inflow density of symptomatic cases, are averaged geometrically instead of arithmetically, is to penalize (in the sense of lowering their BAC index) those census zones that either have a very low inflow positivity, or have a very low density inflow of symptomatic cases. Also, it is worth noting that there is an alternative interpretation we can give to the BAC index:

$$\begin{aligned} BAC_{jd} &= \sqrt{\frac{F^a_{jd}}{F_{jd}} \cdot \frac{F^a_{jd}}{s_j}} \\ &= \frac{F^a_{jd}}{F_{jd}} \sqrt{\frac{F_{jd}}{s_j}}. \end{aligned}$$

Thus, the BAC index can be expressed as the multiplication of the positivity of inflow by the square root of the density of inflow.

Appendix B: Estimated Parameters for Impact Evaluation

Table 2 shows several epidemiological parameters that were used to estimate the impact of the overall implementation of the system. In particular, the table shows: R (which denotes the number of infected people on average by one person with the virus); the number of reported cases, ICU hospitalizations, and fatalities due to Covid-19; and the ratio of the latter two amounts with respect to the reported cases.

We quantified the financial cost associated with an infection through lost working days. The average monthly salary in Chile in 2020 was 794 USD, with 21.5 working days per month, and therefore 10 lost working days during the infection induces a cost of 366 USD per infection. To quantify the cost of an ICU hospitalization, we considered a hospital reimbursement of 37,500 USD per patient (based on Health Ministry Resolution 194 of June 11, 2020).

Table 2 Estimated: R, total cases reported, ICU hospitalizations, fatalities, ICU hospitalization ratio, and fatality ratio. Values are provided for each half-year period (or fraction of a half-year period as the case of 2020 1 and 2022 1), for the whole period of the pandemic (up to January 15, 2022). Source: Products 5, 24, 54, and 91 from `https://github.com/MinCiencia/Datos-COVID19`

	Time Period					
Variable	2020 S1	2020 S2	2021 S1	2021 S2	2022 S1	Total
R	1.31	0.99	1.01	0.98	1.52	1.06
Cases	247,971	329,205	947,140	254,505	60,832	1,839,653
ICU Hospitalizations	10,570	14,221	37,254	10,079	406	80,313
Fatalities	5,688	10,920	15,937	6,570	261	39,376
Case ICU Hospitalization Ratio	4.26%	4.32%	3.93%	3.96%	0.67%	4.37%
Case Fatality Ratio (CFR)	2.29%	3.32%	1.68%	2.58%	0.43%	2.14%

Appendix C: Evaluation of ICU Forecasts

To assess, the accuracy of our forecasts, we conducted a detailed analysis of the results of the first wave, where we compared our one- and two-weeks ahead forecasts against the number of beds that were actually needed at each point in time. Figure 10 plots the observed requirements against the forecasts for Santiago and Valparaiso. These two regions required very different numbers of beds in the first wave. Our numerical analysis indicates that our forecasts outperformed a variety of benchmarks and correctly anticipated when the number of ICU beds would reach its peak. Our predictions achieved average forecasting errors of 4% and 9% for one- and two-week horizons, respectively (Goic et al. 2021), providing a reliable tool to help the Ministry of Health successfully manage critical care capacity for all hospitals in every region.

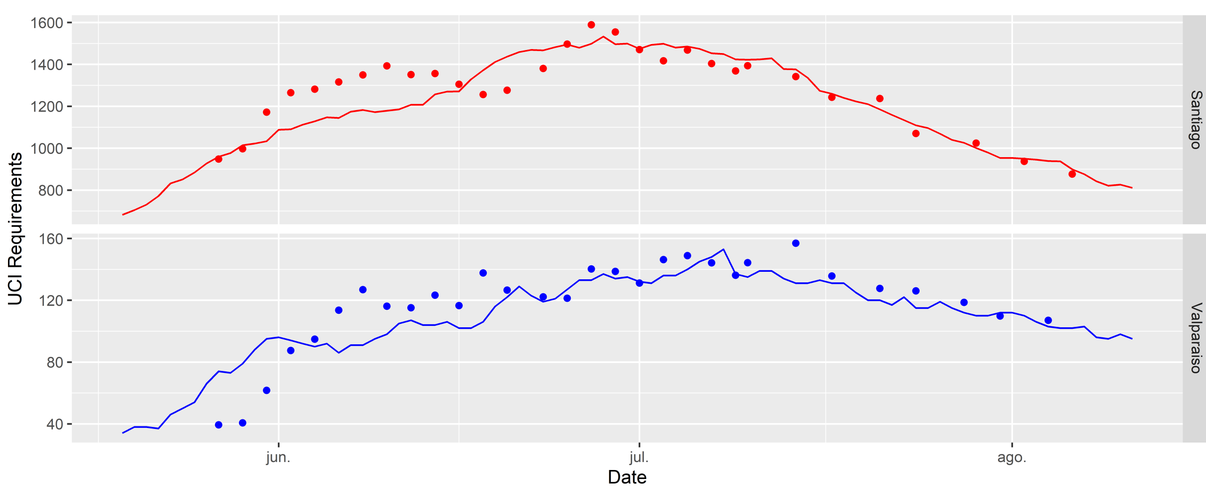


Figure 10 Actual and predicted ICU requirements for two selected regions. Lines represent actual requirements and dots represent forecasts.

To assess the impact of using these forecasts on capacity planning, we compare alternative scenarios with less accurate predictions. We simulate those benchmarks by simulating forecasts with 5% (conservative) and 10% (less conservative) more forecasting errors.

Similar to the newsvendor problem, more precise predictions have asymmetric impacts depending on whether the actual demand is lower or greater than the forecast. The cost of overestimation is associated with a larger financial cost of provision of more beds and the cost of delaying other medical procedures. As we do not access to the monetary cost of converting beds into ICUs and the impact of postponing other surgeries, we exclusively focus on the cost of the underestimations, which are associated with the availability of additional beds for those who needed them. By comparing the actual predictions against the less accurate benchmarks of the simulations, we can compute the number of additional day-beds that were available due to more precise predictions. To translate this additional number of day-beds into patients, we divide it by 26.8 days, the average length of stay in ICU. Finally, we multiply this figure by the fatality rate in ICUs (51.3%) and obtain an estimate for the number of lives saved.

In our estimation we assume that the hospital system had some flexibility to accommodate more patients without converting beds and therefore the forecasts helps only when the utilization surpasses a given threshold. In our more conservative evaluation, we consider that the system did not require additional beds until all available beds before the pandemic were already used. Realizing, that there were always some beds that were used for patients not hospitalized for Covid-19, in the less conservative scenario we assume that additional beds were needed when reaching 90% of the original capacity. Results of these scenarios are displayed in Table C.

Table 3 **Conservative and Less Conservative Evaluations of Impact of ICU Forecast**

	Gained Accuracy	Capacity slack	Deaths
Conservative	5%	100%	467
Less Conservative	10%	90%	1,017

Appendix D: Integer Programming Formulation for Testing-Station Allocation

We use data from the census (INE 2019) and mobility patterns as described in Section 2.2 to formulate the problem of deciding on the location of antibody testing stations in order to maximize the size of a representative sample. Following Section 5.2, here we consider the formulation for a single administrative region.

Let I denote the set of urban census zones from whose population we want to form a representative sample of, and T the set of time blocks (e.g., 10am–1pm and 2–5pm) during which a station can be operating during the planning horizon. Also, let K denote the set of health services (HS) in the Chilean public health-care system.

For $i \in I$, $t \in T$ and $k \in K$, we let $x_{i,t,k} \in \{0,1\}$ denote the decision regarding whether to install a testing station in census zone i during time block t, manned by personnel from HS k ($x_{i,t,k} = 1$), or not ($x_{i,t,k} = 0$). Because HS have a limited personnel, we consider the constraint

$$\sum_{i \in I_k} x_{i,t,k} \leq C_k, \quad k \in K, \quad t \in T, \tag{D-1}$$

where $I_k \subseteq I$ denotes the subset of census zones that fall under the jurisdiction of HS k, and C_k denotes the number of stations that can be operated simultaneously by the HS, for $k \in K$. From the (normalized) mobility data, let $f_{i,j,t} \in [0,1]$ denote a measure of the flow of people originally from census zone i that spend time block t in census zone j. Letting z denote the largest size of a geographically representative sample obtained from operating the testing stations, we impose that

$$z\, r_j \leq \sum_{k \in K} \sum_{i \in I_k} x_{i,t,k}\, f_{i,j,t}, \quad j \in I, \tag{D-2}$$

where r_j denote the percentage of the population of the region that belong to census zone j, taken from INE (2019). The formulation is completed by stating the objective of maximizing the size of the representative sample, z.

Appendix E: Measuring Impact of Booster Shots on Preventing Infections, ICU Hospitalizations, and Deaths

We used data published by the Chilean Ministry of Science reporting infections, ICU hospitalizations, and deaths for different age groups and vaccination schemes. We focused on subjects that had at least two vaccine doses, in order to measure the differential impact of the booster shot.

Define N_{igt} as the number of subjects that in week t from age group g (20–50, 50–70, or older than 70) that completed vaccination scheme i (two doses or booster). Let y_{igt} denote the incidence of infections in this group, defined as the number of infections divided by N_{igt}. Figure 11 shows the weekly incidence of infections, ICU hospitalizations, and deaths during September–December 2021, by age group. The data suggest a substantial reduction in case incidence for all the health outcomes and age groups (the sample includes pairs week–groups with $N_{igt} \geq 300$ thousand subjects).

We build an econometric model to estimate the effect of the booster shot in reducing the incidence of these health outcomes. Let $D_{igt} = 1$ if i corresponds to the booster shot vaccination scheme. We estimate a generalized linear model using a Poisson link function through the following panel-data regression:

$$\log(E(y_{igt})) = \delta_g + \theta_t + \beta_g D_{igt} + \epsilon_{igt}, \tag{E-3}$$

where the coefficient β_g measures the effect of the booster shot on the incidence of health outcome y. The fixed effects δ_g capture differences in outcome incidence across age groups and the week dummy variables θ_t capture temporal variation in the transmission of the virus and other seasonal effects. Regression (E-3) was

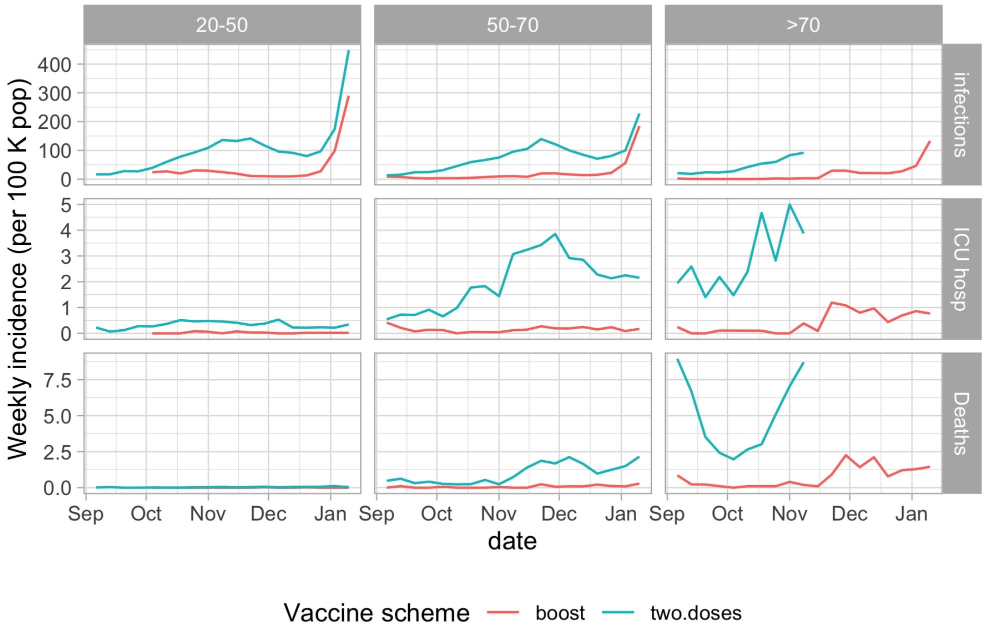


Figure 11 **Incidence of health outcomes (infections, ICU hospitalizations, and deaths) during September–December 2021, for different age groups with and without booster shots. Graph excludes weeks that had fewer than 300 thousand subjects in the sample.**

estimated by changing the dependent variables to ICU hospitalizations, and deaths in order to measure the effect of the booster on those outcomes.

Table 4 shows the estimation results, which are also illustrated in Figure 12. The coefficient estimates suggest that booster shots had a significant impact on reducing infections and ICU hospitalizations for all three age groups. The impact on deaths was significant for the 50–70 and older than 70 age groups but for the younger group the estimate is more imprecise and has lower statistical significance. Figure 13 plots the fitted values of the model (at the 95% confidence interval) against the actual data. We also estimated the model using a negative binomial regression and splines to capture seasonality (instead of week fixed effects) and the results were similar.

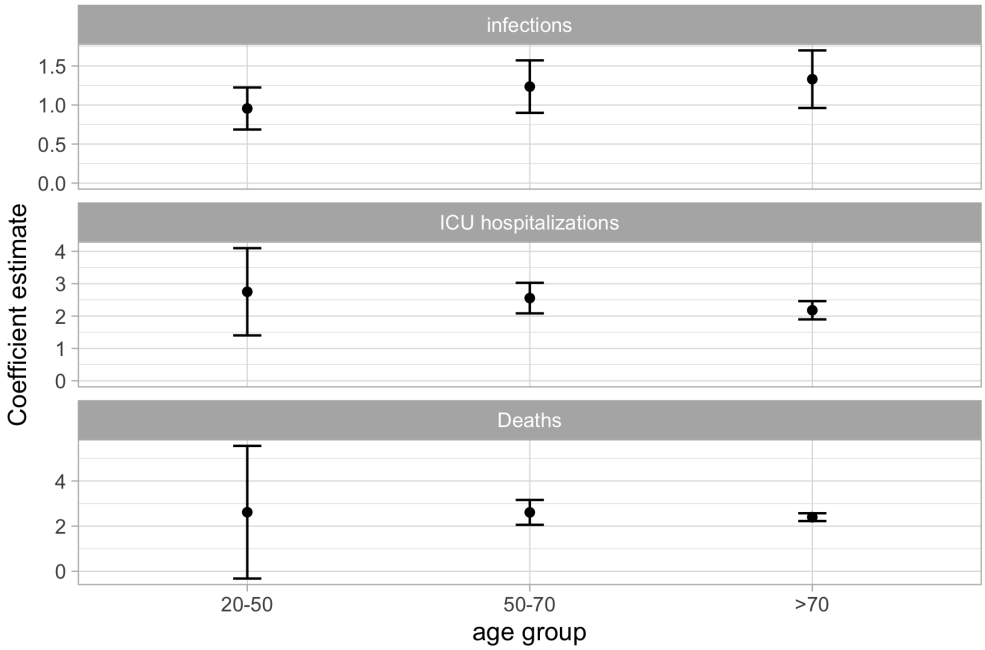


Figure 12 **Estimated coefficients from the Poisson Regression model of the effect of booster on infections, ICU hospitalizations, and deaths for different age groups.**

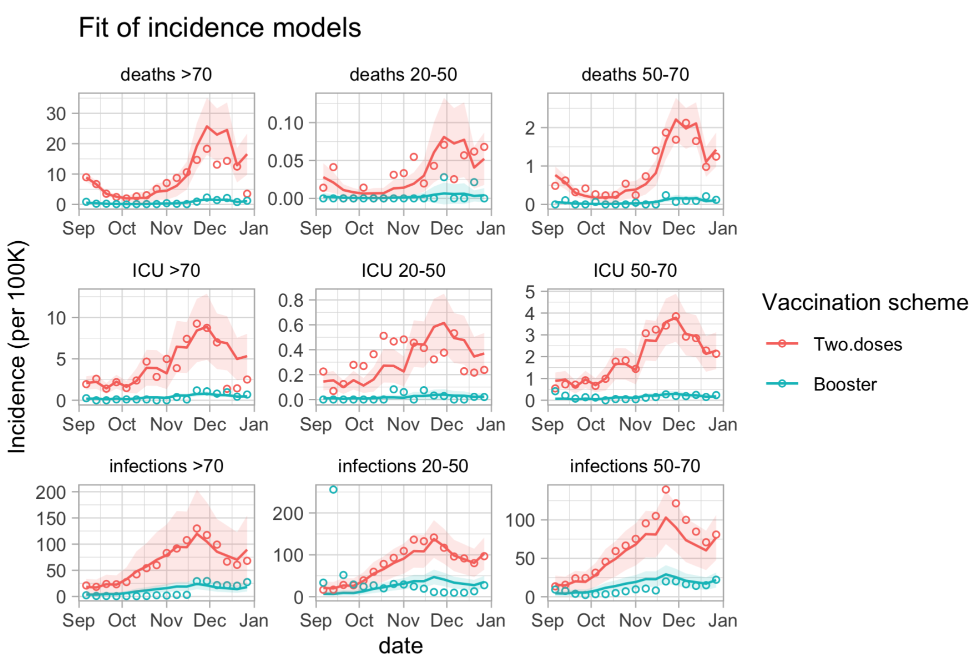


Figure 13 **Fit (solid lines) vs. actual data (dots) of the Poisson regression models. Shaded areas indicate the 95% confidence interval of the prediction and colors indicate the prediction with and without booster shot.**

Table 4 Estimation results of the regression to estimate the effect of booster shot on the incidence of infections, ICU hospitalizations. and deaths.

	Dependent variable:		
	Infections	ICU	Deaths
	(1)	(2)	(3)
age.group50-70	-0.264^{**}	1.837^{***}	3.209^{***}
	(0.105)	(0.162)	(0.342)
age.group>70	-0.375^{***}	2.526^{***}	5.341^{***}
	(0.108)	(0.156)	(0.336)
age.group20-50:booster	-0.955^{***}	-2.751^{***}	-2.616^{*}
	(0.137)	(0.687)	(1.499)
age.group50-70:booster	-1.235^{***}	-2.556^{***}	-2.609^{***}
	(0.171)	(0.241)	(0.281)
age.group>70:booster	-1.330^{***}	-2.180^{***}	-2.397^{***}
	(0.188)	(0.144)	(0.088)
Constant	3.792^{***}	-1.208^{***}	-3.023^{***}
	(0.309)	(0.220)	(0.348)
Observations	129	129	129
Note:	*p<0.1; **p<0.05; ***p<0.01		