Since January 2020 Elsevier has created a COVID-19 resource centre with free information in English and Mandarin on the novel coronavirus COVID-19. The COVID-19 resource centre is hosted on Elsevier Connect, the company's public news and information website.

Elsevier hereby grants permission to make all its COVID-19-related research that is available on the COVID-19 resource centre - including this research content - immediately available in PubMed Central and other publicly funded repositories, such as the WHO COVID database with rights for unrestricted research re-use and analyses in any form or by any means with acknowledgement of the original source. These permissions are granted for free by Elsevier for as long as the COVID-19 resource centre remains active.
Blockchain solution benefits for controlling pandemics: Bottom-up decentralization, automation with real-time update, and immutability with privacy preservation

Xiaoming Li \(^a,^*\), Huigang Liang \(^b\)

\(^a\) Department of Business Administration, Tennessee State University, 330 10\(^{th}\) Ave. N, Nashville, TN 37203, USA

\(^b\) Department of Business Information and Technology, University of Memphis, 3675 Central Avenue, Memphis, TN 38152, USA

**ARTICLE INFO**

Keywords:
- COVID-19 pandemic
- Blockchain
- Epidemic outbreaks
- Risk management

**ABSTRACT**

The current COVID-19 pandemic has created turmoil around the world. To fight this ongoing global crisis and future ones, all stakeholders must collaborate and share timely and truthful information. This paper proposes a blockchain solution based on its inherent technological advantages. We posit that benefits can be derived from three unique blockchain features: bottom-up decentralization, automation with real-time update, and immutability with privacy preservation. A decentralized common platform provides easy access and increases participation in disease surveillance, which reduces the estimation errors of the compartmental model parameters. Automation with real-time update facilitates prompt detection and diagnosis, accurate contact tracing, and targeted mitigation and containment, achieving faster recovery and slower transmission. Being immutable while preserving privacy, the blockchain solution enhances respondents’ willingness to truthfully report their contact history, avoiding false and erroneous data that will cause wrong estimates on pandemic transmission and recovery. Thus, the blockchain solution mitigates three types of risks: sample variance, delay, and bias. Through simulation, we quantify the value of the blockchain solution in these three aspects. Accordingly, we provide specific action plans based on our research findings: before building blockchain solutions for controlling COVID-19, governments and organizations can calculate the blockchain benefits and decide whether or not they should invest in such blockchain solutions by conducting a cost-benefit analysis.

**1. Introduction**

The coronavirus (COVID-19) pandemic has caused millions of confirmed cases and hundreds of thousands of confirmed deaths in more than 200 countries, areas, or territories (Ahmadi, Sharifi, & Khalili, 2021a; WHO Coronavirus Disease COVID-19 Pandemic, 2020). Highly contagious and rapidly spreading, COVID-19 has led to a global crisis and greatly interrupted social and economic activities. Any effective intervention for controlling COVID-19 requires participation from all stakeholders.

However, it is difficult to motivate stakeholders to participate in the battle against COVID-19. Ahmadi et al. (2021b) show that due to privacy and security concerns, most people avoid sharing sensitive data for training. Governments across the world are struggling to contain the pandemic of COVID-19 inside their borders. There were exponential growths of the COVID-19 outbreak in the early stages in many countries like Italy, Spain, and the United States. Moreover, millions of travelers cross country borders every year. For example, 275 million people in 2019 entered the United States through land ports at the U.S.-Mexico border (Department of Transportation Border Crossing Entry Data 2019). The mass transportation of people makes any attempt to contain COVID-19 extremely difficult and requires collaborations from all countries. However, different countries may not share crucial information promptly, may not even share information, or do not necessarily trust each other. World Health Organization (WHO) member states call for an independent inquiry into the global response to the coronavirus pandemic, approved without objection by the WHO’s 194-member annual assembly meeting virtually in Geneva in May 2020, also allowing for the inquiry to look into the health body’s own role (BBC Coronavirus, 2020). The United States ended its WHO membership in 2020 (Huang 2020).

A timely and large-scale survey sheds light on why people do not
trust their governments’ words and actions (COVID-19 Survey 2020): “Rather, many respondents believe that their government is not reacting sufficiently, with 45 % of respondents across the 58 countries holding such beliefs. Similarly, 58 % of respondents perceive that the reaction of their country’s public to the COVID-19 outbreak has been insufficient. Further, 43 % of respondents do not trust that their country’s government is keeping them safe, and 43 % of respondents perceive that their country’s government has not been truthful about the COVID-19 outbreak.”

The COVID-19 pandemic diffuses through different paths, so different governments hold various information at different time points. Governments may not know which information is more relevant and may not release it quickly enough. Some governments may be less transparent than others due to various reasons such as cultures, political systems, self-serving, economic concerns, and social impacts. The failure from one government will adversely disturb other governments’ efforts on the fight against COVID-19. Also, the COVID-19 survey (2020) finds that there is a significant gap between the respondents’ own beliefs and their perceived beliefs of their peers. About 97 % of the respondents believe that social gatherings should not be allowed, but they estimate that only 67 % of their peers will agree.

Key responses for controlling COVID-19 include (1) prompt detection and diagnosis, (2) contact tracing, and (3) mitigation and containment (Böhmer et al., 2020; CDC, 2020; Peak et al., 2020, Xu et al., 2020). First, quick detection and diagnosis are prerequisites to prevent the widespread of COVID-19 (CDC, 2020; Xu et al., 2020). Unless effective surveillance is in place, contact tracing and disease containment cannot be implemented (CDC, 2020). Disease clusters requiring special intervention need to be identified. South Korea excelled in early detection (Kim et al., 2020), conducting 10 times more tests than the US. in early 2020. As a result, COVID-19 was currently well under control in South Korea, with daily new cases below 1,000, compared with over 200,000 daily new cases in the US in early 2021. CDC still uses “antiquated data systems, many of which rely on information assembled by or shared with local health officials through phone calls, faxes and thousands of spreadsheets attached to emails. The data is not integrated, comprehensive or robust enough, with some exceptions, to depend on in real time” (Lipton et al., 2020). Banco (2021) investigates America’s COVID reporting breakdown and identified various problems such as crashing computers, three-week delays in tracking infections, and lab results delivered by snail mails. Some senior administration officials and outside experts were growingly frustrated with the CDC’s slow and sioled approach to sharing data, “which prevented officials across the government from getting real-time information about how the delta variant was bearing down on the United States” in summer 2021 (Abutaleb and Sun, 2021).

Once a person has been tested positive, the subsequent step is contact tracing to notify people who have been in close contact with the infected individual and monitor their whereabouts. It is vital to limit infected individuals’ movements and stop the infection chain (Peak et al., 2020; Böhmer et al., 2020; CDC, 2020). The reduction of pandemic diffusion depends on effective contact tracing. South Korea applies an aggressive contact tracing policy by integrating GPS data, credit card data, and surveillance footage from 28 different sources to perform real-time analysis (Kim et al., 2020). In contrast, the U.S. emphasizes privacy protection and relies on telephone surveys and voluntary disclosure for contact tracing (Kim et al., 2020). Consequently, contact tracing is ineffective due to mistrust and noncompliance in the US. Many do not answer the call from contact tracers, and many refuse to give truthful answers (Khazan, 2020). To make things worse, the US government suffers a trust crisis. A large-scale survey reveals that about 50 % of respondents do not trust the governments’ words and actions (COVID-19 Survey, 2020). Moreover, the health departments are already overburdened and thus do not have enough resources to run large-scale contact tracing. Citing multiple reasons including workload constraints and privacy concerns, the California health agencies refuse to share detailed COVID-19 case-related data with epidemiologists from Stanford University and several University of California campuses, even though the research would lead to more effective interventions (Piller, 2020).

While it is paramount to apply countermeasures to mitigate and contain the outbreak such as personal protective equipment, social distancing, individual quarantine, and lockdown, detection and contact tracing are equally important (Peak et al., 2020). In order to flatten the curves of COVID-19 deaths and new cases, it is essential to know the infected and the potential virus carriers so that the countermeasures can be more effectively applied.

On the technology side, the blockchain has demonstrated its practical value with its unique features, especially in a decentralized manner in a trustless environment (Yli-Huumo et al., 2016; Treiblmaier, 2018; Choi, 2019; Hastig and Sodhi, 2020; Lim et al., 2021; Zutshi et al., 2021). Answering the calls for more blockchain applications and for controlling COVID-19 (Computers & Industrial Engineering Special Issue Announcement, 2021; Computers & Industrial Engineering Special Issue Call for Papers, 2021), this paper proposes a blockchain solution based on its inherent technological advantages. Our research questions include:

1. Which blockchain features can be applied to control COVID-19?
2. How can these features contribute to faster patient recovery and slower disease transmission of COVID-19?
3. How to quantify the blockchain solution benefits?

Our contributions are threefold by addressing the three research questions above. In particular, we integrate an important disruptive technology – blockchain – into fighting the current global crisis – COVID-19 – with bottom-up approach emphases to inspire community participation and grassroots contributions. We apply operational risk management (ORM) and epidemiologic modeling to show how our proposed blockchain solution can alter the spread of pandemics and how its benefits can be quantified. Our findings can thus advance the interventions for controlling COVID-19.

The rest of the paper is organized as follows. In Section 2, we review the literature on operational risk, blockchain, and decentralized applications. Section 3 presents a blockchain solution for controlling pandemics. We then quantify the solution benefits in Section 4. Section 5 concludes our paper.

2. Literature review

2.1. ORM for COVID-19

Araz et al. (2020) develop an ORM framework with three phases: application, analytics techniques, and analytics strategies. In Phase I, researchers need to identify the problem and application domain of ORM, which includes public health risk management (Araz et al., 2020). In Phase II, researchers need to select appropriate analytics techniques. In Phase III, analytics strategies need to be developed and evaluated to apply the analytic techniques in the focal domain through a systematic process consisting of data acquisition, sorting, analysis, interpretation, and exploitation (Chung et al., 2005; Araz et al., 2020).

Applying Araz et al.’s (2020) ORM framework, in Phase I, we identify COVID-19 management as the problem domain; in Phase II, we believe the unique requirement of pandemic management is the collection of representative, truthful, and timely data and the processing capacity for real-time decision making; in Phase III, consistent with Araz et al.’s (2020) recommendation, we propose blockchain as a strategic solution that holds great potential in controlling COVID-19 and demonstrate how a blockchain-based solution provides quantifiable benefits.

In order to stop COVID-19 and save lives, robust and functional information systems are needed. Scholars have attempted to develop information systems to address various challenges of COVID-19. Govindan
et al. (2020) develop a decision support system to break down the COVID-19 propagation chain and mitigate the epidemic outbreaks. After community members (service recipients) input their ages, pre-existing diseases, and COVID-19 symptoms, the decision support system can perform certain diagnoses and classify community members into different classes, following a set of fuzzy inference rules developed by experts and physicians. Each class then links to its corresponding medical treatments. Govindan et al. (2020) thus make unique contributions in applying a decision support system that directly controls COVID-19 propagation and epidemic outbreak in a population.

Sharifi et al. (2021) summarize important Industry 4.0 applications for controlling COVID-19: preventing and decelerating COVID-19 spread, diagnosing and treating COVID-19, equipping and improving healthcare supply chains (SCs), controlling and monitoring medical centers and public places, and reducing virus infection and death. They present a comprehensive review of how key Industry 4.0 technologies (machine learning, intelligent sensor, mobile technology, internet of things, cloud computing, virtual reality, and blockchain) can help control the COVID-19 pandemic.

A technology that is particularly useful for reducing COVID-19 risks is blockchain. Scholars have recognized blockchain’s great potential in addressing various challenges due to its decentralized nature with permanent, traceable, and reliable data (e.g., Choi et al., 2019; Azar et al., 2020; Zutshi et al., 2021). After a systematic literature review, Zutshi et al. (2021) provide thorough procedures to build better digital ecosystems with their blockchain value proposition pyramid representing various levels of complexity and disruption to existing platforms. They categorize five key value propositions: the decentralized data infrastructure, the membership management, the analytics/automation models, the crypto-economic models, and finally the decentralized governance.

Therefore, answering the call by our research community (Computers & Industrial Engineering Special Issue Announcement, 2021; Special Issue Call for Papers, 2021) and following Govindan et al. (2020) and Sharifi et al. (2021), we propose a contemporary blockchain solution that can fulfill the need for real-time information processing and analytics capabilities with automation during a public health crisis. In the next section, we review the blockchain literature and identify relevant blockchain features that can be applied to controlling COVID-19.

2.2. Blockchain Technology

Blockchain is a key new technology in Industry 4.0 (Computers & Industrial Engineering Special Issue Announcement, 2021; Special Issue Call for Papers, 2021). Blockchain applications are still in the early stages, while researchers call to investigate more potential applications (Grillo and Zutshi, 2020; Lim et al., 2021; Zutshi et al., 2021). In a blockchain, data are saved in blocks on a peer-to-peer (P2P) distributed network while cryptography is employed for securing the identity and protecting the privacy of its users (Yli-Huumo et al., 2016; Treiblmaier, 2018). A new block is added to the end of the existing chain in chronological order. It cannot be added unless all nodes on the P2P network have reached a consensus by following their blockchain protocol.

The most important feature of blockchain is its P2P decentralized structure: all participants can communicate and perform transactions with each other without the need of a central governing entity (Lim et al., 2021; Zutshi et al., 2021). Participants can join from different countries around the world. Other key features including permanent data and secure systems enable a variety of SC information sharing and collaborations (e.g., Choi, 2019). In particular, Hastig and Sodhi (2020) recognize blockchain as the focal technology for SC traceability and then find vital business requirements and success factors for SC traceability.

Because of the characteristics of blockchain, we can program business logic with smart contracts and develop various applications on its network. For instance, currently, there are hundreds of decentralized applications (DApps) on the Ethereum network, including finance, gaming, gambling, healthcare, and social (State of DApps, 2020). A blockchain can update and add blocks very quickly. For example, Ethereum adds blocks typically in the tens of seconds (Ethereum, 2020).

Sharifi et al.’s (2021) review of various blockchain applications in food, agricultural, and energy industries provides industrial verification for our research community. They show that COVID-19 causes disruptions to SCs while Industry 4.0 technologies can improve supply chain management (SCM). Varmaghani et al. (2021) present a Distribution-Map-Transfer-Combination distribution method for wireless sensor networks to optimize their energy consumption. Nichol and Brandt (2016) propose a blockchain ecosystem as a healthcare interoperability infrastructure, answering the call by the ten-year plan devised by the Office of the National Coordinator for Health Information Technology. Using their signature, patients have full control of their medical information in the blockchain ecosystem and can select which information to be shared and viewed by third parties such as doctors and insurance companies. This method removes the burden of maintaining patients’ medical records away from doctors.

Some companies plan to apply the blockchain technology for controlling COVID-19. Individuals use encrypted digital wallets on their smartphones to control what they share for what purpose, while organizations design business procedures to determine responses for each health status they review (IBM Digital Health Pass, 2020). Consequently, individuals can protect their privacy while organizations can make timely data-driven decisions.

2.3. Characteristics of blockchain DApps

We summarize similarities and differences between a regular centralized App and a DApp from the literature in Fig. 1 (e.g., Cai et al., 2018; Wu et al., 2019; McCubbin, 2021). The numbers of 1 to 4 show the sequence of activities. A Dapp is a digital application running on a decentralized P2P blockchain network, while a regular centralized App runs by one organization (Cai et al., 2018; Wu et al., 2019; McCubbin, 2021). A regular centralized app typically has three tiers: the presentation tier (i.e., the front end, the user interface, or the client application), the business logic tier (i.e., the application server), and the data storage tier (i.e., the database server). For a DApp, the front end is the same as in the centralized app. Both the centralized app and the DApp use HTML, Cascading Style Sheet (CSS), and JavaScript (JS) to construct the user interface, which are parts of all websites. HTML defines every element of a web page: texts, images, titles, and other tags, CSS provides the look of web pages, and JS adds interactions to web pages. The differences are in business logic and database design.

For a regular centralized app, the database and application servers are controlled by the organization that owns the app. After users use the app to upload their content, the organization possesses the data. For example, Google Map or Facebook possesses all the user data. The data can then be used for various purposes, including user profiling and targeted advertising. The organization can also censor or delete inappropriate data. In sharp contrast, for blockchain-based DApps, the database tier is a decentralized common blockchain platform where smart contracts implement business logic. Next, we explain relevant key characteristics applicable to fighting COVID-19.

2.3.1. Bottom-up decentralized common platform

On blockchain, data are cryptographically stored and distributed in a decentralized ledger, which removes the need of a central authority. In other words, the P2P distributed structure provides a mechanism for all participants to communicate and conduct transactions with each other without central coordination. All communications and transactions are executed and stored in a bottom-up manner, inputted by all participants equally without a hierarchical structure. Consensus algorithms can be proof of work, proof of stake, proof of importance, or proof of activity.
Service providers include government agencies and healthcare organizations that build the blockchain infrastructure with real-time data. People basically have no idea and no control over how the organization will store their data, share their data, and use their data. Based on our assumption or not, they will give out their information to an organization, and what will happen thereafter is a black box. People can conduct business analytics on these data and provide summarized statistics to the blockchain. Then, the service provider can present their data to the blockchain platform.

More specifically, using IoT, a user can access/install a DApp and then send data to and receive data from the blockchain. The users remain anonymous by only sharing their public keys while using their private keys to perform two types of transactions: (1) permission control and (2) data storage and retrieval. As the users sign up for the DApp for the first time, a new shared (user + service) identity is produced and sent to the blockchain along with their permission settings. With embedded smart contracts (i.e., business logic), IoT can automatically collect and upload certain data to the blockchain.

At the next layer, service providers can build various DApps for different purposes. Either the service provider or the user can query the data from the blockchain, which will check the digital signature belonging to the service provider or the user. The service provider can conduct business analytics on these data and provide summarized statistics to the blockchain. Then, the service provider can present the

Fig. 1. Difference between centralized applications and DApps.

Fig. 2 shows how such a blockchain solution works for users and service providers. At the grassroots layer, users can use all IoT (e.g., computers, cell phones, sensors, and mechanical and digital devices) to collect data and upload them onto the blockchain platform. We have various blockchain-IoT combinations that can perform data collection and storage, communications, and transactions (e.g., Christidis and Devetsikiotis, 2016). As in our proposed blockchain solution shown in Fig. 2, it is an IoT to blockchain architecture design. Each IoT device is represented as a node in the blockchain network and communicates directly to the blockchain network. Interactional data between IoT peer devices are logged and captured onto the blockchain.

The key difference from traditional systems is that users own and control their data. Users can give and change permissions to various service providers. Users have complete transparency over what data and when these data are collected and how the data are accessed by choosing different services. As Fig. 2 shows, User 1 uses service A, User 3 uses service B, and User 2 uses both services A and B; services A and B are different but built on the same blockchain platform.

2.3.2. Automation with real-time update
A smart contract is a set of self-executing codes with terms and agreements between blockchain participants. Once written, it cannot be changed. Code is the law. The code controls transactions that are trackable and irreversible. A smart contract permits automation and transactions can be updated in real-time, e.g., blocks on Ethereum are committed every 15 s. Even better, data can be collected and stored automatically with IoT sensors and devices. Smart contracts are executed robotically whenever a certain logic is satisfied.

2.3.3. Privacy-preserving immutability
A blockchain ledger/database remains permanent, indelible, and unaltered. A blockchain platform uses cryptographic hashing to transfer data to a Checksum—an alphanumeric string that serves as a digital signature. The same data input generates the same digital signature, but not vice versa. In other words, a digital signature points to the exact data input, but knowing the digital signature cannot generate the data input. A block contains a digital signature for itself and another one for the preceding block. In this way, all transactions and data on all blocks are chained together and remain immutable.

3. A blockchain solution for controlling pandemics
We propose a blockchain solution for addressing the current COVID-19 challenges, especially in the areas of prompt detection and mitigation, targeted contact tracing, and containment. The advantages of using blockchain here are threefold: transparency, immutability, and no central point of failure. DApps are open to the general public such that every-one can view their codes. Once smart contracts are deployed onto the blockchain, they will be executed in a decentralized manner, and nobody can ever change their codes. Due to the decentralized consensus, DApps can avoid the central point of failure. Without blockchain, people have the rights to share their information or not. If they share, they give out their information to an organization, and what will happen thereafter is a black box. People basically have no idea and no control over how the organization will store their data, share their data, and use their data. Based on our interview results, patients were more likely to report their COVID-19 cases with higher honesty when they can see and control how their data will be used.

The blockchain solution consists of service providers and users. Service providers include government agencies and healthcare organizations that build the blockchain infrastructure with real-time data
key statistics and visualization on a website dashboard. Whenever necessary, the user can change his/her permission settings by transmitting a permission transaction to the blockchain. The next section shows how this blockchain solution can provide prompt detection and diagnosis, accurate contact tracing, and targeted mitigation and containment for controlling COVID-19.

3.1. Blockchain solution benefits

The blockchain solution has three benefits based on three DApp characteristics identified above: (1) bottom-up decentralization, (2) automated real-time updating, and (3) privacy-preserving immutability.

3.1.1. Bottom-up decentralization

Different from typical traditional information systems, the blockchain-based system has no single point of failure. The bottom-up mechanism lifts the usual heavy burden on the central organization, empowered by contributions from all participants. Many data can be collected and uploaded to the platform automatically through participants’ IoT devices (e.g., computers, cell phones, sensors, and smartwatches).

The common blockchain platform provides easy access for all people from all places. Users own their data and control how to use the data. All users have the same transparent end-to-end visibility on the same platform. More users joining the platform will increase the value for existing users, resulting in positive network externalities, which in turn attract more users to join the platform and form a virtuous cycle. This helps to mitigate the risk of small sample size and potential selection biases, which is especially valuable when little is known about a new virus. With larger sample sizes, the key pandemic metrics (e.g., transmission rate and recovery rate) and the dynamics (i.e., evolutions of susceptible, infected, and recovered individuals) of the infectious disease can be more accurately estimated.

3.1.2. Automated real-time updating

Time is critical in disaster management. SCs are transient with ever-changing demand, supply, and resources. A blockchain solution brings forward three benefits from its capability of automated real-time updating.

Prompt detection and diagnosis: A user registers at a blockchain service provider and uploads his/her COVID-19 symptoms and other vital information: temperature, coughing, age, gender, other diseases, etc. The information check and upload/download are facilitated by IoT devices, e.g., cell phones and sensors. A smart contract can perform an initial screening to provide feedback like possibly infected, suspected, and normal. The smart contract can then provide suggestions for the user about what to do next: see a doctor immediately or continue to monitor the symptoms. The smart contract can match users’ testing and diagnostic demand with healthcare capacities: testing sites, testing abilities, available doctors and nurses, and medicines and medical instruments. Business logic in these smart contracts can follow existing literature, e.g., Govindan et al. (2020). All can be done in a highly timely manner. This process thus provides prompt detection and diagnosis, which leads to a faster recovery for patients.

Accurate contact tracing: Users can upload information about their location, travel, and contact to the blockchain. IoT devices such as cell phones and global positioning system (GPS) sensors can help automate contact tracing. Specifically, a smart contract can be developed to show the travel summary of a user and also the aggregated summary of infected and recovered individuals in a region for public health management. The aggregated contact tracing/mapping shows the exact diffusion path of an infectious disease. Once an individual uploads his/her infection status to the platform, all others in close contact in the past, say, two weeks will be notified through the platform to their IoT devices. Timely accurate contact tracing enables us to cut the infection chain earlier, resulting in a lower transmission rate.

Targeted mitigation and containment: A blockchain platform can collect more accurate and transparent data to facilitate decisions on how to mitigate and contain the pandemic. A smart contract can provide relevant and summarized information to support necessary actions such as business/school closing, quarantine, and lockdown. Mitigation and containment mapping data can pinpoint exact people, communities, and locations that need to be in quarantine. Supported data and actions can ensure that the individuals in quarantine have enough food, housing, or
sanitation so they can isolate properly and not continue to spread the disease within and across communities. This benefit causes both faster recovery and lower transmission.

Overall, the benefits of automation with real-time update include a faster recovery rate and a lower transmission rate. Automation with real-time update of blockchain thus reduces delay, which is especially valuable in the early stages of a pandemic cycle.

3.1.3. Privacy-preserving immutability

Data on the blockchain platform are less likely tampered with than those in a centralized database inside an organization. Privacy preservation enhances trust and motivates self-reporting. When privacy is maintained, respondents are more willing to give truthful responses. Moreover, false and erroneous data will cause wrong estimates on both transmission and recovery rates. Forecasting models then will yield incorrect results. Blockchain’s privacy-preserving immutability thus mitigates the risk of biased responses, which motivates people to report private information and helps health management make accurate descriptive and predictive analyses of the pandemic diffusion. The next section shows how to quantify such three blockchain solution benefits.

Our proposed solution is compared to the current centralized solutions, e.g., CDC as in our literature review. Our method is accurate and robust (see details in the following sections). Our parameters are calibrated through real empirical data. We run each simulation over 1000 periods to accommodate various settings and run hundreds of cases in our sensitivity analysis. Our results remain the same for different values of the parameters.

4. Quantifying blockchain benefits

Epidemiologic models have been widely applied for controlling pandemics. In particular, they can simulate how infectious diseases spread and can fit various transmission scenarios ensuring different actions. In this section, we adjust the assumptions and model parameters to match changes in our proposed blockchain solution. By showing how our blockchain solution alters the dynamics of a pandemic outbreak, we quantify its benefits.

4.1. The susceptible-infected-recovered (SIR) model

The SIR model is an important epidemiological model, first used by Kermack and McKendrick in 1927, and has been extensively applied to a variety of diseases, including airborne infectious diseases (e.g., Bertozzi et al., 2020; Westerink-Duijzer et al., 2020; Ahmadi et al., 2021a). The SIR model shows great generality in practice and can be directly extended to other popular epidemic models. \( S, I, \) and \( R \) represent the number of susceptible, infected, and recovered individuals, respectively, and \( N = S + I + R \) is the total population, which is a constant.

The SIR model equations include:

\[
\begin{align*}
\frac{dS}{dt} &= -\frac{\beta}{N}SI, \\
\frac{dI}{dt} &= \frac{\beta}{N}SI - \gamma I, \\
\frac{dR}{dt} &= \gamma I.
\end{align*}
\]

\( S, I, \) and \( R \) denote the number of susceptible, infected, and recovered individuals in the population at time \( t \), respectively. \( \beta \) is the transmission rate constant, \( \gamma \) is the recovery rate constant, and \( \frac{\beta}{\gamma} = \xi \) is the basic reproduction number because the susceptible individuals at the beginning are basically the whole population. The effective reproduction number is \( R_e = \frac{\beta}{\gamma} \xi \) for different susceptible individuals at different times. When \( R_e > 1 \), the infectious disease will spread further; when \( R_e \leq 1 \), the infectious disease will decrease monotonically to 0.

To contain a pandemic, we aim to stop the spread of infectious disease, to lower peaked infected individuals \( I_{\text{max}} \), the so-called flatten-the-curve, due to limited healthcare capacities, and to lower total people infected over the whole cycle. For the standard SIR model (i.e., \( \beta \) and \( \gamma \) do not change over time), the peak of infected individuals and the total number of people infected over the whole cycle are well-known (Kermack and McKendrick, 1927; Ma and Earn, 2006; Miller, 2012):

Dividing Equation (2) by Equation (1) yields:

\[
\frac{dI}{dS} = -1 + \frac{\gamma N}{\beta S}
\]

Integrating both sides, we have,

\[
I_t + S_t - \frac{\gamma N}{\beta} \ln S_0 = I_0 + S_0 - \frac{\gamma N}{\beta} \ln S_0 \quad \text{for all} \quad t.
\]

When \( \frac{\beta}{\gamma} = -1 + \frac{\gamma N}{\beta} = 0 \) or \( S = \frac{\gamma N}{\beta} \), we have \( I_{\text{max}} + \frac{\gamma N}{\beta} \ln \frac{\gamma N}{\beta} = I_0 + S_0 - \frac{\gamma N}{\beta} \ln S_0 \) or \( I_{\text{max}} = I_0 + S_0 - \frac{\gamma N}{\beta} S_0 - \frac{\gamma N}{\beta} + \frac{\gamma N}{\beta} \ln S_0 \).

Since initially the whole population is susceptible to COVID-19,

\[
I_{\text{max}} = N - N \frac{\gamma N}{\beta} \ln N - N \frac{\gamma N}{\beta} + N \frac{\gamma N}{\beta} \ln N = N - N \ln N (1 + \ln R_0),
\]

\[
R_0 = \frac{\beta}{\gamma} \xi \ln N
\]

At the end of an infectious disease cycle, \( I_{\text{End}} = 0 \) and from Equation (4), we have,

\[
S_{\text{End}} - \frac{\gamma N}{\beta} \ln S_{\text{End}} = I_0 + S_0 - \frac{\gamma N}{\beta} \ln S_0.
\]

Since COVID-19 has an initially susceptible population, we have \( S_{\text{End}} - N = N \frac{\gamma N}{\beta} \ln S_0 = \frac{\gamma N}{\beta} \ln S_{\text{End}} \). So,

\[
\ln S_{\text{End}} = \ln S_0 (\frac{\gamma N}{\beta} - 1) \quad \text{or} \quad \frac{\gamma N}{\beta} = e^{-x_0} (1 - \frac{S_{\text{End}}}{S_0}).
\]

The total fraction of the population getting infected over the whole cycle is

\[
\frac{S_{\text{End}}}{S_0} = 1 - \frac{S_{\text{End}}}{S_0}.
\]

We have \( \frac{S_{\text{End}}}{S_0} = 1 - e^{-x_0} \frac{S_{\text{End}}}{S_0} \) (e.g., Ma and Earn, 2006). We now introduce the base case to illustrate how the susceptible, infected, and recovered evolve over time. Without loss of generality, we use the whole population \( N = 1 \). So, the susceptible, infected, and recovered each represent the fraction of the total population in the corresponding category. SIR data for COVID-19 are continuously evolving, and many countries have quite different observations (Vicente and Petsurillo, 2020; Bertozzi et al., 2020; Sanche et al., 2020). Besides SIR estimates, Ahmadi et al. (2021a) present the generalized logistic growth model to assess sub-epidemic waves of the COVID-19 outbreak in Iran, simulating different scenarios of two, three, and four waves in the observed incidence. They then evaluate travel-related risk in inter-provincial travels in Iran.

Using empirical data from three US states (California, New York, and Indiana), Bertozzi et al. (2020) show that \( \gamma \) ranges from 0.06 to 0.19 and \( R_0 \) from 2.1 to 4.4. Many papers estimate that \( R_0 \) lies somewhere between 2 and 3 (Bauch, 2020; Hilton and Keeling, 2020).

We have downloaded the published datasets of daily cases in each Tennessee county (Tennessee Department of Health, 2021). We use the daily cases in Davidson County (i.e., Nashville) from March 3 to June 10, 2020 (the first 100 days in the dataset). The Tennessee government has issued various guidelines during that period: limiting mass gatherings, using alternative business models, and closing schools (Tennessee COVID-19 Timeline, 2020). Estimates of the reproduction number tend to fluctuate considerably in the initial periods because of the small data time step (i.e., daily) and different countermeasures. Plus, data tend to show a weekly pattern, e.g., fewer or no reported cases during weekends. So, we choose a novel, statistically robust analytical method by Cori et al. (2013) for estimating the instantaneous reproduction number.

Following Cori et al. (2013), we use the same \( R_e \) over a time window of length \( \tau \) ending at time \( t \), where we use \( \tau = 7 \) to calculate the
instantaneous reproduction number over a week, depending on the number of incident cases in the time window \([t - r + 1, t]\).

Fig. 3 shows Davidson County’s daily cases and our estimated instantaneous reproduction number using Cori et al.’s tool. The left \(y\)-axis is for estimated instantaneous reproduction number, while the right is for daily cases. The round-dot line shows how the estimated instantaneous reproduction numbers change over time. The estimated instantaneous reproduction number changes a lot at the beginning, reaches its second peak at around 2.5, and finally fluctuates around 1. The vertical bar chart shows daily cases over time.

We set our parameters based on the literature and also our data calibration above. First, in this section, we use \(\beta = 0.25\) and \(\gamma = 0.1\) (or equivalently \(\lambda_0 = 2.5\)) as our base case.\(^1\) Next, we change their values to show various blockchain benefits in the following sections.

### SIR model:

\[ S_{t+1} - S_t = -\beta S_t I_t \quad (5) \]

\[ I_{t+1} - I_t = \beta S_t I_t - \gamma I_t \quad (6) \]

\[ R_{t+1} - R_t = \gamma I_t \quad (7) \]

We use the discrete version of the SIR model because infectious disease data are collected and posted periodically, for example, every day. We set the following initial values in period 0: \(S_0 = 0.9999, I_0 = 0.0001, \text{ and } R_0 = 0\). We end our simulation in period 1000 to accommodate various settings in our sensitivity analysis below. We have \(S_{1000} = 0.1009, I_{1000} = 0, \text{ and } R_{1000} = 0.8991\). Note that \(R_{1000} = 0.8991\) is also the total (i.e., accumulative) infected fraction in the whole cycle. \(I_{max} = 0.2411\) in period 68. \(S, I, \text{ and } R\) from periods 0 to 200 are shown in Fig. 4. The susceptible fraction is always reducing over time while the recovered fraction is increasing over time. Note that the recovered fraction can be viewed as cumulative infected so far. The infected fraction tends to peak in the middle of the cycle.

### 4.2. Quantifying blockchain benefits based on the SIR model

A blockchain solution can change the transmission rate and the recovery rate. It can also reduce estimation errors of these parameters. Below we quantify the blockchain benefits in terms of the SIR model performance improvement in the context of COVID-19.

#### 4.2.1. Benefits of bottom-up decentralization

Holmdahl and Buckee (2020) explain why epidemiologic models are critical for COVID-19 prediction and why we have many uncertainties. They suggest that the lack of knowledge about how many people are, or have been, infected is the most obvious source of uncertainty that affects every model. The bottom-up decentralization of blockchain enables users to own and control their data and provides a transparent platform, which makes collecting data easier. Users in the meantime are more willing to participate. An increased sample size will reduce sampling errors. Below we show how we can reduce the standard errors of the estimators by the maximum likelihood (ML) estimation.

After collecting observational data, Becker and Britton (1999) show that the ML estimators for \(\beta\) and \(\gamma\) are calculated by \(\hat{\beta} = \frac{N}{\sum_{t=1}^{\infty} I_t dt} \) and \(\hat{\gamma}^{-1} = \frac{1}{\sum_{t=1}^{\infty} I_t dt} \), where \(t\) is some finite time when no infectious individuals are present and the epidemic is over, \(N_i = S_0 - S_t\) is the number of individuals infected in \((0, t]\), and \(S_0 = \frac{1}{\beta}\) is the proportion susceptible at \(t\).

Rida (1991) shows that the ML estimators \(\hat{\beta}\) and \(\hat{\gamma}^{-1}\) are consistent and asymptotically independent normal random variables as the population size \(N\) approaches \(\infty\), i.e., sufficiently large. The standard errors of ML estimators are \(SE(\hat{\beta}) = \frac{\hat{\beta}}{\sqrt{\gamma N}}\) and \(SE(\hat{\gamma}^{-1}) = \frac{1}{\sqrt{\gamma N}}\).

When we increase the sample size, the ML estimators \(\hat{\beta}\) and \(\hat{\gamma}^{-1}\) remain the same. But their standard errors become smaller. A confidence interval is constructed by \(\hat{\beta} \pm zSE, \) where \(z\) is the corresponding standard score of a normal distribution. For example, the 95% confidence interval for \(\hat{\beta}\) is \(\left(\hat{\beta} - 1.96 \frac{\hat{\beta}}{\sqrt{\gamma N}}, \hat{\beta} + 1.96 \frac{\hat{\beta}}{\sqrt{\gamma N}}\right)\) and for \(\hat{\gamma}^{-1}\) is \(\left(\frac{1}{\hat{\gamma}} - 1.96 \frac{1}{\sqrt{\gamma N}}, \frac{1}{\hat{\gamma}} + 1.96 \frac{1}{\sqrt{\gamma N}}\right)\). When their standard errors become smaller, for the same confidence level, their intervals are reduced. To say it differently, for the same interval, our confidence level is higher. We are more confident in our estimators \(\hat{\beta}\) and \(\hat{\gamma}^{-1}\) from our observational data.

To show the effects of standard errors, we now use random normal variables for \(\gamma\) and \(\beta\). To avoid negative values, we use truncated normal distributions. \(\hat{\beta}\) follows a normal distribution with a mean 0.25 and a standard deviation (SD) 0.125; truncated at 0 from below; if realized \(\beta < 0\), change to \(\beta = 0\). \(\gamma\) follows a normal distribution with a mean 0.1 and a SD 0.05; truncated at 0 from below: if realized \(\gamma < 0\), change to \(\gamma = 0\). For both \(\beta\) and \(\gamma\), the coefficient of variation is 0.5 to show some variability.

To show the benefits of an increased sample size after using blockchain, we halve the standard deviations of \(\beta\) and \(\gamma\), respectively. Their means remain the same. We have three scenarios: (1) SD of \(\beta\) is halved to 0.0625, corresponding to the sample size of transmission data is 4 times as before blockchain; (2) SD of \(\gamma\) is halved to 0.025, corresponding to the sample size of recovery data is 4 times as before blockchain; and (3) both SD of \(\beta\) is halved to 0.0625 and SD of \(\gamma\) is halved to 0.025, corresponding to both sample sizes are 4 times as before blockchain.

We use MATLAB to run these simulations. Initial conditions are the same as in our base case: \(S_0 = 0.9999, I_0 = 0.0001, \text{ and } R_0 = 0\). First, in each period, \(\beta\) and \(\gamma\) are randomly drawn from the above normal distributions and then truncated. Then, the susceptible, infected, and recovered are calculated based on Equations (5)–(7). We calculate 1000 periods for each round. Second, we run 100 rounds. Fig. 5 shows their evolutions for the above 4 scenarios with different standard deviations: (a) SD(\(\beta\)) is 0.125 and SD(\(\gamma\)) is 0.05; (b) SD(\(\beta\)) is 0.125 and SD(\(\gamma\)) is 0.025; (c) SD(\(\beta\)) is 0.0625 and SD(\(\gamma\)) is 0.05; and (d) SD(\(\beta\)) is 0.0625 and SD(\(\gamma\)) is 0.025. When SDs of \(\beta\) and \(\gamma\) are larger, the susceptible, infected, and recovered fluctuate more over time, indicating lower accuracy.

For the ending recovered on 100 rounds, we show VaR\(_{95\%}\) and CVaR\(_{95\%}\) in Table 1. VaR measures a percentile of a random variable, while CVaR is the conditional expectation of the random variable in the tail beyond the cut-off level in VaR. VaR\(_{95\%}\) is the 6th largest total infected among 100 rounds, indicating the total infected fraction at 95% worst-case level. VaR\(_{95\%}\) measures the 95th percentile of the total infected in 100 rounds. CVaR\(_{95\%}\) is the average of the largest five total infected among 100 rounds; the conditional expectation of the total infected fraction crossing the 95% worst-case threshold. VaR\(_{95\%}\) shows that with 95% confidence, the total infected over the whole cycle is not more than 0.9438 for the case SD(\(\beta\)) = 0.125 and SD(\(\gamma\)) = 0.05. 0.9438 is the largest or the worst performance among the four cases. For the case SD(\(\beta\)) = 0.0625 and SD(\(\gamma\)) = 0.025, with 95% confidence, the total infected over the whole cycle is not more than 0.9217, which is the best performance among all four cases. The total infected are somewhere in between for two other cases. So, reducing SDs of \(\beta\) and \(\gamma\) improves the worst performance at a given confidence level.

CVaR\(_{95\%}\) is the conditional expectation of total infected above 95% confidence. CVaR\(_{95\%}\) in four cases show the same pattern as VaR\(_{95\%}\). In words, reducing SDs of \(\beta\) and \(\gamma\) improves the expected worst performance at a given confidence level. This result remains the same when we use other values of confidence levels and SDs of SDs of \(\beta\) and \(\gamma\).

\(^1\) We also run tests with many other different values of \(\beta\) and \(\gamma\), and the results are highly similar.
4.2.2. Benefits of automation with real-time update

As discussed in Sections 3.1.1 and 3.1.2, the blockchain benefits of automation with real-time update include faster recovery, lower transmission, and both. The base case sets $\beta = 0.25$ and $\gamma = 0.1$ (or equivalently $R_0 = 2.5$) as discussed earlier. In this section, we change $\beta$ and $\gamma$ to reflect three blockchain solution benefits: (1) Faster recovery due to prompt detection and diagnosis by blockchain: We increment the recovery rate $\gamma$ from 0.1 to 0.2 with a step size 0.01, resulting in 11 values: 0.10, 0.11, 0.12, 0.13, 0.14, 0.15, 0.16, 0.17, 0.18, 0.19, and 0.20. (2) Lower transmission due to improved contact tracing and quarantine by blockchain: We decrease the transmission rate $\beta$ from 0.25 to 0.125 with a step size 0.0125, which result in 11 values: 0.25, 0.2375, 0.225, 0.2125, 0.2, 0.1875, 0.175, 0.1625, 0.15, 0.1375, and 0.125. (3) Faster recovery and lower transmission due to better treatment and containment by blockchain: we simultaneously increase the recovery rate $\gamma$ and reduce the transmission rate $\beta$ with the following ten cases: $\gamma = 0.11$ and $\beta = 0.2375$, $\gamma = 0.12$ and $\beta = 0.225$, $\gamma = 0.13$ and $\beta = 0.2125$, $\gamma = 0.14$ and $\beta = 0.2$, $\gamma = 0.15$ and $\beta = 0.1875$, $\gamma = 0.16$ and $\beta = 0.15$.

Fig. 3. Davidson County’s daily cases in 2020 and our estimated instantaneous reproduction number.

Fig. 4. The evolution of susceptible, infected, and recovered fractions over time.
Facing a crisis, organizations may react promptly or slowly. Also, for the distributed issue, we cannot achieve real-time update if the blockchain cannot update itself quickly. The good news is that today’s blockchain platforms update rather quickly, e.g., 10–20 s for Ethereum (Etherchain, 2020). To show the distributed issue and the time value of different responsiveness, we choose periods from 11 to 101 with an increment of 10 to cover various early and late updates in the whole epidemic cycle. In other words, we change $\beta$ and $\gamma$ starting in different periods: 11, 21, 31, 41, 51, 61, 71, 81, 91, and 101, respectively. For instance, for starting period 11, $\beta = 0.25$ and $\gamma = 0.1$ remain the same from the beginning to period 10, while we increase $\gamma$ and/or decrease $\beta$ for faster recovery and/or lower transmission from period 11 to the end. Similarly, for starting period 21, we make the above changes from period 21 to the end.

Fig. 6a-c show the response surfaces of the total infected fractions of the population at the end of the epidemic cycle with different recovery rates ($\gamma$) and different starting times ($t$), assuming $\beta = 0.25$. From back to front, the lines correspond to different $\gamma$ values from 0.11 to 0.20. From right to left, the lines correspond to different starting times from 11 to 101. At each starting time, the total infected fractions are linearly correlated with $\gamma$. There is a clear change pattern in the linear

Table 1

| Periods | Recovered | SD($\beta$) = 0.125 | SD($\gamma$) = 0.05 | SD($\beta$) = 0.0625 | SD($\gamma$) = 0.025 | SD($\beta$) = 0.0625 | SD($\gamma$) = 0.025 |
|---------|-----------|---------------------|---------------------|---------------------|---------------------|---------------------|---------------------|
| 11-100  | 0.9438    | 0.9364              | 0.9323              | 0.9217              | 0.9575              | 0.9436              | 0.9428              | 0.9263              |

$0.1750, \gamma = 0.17$ and $\beta = 0.1625$, $\gamma = 0.18$ and $\beta = 0.1500$, $\gamma = 0.19$ and $\beta = 0.1350$, and $\gamma = 0.20$ and $\beta = 0.1250$, respectively.

Fig. 5. The 100-round evolutions of susceptible, infected, and recovered fractions in four scenarios.
relationship. When the starting time is early, there is a strong negative relationship – the larger the $\gamma$ value, the smaller the total infected fraction. However, as the starting time gets late, the negative relationship gradually weakens. If the starting time is over 91, the correlation disappears – that is, the total infected fraction is constant regardless of how $\gamma$ changes.

Taking a different perspective, at each $\gamma$ value, the relationship between starting time and the total infected fraction is contingent on the specific $\gamma$ value. When $\gamma$ is low, the correlation is linear and slightly positive; that is, having an early start time only has a minor effect on lowering the total infected fraction. However, as $\gamma$ increases, the relationship slowly shows a distinct S-shaped curve: the marginal change of the total infected fraction does not change much. The marginal value of time is the largest when starting time is in the mid-stage of the cycle. During this stage, the earlier the blockchain solution is implemented, the greater a reduction of the total infected fraction can be gained. It suggests that in order for a blockchain solution to be effective, it needs to have a sufficiently low $\gamma$ and be implemented in the early stage of the epidemic cycle.

We also create the response surface of the total infected fraction of the population at the end of the epidemic cycle with different transmission rates ($\beta$) and different starting times ($t$), assuming $\gamma = 0.1$. As Fig. 6d-f shows, the pattern of benefit gains of slower transmission is similar to that of faster recovery. Moreover, we create the response surface to explore how the simultaneous recovery rate increment and transmission rate decrement can lead to changes in the total infected fraction. As Fig. 6c shows, the pattern is still largely similar. The only difference is that when $\gamma \geq 0.18$ and $\beta \leq 0.15$, the total infected fractions fall below 5% when the starting period is earlier than 31. Overall, our findings suggest that the blockchain solution is most effective when it can simultaneously reduce transmission and enhance recovery and it also needs to be put in use at the early stage of the epidemic outbreak, particularly no later than the critical time window when the epidemic outcome will exponentially deteriorate if it is not intervened.

In addition, we visualize the blockchain benefits for the maximum infected fraction in the epidemic lifecycle. As Fig. 4 shows, the maximum infected fraction occurs in period 68 in our base case, so if the blockchain solution is implemented after the 68th period, the maximum infected fraction will not change. Except for this, the patterns for lowering the maximum infected fraction are similar to those for controlling the total infected fraction.

4.2.3. Benefits of privacy-preserving immutability

As the data saved in blockchain are immutable, people are assured that their data are secure and their privacy is effectively protected. As a result, they are more willing to honestly report COVID-19 related information to the system. Without such assurance, people are reluctant to report truthfully due to concerns that disclosing sensitive personal health information could cause inconveniences to their personal lives. Currently, there is a massive underreporting of COVID-19 cases in the U.
S. (Reese et al., 2020), which will lead to a biased understanding of the epidemic. In the meantime, governments implement data-driven guidance for critical decisions such as closing and reopening schools and businesses. For example, New York schools will close if the regional daily infection rate rises above 9% using a 7-day average and will open if the daily infection rate remains below 5% using a 14-day average (Cuomo, 2020). If the infection rate is underestimated, such decisions could put people’s health and lives in jeopardy.

Next, we demonstrate how underreporting can impair epidemic control. Assume that in reality $\gamma = 0.10$ and $\beta = 0.25$. We use the same initial values in period 0 as in our previous sections: $S_0 = 0.9999$, $I_0 = 0.0001$, and $R_0 = 0$. The evolutions of the susceptible, infected, and recovered fractions follow Equations 5, 6, and 7, assuming 100% honest reporting from the infected individuals. Adopting the New York policy (Cuomo, 2020), if the infection fraction ($I_t$) rises above 9 percent in a period based on the 7-period average, lockdown orders will be issued.
We assume that lockdown can reduce the transmission rate $\beta$ by half, from 0.25 to 0.125 in the next period. In other words, if $I_{t-7} = \frac{1}{2}I_{t-6} + \frac{1}{2}I_{t-5} + \frac{1}{2}I_{t-4} + \frac{1}{2}I_{t-3} + \frac{1}{2}I_{t-2} > 9\%$, lockdown will be started and the transmission rate $\beta$ drops to 0.125 in period $t + 1$. Similarly, once the infection fraction drops to below 5 percent based on the 14-period average, the lockdown is removed and the transmission rate $\beta$ rises back to 0.25 in the next period. In our simulation, when $I_{t-14} = 9.61\%$ (above 9 %) in period 54, the lockdown is in effect from period 55 till period 99, when $I_{t-14} = 4.88$ (below 5 %). All the remaining periods stay normal since $I_{t-14}$ never rises above 9 % again. In Fig. 7, the three solid lines show how the susceptible, infected, and recovered fractions evolve over the whole cycle. Note that the curve for infected fractions from periods 54 to 99 is flattened due to lockdown.

Governments post new cases daily. We define the newly infected fraction in period $t$ as $i_t = \beta S_t I_t$. Note that $i_t$ is different from the infected fraction $I_t$ and the two have the following relationship: $I_{t+1} - I_t = \beta S_t I_t - \gamma I_t = I_t - \gamma I_t$, while initially $i_0 = I_0$. That is, $I_{t+1}$ (the infected fraction at the beginning of period $t + 1$ or equivalently at the end of period $t$) is equal to $I_t$ (the infected fraction at the beginning of period $t$)

![Figure 7](image-url)
plus $i_t$ (the newly infected fraction during period $t$) minus $\gamma I_t$ (the recovered fraction during period $t$).

We introduce a discount factor $\alpha$ ($0 \leq \alpha \leq 1$) for underreporting. The reported newly infected fraction in period $t$ is $i'_t = \alpha i_t = \alpha i_t S_t I_t$. Initially $i'_0 = \alpha i_0$. The evolutions of the reported susceptible, infected, and recovered fractions are as follows:

$$S'_{t+1} - S'_t = -i'_t$$

$$I'_{t+1} - I'_t = \alpha i'_t - \gamma I'_t$$

$$R'_{t+1} - R'_t = \gamma I'_t$$

At the beginning, $S'_0 = 1 - i'_0$, $I'_0 = i'_0$, and $R'_0 = 0$. Decisions are made on the reported infected fractions $I'_t$ while actual evolutions are determined by the real susceptible, infected, and recovered fractions, i.e., $S_t$, $I_t$, and $R_t$. We first use $\alpha = 50\%$ to examine the impact of underreporting. (We conduct a sensitivity analysis of various $\alpha$ in the next section.) As the dashed lines in Fig. 7 show that in period 62, $\frac{i'_{62}}{i_{62}} = 9.25\%$ rises to above 9\%, triggering the implementation of lockdown from period 63 till period 89, when $\frac{i_{89.14}}{i_{89.14}} = 4.81$ drops to below 5\%. All the remaining periods stay open since $I_{t+1}$ never rises above 9\% again. Because of underreporting, the lockdown order is started later and removed earlier than it should. Compared with the solid lines, more people got infected between periods 54 and 95 and the overall infection rate is also higher.

However, if the government is not aware of the underreporting problem, it assumes the reported fractions are real fractions. The SIR model constructed based on the underreported fractions will provide a falsely optimistic assessment of the epidemic evolution. As the dotted lines in Fig. 7 show that such an assessment greatly underestimates the total infection fraction, which could lead to complacency and missed opportunities in controlling the epidemic. This is a typical problem of “we don’t know what we don’t know.” When it happens in an epidemic crisis like COVID-19, it will put many people’s lives in danger. Fig. 8 shows the ending recovered fraction (i.e., the total infected fraction) in the three scenarios. The relative deviation is compared to the 100\% reporting case. For 50\% reporting, the real recovered fraction is 4.24\% higher than that of 100\% reporting. In other words, underreporting makes the lockdown intervention 4.24\% less effective. However, the reported fraction based on underreported data is 47.88\% lower, indicating a false sense of security.

4.3. Sensitivity analysis

We conduct a sensitivity analysis to see how the ending recovered fractions will change due to different levels of underreporting and transmission rates. In the analysis, we let the reporting level change from 10\% to 100\% with a step size of 10\%, resulting in 10 levels. We also let the transmission rate change from 0.15 to 0.375 with a step size of 0.025, resulting in 10 rates. We then create a 10-by-10 matrix to get the 100 combinations of these two arithmetic sequences. The real and reported ending recovered fractions are calculated for each combination, respectively, and two response surfaces are drawn. As Fig. 9 shows, when the transmission rate is very small ($\beta = 0.15$), there is no lockdown, and all real ending recovered fractions are the same regardless of the reporting level. In general, a higher $\beta$ will cause a higher real ending
a) Real ending recovered fraction

![Graph showing real ending recovered fraction](image)

b) Reported ending recovered fraction

![Graph showing reported ending recovered fraction](image)

Fig. 9. The ending recovered (total infected) fractions at different levels of underreporting and transmission rates.

recovered fraction for a given reporting level. Out of the 100 scenarios, there are four exceptions when a higher $\beta$ triggers a lockdown with longer periods and then a smaller real ending recovered fraction. For 10 % and 20 % reporting, the real ending recovered fractions are the same because no lockdown is implemented. For a given transmission rate $\beta$, a higher reporting level tends to cause an earlier and longer lockdown, resulting in a smaller real ending recovered fraction.

The reported ending recovered fractions are the multiplications of real ending recovered fractions and their reporting levels. The reported ending recovered fractions are larger when the transmission rates are larger and/or the reporting levels are higher, with a few exceptions due to different lockdown duration. The reported ending recovered fractions range from 5.9 % to 86 %, which fluctuate much more than the real ending recovered fractions (from 59 % to 86 %). It is obvious that the higher the underreporting, the more severe the real recovered fractions are underestimated.
5. Conclusions

Our contributions are threefold. First, we identify three key blockchain features that can be applied to fight COVID-19: bottom-up decentralization, automation with real-time update, and immutability with privacy preservation. Second, we propose a blockchain solution that utilizes the above three features and leads to faster patient recovery and slower disease transmission of COVID-19. Third, we quantify the blockchain solution benefits for controlling COVID-19.

Bottom-up decentralization enables easy access and increases broad participation, which reduces estimation errors of the compartmental model parameters and reduces the variations in estimating the pandemic's evolution. The decentralized platform removes the risk of single-point failure by the principal organization in a centralized approach. Transparency, data ownership, and network externalities encourage more bottom-up participation from users. Automation with real-time update provides prompt detection and diagnosis, accurate contact tracing, and targeted mitigation and containment, thus helping to achieve faster recovery and lower transmission faster. Because the blockchain's immutability ensures privacy preservation, people are more willing to give truthful responses, avoiding false and erroneous underreported data that will lead to wrong estimates and derail interventions.

Under the ORM framework, the blockchain solution mitigates three specific types of risks: sample variance, delay, and bias. We quantify the corresponding benefits by using the SIR Model. When we increase the sample size, the standard errors of the transmission and recovery estimators become smaller, which decreases the variations of susceptible, infected, and recovered fractions over the cycle. Hence, the total infected fraction decreases at a given probability in terms of the widely used risk measures, VaR and CVaR. If delay can be avoided and interventions start on time, the benefits will be larger: the less total infected and the less maximum infected fraction over the whole cycle. This confirms that quick response can improve performance by reacting promptly to changes. When input data are biased, our estimated evolutions of susceptible, infected, and recovered fractions are wrong. All decisions based on the biased data are off-target, which impedes the efforts to control the crisis and put many people's lives in danger. In particular, because of underreporting, the severity of a crisis will be greatly underestimated, which leads to complacency and missed opportunities, a typical problem of "we don’t know what we don’t know." In sum, the blockchain solution is especially valuable when more participation from the public is needed, a high level of responsiveness to the crisis is necessary, and when people have serious concerns regarding their privacy in disclosing health information.

6. Discussion and managerial implications

Our work has important practical implications because it integrates a major disruptive technology in Industry 4.0 - blockchain - into fighting our current biggest challenge - COVID-19. Blockchain is in a phase of fast growth and has demonstrated versatility in various industries. We advance the solution method in two perspectives: (1) transform blockchain technological features into better COVID-19 mitigation and control performances and (2) provide detailed analytical analysis and quantitative benefits.

Regarding optimization, we have two conclusions. First, we can always achieve better performances if we can increase the sample size (i.e., more participation from users on the blockchain platform), if we can simultaneously reduce transmission and enhance recovery as early as possible, or if we can reduce the proportion of the underreporting. Second, our quantitative benefits can help to justify the additional investment by making it possible to conduct a cost-benefit analysis. It tends to cost more for more participation, earlier interventions, and less underreporting.

We recommend three specific action plans based on our research findings: (1) governments and organizations can build ready-to-use blockchain solutions for controlling COVID-19 if it is cost-effective, (2) but they should first quantify the blockchain benefits, and (3) then justify whether or not they should invest such blockchain solutions by conducting a cost-benefit analysis. Our findings support the current industrial practice (i.e., industrial verification) (Sharifi et al., 2021).

Our blockchain-enabled approach has valuable managerial implications. To cope with a global crisis, we need community participation and grassroots contributions with bottom-up initiatives. The blockchain technology fits naturally and can play a significant role in fighting a crisis that affects every-one in the global community. While letting users control their own data and privacy, our approach can help to make better decisions based on more truthful data in a timely manner from community users.
