https://doi.org/10.1016/j.ejor.2021.11.012

Efficient social distancing during the COVID-19 pandemic: integrating economic and public health considerations

Kexin Chen[a,b], Chi Seng Pun[c,*], Hoi Ying Wong[b,*]

[a] *Department of Applied Mathematics, The Hong Kong Polytechnic University, Hong Kong*
[b] *Department of Statistics, The Chinese University of Hong Kong, Shatin, N.T., Hong Kong*
[c] *School of Physical and Mathematical Sciences, Nanyang Technological University, Singapore*

Abstract

Although social distancing can effectively contain the spread of infectious diseases by reducing social interactions, it may have economic effects. Crises such as the COVID-19 pandemic create dilemmas for policymakers because the long-term implementation of restrictive social distancing policies may cause massive economic damage and ultimately harm healthcare systems. This paper proposes an epidemic control framework that policymakers can use as a data-driven decision support tool for setting efficient social distancing targets. The framework addresses three aspects of the COVID-19 pandemic that are related to social distancing or community mobility data: modeling, financial implications, and policy-making. Thus, we explore the COVID-19 pandemic and concurrent economic situation as functions of historical pandemic data and mobility control. This approach allows us to formulate an efficient social distancing policy as a stochastic feedback control problem that minimizes the aggregated risks of disease transmission and economic volatility. We further demonstrate the use of a deep learning algorithm to solve this control problem. Finally, by applying our framework to U.S. data, we empirically examine the efficiency of the U.S. social distancing policy.

*Corresponding authors
Email addresses: kexinchen@polyu.edu.hk (Kexin Chen), cspun@ntu.edu.sg (Chi Seng Pun), hywong@cuhk.edu.hk (Hoi Ying Wong)

Keywords: OR in Health Services, Stochastic SIRD Model, Google Mobility Indices, Economic Modeling, Stochastic Controls, Deep Learning

1. Introduction

COVID-19, an infectious disease caused by a novel coronavirus, was first identified in Wuhan, China in December 2019 and has since spread worldwide, resulting in a global pandemic. Since February 2020, countries around the globe have faced unexpected challenges related to the COVID-19 pandemic, including hundreds of thousands of deaths, severe shortages of resources, and massive blows to economic systems. Globally, disruptions related to the COVID-19 pandemic have prompted individuals to change their behavior and governments to implement administrative policies imposing restrictions on human interactions, including bans on public events, shutdowns of schools and non-essential workplaces, and limits on public transportation; see Kissler et al. (2020). These non-pharmaceutical interventions and social distancing policies have been implemented to minimize the spread of the virus by encouraging physical distancing and decreasing social interactions; see Habersaat et al. (2020). The impact of these measures has been quantified using community mobility indices, which are aggregated spatial statistics of people's daily movements. Companies such as Google, Apple, and Cuebiq are using their users' location data to deliver useful mobility trend information; see Google (2020); Apple (2020); Cuebiq (2020). Many recent studies on COVID-19 focus on the role of mobility in disease prediction (Liu et al. (2020); Miller et al. (2020)), patterns of change in mobility (Warren & Skillman (2020); Ghader et al. (2020); Gao et al. (2020)), estimates of the effectiveness of social distancing strategies (Liu et al. (2020); Soucy et al. (2020)), and the economic consequences of mobility restrictions (Bonaccorsi et al. (2020)).

Recent studies present evaluations of and justification for the efficacy of these intervention policies; however, the design and implementation of such measures requires trade-offs between health outcomes and economic effects. Consequently,

policymakers must grapple with the difficult task of balancing stringent public health demands while preserving economic functions. A key lesson learned from recent pandemics is that the economy should be considered when developing pandemic mitigation policies. There is substantial evidence that governmental restrictions on commercial activities lead to extensive economic damage, especially in countries with service-oriented economies, such as the United States; see Baker et al. (2020). According to McKee & Stuckler (2020), a global failure to protect the economy during the COVID-19 pandemic could lead to costs beyond those directly attributable to the disease and long-term damage to healthcare systems. Some studies even predict an impending global recession as a consequence of the COVID-19 pandemic (see Wren-Lewis (2020); Fernandes (2020); Makridis & Hartley (2020)) and suggest that the combined impact of the public health crisis and related intervention policies may sharply exacerbate the contraction of economic activity (see Barua (2020)). This finding is significant because it suggests that policymakers may fail to implement adequate and sustainable contagion control strategies due to a lack of practical, well-balanced objective planning. Therefore, adaptive policies that use quantitative tools to balance health outcomes and economic impacts, such as mathematical modeling of contagion mechanisms, simulation of different epidemic scenarios, and suitable measurements of policy efficiency, are urgently needed.

Hereafter, we define a social distancing policy broadly as the measures implemented to reduce interactions among populations. In this category, lockdown measures could be considered an extreme example. In this study, we only focuses on Google's community mobility data which quantitatively assess the impact of social distancing policies on peoples' movements in each geographic region. Integrating a classic epidemic model with modern techniques such as machine learning, we develop a novel data-driven decision support tool that policymakers can use to set *efficient* community mobility targets (i.e., social distancing targets). Specifically, we introduce the concept of an efficient social distancing policy (ESDP), defined as a community mobility target that minimizes the aggregated risks of disease transmission and financial market instability. We then

optimize the computation of ESDPs through deep learning, using this process to formulate our problem as a stochastic control problem with constraints. We implement deep learning conveniently and efficiently using the Python-based TensorFlow platform. We apply our framework to daily sampled financial data by considering financial market instability as a measure of the pandemic's effect on the economy; mobility indices and numbers of infections, recoveries, and deaths are also updated daily.

To demonstrate the utility of our framework, we apply it to COVID-19 data from the United States. In comparison to related studies Benzell et al. (2020); Block et al. (2020), which focus on qualitative discussion and offer guidance on implementing social distancing, we offer a pioneering framework for developing efficient social distancing policies. Our framework can be applied regardless of the statistical models used to predict the pandemic's evolution and the quantification methods used to model uncertainty.

Our epidemic control framework for ESDPs is the major contribution of our work. We show that statistical community mobility data are key to both predicting pandemics and developing related policy. The key empirical findings of our work are as follows.

- Mobility data improve pandemic predictions, demonstrating the utility of community mobility obtained by big-data analytics (Section 3.2).
- The stock market is sensitive to pandemic situations and community mobility patterns, especially during pandemics (Section 4.1).
- Deep learning is a useful technique for computing data-driven optimal policies in complicated dynamic systems. Benchmarking against the suggested ESDP reveals a decrease in the efficiency of social distancing policy in the U.S. in June 2020 (Section 5.3).

The rest of the paper is organized as follows. Section 2 contains a qualitative discussion of the conceptual framework of our methodologies. Section 3 introduces the SIRD model and explains our extension, which incorporates stochastic

coefficients. We then statistically determine the coefficients as functions of the observed mobility data and provide a simulation to illustrate the effect of mobility. In Section 4, we conduct regression analysis to determine the effect of the pandemic on a stock market index. Section 5 details the formulation of the optimization problem used to determine the ESDP and measure the efficiency of a given social distancing policy. The empirical results calculated using U.S. data are presented in Section 5.3. Concluding remarks are made in Section 6.

2. The Conceptual Framework

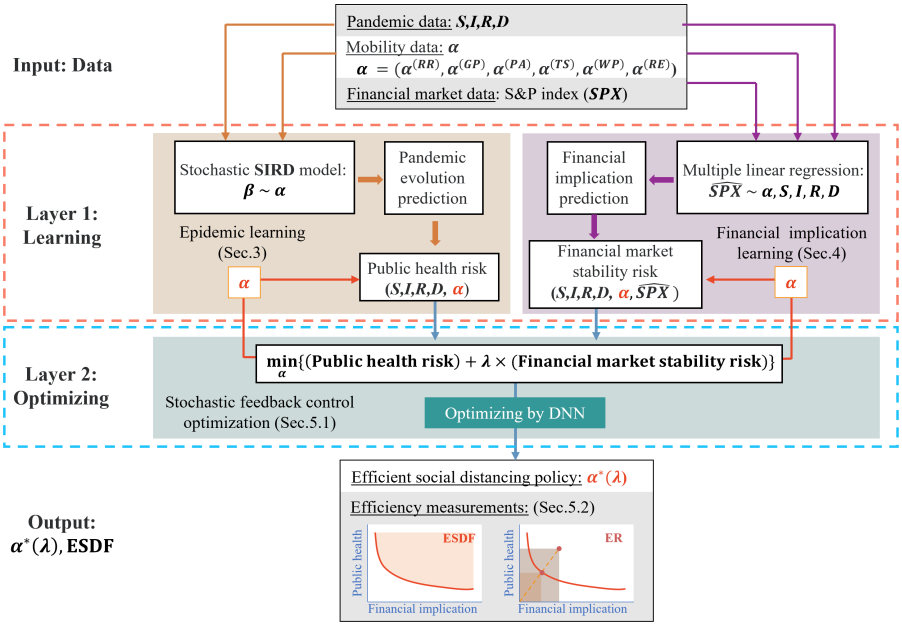
We first outline our proposed methodological framework and the underlying rationale and then introduce the technical details. We also discuss the strengths and weaknesses of our approach relative to related studies.

Table 1: Data and Notation

Available Data			Notation	
Pandemic data	Publicly accessible data	Daily reported cases: • S: susceptible; • I: infected; • R: recovered; • D: deceased.	Population fraction at time t	Stochastic process: • S_t: susceptible; • I_t: infected; • R_t: recovered; • D_t: deceased; • C_t: cumulative confirmed.
	(Micro - perspective) Not used in this study	E.g., hygiene measures, population behavior, geographic stratification, age stratification, hospital usage.	Rates of period $[t, t+1]$	• $f_\beta(\beta_{t+1})$: infection rate; • $f_\gamma(\gamma_{t+1})$: recovery rate; • $f_\delta(\delta_{t+1})$: death rate; • β_{t+1}: log odds of infection; • γ_{t+1}: log odds of recovery; • δ_{t+1}: log odds of death.
Mobility data	Google Mobility	Six categories: • RR: retail & recreation; • GP: grocery & pharmacy; • PA: parks; • TS: transit stations; • WP: workplaces; • RE: residential.	Mobility control	**(1)** Mobility control vector: α_t, Optimal mobility target: α_t^*; **(2)** Moving average mobility: $\bar{\alpha}$; **(3)** Principal component analysis: $P = \hat{A}\alpha$ • $P \in \mathbb{R}^6$: historical PCs; • $\hat{A} \in \mathbb{R}^{6\times 6}$: determinant matrix; • $\hat{L} \in \mathbb{R}^6$: lower bound; • $\hat{U} \in \mathbb{R}^6$: upper bound.
Financial market data	Stock market indices	S&P 500 index: SPX	Learned SPX component associated with pandemic and mobility	$\widehat{SPX}_t \sim S_t, I_t, R_t, D_t, \alpha_t$; λ: weight parameter.

We develop a system that uses publicly available data as the input and generates ESDPs as the output. The ESDPs generated using this system can be used to assess the efficiency of real-life social distancing policies. Table 1 summarizes the types of data and notation used in this study, expands on the conceptual framework illustrated in Figure 1, and enhances understanding of the technical content in this paper.

Figure 1: An overview of an efficient social distancing framework



The proposed framework consists of two layers, learning and optimizing, as shown in Figure 1. In the learning layer, the pandemic and its financial implications are learned as functions of historical pandemic data and mobility control. We separate epidemic and financial implication learning into two procedures in this layer. For epidemic learning, we extend the classic susceptible–infected–recovered–deceased (SIRD) model to incorporate stochastic infection, recovery, and death rates and learn them as functions of pandemic and mobility data. By varying the mobility controller α, we can predict the future epidemic situation conditional on a given mobility control (or social distancing)

and link this prediction to a pandemic risk function. We use the transmission rate to represent pandemic risk. Similarly, for financial implication learning, we filter the COVID-19 driven component (i.e., pandemic component) of a stock index (the S&P 500 index, SPX) as a function of pandemic and mobility data. We then link this pandemic component in the SPX to a risk function, which is realized as market volatility and thus predicts the pandemic effect reflected in the market rather than the market itself. By varying the mobility controller α, the system can predict the effect of different social distancing policies on financial market stability.

In the optimizing layer of the proposed framework, we minimize the aggregated risk over a time period by constructing a time-varying mobility target or ESDP, α^*, based on the following three principles.

- **Striking a balance between public health risks and financial implications**: Social distancing policies involve a trade-off between lowering the public health risk by reducing the infection rate and mitigating the economic costs of social distancing interventions (e.g., consequent financial market instability). Such a trade-off requires the selection of a sensible aggregated risk function that integrates both public health risk and financial implication measure.

- **Data-driven feedback policy**: Policymakers should update their policies based on observed data to ensure that ESDPs are highly responsive to dramatic shifts in the pandemic situation. Optimization should be used to determine policies at different future time points while incorporating feedback from updated observations. In other words, optimization is essentially a stochastic feedback control problem.

- **Stable over a preset time period**: Once announced, pandemic-related policies are usually implemented for a defined period of time (e.g., 1 or 2 weeks). Policymakers aim to avoid frequent changes to such policies except in response to abrupt changes in the pandemic situation. This

reluctance to make policy changes places an additional constraint on the mobility controller.

Finally, we obtain a constrained stochastic feedback control problem, which can be solved numerically using convenient, recently developed deep-learning techniques. Specifically, the stochastic feedback control problem optimizes a convex combination of public health and financial market stability risks, as follows:

$$\min_{\alpha_{t+1},\cdots,\alpha_{t+h}} \{(\text{public health risk}) + \lambda(\text{financial market stability risk})\}, \quad (1)$$

where λ is the weight parameter, which reflects a policymaker's level of concern regarding financial market stability, and $\alpha_{t+1}, \cdots, \alpha_{t+h}$ constitute the data-driven feedback mobility control policy from day $t+1$ to day $t+h$. For any given value of λ, our system outputs an optimal community mobility control or ESDP, $\alpha^*(\lambda)$. By varying λ, the generated ESDPs can be visualized as an efficient frontier that explains the trade-off between public health and market stability risks, as shown in Figure 1. Based on this frontier, we can benchmark the efficiency of an actual social distancing policy against that of the optimal policy.

2.1. Strengths and Weaknesses

Next, we discuss the strengths and weaknesses of our framework in three aspects, namely **modeling**, **financial implications**, and **policy-making**. These aspects echo the three key empirical findings stated in Section 1. The weaknesses mainly reflect the assumptions made when facing challenges in different aspects, while the strengths reflect the rationale behind our proposals.

2.1.1. Modeling (Section 3)

Quantitative analysis of a pandemic requires modeling of its spread and impact. We consider the SIRD model, which is arguably the simplest epidemic model. This model is limited by challenges related to estimation (see Ahumada et al. (2020)) and places a minimum requirement on the input data,

which are basically available to the public. However, we can easily extend our proposed framework to a more complicated model, such as the SEIR (susceptible–exposed–infected–recovered) model (see Mwalili et al. (2020)) and many other compartmental models; see Brauer (2008). Notably, although many studies consider compartmental models with *fixed* parameters for illustrative purposes, only a few integrate both modeling of the available data and data randomness. For the sake of simplicity and data availability, we choose the SIRD model as our base model and incorporate both mobility data and randomness.

Like other compartmental models, empirical analyses of the SIRD model demonstrate the presence of time-varying coefficients in many applications, and modeling of the COVID-19 pandemic is no exception; see Calafiore et al. (2020). Several modifications are proposed, including a basic expansion of coefficients with sparse identification (Calafiore et al. (2020)), application of dynamic density functional theory (te Vrugt et al. (2020)), and adoption of a rolling regression approach (Rubio-Herrero & Wang (2021)). In this paper, we address time variation in the SIRD model by making two crucial assumptions.

1. Time variation in the infection rate is explained by the time-series factors of U.S. mobility indices, which are released by Google and aggregated at the country level as listed in Table 1.

2. The log odds of infection, recovery, and death rates are stochastic.

Our first assumption is consistent with many papers that address the effects of social distancing on the spread of the disease, including those of Mwalili et al. (2020); Dashtbali & Mirzaie (2021), who include more compartments; te Vrugt et al. (2020); Willis et al. (2020), who add diffusion or differential equations; and Zheng et al. (2021), who include piece-wise infection rates. All of these papers characterize the attribution of community mobility in different ways. Some studies even investigate the mobility factor, which is usually aggregated as a single user-specified measure with limited practical applicability. *In this paper, we take advantage of Google's community mobility data when quantifying social distancing policies.* Sulyok & Walker (2020) empirically shows a signifi-

cant correlation between mobility data and the case incidence or infection rate. Hence, to model the spread of COVID-19, we extend the SIRD model by incorporating community mobility data into the log odds of infection. In Layer 1 (learning), the mobility data act as factors, and their historical data are fitted into model training. In Layer 2 (optimizing), the mobility indices are used as the control variates to characterize social distancing policies.

Our second assumption fixes the model structure through adoption of the sigmoid function; however, the structure is extendable to any other transform function that maps a real line to $[0, 1]$. The benefit of this specification is immediate, as it allows us to exclude non-negativity constraints on the infection, recovery, and death rates during learning and optimizing. It is noteworthy that most of the literature ignores these constraints. Although the constraints may be less pronounced in a study that focuses on either learning or optimizing, we must address this issue at the cost of additional complexity in policy-making; this problem is solved by the use of deep learning. However, we admit that specification of the transform function should affect the model and the resulting optimal policy.

The log odds of infection are modeled using a factor model with noise, whereas those of recovery and death are modeled using multiplicative models examined through a few model structures. As our focus is social distancing, we naturally consider that community mobility may directly affect infection but not recovery or death. Hence, we do not consider the micro perspectives exemplified in Table 1 and consider them to be absorbed as noise. Although our model structure partially addresses time variation in the SIRD model via time-series mobility factors, our parameter estimation is based on a static rolling-window regression, which is effective when the environment within the window remains stationary. A remedy for non-stationarity of time series data is to consider time-varying coefficients as described in Calafiore et al. (2020); however, such consideration is beyond the scope of this paper.

Our model is practically relevant, as it accounts for both publicly available data and randomness. Once calibrated, we use our model to make scenario

projections based on representative past mobility data. Our simulation studies show that including mobility data improves predictions regarding pandemic evolution and ensures a good fit for the log odds of infection.

2.1.2. Financial implications (Section 4)

To characterize the impacts of the pandemic on the financial market, we formulate a supervised learning problem focused on a major financial index, S&P 500, that includes the features of mobility indices and the statistics of COVID-19 processes. Using this explanatory model, we *approximate the economic effects on the financial market of social distancing policies and the pandemic.* Although we use multiple linear regression simply as a learning tool, the adjusted R^2 value is very high, indicating strong correlations between the financial market data, pandemic data, and mobility indices. It should be noted that the learning or statistical estimation is *ex-post* rather than *ex-ante*. Our objective is not to predict the index price for profits but rather to identify a proxy function that maps the community mobility indices and pandemic statistics to the market's performance.

It is clear that ours is not the only way to characterize the financial implications of a pandemic. Our framework may be affected fundamentally in two directions. First, we can identify another measure of the economy or market performance that might even be vector-valued. Second, we can propose many other supervised learning algorithms or statistical estimations to improve mapping. As with most data-driven approaches, our framework is most effective if the impacts of the pandemic and mobility on market performance during the policy-making period are identical to those identified from historical data. These weaknesses should be addressed in further studies.

2.1.3. Policy-making and technology (Section 5)

Numerous studies investigate optimal policymaking and control methods in response to the COVID-19 pandemic. Benzell et al. (2020) and Block et al. (2020) provide qualitative discussions and offer guidance on implementing so-

cial distancing policies. In some studies, such as those by Cont et al. (2020); Alvarez et al. (2020); Jones et al. (2021), real data are used to build various epidemic models, and different scenarios of distancing policies are evaluated through simulation studies. From a mathematical perspective, Choi & Shim (2021) and Lesniewski (2020) develop mathematical models and apply optimal control theory to identify the optimal strategies under various epidemiological conditions. More recently, multiple works include a social distancing control variable in the epidemic model and optimize this variable using an open-loop optimal control framework; see Piguillem & Shi (2020); Alleman et al. (2020); Morato et al. (2020); Köhler et al. (2021). In these studies, control inputs are commonly used to represent various levels of social distancing policies. In contrast, our paper incorporates community mobility indices as control variables when modeling the pandemic and its financial implications. Assuming that the government can control community mobility by imposing social distancing policies, the search for optimal community mobility can be formulated as a feedback stochastic control problem.

For the sake of practicality, additional constraints on the optimization are imposed. For example, correlated mobility indices cannot be controlled individually, as restricting one mobility category affects the other categories. We perform principal component analysis (PCA) of historical community mobility data and set linear inequality constraints on the feasible mobility indices, based on a sample PCA of the historical minimums and maximums. We also consider a dwell-time constraint to ensure that the social distancing policy remains stable for at least h days, thus avoiding frequent transitions that would confuse the public. To address this constraint, we introduce a penalty function to the objective function of the daily mobility indices adjustment. Our empirical results show that *these constraints yield more reasonable social distancing policies than unconstrained optimization.*

However, the complex constraints imposed on the control set make it difficult to derive an analytical solution. Even traditional numerical methods, such as the backward regression-based simulation algorithm in Li et al. (2020), cannot

guarantee a numerical solution to our problem. To tackle this numerical problem practically, we are motivated to use a deep learning approach recently investigated by Tsang & Wong (2020). Thus, the stochastic optimal control component of our problem can be transformed into a deep neural network (DNN) training problem. In a pandemic situation (e.g., COVID-19), DNN software is an important tool that can be used to compute optimal control policies in real time. Regarding our problem, we find that the TensorFlow platform provides subroutines for DNN implementation that efficiently correspond to our problem.

Various aspects of our optimization framework can be improved; for example, robustness and worst-case scenario optimization can be considered as in Köhler et al. (2021); Morato et al. (2020). Further studies to address these weaknesses should be considered.

3. Stochastic modeling of COVID-19

We model the COVID-19 process based on a compartmental epidemiological framework derived from Kermack & McKendrick (1927). Specifically, the population is assigned to compartments with the labels S, I, R, and D, which represent the fractions of individuals who are *susceptible* to the infectious disease, *infected*, *recovered*, and *deceased*, respectively. The study of several aspects of this process is enabled by many extensions of this framework developed to include additional or modified compartments (see, e.g., Hethcote (2000)) and increasing mathematical complexity enabling the study of several aspects of the epidemic process. For example, Karako et al. (2020) and Bardina et al. (2020) model an additional individual action by including periods of latency and the presence of asymptomatic patients. Several studies (He et al. (2020); Chang et al. (2020); Moore & Okyere (2020)) model COVID-19 transmission by considering governmental control measures, while others (de Souza et al. (2020)) incorporate epidemiological and clinical characteristics into the modeling process. However, the applicability of the aforementioned studies depends heavily on the available data. Access to data on the most infectious diseases, especially

COVID-19, is limited, and processing of enormous sets of data compiled daily raises concerns about accuracy. Therefore, we focus on the SIRD model, which requires only the numbers of reported cases, active cases, and deaths.

In this paper, we present a two-part stochastic modeling procedure. In the first part, we build a stochastic version of the SIRD model that can fit real COVID-19 data from the United States. Although the SIRD model is usually presented using differential equations, we consider a discrete-time SIRD model that is more intuitive in practice. To align with the frequency of data collection, the time interval is set at 1 day and indexed by $t \in \mathbb{N}$. Fernández-Villaverde & Jones (2020) also use a standard SIRD epidemiological model to inversely deduce a time-varying virus reproduction rate and thus capture asynchronous changes in behavior and policy. However, we use real data to estimate the stochastic process of reproduction. Our research demonstrates the value of modeling changing infection rates under social distancing practices. Our model of the COVID-19 situation is distinct because it incorporates mobility indices available from Google (2020). We show that these indices are highly correlated with infection rates and that, indeed, infectious disease modeling is accurate only if community mobility is considered.

3.1. The SIRD model with stochastic log odds

Here, S, I, R, and D are treated as (stochastic) processes indexed by t. A discrete-time classical SIRD model is given by

$$\begin{cases} S_{t+1} = 1 - I_{t+1} - R_{t+1} - D_{t+1}, \\ I_{t+1} = I_t \left(1 + S_t f_\beta(\beta_{t+1}) - f_\gamma(\gamma_{t+1}) - f_\delta(\delta_{t+1})\right), \\ R_{t+1} = R_t + I_t f_\gamma(\gamma_{t+1}), \\ D_{t+1} = D_t + I_t f_\delta(\delta_{t+1}), \end{cases} \tag{2}$$

where $f_\beta(\beta_{t+1})$, $f_\gamma(\gamma_{t+1})$, and $f_\delta(\delta_{t+1})$ represent the rates of *infection*, *recovery*, and *death*, respectively, over the period $[t, t+1]$. Here, we consider all f functions as sigmoid functions, i.e.,

$$f_\beta(x) = f_\gamma(x) = f_\delta(x) = \sigma(x) := \frac{1}{1 + e^{-x}},$$

such that β_{t+1}, γ_{t+1}, and δ_{t+1} are interpreted as the *log odds* of *infection*, *recovery*, and *death*, respectively. In general, inhomogeneous functions can be considered as f functions and modified as $1/(K + e^{-x})$, such that the corresponding rate range is $(0, 1/K)$ for any constant $K > 0$. Although these rates, especially the infection rate during the early outbreak period, may exceed 1, we found that using the sigmoid function ($K = 1$) consistently yields decent fitting results for U.S. COVID-19 data. Therefore, we continue to use the standard sigmoid function. The specific sigmoid function $f(x) = 1/(1+\exp(-x))$ is considered to alleviate positivity constraints on controlled infection and other rates. In contrast to most studies, which assume constant rates and disregard the ranges (or positivity) of the rates, our formulation allows us to study the log odds on the real line, yielding rates always within the range $(0, 1)$. This feature is crucial for later analyses in which we introduce mobility controls to the log odds, as both stochastic modeling (of the log odds) and the controls are rendered free of constraints. Using the data set $\{(I_t, R_t, D_t)\}_{t=0}^{T}$ and the model (2), we can compute the log odds as follows: for $t = 0, 1, \ldots, T-1$,

$$\beta_{t+1} = \operatorname{logit}\left(\frac{C_{t+1} - C_t}{I_t(1 - C_t)}\right),\ \gamma_{t+1} = \operatorname{logit}\left(\frac{R_{t+1} - R_t}{I_t}\right),\ \delta_{t+1} = \operatorname{logit}\left(\frac{D_{t+1} - D_t}{I_t}\right), \tag{3}$$

where $C_t = I_t + R_t + D_t$ represents the total confirmed cases up to time t, and $\operatorname{logit}(x) := \sigma^{-1}(x) = \ln\left(\frac{x}{1-x}\right)$.

One simple but key observation is that the collected data are usually affected by environmental changes. For example, the government and medical authorities may attempt to reduce the numbers of infectious cases and deaths by isolating infected individuals, implementing social distancing measures, and/or promoting vaccination (if available); see Cheng et al. (2020). Most countries, including the United States, implemented such measures once they became aware of the seriousness of the domestic COVID-19 situation. Hence, these factors should be taken into account when the data are fed into the SIRD model (2) or (3). This paper focuses only on COVID-19 mitigation attempts involving mobility controls (isolation and social distancing). Assuming that no vaccine is available,

it is natural to consider the recovery log odds γ and death log odds δ in (3) as specific to the virus and unaffected by exogenous factors. We do not consider viral variants but rather assume that a single major virus strain affects each country. However, the infection rate β depends on community mobility. We recognize that the landscape of the COVID-19 pandemic has changed continuously from the early period, through the time this study was prepared, to the time of writing. Although multiple vaccines have become available during the revision of this paper, outbreaks continue partly due to constraints in vaccine supplies and distribution capacity and the emergence of new variants of concern (VoC). The SIRD model under consideration is arguably the simplest epidemic model and does not take into account the impact of vaccines or viral variants. Although this omission may be a limitation, it does not materially affect our proposed conceptual framework, which is expected to be extended to accommodate such practical considerations by examining more sophisticated models. One future research direction could incorporate the impacts of vaccination and VoC by adjusting the infection, recovery, or mortality rates within the model (see Faria et al. (2021) for VoC analyses, Fudolig & Howard (2020) for modified multi-strain SIR model analysis and Moghadas et al. (2021) for COVID-19 vaccine analysis). Nonetheless, we maintain that social distancing remains an important pandemic control and mitigation strategy and we are continuously updating our results at `https://sites.google.com/view/esd-covid19/`.

Google, a subsidiary of Alphabet Inc., recently published COVID-19 community mobility reports based on data collected via their products (e.g., Google Maps) to contribute to COVID-19 mitigation efforts and help public health officials make critical decisions; see Google (2020). Google mobility indices are considered feasible controls because they are linked directly to governments' social distancing policies. This paper makes a major contribution to the field by deriving optimal mobility controls (i.e., social distancing). Google mobility indices cover changes in mobility relative to the baseline [1] in the areas of retail &

[1] The baseline values are computed as the median values of the indices during the *pre-virus*

recreation (RR), grocery & pharmacy (GP), parks (PA), transit stations (TS), workplaces (WP), and residential (RE), as denoted by vector $\alpha = (\alpha^{(RR)}, \alpha^{(GP)}, \alpha^{(PA)}, \alpha^{(TS)}, \alpha^{(WP)}, \alpha^{(RE)})^\top$.

Based on the discussion above, the stochastic modeling of β uses mobility indices as independent variables, whereas that of γ and δ does not use any independent variables. To investigate the dynamics of the log odds, we consider three forms of the dependent variables. Specifically, for $\xi_t = \beta_t$, γ_t, or δ_t, we consider the dependent variables of the following forms: 1) ξ_t; 2) $\xi_t - \xi_{t-1}$; and 3) $\xi_t/\xi_{t-1} - 1$. When the residuals are fitted with normality, the process for these three forms of dependent variables is essentially modeled as independent normal random variables, Brownian motion, and geometric Brownian motion, respectively. When we attempt to model the relationships between the infection rate (log odds) and the mobility indices, our preliminary studies show that using *5-day moving average* data on each mobility index produces better results. From an epidemiological perspective, infection is related to the number of contacts over a certain incubation period. We denote the moving average mobility indices as $\bar{\alpha} = (\bar{\alpha}^{(RR)}, \bar{\alpha}^{(GP)}, \bar{\alpha}^{(PA)}, \bar{\alpha}^{(TS)}, \bar{\alpha}^{(WP)}, \bar{\alpha}^{(RE)})^\top$. These indices are used as the explanatory variables for β.

3.2. U.S. COVID-19 data

In this subsection, we present the analytical and empirical results related to epidemic learning in Layer 1 of our framework. Our data on coronavirus cases, including the numbers of reported cases, active cases, and deaths, are derived from Worldometer.info; see Worldometers.info (2020). The data on I, R, and D correspond to the 114-day period from March 1 to June 21, 2020, as reliable Google mobility index data are not available before February 15, 2020. As we focus on the weekly mobility controls in place during June 2020, we conduct model fitting using data from each Sunday in that month, which enables us to

period of January 3 to February 6, 2020. A zero control vector represents a baseline day. Note that the baseline day is defined statistically and did not actually exist.

investigate the sensitivity of the estimated parameters with respect to the data input. Using (3), we can compute the log odds of infection, recovery, and death from March 2, 2020 ($t = 1$) to June 21, 2020 ($t = T := 113$). Although this paper illustrates our methodologies using data collected up to June 21, 2020, the stochastic modeling and suggested ESDPs (since June 21, 2020) are regularly updated at `https://sites.google.com/view/esd-covid19/`.

Before presenting the results of our data analysis, we note that the officially released COVID-19 statistics during the early outbreak period are subject to uncertainty and possible inaccuracy. For example, in the dataset from Worldometer.info, we observe occasional decreases in the total number of recovery cases up to March 19, 2020 and values of R and D that are smaller in magnitude relative to I. Therefore, when we analyze the $\{\gamma_t\}$ and $\{\delta_t\}$ processes, we only feed in data starting from March 19, 2020. However, as the controversial data points before March 19 do not have a large impact on I, we continue to study $\{\beta_t\}$ using data starting from March 1, 2020. After considering all of these factors, we consider it reasonable to believe that the official statistics are highly relevant to community mobility. First, these official data are publicly available and can affect community mobility by prompting voluntary behavioral changes. Second, governments use these statistics to develop social distancing policies, which naturally influence community mobility. In light of these considerations, we choose to model the rates of infection, recovery, and death as stochastic functions of time and mobility.

Initially, we attempt to fit the three transformations of $\xi_t = \beta_t$, γ_t, or δ_t with normality. We subject the log odds to Shapiro–Wilk normality tests (see Shapiro & Wilk (1965)). The p-values are reported in Table 2, where each row represents testing of the data available up to the indicated date. When the p-value is large (i.e., exceeds the significance level), we cannot reject the assumption of normality of the data. $\{\gamma_t/\gamma_{t-1} - 1\}$ can be assumed to be normally distributed, and $\{\delta_t/\delta_{t-1} - 1\}$ has a comparably normal distributed[2].

[2]The death rates are likely to be affected by other factors across time that are not captured

$\{\beta_t - \beta_{t-1}\}$ may also be considered to exhibit a marginally normal distribution. However, as discussed earlier, the infection rate (or log odds) generally depends on the country-level mobility indices.

Table 2: Statistical results of a Shapiro–Wilk normality test of the log odds of the three forms. The largest p-value in each cluster of log odds is highlighted in bold.

Date	β_t	$\beta_t - \beta_{t-1}$	$\beta_t/\beta_{t-1} - 1$	γ_t	$\gamma_t - \gamma_{t-1}$	$\gamma_t/\gamma_{t-1} - 1$	δ_t	$\delta_t - \delta_{t-1}$	$\delta_t/\delta_{t-1} - 1$
May 31	<1e-04	**0.0506**	<1e-04	<1e-04	0.0009	**0.5587**	0.0012	0.0216	**0.0355**
June 7	<1e-04	**0.0275**	<1e-04	<1e-04	0.0004	**0.6304**	0.0014	0.0237	**0.0388**
June 14	<1e-04	**0.0183**	<1e-04	<1e-04	0.0003	**0.6140**	0.0012	0.0116	**0.0222**
June 21	<1e-04	**0.0202**	<1e-04	<1e-04	0.0001	**0.6895**	0.0012	0.0047	**0.0102**

Next, we regress each transformation of the log odds of infection on the 5-day moving average mobility indices $\bar{\alpha} = (\bar{\alpha}^{(RR)}, \bar{\alpha}^{(GP)}, \bar{\alpha}^{(PA)}, \bar{\alpha}^{(TS)}, \bar{\alpha}^{(WP)}, \bar{\alpha}^{(RE)})^\top$. In other words, we consider the regression model $\text{LHS}_{t+1} = c_0 + c^\top \bar{\alpha}_t + \epsilon_{t+1}$ for $t = 1, \ldots, T-1$, where $\text{LHS}_{t+1} = \beta_t$, $\beta_t - \beta_{t-1}$ or $\beta_t/\beta_{t-1} - 1$; c_0 and c are regression coefficients; and ϵ_{t+1} is the error term at time $t+1$. The regression results for each period are presented in Table 3[3]. Our finding that the fit is greatly improved when the mobility indices are used as independent variables is promising. The best models consistently use β_t as a dependent variable; in this case, all of the coefficients of determination (R^2) are greater than 0.84, and all of the residuals pass the Shapiro–Wilk normality test with a significance level greater than 9.8%.

using the mobility indices considered in this paper. However, we adopt this basic model for simplicity.

[3]For fittings in which $\beta_t - \beta_{t-1}$ or $\beta_t/\beta_{t-1} - 1$ is used as a dependent variable, we observe that truncation of the problematic early data points (i.e., the first 20 points) improves the model fit. However, truncation does not improve the model fit when β_t is used as a dependent variable. Table 3 presents the best results with each potential dependent variable and shows that β_t remains superior.

Table 3: Results of the regression used to fit the log odds of infection with the moving average mobility indices. Standard errors are in curved brackets under the coefficient estimates. The p-values obtained from a Shapiro–Wilk normality test of the residuals are reported in the last column. The codes of significance are [0 '***' 0.001 '**' 0.01 '*' 0.05 '#' 0.1 ' ' 1]. The date in the header indicates the fitting date. The best fitting value in each period is highlighted in bold.

LHS	(Intercept)	$\bar{\alpha}^{(RR)}$	$\bar{\alpha}^{(GP)}$	$\bar{\alpha}^{(PA)}$	$\bar{\alpha}^{(TS)}$	$\bar{\alpha}^{(WP)}$	$\bar{\alpha}^{(RE)}$	R^2	p-value
May 31									
$\boldsymbol{\beta_t}$	**0.5824***	**-20.6616*****	**5.2851*****	**-2.0522****	**32.4049*****	**3.4093**	**34.9708*****	**0.8506**	**0.1706**
	(0.2527)	(3.4974)	(1.2411)	(0.7075)	(5.0816)	(4.3353)	(7.9842)		
$\beta_t - \beta_{t-1}$	0.1378	3.5887***	-1.5294#	-0.5604#	-2.5342	-1.9667	-6.3625**	0.3470	0.0413
	(0.5303)	(1.0015)	(0.8152)	(0.3157)	(2.3434)	(1.2252)	(1.8915)		
$\beta_t/\beta_{t-1} - 1$	0.5811	-3.7691***	1.5588*	-0.0191	4.9242**	0.3565	3.9427**	0.5170	<1e-04
	(0.3865)	(0.7298)	(0.5940)	(0.2301)	(1.7077)	(0.8928)	(1.3784)		
June 7									
$\boldsymbol{\beta_t}$	**0.4205#**	**-18.0781*****	**4.9546*****	**-1.9228****	**28.8964*****	**5.2908**	**36.7253*****	**0.8549**	**0.2183**
	(0.2288)	(3.2160)	(1.1862)	(0.6214)	(4.4397)	(3.6844)	(7.6324)		
$\beta_t - \beta_{t-1}$	0.1514	3.5472***	-1.5934*	-0.5584#	-2.4748	-1.9576#	-6.3926***	0.3447	0.0533
	(0.5216)	(0.8950)	(0.7875)	(0.3061)	(2.2280)	(1.0967)	(1.8285)		
$\beta_t/\beta_{t-1} - 1$	0.5617	-3.425***	1.5624**	-0.0369	4.487**	0.4889	3.9627**	0.5065	<1e-04
	(0.3713)	(0.6371)	(0.5606)	(0.2179)	(1.5859)	(0.7806)	(1.3016)		
June 14									
$\boldsymbol{\beta_t}$	**0.3180**	**-16.1974*****	**4.5635*****	**-1.7821****	**26.6568*****	**5.8699**	**36.6619*****	**0.8526**	**0.1589**
	(0.2268)	(3.1344)	(1.1792)	(0.5747)	(4.3151)	(3.6084)	(7.6185)		
$\beta_t - \beta_{t-1}$	0.0273	3.7934***	-1.3501**	-0.5879	-3.2047**	-1.8478	-6.6743**	0.3768	0.022
	(0.4811)	(0.8261)	(0.6988)	(0.2917)	(2.0765)	(1.0457)	(1.7541)		
$\beta_t/\beta_{t-1} - 1$	0.6694*	-3.2298***	1.3405**	-0.0853	4.6834**	0.3634	3.8151**	0.4992	<1e-04
	(0.3417)	(0.5869)	(0.4964)	(0.2072)	(1.4751)	(0.7428)	(1.2460)		
June 21									
$\boldsymbol{\beta_t}$	**0.3282**	**-16.9957*****	**4.4961*****	**-1.4419****	**28.5379*****	**3.7798**	**34.6980*****	**0.8413**	**0.0988**
	(0.2313)	(3.0850)	(1.2010)	(0.5302)	(4.3214)	(3.5796)	(7.7061)		
$\beta_t - \beta_{t-1}$	0.0965	3.6512***	-1.6301*	-0.5173#	-2.7958	-1.6731	-6.0067**	0.3534	0.0388
	(0.4690)	(0.8019)	(0.6658)	(0.2959)	(1.9596)	(1.0068)	(1.7752)		
$\beta_t/\beta_{t-1} - 1$	0.6708*	-3.0629***	1.357**	-0.1323	4.5882***	0.2053	3.5132**	0.4871	<1e-04
	(0.3170)	(0.5420)	(0.4500)	(0.2000)	(1.3245)	(0.6805)	(1.1998)		

Therefore, a suitable stochastic model of the log odds within the context of the COVID-19 pandemic in the United States is proposed as follows:
for $t = 1, \ldots, T-1$,

$$\beta_{t+1} = c_0 + c^\top \bar{\alpha}_t + \sigma_\beta Z^\beta_{t+1},\ \gamma_{t+1} = \gamma_t(1+\mu_\gamma+\sigma_\gamma Z^\gamma_{t+1}),\ \delta_{t+1} = \delta_t(1+\mu_\delta+\sigma_\delta Z^\delta_{t+1}), \tag{4}$$

where $\bar{\alpha}_t = \frac{1}{5}\sum_{t-4}^{t} \alpha_t$ and $Z^\beta_{t+1}, Z^\gamma_{t+1}, Z^\delta_{t+1}$ are independent and identically distributed (iid) $N(0,1)$ random variables for $t = 1, \ldots, T-1$. We highlight

Table 4: Estimated parameters of the model in (4) for different periods in 2020 using historical data.

Date	μ_γ	σ_γ	μ_δ	σ_δ	σ_β
May 31	0.0107	0.1987	0.0058	0.0432	0.5036
June 7	0.0178	0.1996	0.0062	0.0433	0.4988
June 14	0.0173	0.1963	0.0061	0.0444	0.5032
June 21	0.0176	0.1919	0.0061	0.0451	0.5166

the noise-like randomness observed when modeling β. Although δ and γ are modeled with stochastic process-like uncertainty, their magnitude and volatility are small compared to those of β. The intercept c_0 and coefficients $c \in \mathbb{R}^6$ for fittings on different dates can be found in Table 3, and other parameters (mean μ and standard deviation σ) can be easily estimated using historical data, as given in Table 4. Table 3 and Table 4 show that the parameter estimates remain highly stable throughout June. Although the intercept c_0 and the coefficient corresponding to $\bar{\alpha}^{(WP)}$ are sometimes statistically insignificant, we retain these in the model to maintain consistency and capture all of the effects of mobility. The regression coefficients in Table 3 should be interpreted cautiously because the mobility indices are highly correlated with each other, a phenomenon known as multicollinearity. However, we use the regression model only to predict the drift of β_{t+1} in (4) and, thus, our prediction is not affected by multicollinearity.

3.3. Responses to COVID-19 and the implications in simulation studies

The SIRD model (2), together with the stochastic processes of (4), can be used to describe the evolution of an infectious disease at the country level. In this subsection, we conduct simulation studies to analyze some scenarios of COVID-19 in the United States.

Note that a country's daily mobility indices fluctuate, and different indices are highly correlated with each other. It is very difficult to fix community mobility, α, as an arbitrary vector. In this subsection, we investigate real historical representative mobility data and how the variability in these data might

affect the evolution of the COVID-19 pandemic. In particular, community mobility reflects the different social distancing measures and restrictions on public gatherings imposed by the government. Since the initial outbreak, the U.S. government's response to the COVID-19 pandemic has varied over time; see Cheng et al. (2020) and Ritchie et al. (2020). Based on the government's responses, we consider four periods and consider the median mobility indices in each period as representative of community mobility in the corresponding period. The results are presented in Table 5.

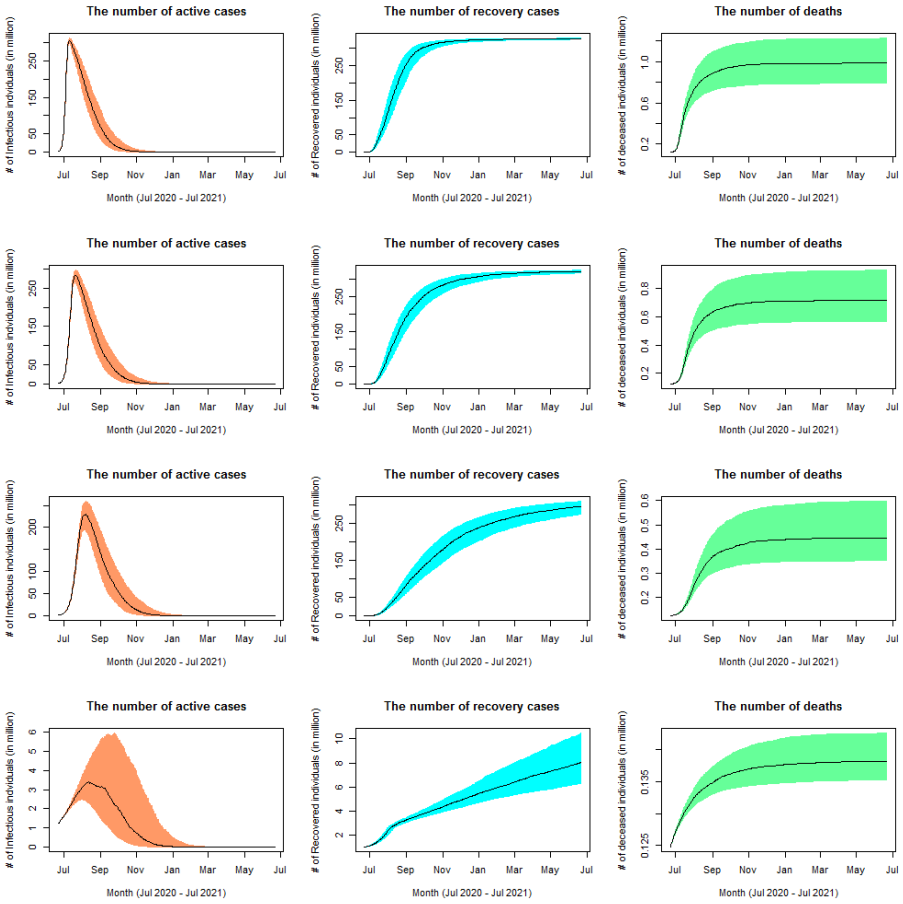
Table 5: Median mobility indices over different periods corresponding to different major U.S. governmental pandemic measures.

Period	Government's responses	Median mobility indices
January 3–February 6	Baseline	$\alpha^{(0)} = (0, 0, 0, 0, 0, 0)^\top$
February 15–March 4	Alerts	$\alpha^{(1)} = (0.07, 0.02, 0.12, 0.02, 0.02, -0.01)^\top$
March 5–March 18	School closures	$\alpha^{(2)} = (0.055, 0.09, 0.15, -0.045, -0.015, 0.01)^\top$
March 19–June 21	School & workplace closures	$\alpha^{(3)} = (-0.3, -0.07, 0.06, -0.42, -0.4, 0.15)^\top$

Next, we simulate the SIRD model (2) and (4) using parameters estimated on June 21, 2020 and the mobility indices listed in Table 5. Although the mobility indices are merely median values, we find that $\alpha^{(1)}$, $\alpha^{(2)}$, and $\alpha^{(3)}$ are close to the indices calculated on February 27, March 10, and May 6, respectively, in terms of Euclidean distance. For each measure of community mobility, we conduct 10,000 simulations and compute the median curves of I, R, and D over the subsequent 1-year period (June 22, 2020–June 21, 2021). We also include the 0.45- and 0.55-quantile curves to visualize the volatility of the data. The results calculated using different mobility controls are presented in Figure 2.

In Figure 2, we can see that choosing the mobility index during January 3–March 18, 2020 results in many cases of infection and death. At this level of baseline mobility, more than 90% of U.S. population would be infected. Following the responsive public health measures imposed in early March, the numbers of cases of infection and death are reduced, although the absolute values remain high. These measures lead to a significant decrease in the numbers of cases of infection and death when the mobility index $\alpha^{(3)}$ is used.

Figure 2: Simulated U.S. COVID-19 cases with mobility controls $\bar{\alpha}_t \equiv \{\alpha^{(0)}, \alpha^{(1)}, \alpha^{(2)}, \alpha^{(3)}\}$ (from top to bottom). The black curves represent the median values, and the colored shadow areas are bounded by the 0.45- and 0.55-quantiles.



We also observe that the variances in I, R, and D are time-inhomogeneous. Specifically, these variances are positively correlated with the value of I, which represents current cases of infection. This relationship can be explained by a specification in the SIRD model (2). Hence, controlling mobility not only lowers the infection rate but also stabilizes the pandemic situation.

Again, we emphasize that the model used in this paper does not take into

account viral variants but rather represents single-pattern outbreaks of a major virus strain. As a result, the suggested method for developing optimal social distancing policy should be online based on updated estimations of parameters and re-initiation of the optimization procedure due to the limited time window for the analysis of epidemic evolution. We note that the simulation results presented in Figure 2 are based on data up to June 21, 2020 and do not coincide with currently available real U.S. COVID-19 data. This discrepancy is partly attributable to (i) the continuously changing landscape of COVID-19, particularly with the emergence of VoC since September 2020 [4], and partly to (ii) changes in the U.S. community mobility. In light of these considerations, Figure 2, which presents simulated cases, emphasizes the importance of persistent implementation of social distancing policies. If the infection, recovery, and death data are further classified by virus strain, extending our approach to a scenario with multiple strains is feasible, given the newly proposed multi-strain SIR models (Fudolig & Howard (2020)). However, such data are not currently available.

Different governmental responses or community mobility patterns are expected to lead to different epidemic processes. However, a natural question arises: can we determine the best mobility control (in terms of the mobility indices) with respect to lowering the infection rate and maintaining financial market stability? To this end, we must investigate the effects of the pandemic on the national economy.

4. Exploring the economic impact of COVID-19 by leveraging financial market data

Although restrictive interventions can help control disease transmission, they may have a negative economic impact. In theory, ideal social distancing policies

[4] See the World Health Organization website for more information: `https://www.who.int/en/activities/tracking-SARS-CoV-2-variants/`

would balance the benefit of reducing public health risks by reducing the infection transmission mortality rate with the benefit of mitigating the aggregated economic costs of these interventions. In practice, the aggregated economic costs increase the risks of economic downturns, disease spreads, and social dissatisfaction. Recent studies provide many insights on this topic. For example, the cost is considered as a loss of consumption in Hall et al. (2020), quantified using multiple indicators based on economic and consumer-based contributions in Benzell et al. (2020), and expressed in terms of restrictions on social contacts in Cont et al. (2020). In a study from 2021 Tanrisever et al. (2021), the cost is considered as the negative impact on both economic activity and mortality. We further aim to understand the COVID-19-driven aggregated economic cost from a different perspective, namely as a function of the pandemic and mobility data, by exploring the corresponding financial market data. Following empirical studies on the 2003 SARS epidemic (Chou et al. (2004); Chen et al. (2007); Wong (2008)), we share the perspective that financial markets were highly influenced by early COVID-19 outbreaks, which reflects investors' reactions to the crisis and their economic expectations.

In this section, we explore the relationships between COVID-19 statistics, mobility controls, and the market index during the early stage of the COVID-19 pandemic. In this paper, we choose the S&P 500 as a representative market index. As the most important weighted-average market capitalization index in the United States, the S&P 500 is considered the index most representative of the U.S. stock market and consequently is widely regarded as a key indicator of the U.S. economy (Parnes, 2020). The National Bureau of Economic Research (NBER) also classifies this index as a leading indicator of U.S. business cycles. Heiberger (2018) empirically demonstrates a strong correlation between the S&P 500 and normal growth of the U.S. economy. Once the relevant relationships are identified, we postulate that the economic impact of the COVID-19 pandemic will be reflected in market stability during the crisis period. The aim of efficient mobility control policy is to achieve public health goals (e.g., a target infection rate less than 1) with a minimum impact on financial market stability. To con-

ceptualize efficiency, we seek the optimal set of mobility control policies that minimizes the acquired cost functional while incorporating both public health outcomes and the financial market stability risk.

4.1. S&P 500 prices data

In this subsection, we present our data analysis and empirical results related to financial implication learning, which occurs in Layer 1 of our framework. We use S&P 500 prices downloaded from Yahoo!Finance. We conduct an ex-post regression of daily S&P 500 closing prices (SPX_t) on COVID-19 statistics and mobility indices. Specifically, we postulate that

$$\mathrm{SPX}_t = \kappa_0 + \kappa^\top \alpha_t + \kappa_I I_t + \kappa_R R_t + \kappa_D D_t + \epsilon_t^{SPX}$$

for $t = 1, \ldots, T$, where κ_0, $\kappa \in \mathbb{R}^6$, κ_I, κ_R, and κ_D are the regression coefficients, and ϵ_t^{SPX} is the error term at time t. Note that we use α instead of the moving average $\bar{\alpha}$ as an independent variable because the market's reaction to community mobility is immediate. The regression results are reported in Table 6. Each model fitting is based on daily data (business days) from March 9 to the fitting date indicated in the table.

Table 6 shows that all of the model fittings are good, as all of the R^2 values exceed 0.89 and the residuals pass the Shapiro–Wilk normality test at any significance level greater than 38%. Therefore, the disease statistics and mobility indices have high explanatory power with respect to the SPX price. Although some independent variables are insignificant, we retain all of them because we seek a high explanatory power. Again, the coefficients must be interpreted cautiously due to multicollinearity. However, this limitation does not affect our use of a regression model as a proxy for COVID-19 driven component derived from the S&P 500 prices.

5. Efficient social distancing

5.1. The optimization problem of searching for an efficient policy

In this subsection, we detail the formulation of stochastic feedback control

Table 6: Regression of S&P 500 index prices fitted with U.S. COVID-19 statistics and mobility indices in 2020. Standard errors are in curved brackets under the coefficient estimates. The R-squared values and p-values calculated via the Shapiro–Wilk normality test of the residuals are reported next to the fitting dates in the headers. The codes of significance are [0 '***' 0.001 '**' 0.01 '*' 0.05 '#' 0.1 ' ' 1].

κ_0 (Intercept)	$\kappa^{\top}_{\alpha^{(RR)}}$	$\kappa^{\top}_{\alpha^{(GP)}}$	$\kappa^{\top}_{\alpha^{(PA)}}$	$\kappa^{\top}_{\alpha^{(TS)}}$	$\kappa^{\top}_{\alpha^{(WP)}}$	$\kappa^{\top}_{\alpha^{(RE)}}$	κ_I	κ_R	κ_D
May 31 ($R^2 = 0.8903$, p-value= 0.6291)									
2919.7***	2.5	-832.7***	-131.1	1847.4	-64.0	830.4	5.958e05***	3.943e05*	-6.493e06**
(96.1)	(722.6)	(235.9)	(197.6)	(1049.7)	(824.2)	(918.0)	(1.374e05)	(1.485e05)	(1.972e06)
June 7 ($R^2 = 0.9075$, p-value= 0.3995)									
2949.2***	-71.5	-811.1***	-233.6	2202.0*	-288.6	790.6	6.508e05***	5.578e05***	-7.668e06***
(89.1)	(667.8)	(230.6)	(182.0)	(978.5)	(793.7)	(894.4)	(1.304e05)	(1.128e05)	(1.815e06)
June 14 ($R^2 = 0.9039$, p-value= 0.3899)									
2917.0***	601.5	-943.5***	-314.3	1656.3	-380.7	855.7	4.925e05***	4.008e05***	-5.361e06**
(92.9)	(655.3)	(243.2)	(190.9)	(1011.5)	(841.8)	(945.1)	(1.276e05)	(9.899e04)	(1.741e06)
June 21 ($R^2 = 0.8979$, p-value= 0.4141)									
2875.8***	1108.1#	-1100.5***	-301.0	1001.4	-179.1	1195.2	2.772e05*	1.957e05*	-2.217e06
(91.4)	(623.6)	(246.3)	(193.9)	(982.4)	(862.5)	(963.2)	(1.111e05)	(7.793e04)	(1.470e06)

optimization in Layer 2 of our framework. Searching for an efficient policy is formulated as a stochastic feedback control optimization problem with a convex combination of the public health risk and financial market stability risk as the objective:

$$\min_{\alpha_{t+1},\cdots,\alpha_{t+h}} \{(\text{public health risk}) + \lambda(\text{financial market stability risk})\}, \quad (5)$$

where λ is a weight parameter subjective to political choice, and $\alpha_{t+1}, \cdots, \alpha_{t+h}$ constitute the data-driven feedback mobility control policy from day $t+1$ to day $t+h$. When $\lambda = 0$, the policymaker only cares about public health, regardless of the financial market stability. This condition could lead to severe lockdown. In contrast, when λ is positive, the policymaker is concerned about financial market stability. However, an extremely large λ indicates that public health is neglected, resulting in accommodative social distancing policies. λ is subject to policy policymakers' judgment, which is influenced by political issues and national conditions. If the conceptual problem (5) can be solved for any fixed $\lambda \geq 0$, then the optimal mobility control $\alpha^*(\lambda)$ is called an ESDP. By varying

λ, the set of ESDPs can be visualized as an efficient frontier that explains the trade-off between the public health and market stability risks. We define this efficient social distancing frontier (ESDF) as the set of all ESDPs: ESDF = $\{\alpha^*(\lambda) : \lambda \geq 0\}$.

During the COVID-19 pandemic, infection containment via social distancing is an obvious public health target. From a statistical perspective, we focus on the probability of an increasing infection rate $f_\beta(\beta_t)$ in response to a social distancing policy versus the probability of a decreasing infection rate. Essentially, the ratio between these two probabilities is the odds of infection. In Section 3, we use an empirical dataset and expend considerable effort to determine statistically how the log odds of infection, β_t, is correlated with the stochastic epidemic model with mobility controls. Therefore, as the public health target, we aim to minimize the expected log odds of infection, $\min E[\beta_t]$, at a future time t. In other words, we set the expected future log odds of infection as the public health risk. Because the log odds of infection is an increasing function of the infection rate, it is more sensitive to changes in social distancing policy, thus facilitating numerical optimization efficiency.

We model economic impact as the stability of the financial market value associated with the epidemic model and social distancing policy. Using the predicted SPX over the indicated COVID-19 crisis period, we statistically determine the COVID-19-driven SPX value at a given time t as

$$\widehat{\text{SPX}}_t = \hat{\kappa}_0 + \hat{\kappa}^\top \alpha_t + \hat{\kappa}_I I_t + \hat{\kappa}_R R_t + \hat{\kappa}_D D_t, \tag{6}$$

where $\hat{\kappa}_0$, $\hat{\kappa}$, $\hat{\kappa}_I$, $\hat{\kappa}_R$, and $\hat{\kappa}_D$ are the estimated parameters as presented in Table 6. The financial market stability risk is defined as the tracking error between COVID-19-driven SPX growth and reasonable growth. Mathematically, at a given time t, the tracking error over 1 day is calculated as

$$\left(\frac{\widehat{\text{SPX}}_{t+1} - \widehat{\text{SPX}}_t}{\widehat{\text{SPX}}_t} - r_t \right)^2 \propto \left(\widehat{\text{SPX}}_{t+1} - \text{Target}_{t+1} \right)^2,$$

where r_t is the target daily growth rate, and $\text{Target}_{t+1} = \widehat{\text{SPX}}_t(1 + r_t)$ is the target SPX value at time $t + 1$. A reasonable growth rate r_t can represent the

risk-free interest rate, the inflation rate, or the GDP growth rate. However, when the period is small (e.g., in the near future), all of the target rates are close to zero. In fact, $r_t = 0$ is used to minimize the expected tracking error with the aim of reducing the risk of a significant financial market downturn without emphasizing growth.

Therefore, we model our conceptual framework, which integrates public health and financial market stability as in (5), using the (running) cost functional:

$$c(t+1, X_{t+1}, \lambda; \alpha_t) = \beta_{t+1} + \lambda(\widehat{\mathrm{SPX}}_{t+1} - \mathrm{Target}_{t+1})^2, \tag{7}$$

where $X_t := (I_t, R_t, D_t, \beta_t, \gamma_t, \delta_t)^\top$. In reality, a policymaker seldom considers the effect on a single future day but rather the effect on a future time period h. For instance, if a policymaker considers the projection for 5 business days in a week, then $h = 5$, which is also the period selected in our numerical study. Therefore, we formulate and quantify the efficient social distancing problem (5) as a stochastic control problem:

$$\begin{aligned} \min_{\alpha_s \in \mathcal{A}_s,\ s=t,\ldots,t+h-1} \quad & J(t, X_t; \{\alpha_s\}_{s=t}^{t+h-1}) := \mathbb{E}\left[\sum_{s=t+1}^{t+h} c(s, X_s, \lambda; \alpha_{s-1})\right] \\ \text{s.t.} \quad & \{X_s\}_{s=t+1}^{t+h} \text{satisfies (2) and (4)},\ \widehat{\mathrm{SPX}}_t \text{satisfies (6)}, \end{aligned} \tag{8}$$

where the cost functional c is given in (7), and $\mathcal{A}_s$ is the feasible set of mobility controls determined at time s. As discussed above, the control, α, must be constrained in the optimization procedure to obtain reasonable results, as the mobility indices are intercorrelated. This inherent correlation structure makes it impractical to simply impose element-wise constraints using the initial data field. Therefore, we aim to find a transformation of α that yields decorrelated elements on which we can impose element-wise constraints. Specifically, we use historical mobility index data to determine the transformation $\hat{A}$ from indices α to their PCs and obtain $P = \hat{A}\alpha$, which can also be interpreted as a Karhunen–Loéve transform (KLT) in decorrelation operations (see e.g., Dony (2001)). We then impose element-wise constraints on $\hat{A}\alpha$. To determine these constraints, we compute the historical PCs, $P \in \mathbb{R}^6$, and obtain the lower and upper bounds

of each historical PC entry, namely $\hat{L} \in \mathbb{R}^6$ and $\hat{U} \in \mathbb{R}^6$, respectively. The definition of α requires each α entry to fall within the range of $[-1, 1]$. Let $\hat{A}_t$, $\hat{L}_t$, and $\hat{U}_t$ represent the estimates calculated using data up to time t. Then, the feasible set $\mathcal{A}_s$ at time $s \in [t, t+h-1]$ is given by

$$\mathcal{A}_s := \{\alpha_s(w_s) : \mathbb{R}^{30} \to \mathbb{R}^6 \mid \hat{L}_t \leq_e \hat{A}_t \alpha_s \leq_e \hat{U}_t \text{and} -\mathbf{1} \leq_e \alpha_s \leq_e \mathbf{1}\}, \quad (9)$$

where $\leq_e$ is the element-wise less than or equal to operator, and $\mathbf{1} = (1, \ldots, 1)^\top \in \mathbb{R}^6$. Note that the feasible set (9) is an approximation of the full feasible set. However, our numerical studies show that this approximation works well in practice and yields ESDPs similar to some realized mobility indices.

To incorporate the PC constraints on α_s into the objective function, we use a Lagrangian approach, such that the revised cost is

$$\overline{c}^{\lambda}_{s+1} = c_{s+1} + \sum_{i=1}^{6} \overline{\lambda}_i \left[(\hat{A}_t \alpha_s - \hat{L}_t)^+_i - (\hat{A}_t \alpha_s - \hat{U}_t)^-_i \right]^2, \quad t \leq s \leq t+h-1, \quad (10)$$

where $(a)_i$ denotes the i^{th} entry of vector a, and $\overline{\lambda}_i$ is a penalty coefficient of $i = 1, 2, \ldots, 6$. Using the available dataset, we perform cross-validation to select $\overline{\lambda}_i$ and thus ensure that the learned optimal controls strictly satisfy all of the constraints. Cross-validation is a well-accepted classic machine and statistical learning technique used to identify the most useful opportunity set.

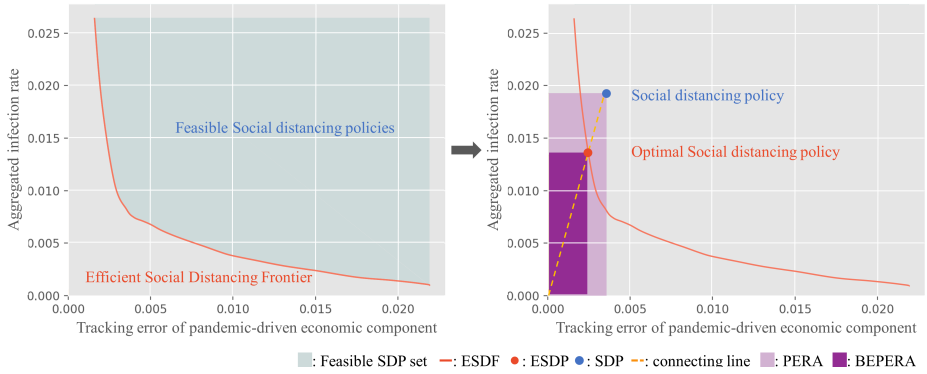
We then solve (8) numerically using the deep learning framework in Tsang & Wong (2020). The implementation procedure is detailed in Appendix A.

5.2. Efficiency measurement

After solving the ESDP from the stochastic feedback control problem, we can compare the efficiency of an existing social distancing policy with that of the generated ESDP. We propose a simple measure for this purpose.

First, we visualize the ESDF using a two-dimensional coordinate system. In the graph, the $y-$axis denotes the aggregated infection rate, which is transformed from the log-odds of infection as $\sigma(\sum_{h=1}^{5} \beta_{t+h}/5)$. The $x-$axis denotes the tracking error (TE) of the COVID-19-driven SPX value. To match the or-

Figure 3: Illustration of the efficiency ratio (ER)



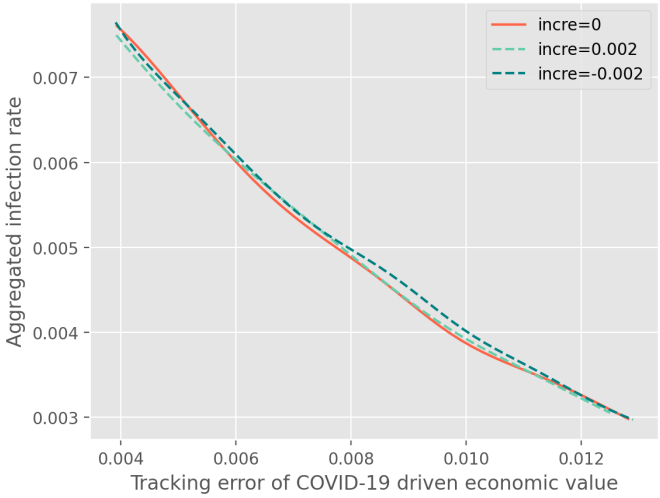
der of magnitude between the infection rate and the TE, we report the latter in terms of the root mean square error (RMSE) in returns. Specifically,

$$\mathrm{TE} = \sqrt{\frac{1}{h}\sum_{s=t+1}^{t+h}\left(\frac{\widehat{\mathrm{SPX}}_s - \widehat{\mathrm{SPX}}_{s-1}}{\widehat{\mathrm{SPX}}_{s-1}} - r\right)^2}.$$

The graph on the left side of in Figure 3 depicts a typical ESDF, which is computed for the week containing June 29, 2020 by setting the target daily increment rate r as 0. In other words, the identified COVID-19-driven SPX value during the upcoming week should remain near the initial level calibrated on June 26, 2020. A sensitivity analysis comparing the ESDF against the target rate r is shown in Figure 4. It is clear that a reasonable choice of r_t does not materially affect the ESDF, which consistently takes a bullet-shaped curve lower than that of the target rate. In the left panel of Figure 3, the shaded area enclosed from below by the ESDF is the region of epidemic-economic outcomes that may result from feasible social distancing policies. The lower bullet-shaped ESDF curve is drawn by varying the value of λ.

Any social distancing policy within the feasible set corresponds to a point (TE, aggregated infection rate) in the shaded region, as illustrated by the blue point in Figure 3. To quantify the efficiency of the indicated policy, we define the public-economic risk area (PERA) as the product of the infection rate multiplied

Figure 4: ESDFs with different increment rates



by the TE (i.e., the product of the coordinates of the blue point), as illustrated by the large purple rectangle in Figure 3. Clearly, efficiency decreases as the PERA increases. The origin (0,0) reflects the idealized situation, in which both the aggregated infection rate and financial market stability risk are zero. We draw a straight line connecting the feasible point and the origin, illustrated by the yellow dashed line, to indicate the direction of efficiency improvement. The point where this line intersects the ESDF $(\text{TE}^*, \text{infection rate}^*)$, is illustrated by the orange point in Figure 3 and represents our benchmarking ESDP. Similarly, we measure the PERA of the benchmarking ESDP, which is designated the BEPERA and illustrated by the small dark purple rectangle in Figure 3. We then use the ratio of PERA and BEPERA to measure the efficiency of a candidate social distancing policy (i.e., the blue point). Hence, the efficiency ratio (ER) is defined as

$$\text{ER} = \frac{\text{BEPERA}}{\text{PERA}} = \frac{\text{TE}^* \times \text{infection rate}^*}{\text{TE} \times \text{infection rate}} \in [0, 1].$$

The greater the ER, the higher the efficiency of a candidate social distancing

policy. The ESDP has an ER equal to 100%. We adopt this measurement due to its simplicity in calculation and comprehensibility.

5.3. Results and discussion

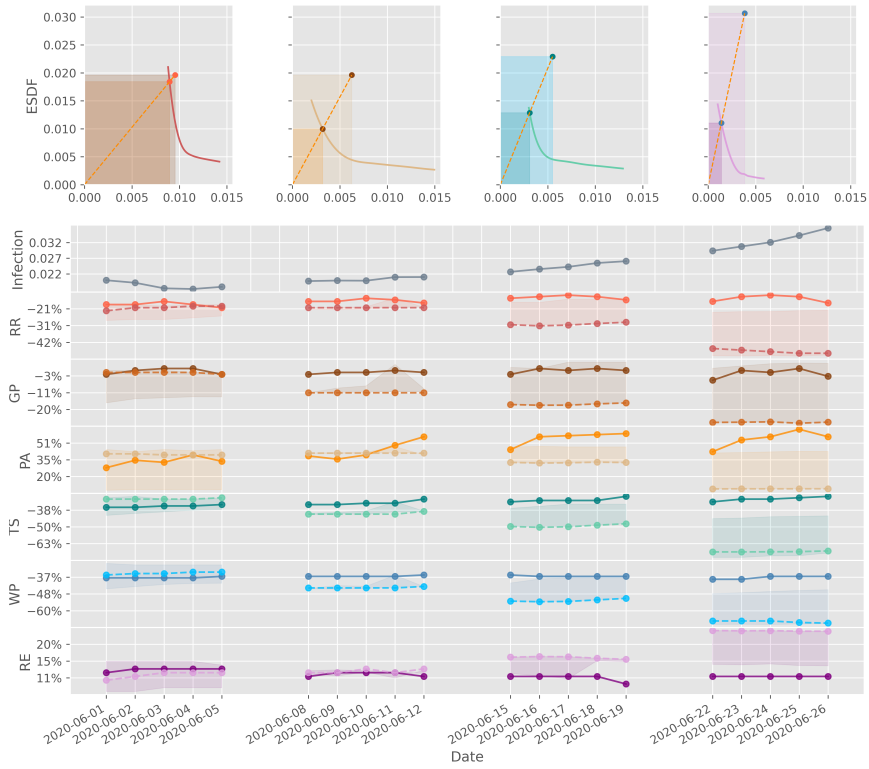
In this subsection, we discuss the empirical results related to the ESDF and efficiency measurements as the output of our framework. We emphasize that the discussions and speculations arising from this empirical study are based on relatively simple models, and that the data may be subject to uncertainty. We suggest that all policy discussions based on relevant numerical findings should be conducted cautiously. Nevertheless, we attempt to present some insights regarding the efficiency of historical implemented policies and potential improvements to social distancing measures.

Figure 5 displays the ESDPs for the validation dataset ranging from June 1 to June 26, 2020. We set $r = 0$ and compute ESDPs over the range of $\lambda \in (0.001, 0.1)$. Note that all of the computed ESDPs satisfy the constraints on the control. As Figure 5 contains multiple types of information, we explain each part below:

1. We aim to visualize the efficiency of historical social distancing policies by benchmarking existing policies against our ESDPs. In the top panel of Figure 5, the graph shows the weekly ESDFs (curves), historical epidemic-economic points (corresponding to historical social distancing policies), and benchmarking ESDPs over a 4-week period in June 2020. Thus, we can compute the ERs corresponding to this 4-week period. Both our inspection and the ER number demonstrate that the policy during the first week is most efficient. In fact, the ER number decreases throughout June, implying a loss of efficiency of historical social distancing policies over the month. The second panel shows the historical trends in infection rate over the same 4-week period. Although we recognize a slight decreasing trend in the first week, an increasing trend emerges in the second week, and infection rate continues to increase throughout the month. This trend seems to be related to the efficiency of the social distancing scheme.

2. The third panel of Figure 5 presents the historical mobility indices, namely

Figure 5: U.S. mobility index controls. In each panel, the shaded area depicts the range of average efficient mobility controls over the validation data set, and the solid line represents historical mobility.



RR, GP, PA, TS, WP, and RE, and the ESDPs. Solid lines represent the realized indices in the dataset. As these historical indices are typically more volatile over weekends than on weekdays, we only report the results for the latter (business days). Although all of the historical mobility indices appear to remain stable over the studied period, we identify slight increases in RR and TS and a clear increasing trend in PA over the last 3 weeks relative to the first week, indicating a certain level of relaxation in mobility constraints. The shaded area in each mobility index graph represents the range of ESDPs over the range of $\lambda \in (0.001, 0.1)$. The dashed lines represent the recommended ESDPs, which

are obtained as follows. First, we use the mobility indices from the previous Friday and compute the TE (infection rate) of the current week from the training sample. Second, we identify the $\hat{\lambda}$ at which the ESDP offers the same TE to obtain our recommended ESDP. Here, we aim to maintain the financial market stability risk level while reducing the infection rate to the efficient level. Based on the results of training samples in primary tests, a suggested range of $\lambda \in [0.005, 0.1]$ is identified; this range is applicable in practical settings and can always cover the recommended ESDP $\hat{\lambda}$. Figure 5 shows that the policy during the first week is very close to the recommended ESDP. In subsequent weeks, the implemented policies gradually drift from the recommended ESDPs, which state that social restrictions should not be lifted and that measures in some categories should be reinforced. Such a widening gap between the implemented policies and recommended ESPD appears to be associated with the observed increasing trend in the infection rate.

6. Conclusion

The major contribution of this research is the proposal of an operational research framework to determine an ESDP. We demonstrate the utility of this framework by applying it to U.S. COVID-19 data. We note that we only consider publicly available data because we are focused on social distancing. This is a clear limitation of our study because we are unable to obtain various types of data, including daily asymptomatic infection case numbers and mask-wearing, hospitalization, and ICU use statistics. Our dynamic models of the COVID-19 pandemic takes into account both data availability and randomness. The resulting stochastic model is a SIRD-type equation that includes community mobility indices generated via big data analysis from Google as new factors to account for the population response to social distancing policies. Using U.S. COVID-19 data, we find that the fittings of our stochastic SIRD model improve significantly with the inclusion of mobility data. We also find that during the pandemic crisis, the power of the mobility indices to explain market index (S&P

500) responses is high. Our stochastic modeling of COVID-19 in the U.S. allows us to simulate possible future scenarios in response to specific mobility indices and thus verify the effectiveness of social distancing policies. We formulate the optimization problem of a search for efficient policy as a stochastic feedback control problem that objectively integrates public health outcomes and financial market impacts. We then pose a realistic problem with analytical challenges and solve it using a deep learning approach.

Under our framework, modeling can be improved in various ways to improve accuracy. For example, the epidemic process modeled using our stochastic SIRD model does not consider micro perspectives, such as population behavior and hygiene measures (e.g., mask-wearing, hand-washing). Our model also ignores subcategories of confirmed cases, such as serious or non-serious illnesses and asymptomatic infections, and does not stratify patients by age due to data unavailability. This limitation implies a promising direction for future research once these data become available. The decision to use a market index or the SPX as a proxy for the national economy is debatable, and the trade-off between the economy and public health is specified by policymakers; therefore, a single "correct" answer applicable to different governments cannot be achieved. Because the models considered in this paper include the statistical modeling of noise, they may be oversimplified. We expect that more sophisticated models will be examined in the future. Although the quantification of uncertainty may vary and even yield improved results, the framework that we use to explore the efficiency of social distancing policies in this paper signifies a contribution of blended operational research methods to the field of pandemic risk management. In particular, we demonstrate the importance of modern techniques, such as big-data analytics for generating community mobility and machine learning algorithms for understanding and controlling pandemic situations. Our proposed framework can be extended to accommodate more practical considerations, such as the effect of vaccination and viral variants on epidemiological features, the latency period and asymptomatic cases as well as geographical and age stratification.

Acknowledgements

We thank the Editor and Associate Editor for handling this paper and constructive comments from four anonymous referees. H.Y. Wong acknowledges the Research Matching Grant (RMG project code: 8601495) received from the Research Grants Council of Hong Kong.

Appendix A. Deep learning solution

As explained in the main text, identifying efficient adaptive polices is formulated as a stochastic feedback control problem with constraints. Analogous problems without constraints can be solved analytically using the dynamic programming principle. However, our problem imposes complex constraints on the control set, making analytical solutions impossible, and even traditional numerical methods do not guarantee a numerical solution. Two numerical algorithms in the literature might yield solutions. The first is a backward regression-based simulation algorithm, which is described in detail in Li et al. (2020). However, empirical studies suggest that this algorithm has significant bias. Alternatively, a deep learning approach recently investigated by Tsang & Wong (2020) offers a practical solution to this type of numerical challenge. Using such an algorithm, we show empirically that the stochastic gradient descent method can link to a deep neutral network training problem (DNN) and thus compute numerical solutions efficiently. As the DNN is closely related to the Markov decision process (MDP), we transform the stochastic control problem (8) into a discrete-time MDP, as considered in Tsang & Wong (2020). Specifically, we write the problem as follows:

$$\begin{aligned} V_t(w) = &\min_{\alpha_s \in \mathcal{A}_s,\ s=t,\dots,t+h-1} J(t, X_t; \{\alpha_s\}_{s=t}^{t+h-1}), \\ &\text{s.t.} \qquad w_s = \phi(w_{s-1}, \alpha_{s-1}(w_{s-1}), \eta_s),\ s = t+1, \cdots, t+h,\ w_t = w, \end{aligned} \tag{A.1}$$

where $w_s := (X_s, \widetilde{\alpha}_s) \in \mathbb{R}^{30}$ is the state variable; $\widetilde{\alpha}_s := (\alpha_i)_{i=s-4}^{s-1} \in \mathbb{R}^{24}$ is a vector of the mobility indices over time steps from $s-4$ to $s-1$; ϕ denotes the transition function given by the model (2) and (4) wherein $\widetilde{\alpha}_{s+1}$ updates $\widetilde{\alpha}_s$ by

deleting α_{s-4} and adding α_s; and $\eta_s := (Z_s^\beta, Z_s^\gamma, Z_s^\delta) \in \mathbb{R}^3$ is a vector of noisy information at time s. The model is fully characterized by the state variable w_s, and thus the optimal control α_s depends only on the current state w_s. To simplify the notation, we define the cumulative cost $C_\tau^\lambda = \sum_{s=t+1}^{\tau} \bar{c}^\lambda(s, X_s, \lambda; \alpha_{s-1})$ for $t+1 \leq \tau \leq t+h$, where $\bar{c}^\lambda$ is defined in (10). Next, we detail the deep learning solution to the MDP (A.1). This solution agrees with that proposed in Tsang & Wong (2020), in which the theoretical convergence of the deep learning solution is established and a computation up to 100 dimensions is demonstrated. Our problem also encounters high dimensionality but contains a maximum of 30 dimensions, including the moving average controls. As mentioned before, the optimal control, α, is a function of the current state w. We represent this dependence using a multilayer feedforward neural network (FNN) with L hidden layers in the following form:

$$w_s \in \mathbb{R}^{30} \to g_s^{L+1}(\theta_s) \in \mathbb{R}^6 \to \alpha_s, \quad s = t, \cdots, t+h-1, \tag{A.2}$$

where θ_s denotes the parameters of the network. Let $\mathcal{D}^L = \{g^{L+1}(\theta) : \mathbb{R}^{30} \to \mathbb{R}^6\}$ be the class of functions computed by the standard FNN. $\mathcal{D}^L$ can be represented as follows:

$$\begin{aligned} g^0 &= w \\ g^\ell &= \sigma_\ell(B^{\ell-1} g^{\ell-1} + M^{\ell-1}),\ \ell = 1, \cdots, L, \\ g^{L+1} &= B^L g^L + M^L, \end{aligned} \tag{A.3}$$

where $\sigma_\ell(\cdot)$ are non-linear activation functions, B^ℓ is the matrix weight, and M^ℓ is the vector bias and we denote $\theta = (B^\ell, M^\ell)_{\ell=0}^L$. In this FNN, $g^0 \in \mathbb{R}^{30}$ is the input layer, $g^{L+1} \in \mathbb{R}^6$ is the output layer, and $\{g^\ell\}_{\ell=1}^L$ are hidden layers to which the nonlinear activation functions are applied. Many activation function options for $\sigma_\ell(\cdot)$ are available, including the tanh activation function, rectified linear unit (ReLU), and softmax activation function.

To impose control bounds on the last step, we apply $\tanh(x) := \frac{2}{1+e^{-2x}} - 1 \in [-1, 1]$ to the output layer g^{L+1} in an element-wise manner. The class of control

functions computed by (A.2) at time $s \in [t, t+h-1]$ is defined as

$$\mathcal{G}_s = \{\alpha_s : \mathbb{R}^{30} \to \mathbb{R}^6 : \alpha_s(w_s; \theta_s) = \tanh(g_s^{L+1}),\ g_s^{L+1} \in \mathcal{D}^L\}. \tag{A.4}$$

As we consider the feedback control policies $\alpha \in \mathcal{G}$, the relaxed problem becomes

$$\begin{aligned} V_t^*(w) &= \min_{\alpha_s \in \mathcal{G}_s,\ s=t,\cdots,t+h-1} J(t, X_t; \{\alpha_s\}_{s=t}^{t+h-1}) \\ &= \min_{\theta_s,\ s=t,\cdots,t+h-1} J(t, X_t; \{\alpha_s\}_{s=t}^{t+h-1}). \end{aligned} \tag{A.5}$$

The approximation of $V_t(w)$ by $V_t^*(w)$ is consistent with that of Tsang & Wong (2020). Based on the above relaxed problem, the DNN architecture is illustrated as follows.

Algorithm 1 Deep neural network architecture

Require:

Current state variable w_t, sampled random vectors $\{\eta_s\}_{s=t+1}^{t+h}$

Ensure:

Total cost C_{t+h}^{λ}

for all timestep $s \in [t, t+h-1]$ **do**

Build FNN learning layers of the control: $w_s \to g_s^0 \to g_s^1 \to \cdots \to g_s^{L+1} \to \alpha_s$:

$g_s^0 \leftarrow w_s$

for all $\ell = 1 : L$ **do**

$g_s^{\ell} \leftarrow \sigma_\ell(B_s^{\ell} g_s^{\ell-1} + M_s^{\ell-1})$

end for

$g_s^{L+1} \leftarrow B_s^L g_s^L + M_s^L$

$\alpha_s \leftarrow \tanh(g_s^L)$

Update the state variable using the transition function $(w_s, \alpha_s, \eta_{s+1}) \to w_{s+1} : w_{s+1} = \phi(w_s, \alpha_s, \eta_{s+1})$

Compute the cumulative cost function $(w_{s+1}, \alpha_s, C_s^{\lambda}) \to C_{s+1}^{\lambda} : C_s^{\lambda} + \overline{c}^{\lambda}(s+1, X_{s+1}, \lambda; \alpha_s)$

end for

We emphasize that our empirical computation includes PC constraints on α_s

to enhance computational efficiency. In other words, we optimize the cumulative cost $\overline{C}_t^{\lambda}$, for which we perform cross-validation to identify the best fitting penalty coefficients $\overline{\lambda}_i$ for $i = 1, 2, \cdots, 6$. Using the above architecture, we sample $\{\eta_s\}_{s=t+1}^{t+h}$ as the input data and train the DNN using the standard stochastic gradient descent (SGD) method. The training algorithm can be implemented easily using common libraries (e.g., TensorFlow Abadi et al. (2015)).

We use Keras in TensorFlow to implement our algorithm and the Adam optimizer Kingma & Ba (2015) to optimize the parameters. For the FNN architecture, we set the number of hidden layers as 2 (i.e., $L = 2$). We choose the ReLU as our activation function for the hidden layers. We set the number of neurons in each hidden layer as 64 and the initial learning rate as 0.0001. We train the model 100 times using 20,000 simulated training samples. After training, we apply a set of simulated testing samples to the model to obtain a set of optimized controls.

References

Abadi, M., Agarwal, A., Barham, P., Brevdo, E., Chen, Z., Citro, C., Corrado, G. S., Davis, A., Dean, J., Devin, M. et al. (2015). TensorFlow: Large-scale machine learning on heterogeneous systems. URL: `https://www.tensorflow.org/` software available from tensorflow.org.

Ahumada, H., Espina, S., & Navajas, F. (2020). COVID-19 with uncertain phases: estimation issues with an illustration for Argentina. doi:`10.2139/ssrn.3633500` available at SSRN 3633500.

Alleman, T., Torfs, E., & Nopens, I. (2020). COVID-19: from model prediction to model predictive control. Available online `https://biomath.ugent.be/sites/default/files/2020-04/Alleman_etal_v2.pdf`. (Accessed 20 Oct 2021).

Alvarez, F. E., Argente, D., & Lippi, F. (2020). A simple planning problem

for COVID-19 lockdown. doi:10.3386/w26981 working Paper 26981 National Bureau of Economic Research.

Apple (2020). Mobility trends reports. URL: https://www.apple.com/covid19/mobility.

Baker, S. R., Bloom, N., Davis, S. J., Kost, K., Sammon, M., & Viratyosin, T. (2020). The unprecedented stock market reaction to COVID-19. *The review of asset pricing studies*, *10*, 742–758. doi:10.1093/rapstu/raaa008.

Bardina, X., Ferrante, M., & Rovira, C. (2020). A stochastic epidemic model of COVID-19 disease. *AIMS Mathematics*, *5*, 7661–7677. doi:10.3934/math.2020490.

Barua, S. (2020). Understanding coronanomics: The economic implications of the coronavirus (COVID-19) pandemic. doi:10.2139/ssrn.3566477 available at SSRN 3566477.

Benzell, S. G., Collis, A., & Nicolaides, C. (2020). Rationing social contact during the COVID-19 pandemic: Transmission risk and social benefits of US locations. *Proceedings of the National Academy of Sciences*, *117*, 14642–14644. doi:10.1073/pnas.2008025117.

Block, P., Hoffman, M., Raabe, I. J., Dowd, J. B., Rahal, C., Kashyap, R., & Mills, M. C. (2020). Social network-based distancing strategies to flatten the COVID-19 curve in a post-lockdown world. *Nature Human Behaviour*, *4*, 588–596. doi:10.1038/s41562-020-0898-6.

Bonaccorsi, G., Pierri, F., Cinelli, M., Flori, A., Galeazzi, A., Porcelli, F., Schmidt, A. L., Valensise, C. M., Scala, A., Quattrociocchi, W. et al. (2020). Economic and social consequences of human mobility restrictions under COVID-19. *Proceedings of the National Academy of Sciences*, *117*, 15530–15535. doi:10.1073/pnas.2007658117.

Brauer, F. (2008). Compartmental models in epidemiology. In *Mathematical Epidemiology* (pp. 19–79). Springer Berlin Heidelberg volume 1945. doi:10.1007/978-3-540-78911-6_2.

Calafiore, G. C., Novara, C., & Possieri, C. (2020). A time-varying SIRD model for the COVID-19 contagion in Italy. *Annual Reviews in Control*, *50*, 361–372. doi:10.1016/j.arcontrol.2020.10.005.

Chang, S. L., Harding, N., Zachreson, C., Cliff, O. M., & Prokopenko, M. (2020). Modelling transmission and control of the COVID-19 pandemic in Australia. *Nature communications*, *11*, 1–13. doi:10.1038/s41467-020-19393-6.

Chen, M.-H., Jang, S. S., & Kim, W. G. (2007). The impact of the SARS outbreak on Taiwanese hotel stock performance: an event-study approach. *International Journal of Hospitality Management*, *26*, 200–212. doi:10.1016/j.ijhm.2005.11.004.

Cheng, C., Barceló, J., Hartnett, A. S., Kubinec, R., & Messerschmidt, L. (2020). COVID-19 government response event dataset (CoronaNet v. 1.0). *Nature human behaviour*, *4*, 756–768. doi:10.1038/s41562-020-0909-7.

Choi, W., & Shim, E. (2021). Optimal strategies for social distancing and testing to control COVID-19. *Journal of theoretical biology*, *512*, 110568. doi:10.1016/j.jtbi.2020.110568.

Chou, J., Kuo, N.-F., & Peng, S.-L. (2004). Potential impacts of the SARS outbreak on Taiwan's economy. *Asian Economic Papers*, *3*, 84–99. doi:10.1162/1535351041747969.

Cont, R., Kotlicki, A., & Xu, R. (2020). Modelling COVID-19 contagion: risk assessment and targeted mitigation policies. *Royal Society open science*, *8*, 201535. doi:10.1098/rsos.201535.

Cuebiq (2020). COVID-19 mobility insights. URL: https://www.cuebiq.com/visitation-insights-covid19/.

Dashtbali, M., & Mirzaie, M. (2021). A compartmental model that predicts the effect of social distancing and vaccination on controlling COVID-19. *Scientific Reports*, *11*. doi:10.1038/s41598-021-86873-0.

Dony, R. (2001). Karhunen-loeve transform. In *The transform and data compression handbook*. CRC Press Boca Raton.

Faria, N. R., Mellan, T. A., Whittaker, C., Claro, I. M., Candido, D. d. S., Mishra, S., Crispim, M. A., Sales, F. C., Hawryluk, I., McCrone, J. T. et al. (2021). Genomics and epidemiology of the P. 1 SARS-CoV-2 lineage in Manaus, Brazil. *Science*, *372*, 815–821. doi:10.1126/science.abh2644.

Fernandes, N. (2020). Economic effects of coronavirus outbreak (COVID-19) on the world economy. doi:10.2139/ssrn.3557504 available at SSRN 3557504.

Fernández-Villaverde, J., & Jones, C. I. (2020). Estimating and simulating a SIRD model of COVID-19 for many countries, states, and cities. doi:10.3386/w27128 working Paper 27128 National Bureau of Economic Research.

Fudolig, M., & Howard, R. (2020). The local stability of a modified multi-strain sir model for emerging viral strains. *PloS one*, *15*, e0243408. doi:10.1371/journal.pone.0243408.

Gao, S., Rao, J., Kang, Y., Liang, Y., & Kruse, J. (2020). Mapping county-level mobility pattern changes in the United States in response to COVID-19. *SIGSpatial Special*, *12*, 16–26. doi:10.1145/3404820.3404824.

Ghader, S., Zhao, J., Lee, M., Zhou, W., Zhao, G., & Zhang, L. (2020). Observed mobility behavior data reveal social distancing inertia. ArXiv preprint arXiv:2004.14748.

Google (2020). COVID-19 community mobility reports. URL: https://www.google.com/covid19/mobility/.

Habersaat, K. B., Betsch, C., Danchin, M., Sunstein, C. R., Böhm, R., Falk, A., Brewer, N. T., Omer, S. B., Scherzer, M., Sah, S. et al. (2020). Ten con-

siderations for effectively managing the COVID-19 transition. *Nature human behaviour*, *4*, 677–687. doi:10.1038/s41562-020-0906-x.

Hall, R. E., Jones, C. I., & Klenow, P. J. (2020). Trading off consumption and COVID-19 deaths. doi:10.3386/w27340 working Paper 27340 National Bureau of Economic Research.

He, S., , Tang, S., & Rong, L. (2020). A discrete stochastic model of the COVID-19 outbreak: forecast and control. *Mathematical Biosciences and Engineering*, *17*, 2792–2804. doi:10.3934/mbe.2020153.

Heiberger, R. H. (2018). Predicting economic growth with stock networks. *Physica A: Statistical Mechanics and its Applications*, *489*, 102–111. doi:10.1016/j.physa.2017.07.022.

Hethcote, H. W. (2000). The mathematics of infectious diseases. *SIAM Review*, *42*, 599–653. doi:10.1137/S0036144500371907.

Jones, C., Philippon, T., & Venkateswaran, V. (2021). Optimal mitigation policies in a pandemic: social distancing and working from home. *The Review of Financial Studies*, *34*, 5188–5223. doi:10.1093/rfs/hhab076.

Karako, K., Song, P., Chen, Y., & Tang, W. (2020). Analysis of COVID-19 infection spread in Japan based on stochastic transition model. *BioScience Trends*, *14*, 134–138. doi:10.5582/bst.2020.01482.

Kermack, W. O., & McKendrick, A. G. (1927). A contribution to the mathematical theory of epidemics. *Proceedings of the Royal Society of London (Series A)*, *115*, 700–721. doi:10.1098/rspa.1927.0118.

Kingma, D. P., & Ba, J. L. (2015). Adam: A method for stochastic optimization. In *The 3rd International Conference for Learning Representations (ICLR 2015)*. San Diego. URL: http://arxiv.org/abs/1412.6980.

Kissler, S., Tedijanto, C., Lipsitch, M., & Grad, Y. H. (2020). Social distancing strategies for curbing the COVID-19 epidemic. doi:10.1101/2020.03.22.20041079 medRxiv preprint.

Köhler, J., Schwenkel, L., Koch, A., Berberich, J., Pauli, P., & Allgöwer, F. (2021). Robust and optimal predictive control of the COVID-19 outbreak. *Annual Reviews in Control*, *51*, 525–539. doi:10.1016/j.arcontrol.2020.11.002.

Lesniewski, A. (2020). Epidemic control via stochastic optimal control. ArXiv preprint arXiv:2004.06680.

Li, Z., Tsang, K. H., & Wong, H. Y. (2020). Lasso-based simulation for high-dimensional multi-period portfolio optimization. *IMA Journal of Management Mathematics*, *31*, 257–280. doi:10.1093/imaman/dpz013.

Liu, P., Beeler, P., & Chakrabarty, R. K. (2020). COVID-19 progression timeline and effectiveness of response-to-spread interventions across the United States. doi:10.1101/2020.03.17.20037770 medRxiv preprint.

Makridis, C., & Hartley, J. (2020). The cost of COVID-19: A rough estimate of the 2020 US GDP impact. doi:/10.2139/ssrn.3570731 available at SSRN 3570731.

McKee, M., & Stuckler, D. (2020). If the world fails to protect the economy, COVID-19 will damage health not just now but also in the future. *Nature Medicine*, *26*, 640–642. doi:10.1038/s41591-020-0863-y.

Miller, A. C., Foti, N. J., Lewnard, J. A., Jewell, N. P., Guestrin, C., & Fox, E. B. (2020). Mobility trends provide a leading indicator of changes in SARS-CoV-2 transmission. doi:10.1101/2020.05.07.20094441 medRxiv preprint.

Moghadas, S. M., Vilches, T. N., Zhang, K., Nourbakhsh, S., Sah, P., Fitzpatrick, M. C., & Galvani, A. P. (2021). Evaluation of COVID-19 vaccination strategies with a delayed second dose. *PLoS biology*, *19*, e3001211. doi:10.1371/journal.pbio.3001211.

Moore, S. E., & Okyere, E. (2020). Controlling the transmission dynamics of COVID-19. ArXiv preprint arXiv:2004.00443.

Morato, M. M., Bastos, S. B., Cajueiro, D. O., & Normey-Rico, J. E. (2020). An optimal predictive control strategy for COVID-19 (SARS-CoV-2) social distancing policies in Brazil. *Annual reviews in control*, *50*, 417–431. doi:10.1016/j.arcontrol.2020.07.001.

Mwalili, S., Kimathi, M., Ojiambo, V., Gathungu, D., & Mbogo, R. (2020). SEIR model for COVID-19 dynamics incorporating the environment and social distancing. *BMC Research Notes*, *13*. doi:10.1186/s13104-020-05192-1.

Parnes, D. (2020). Exploring economic anomalies in the S&P500 index. *The Quarterly Review of Economics and Finance*, *76*, 292–309. doi:10.1016/j.qref.2019.09.012.

Piguillem, F., & Shi, L. (2020). Optimal COVID-19 quarantine and testing policies. Available at SSRN 3594243.

Ritchie, H., Mathieu, E., Rodés-Guirao, L., Appel, C., Giattino, C., Ortiz-Ospina, E., Hasell, J., Macdonald, B., Beltekian, D., & Roser, M. (2020). Coronavirus pandemic (COVID-19). Available on https://ourworldindata.org/coronavirus (Accessed 20 Oct 2021).

Rubio-Herrero, J., & Wang, Y. (2021). A flexible rolling regression framework for time-varying SIRD models: Application to COVID-19. ArXiv preprint arXiv: 2103.02048.

Shapiro, S. S., & Wilk, M. B. (1965). An analysis of variance test for normality (complete samples). *Biometrika*, *52*, 591–611. doi:10.2307/2333709.

Soucy, J.-P. R., Sturrock, S. L., Berry, I., Westwood, D. J., Daneman, N., MacFadden, D. R., & Brown, K. A. (2020). Estimating effects of physical distancing on the COVID-19 pandemic using an urban mobility index. doi:10.1101/2020.04.05.20054288 medRxiv preprint.

de Souza, W. M., Buss, L. F., da Silva Candido, D., Carrera, J.-P., Li, S., Zarebski, A. E., Pereira, R. H. M., Prete, C. A., de Souza-Santos, A. A.,

Parag, K. V. et al. (2020). Epidemiological and clinical characteristics of the COVID-19 epidemic in Brazil. *Nature Human Behaviour*, *4*, 856–865. doi:10.1038/s41562-020-0928-4.

Sulyok, M., & Walker, M. (2020). Community movement and COVID-19: A global study using Google's community mobility reports. *Epidemiology and Infection*, *148*, e284. doi:10.1017/s0950268820002757.

Tanrisever, F., Shahmanzari, M., Eryarsoy, E., & Şensoy, A. (2021). Optimizing disease containment measures during a pandemic. Available at SSRN 3795643.

Tsang, K. H., & Wong, H. Y. (2020). Deep-learning solution to portfolio selection with serially dependent returns. *SIAM Journal on Financial Mathematics*, *11*, 593–619. doi:10.1137/19M1274924.

te Vrugt, M., Bickmann, J., & Wittkowski, R. (2020). Effects of social distancing and isolation on epidemic spreading modeled via dynamical density functional theory. *Nature Communications*, *11*, 5576. doi:10.1038/s41467-020-19024-0.

Warren, M. S., & Skillman, S. W. (2020). Mobility changes in response to COVID-19. ArXiv preprint arXiv:2003.14228.

Willis, M. J., Díaz, V. H. G., Prado-Rubio, O. A., & von Stosch, M. (2020). Insights into the dynamics and control of COVID-19 infection rates. *Chaos, Solitons & Fractals*, *138*, 109937. doi:10.1016/j.chaos.2020.109937.

Wong, G. (2008). Has SARS infected the property market? evidence from Hong Kong. *Journal of Urban Economics*, *63*, 74–95. doi:10.1016/j.jue.2006.12.007.

Worldometers.info (2020). COVID-19 coronavirus pandemic. Published online at Worldometers.info. URL: https://www.worldometers.info/coronavirus/.

Wren-Lewis, S. (2020). The economic effects of a pandemic. In R. Baldwinand, & B. W. di Mauro (Eds.), *Economics in the Time of COVID-19* (pp. 109–112). Centre for Economic Policy Research (CEPR) Press.

Zheng, Z., Xie, Z., Qin, Y., Wang, K., Yu, Y., & Fu, P. (2021). Exploring the influence of human mobility factors and spread prediction on early COVID-19 in the USA. *BMC Public Health*, *21*, 615. doi:10.1186/s12889-021-10682-3.