The Weakest Goes to the Wall: The Impact of the Covid-19 Pandemic on Psychiatric Acute Care

Hossein Hejazian
Harvard Kennedy School, Harvard University, Cambridge, MA 02138,
hossein.hejazian@mail.mcgill.ca

Javad Nasiry
Desautels Faculty of Management, McGill University, 1001 Sherbrooke West, Montreal, Canada, H3A 1G5,
javad.nasiry@mcgill.ca

Vedat Verter
Smith School of Business, Goodes Hall, Queen's University, Kingston, Ontario, Canada, K7L 3N6,
vedat.verter@queensu.ca

Daniel Frank
Continuing Care Clinic, Jewish General Hospital, 4333 Chem. de la Côte-Sainte-Catherine, Montreal, Canada, H3T 1E4,
daniel.frank.med@ssss.gouv.qc.ca

We study the impact of the COVID-19 pandemic on psychiatric patient outcomes. We design a Difference-in-Differences (DID) framework to assess the causal effects of the pandemic on these outcomes. This study specifically examines changes in the probability of admission and the timing of return visits to the emergency department (ED), informed by a dataset tracking individual visits daily across a three-year span. The analysis is segmented by various stages of the pandemic, with the Oxford COVID-19 Government Response Tracker (OxCGRT) providing insights into the impact of policy interventions. The results reveal that the pandemic reduced the ED return time of psychiatric patients by lowering their hospitalization likelihood. The study underscores the complex interplay between a reduction in ED visits, likely due to concerns of contracting the virus, and an increased demand for emergency psychiatric care as mental health crises intensified amidst pandemic conditions.

Key words: mental health care; COVID-19 pandemic; ED return time; admission; length of stay; mediation analysis

---

1. Introduction

Emergency departments (EDs) are always crowded and hectic (Hoot & Aronsky 2008). ED visits aim to take care of acute cases that need immediate action (Durand et al. 2011). This procedure speeds up if the diagnostic tasks are done accurately and promptly (Croskerry 2012). Patients with mental disorders are one of the hardest cases in this regard as they may need more diagnostic tests/procedures than non-psychiatry patients (Koekkoek et al. 2006). Also, the diagnostics for psychiatric patients may not be accurate. These issues can contribute to the relapse of psychiatric patients, which in turn, increases their usage of EDs and adds to EDs' crowdedness (Huang et al. 2003). That said, mental patients are one of the heavy consumers of ED resources. In Quebec, psychiatric patients with frequent ED visits are recognized to constitute almost half of the ED visits for mental health conditions. This means that around half of the ED visits for mental conditions are from returning patients, who had a prior ED visit in the recent past (Gaida 2023).

The COVID-19 pandemic had a major impact on various aspects of people's lives. A surge of mental conditions resulted from people developing levels of panic, anxiety, depression, and mood-related disorders. In particular, the pandemic significantly altered social interaction patterns. Social distancing measures, lockdowns, and restrictions on gatherings imposed to flatten the curve of the virus spread changed profoundly how individuals interacted with one another. These measures deeply affected the way individuals socialized with one another, which took a toll on their mental health and well-being (Brooks et al. 2020). Prolonged isolation and lack of in-person social support exacerbated feelings of loneliness, stress, and vulnerability, further compounding mental health challenges. On the other hand, access to mental health care was limited (Duden et al. 2022), making it more difficult for individuals to seek assistance and care during this critical time.

The COVID-19 pandemic extremely impacted organizational structures globally, particularly healthcare systems. The rapid spread of the virus and increasing public health concerns prompted healthcare facilities to implement various protective measures to mitigate infection transmission. These measures included reallocating resources, restructuring care delivery models, and making fundamental changes to hospitalization protocols. In healthcare delivery, the pandemic accelerated the adoption of telemedicine to ensure continuity of care while minimizing in-person contact and reducing the risk of virus transmission (Hollander & Carr 2020). However, the usefulness of this alternative was not completely conceivable (Duden et al. 2022).

Patients with psychiatric conditions often return to EDs at higher rates, primarily due to the chronic nature of their disorders. Consequently, metrics such as hospital readmission rates and ED return times serve as standard measures of outcomes for these individuals (Hejazian et al. 2023). The COVID-19 pandemic, however, introduces potential shifts in these metrics through two opposing mechanisms. On one hand, the general public's reluctance to visit EDs, seen as potential hotspots for infections, may have resulted in fewer non-COVID patients seeking emergency care during the pandemic. On the other hand, the pandemic's broader implications—such as enforced isolation and community care interruptions stemming from social contact restrictions—may have exacerbated the mental health challenges for patients, increasing the need for acute care services.

The pandemic outbreak presented hospitals with a dual challenge: effectively managing COVID-19 cases while maintaining critical care for patients with non-COVID disorders. Among those impacted, individuals with mental disorders faced significant challenges. The pandemic notably influenced admission decisions for psychiatric patients, leading to a decreased likelihood of hospitalization (Xiong et al. 2020). This reduced hospitalization rate, coupled with disruptions in community care, exacerbated patient outcomes for individuals with psychiatric conditions (Duden et al. 2022). Moreover, the isolation and heightened stress levels experienced during the pandemic exacerbated the mental health challenges confronting these patients. This situation made care access and treatment effectiveness more critical than ever underscoring the urgent need for innovative solutions to provide continuous mental health support amidst healthcare system constraints. Addressing these issues became a matter of utmost urgency to mitigate the detrimental impacts on this vulnerable population.

In this paper, we develop a year-to-year difference-in-differences (Y2Y-DID) to study how the pandemic impacted the ED return time of psychiatric patients. We consider hospitalization as a possible mediator of the relationship between the pandemic and ED return time. We investigate how the admission likelihood was affected after the pandemic onset and how it impacted ED return time during the pandemic. We also utilize the Oxford COVID-19 Government Response Tracker (OxCGRT) data as a robustness check for the pandemic effects. This data helps us analyze the government responses/policies to the pandemic. These policies are possible mechanisms for explaining the impacts of the pandemic on patient outcomes.

This study reveals that the pandemic reduces the ED return time of psychiatric patients significantly compared to the pre-pandemic period. On the other hand, we illustrate that the pandemic reduces the admission likelihood of patients in psychiatry wards which in turn decreases the ED return time. This can serve as a possible mechanism to explain why and how the pandemic decreases the ED return time. Furthermore, we capture the impact of the pandemic by incorporating the government's response and COVID-19 policies. We highlight that economic support of households provided by the government during the pandemic can contribute to decreasing the burden on health systems by keeping patients away from EDs.

The paper is organized as follows. Section 2 provides a literature review on the topics of this study. Section 3 discusses the impact of the pandemic on patient outcomes and how these outcomes affect each other. Section 4 explains the data sets used for this study. Section 5 elaborates on the DID methodology and illustrates the details of the text analytics implemented on the triage notes. Section 6 provides the empirical models developed for this study and Section 7 shows the results. Section 8 provides a set of robustness checks to validate the results. Section 9 discusses the findings and Section 10 concludes the paper.

2. Literature Review

The ED revisit is considered a problem in the medical settings, and in particular, in psychiatric care (Hejazian et al. 2023). The COVID-19 pandemic is reported to exacerbate this problem (Duden et al. 2022). This section provides a summary of the works in both medical and operations literature.

2.1. ED Utilization and Revisit in Medical Literature

The ED revisits might be due to inaccurate diagnosis, ineffective treatment, or lack of post-discharge follow-up for the discharged patients. The impacts of these factors can be more pronounced for vulnerable patients (Sabbatini et al. 2016), and in particular, for psychiatric patients. Community care and accessibility of care providers outside hospitals are among the most important factors that can prevent ED revisits and readmissions. On the other hand, hospital characteristics and ED structures/systems can impact the quality of care and patient outcomes, including ED revisits (Warner et al. 2018). That said, the practices of ED providers and the operational systems of care within ED facilities can be focal points for interventions to decrease ED revisits and readmission rates (Singh et al. 2015).

The global health crisis of the COVID-19 pandemic has unquestionably reshaped emergency healthcare landscapes. It has triggered an influx of patients that, in turn, has overwhelmed EDs and exerted enormous pressure on healthcare systems (Duden et al. 2022). As a result, hospital resources became thinly spread, obliging healthcare providers to meticulously balance between attending to COVID-19 patients and delivering care for individuals with other pressing medical requirements (Johnson et al. 2021, Shobassy et al. 2022, Deshpande et al. 2022). Amid the growing concerns surrounding virus transmission, hospitals rolled out stringent infection control protocols and precautionary steps, which influenced admission decisions, particularly for psychiatric patients (McGuire et al. 2021, Duden et al. 2022). Given the pandemic's disruptive effect on healthcare operations and resource allocation, it becomes imperative to scrutinize its impact on hospitalization rates and the accessibility of psychiatric care.

Psychiatric care, in particular, encountered distinct challenges during the COVID-19 pandemic (Goldenberg & Parwani 2021). Mental health disorders and distress have been amplified by the multitude of stressors associated with the pandemic, such as social isolation, financial hardship, and anxiety about the virus. Conversely, the pandemic has imposed obstacles to access psychiatric care, including a contracted capacity in psychiatric units, constrained availability of outpatient services, and disruptions to traditional therapy modalities (Yalçın et al. 2021). The escalated pressure on mental health professionals and the overall healthcare workforce has further complicated psychiatric care delivery, with an uptick in staff burnout rates and sick leaves (Phillips 2020, Weibelzahl et al. 2021). Therefore, investigating the pandemic's impact on psychiatric care utilization and understanding the consequent alterations in hospitalization decisions for patients with psychiatric disorders is of paramount importance. This knowledge is crucial for addressing the dynamic mental healthcare needs of the population during the unprecedented times.

2.2. ED Utilization and Revisit in Operations Literature

The operations literature has not addressed the ED revisit problem during the pandemic. However, the problem of ED revisits along with admissions to the hospital has been studied vastly from several perspectives, including admission prediction during the ED visit and readmission prediction after discharge. An important factor in reducing the ED revisit frequency is hospitalization. Various studies highlight the significant impact of hospitals and their associated community networks, including local physicians and support services, on the likelihood of hospital admissions.

The prediction of patient outcomes following ED departure represents a critical issue in healthcare operations management (Saghafian et al. 2012, 2014). At the earliest juncture, the goal is to establish whether a patient will necessitate hospitalization, thus facilitating efficient bed allocation and transfer arrangements (Chen et al. 2022). Determining the probability of admission at the triage stage carries considerable operational ramifications, influencing patient ED length of stay (LOS), inpatient ward bed distribution, and staff management strategies (Parker et al. 2019). Although this issue might appear to have been thoroughly examined in operations management literature, we discern a significant gap in the existing body of research.

In recent years, applying predictive techniques, such as machine learning and natural language processing (NLP), to augment the precision of hospital admission predictions during the ED triage stage has piqued considerable interest (Hong et al. 2018). These models strive to estimate individual patient admission likelihood by analyzing an array of data sources, encompassing triage severity scores, chief complaints, vital signs, and demographic details (Feizi et al. 2022). The wealth of data gathered during triage provides a unique opportunity to harness advanced analytics to forge predictive models beneficial for early decision-making concerning bed distribution and resource management. Incorporating such predictive models into ED workflows can streamline patient flow, enhance resource allocation, and bolster overall care delivery efficiency.

Researchers and administrators can use advanced analytical techniques and data-driven methodologies to sift through vast volumes of healthcare data, extracting meaningful patterns, trends, and correlations. These insights catalyze informed decision-making across a plethora of healthcare aspects, including wait times, admissions, hospital performance, and other pressing operational issues. Wowak et al. (2023) present a review of business analytics in the healthcare literature across operations management, information systems, healthcare, and analytics. Infusing analytics into healthcare operations management promotes evidence-based decision-making, optimizes resource allocation, and fosters improved patient outcomes. By employing these analytic tools and methodologies, researchers and administrators can pinpoint areas needing enhancement, streamline processes, and bolster overall efficiency and the quality of healthcare delivery.

3. Impacts of the COVID-19 Pandemic on Psychiatric Patients

The effect of the COVID-19 pandemic on psychiatric patients is supposed to be intuitive. The pandemic outbreak changed every aspect of people's lives across the globe. World Health Organization (WHO) reported that the prevalence of depression and anxiety increased about 25% globally (Kupcova et al. 2023). In Canada, major depressive disorder (MDD) is reported to be more than two times higher during the pandemic compared to the pre-pandemic period (Shields et al. 2021). However, these effects are mostly related to the general well-being of the population.

The pandemic's effects on people's well-being have been studied vastly in the recent ever-growing literature. However, its impact on patients with severe mental disorders is not entirely understood (Alsaeed et al. 2023, Yang et al. 2020, Brooks et al. 2020, Roy et al. 2020, Krishnamoorthy et al. 2020, Bueno-Notivol et al. 2021). Moreover, the pandemic's effects on the healthcare systems continue to persist until today. For example, the workload increase in mental healthcare amid the pandemic outbreak did not get back to the pre-pandemic level in Canada.[1] To this end, we aim to address two main research questions in this study: 1) how does the COVID-19 pandemic impact the ED return time of psychiatric patients? 2) how does the COVID-19 pandemic impact the hospitalization likelihood of psychiatric patients?

3.1. Impacts of the COVID-19 Pandemic on Returning Time to ED

Amid the pandemic outbreak, patients with mental illnesses developed two different concerns: on one hand, they were reluctant to visit EDs due to their fear of infection and hypochondria (Shobassy et al. 2022, Duden et al. 2022), and on the other hand, they experienced a deterioration in their pre-existing conditions caused by the pandemic (e.g., lockdown, anxiety, and loneliness) yielding a surge in the population mental health problems (Johnson et al. 2021, Duden et al. 2022).

An adverse outcome commonly associated with the quality of care provided in the ED is the patient's rapid return to the ED, sometimes termed as 'bouncebacks.' The goal of an ED visit, besides immediate stabilization, is to provide a care plan that prevents short-term return to the ED. This outcome is often construed as a failure of the system to manage the patient's condition appropriately. This is particularly relevant for psychiatric patients, who might be more likely to return due to a lack of appropriate follow-up options or a failure to fully manage their psychiatric crisis during the initial visit (Hejazian et al. 2023).

In the context of the COVID-19 pandemic, we surmise that these short-term ED returns may have been exacerbated. Health systems were struggling with a shortage of staff even before the pandemic onset. This issue was worsened amid the pandemic outbreak due to several physical and mental strains affecting the health staff. These strains include staff's concern for the loneliness of

[1] https://www.cbc.ca/news/health/canada-mental-health-crisis-covid-19-pandemic-1.6382378

their patients, themselves, and their families. Moreover, staff was worried about the pandemic's consequences, the possibility of getting infected, quarantined, and being burned out (Sibeoni et al. 2021, Shaw et al. 2021, Duden et al. 2022). On the other hand, some psychiatric services shifted from in-person to online mode (telepsychiatry) during the pandemic. That may have resolved the issue of losing access to mental care due to health/infection concerns. However, Vaičekauskaitė et al. (2021) reported that psychologists found telework exhausting. Moreover, there were several issues in maintaining the protocols and procedures for protecting individuals from getting infected in nursing homes and community dwellings. For instance, Johnson et al. (2021) show that more than half of staff in residential facilities experienced difficulties in sustaining infection prevention plans, partly due to lack of protective equipment (Duden et al. 2022).

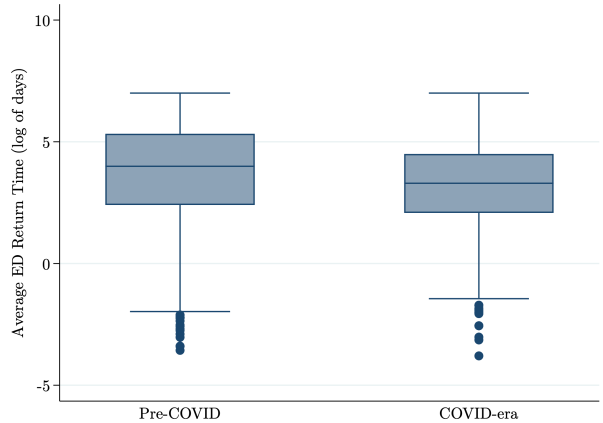
The pandemic disproportionately impacted psychiatric patients, who often rely heavily on consistent and continuous care. These patients experienced disruptions in their regular care routines due to pandemic-related restrictions or the shifts in resource allocation to keep up with the surge of COVID-19 cases. In particular, community care and social support were reduced, often due to being conceived as non-essential (Jurcik et al. 2021). This resulted in challenges in discharging patients from hospitals due to the lack of outpatient care (Duden et al. 2022). For instance, social workers who once visited patients with severe mental disorders at their residences refrained from doing so out of infection concerns, further limiting psychiatric patients' access to community support (Nicholas et al. 2023). Helplines appeared as a replacement for acute care services Johnson et al. (2021), Nair et al. (2021), however, these attempts were insufficient as community care remains not a priority in mental health services. World Health Organization (WHO) reports that the majority of countries dedicate less than 20% of their budget to these services (Duden et al. 2022).

As discussed, the pandemic has created a variety of stressors on the healthcare system, including increased crowding, reduced capacity, and altered admission patterns, all of which could impact the quality of care delivered and, hence, the rate of short-term returns. Shobassy et al. (2022) showed that the pandemic impaired patients' access to mental healthcare increasing their need to seek psychiatric emergency care. That said, the ED return time of psychiatric patients may be shorter in the COVID period compared to the pre-COVID period, indicating a potentially higher frequency of ED visits. Figure 1(a) shows how the return time to the ED reduces after the pandemic. Interestingly, Figure 1(b) shows this effect escalated throughout the pandemic waves.

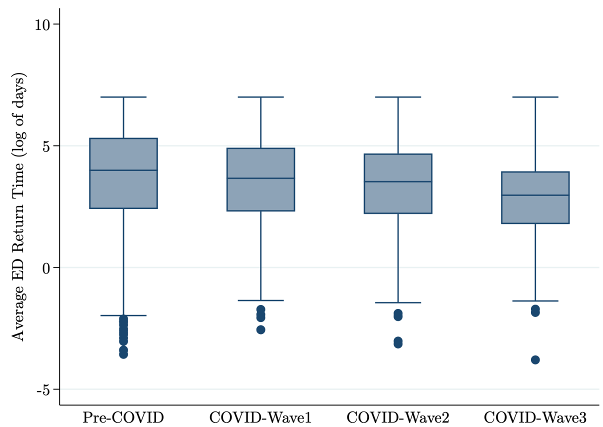
3.2. Impacts of the COVID-19 Pandemic on Admission Likelihood

Admission decisions are intrinsically linked to the future healthcare trajectory and patient outcomes. The COVID-19 pandemic has imposed substantial shifts in hospital admission decisions, specifically for patients with psychiatric disorders, potentially leading to underutilization or overutilization of mental healthcare services. As the pandemic progressed, EDs and hospitals wrestled with

Figure 1 Boxplots of Returning Times for pandemic periods and waves.



(a) Pandemic Periods



(b) Pandemic Waves

overcrowding and resource strain issues, exacerbating the overutilization of their available resources and necessitating more resources in healthcare delivery. It is reported that hospital admissions and hospital length of stay for psychiatric patients decreased after the pandemic outbreak in several places of the world (Carpiniello et al. 2020, McGuire et al. 2021, Ornell et al. 2021, Roncero et al. 2020, Fasshauer et al. 2021a,b, Duden et al. 2022). This decrease is attributed to a reduction in psychiatry inpatient capacity. On the other hand, health systems may have visited fewer patients to enforce social distancing measures and limit virus spread, thus contributing to the underutilization of their original capacities (Duden et al. 2022, Shobassy et al. 2022). Moreover, the demand for acute care underwent a decline as some patients avoid seeking acute care due to their fear of getting infected by the virus.[2]

The pandemic's surge effects inevitably influence care utilization, mirrored in patients' disposition at the end of their ED visit. Hospitalization decisions—whether to hospitalize a patient

[2] https://www.nytimes.com/2020/05/25/health/coronavirus-cancer-heart-treatment.html

or discharge them home—depend on the hospital's capacity and the occupancy level of inpatient wards, amongst other factors. We conjecture that the pandemic has significantly impacted these two key factors, leading to a measurable discrepancy between the periods before and after the pandemic outbreak. That said, the COVID-19 pandemic reduced the hospitalization likelihood for psychiatric patients.

3.3. Effects of Hospitalization on ED Return Time

As Figure 1 shows that the ED return time reduces in the COVID-era period, we elaborate further on this in Figure 2 by illustrating the average and confidence intervals of ED return time for each triage score. We observe no difference between the ED return time of admitted and discharged patients for all the triage scores except level 1 (resuscitation) in the pre-COVID period. However, in the COVID-era, the admitted patients have an evident higher ED return time than discharged patients across all the triage scores.

Figure 2 shows that the ED return time for discharged patients with a triage score of 5 is lower than that of a triage score of 4. To explain this phenomenon, we refer to a report showing Montreal's EDs were struggling to stick to the Canadian triage guidelines: "For P5 [triage score of 5] patients, who should be seen within two hours, the ED response times in the Montreal area were shorter than for P3s [triage score of 3] and P4s [triage score of 4], suggesting that some doctors may be examining them before the more urgent cases. The veteran triage nurse confirmed that this does sometimes occur."[3]

The above evidence makes us believe that ED care can be augmented by hospitalizing a patient in an inpatient ward. In this case, the low quality of care provided to the patient might be compromised by the care the patient receives in the hospital. Therefore, the COVID-19 pandemic could intensify the discrepancy between the ED return time of admitted and discharged patients. In other words, the discharged patients return to the ED sooner than the admitted patients in the COVID-era.

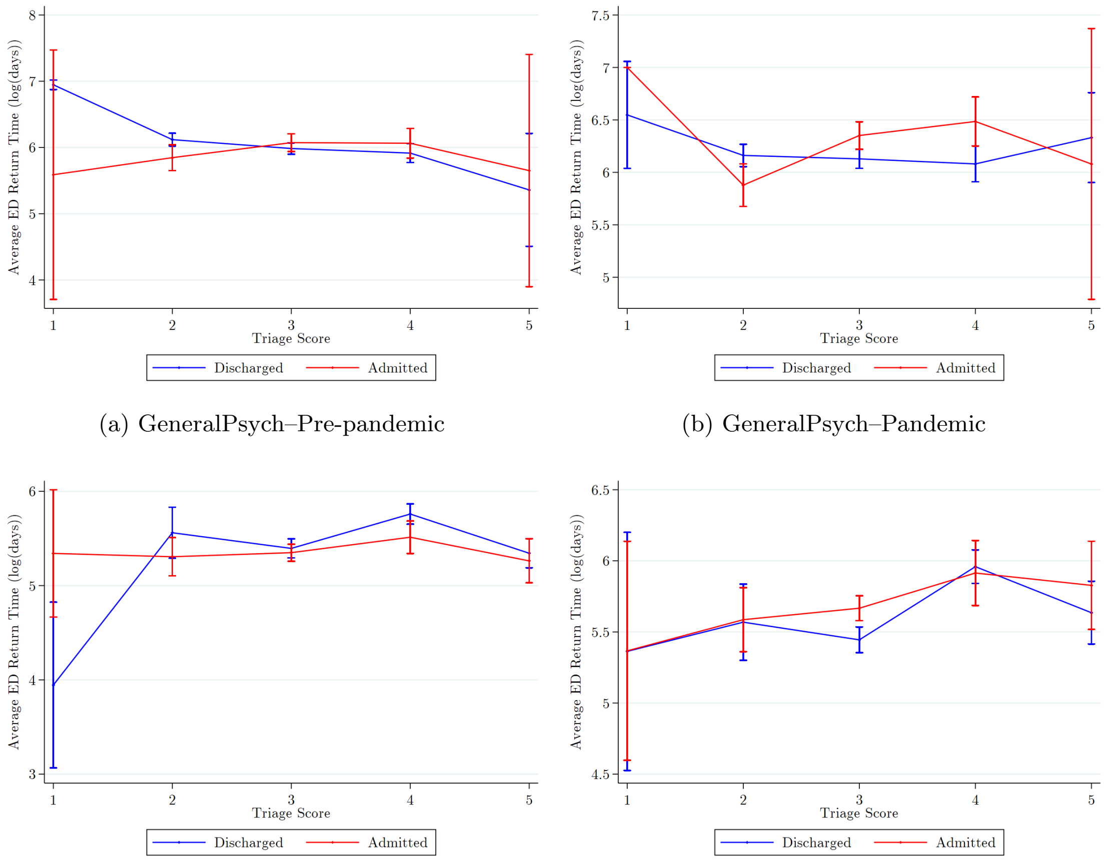
3.4. Effects of Govermental COVID-19 Policies on Psychiatric Patient Outcomes

The outbreak of the COVID-19 pandemic created an unparalleled situation, leading to profound changes in every facet of daily life. Governments worldwide implemented a myriad of measures in response, including restrictions on gatherings, closures of recreational and social venues, lockdowns, and curfews. These policies aimed to contain the spread of the virus, flatten the infection curve of the population, and control the overall situation to prevent crises.

Given Canada's federal system, each province, with its own governing body, shaped its unique response to the challenges of the pandemic. This resulted in a variety of approaches with varying stringency across cities and provinces throughout Canada (Breton et al. 2020). We utilize

[3] https://montrealgazette.com/news/local-news/montreal-ers-fall-far-short-of-canadian-triage-guidelines-report

Figure 2 Average ED Return Time by Triage Levels in the Pre-COVID and COVID-era Periods.



(a) GeneralPsych–Pre-pandemic

(b) GeneralPsych–Pandemic

(c) SpecialPsych–Pre-pandemic

(d) SpecialPsych–Pandemic

the Oxford COVID-19 Government Response Tracker (OxCGRT) to analyze the governmental responses in detail. This comprehensive data set captures a wide range of government measures, spanning "closure and containment restrictions, health policies, economic support measures, and vaccination prioritization, delivery, funding, and requirements" (Hale et al. 2021). Notably, it includes data for subnational jurisdictions of some countries, Canada being one of them. The data set recorded the government responses to COVID-19 in 19 policy areas, aggregated into four main policy metrics. These are Government Response Index (GRI), Stringency Index (SI), Containment and Health Index (CHI), and Economic Support Index (ESI). The compositions of these indexes are explained in Appendix EC.1.

We do not use these indices to study the effectiveness of the government's COVID-19 responses. Instead, we use them to study how the macro-level policies could impact the patient outcomes of psychiatric patients. Drawing on this data, we delineate the nature and chronology of policies in place. Integrating these policy indicators with our data sets allows us to elaborate on the empirical consequences of policy responses on psychiatric patients seeking acute care in emergency departments during the spread of COVID-19.

4. Data Set and Data Preparation

4.1. Two Data Sets

A consortium of Quebec psychiatric researchers was established to measure the impact of the COVID-19 pandemic on the care provided to psychiatrically ill patients in Quebec and any resultant effects on their mental health. Being part of this consortium, we obtained two sets of data from two different hospitals in Montreal. The first data set contains ED visits of psychiatric patients from one of Montreal's psychiatric hospitals that serve only psychiatric patients. We call this hospital "SpecialPsych" in our study. The second data set consists of ED visits of patients with acute mental disorders from one of Montreal's general hospitals. We call this hospital "GeneralPsych".

The advantage of our data is the detailed information about patients. These two data sets are constituted of similar fields, including patient flow information (arrival/discharge date to/from the ED and hospital admission/discharge date), patients' sociodemographics (age, gender, marital status, employment status, living status, and family status), physical condition (blood pressure, heartbeat rate, etc.), and patients' clinical information (triage score, chief complaint, principal diagnosis codes, medications, etc.). However, the two data sets have some discrepant fields. The data set GeneralPsych is richer regarding patient flow times, while the triage notes are available only in the data set SpecialPsych. The triage notes are another advantage of our data that illustrates patients' conditions in unstructured text formats. We use NLP methodology to exploit the information available in these notes that is not usable otherwise. We elaborate on these notes later in Sections 4.3 and 5.2.

Despite the few differences, both data sets contain three years of patient records from September 14, 2018 to September 14, 2021. We consider March 14, 2020 as the commencement date of the COVID-19 pandemic, when the government of Quebec officially declared a public health emergency to last ten days.[4] This measure granted the premier powers under the Public Health Act to enforce mitigation measures.[5] March 14, 2020 splits our data in two equal shares such that we obtain patient records for 18 months before and 18 months after the pandemic onset.

We should note that the pandemic outbreak was not a one-time shot. Instead, it had manifold impacts on the societies that started and ended at different time frames. Thus, we devise different time frames to capture the effects of different waves of the pandemic and the corresponding policies initiated by the authorities.

The data set SpecialPsych contains 13,422 visits from 7,702 patients, and GeneralPsych contains 9,719 visits from 7,148 patients. Table 1 shows the summary statistics.

[4] `https://web.archive.org/web/20200316155253/https://www.cbc.ca/news/canada/montreal/covid-19-march14-1.5497961`

[5] `https://en.wikipedia.org/wiki/Timeline_of_the_COVID-19_pandemic_in_Quebec`

Table 1 The Summary Statistics across the two hospitals.

	SpecialPsych		GeneralPsych		Difference	
	Mean	S.D.	Mean	S.D.	b	t
Age	35.98	16.58	38.42	20.91	-2.44***	(-9.52)
Gender (male)	54.9%		50.4%			
Triage Score	3.34	0.82	2.85	0.72	0.49***	(47.82)
Admitted	0.41	0.49	0.22	0.41	0.20***	(32.63)
Ward LOS (day)	15.90	58.51	6.86	24.71	9.04***	(16.03)
Observations	13412		9719		23131	

The mean (S.D.) shows the average (standard deviation) of variables.
$^{+}$ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

4.2. Outcome Variable

The outcome variable *ED_ReturnTime* represents the duration (in days) between a patient's successive ED visits. This metric reflects the frequency of ED revisits by psychiatric patients. For instances where patients do not make a return visit to the ED following their initial visit, we consider the return time as 1100 days (a few days more than three years), a duration exceeding our study horizon. We posit that this approach does not impair the validity of our results, given that we apply a log transformation (with a natural base) to the return time to correct distribution skewness. However, to ensure the validity of this approach, we perform a robustness check on the ED return time of non-returning psychiatric patients in Section 8.8.

4.3. Translation of French Triage Notes

The data records we have in SpecialPsych are in French. We are confronted with two options: to process the NLP analysis with French text or to translate the text into English first and then proceed with the analysis. We opt for the second approach. For all our NLP tasks, we utilize HuggingFace Transformers. In particular, we employ the pre-trained model "Helsinki-NLP/opus-mt-fr-en" for translation. We use this pre-trained model to construct a translator that converts batches of French text into English. The Marian model from HuggingFace's Transformers library, pre-trained on a vast multilingual data corpus offering highly sophisticated translation capabilities, is employed for this purpose. The translation process begins with the model converting the input text into an understandable format, a process known as tokenization. The model is then presented with these tokenized texts and produces the corresponding translations. Initially in tokenized form, these translations are decoded back into readable text. The final output of the translator is a list of translated texts.

Assessing the performance of a translation model can be pretty tricky, particularly in the absence of "ground truth" translations for comparison. However, automated metrics can offer an approximate evaluation of the model's performance. For example, the BLEU (Bi-lingual Evaluation Understudy) score measures machine-generated translations' similarity to human-generated reference translations. CHRF is another evaluation metric that calculates the F-score using character n-grams, mostly bounded between 0 and 1. Note that no single metric is perfect, as all metrics have limitations. When evaluating translation quality, considering multiple metrics and comparing their results is often helpful. Additionally, the quality of human evaluation is still essential in assessing the quality of translations.

To verify this, we randomly selected a sample of translated patient notes to be translated independently by a research assistant for comparison. We perform stratified sampling to draw examples from our data that maintain the distributions of critical categorical variables in the sample as in the main data set. These variables are pre-COVID and COVID-era visits, sex, triage severity score, marital status, and ED disposition destination (if admitted to the hospital or discharged to home). We also ensure each patient is only represented once in the sample. This sampling approach aims to ensure that the sample accurately represents the overall population across the specified categorical variables, providing a sound basis for our performance evaluation.

The BLEU score of our translator is 0.19, and its CHRF score is 0.85. We are content with the results and conclude that our translator is an efficient tool for translating sizable amounts of text data, leveraging advanced machine translation models, batch processing, and GPU computations.

5. Methodology

We use a difference-in-differences (DID) framework to study how the COVID-19 pandemic affected psychiatric patients and acute mental healthcare. The outcome of interest is the return time of psychiatric patients after their ED visit. Our proposed methodology has two main components. The first and primary component embodies the foundation of our empirical methodology discussed in Section 5.1. The second component deals with the evaluation of the admission likelihood using language models and machine learning, explained in Section 5.2. This component provides us with the instrumental variable used for correcting endogeneity in our empirical models in Section 6.

5.1. Difference-in-Differences

5.1.1. Basics. The DID framework was introduced around 170 years ago to find out that the water supply was the source of cholera waves in London in the mid-1800s (Caniglia & Murray 2020). Since then, DID literature has been silent for more than a century. However, it has received much attention in the past few decades following the paper written by Ashenfelter & Card (1984) that coined the term difference in differences. Currie et al. (2020) showed an increasing trend in using DID methods in top economic journals, recently.

The purpose of DID is to find the causal effect of a treatment on a sample of individuals. In the DID framework, we need two distinct groups of individuals, named control and treatment, for which we should have data before and after a specific event date. This is called a 2-by-2 design which requires two assumptions to hold: *parallel trends* and *no anticipation*. If the DID setting satisfies these assumptions, one can claim that the DID estimates the average treatment effects for the treated units (ATT). These assumptions make the DID resemble a natural experiment by *"tidy[ing] up confounding factors that Nature has not controlled for us"* (Cunnigham 2021). In the following, we briefly discuss these assumptions and how to ensure they are satisfied.

No anticipation assumption refers to a situation where individuals do not predict that the event is coming. In our case, the DID estimator will be biased if the people anticipate that the COVID-19 pandemic will happen. This is because the people, specifically patients and the health system, will change their behaviors in advance before the pandemic occurs. As a result, we fail to identify the treatment effect by distinguishing between patients in the pre- and post-pandemic commencement period. Although the COVID-19 pandemic was an unfortunate event for the whole world, fortunately, it occurred suddenly against most people's expectation (to the best of our knowledge).

We set the pandemic commencement date on March 14, 2020. One may argue that people, the health system, and the government were aware of the pandemic coming in early March, or even in late February. We address this concern by a robustness check in which we shift the pandemic outbreak date to a few weeks earlier and later than March 14, 2020. The results are shown in Section 8.5 indicating our model is robust concerning the pandemic outbreak. That said, we conclude that our empirical methodology satisfies the no anticipation assumption.

The parallel trend (PT) assumes that if the treatment never happened, then the outcome of the treated units would not differ from the outcomes of the control units. This is required to estimate the causal ATT by comparing the outcomes of the treatment group with the control group. Note that PT is only about the outcome of untreated units of the two groups (treatment group and control group). This means that PT has nothing to say about the treated outcomes. Also, note that confirming PT does not guarantee that the outcomes of the treated units in case of no treatment would be similar to their counterparts in the control units. Instead, failing to confirm the PT assumption makes the treatment effect contaminated with selection bias and invalidates the DID results. Thus, PT is used along with a counterfactual assumption.

5.1.2. Year-to-Year DID. In our setting, we consider the pandemic outbreak as the treatment and aim to identify its causal effect on the ED return time. In the typical DID framework, there are two groups of individuals and two periods. However, our setup differs from the typical DID framework because the groups and the periods are mixed, i.e., there is no untreated individual after March 14, 2020, when the pandemic hit all the people without exception. To resolve this issue, we follow Bandiera et al. (2005), Cao et al. (2022) and introduce year-to-year DID (Y2Y-DID).

A Y2Y-DID considers the intervention date as the treatment and differentiates between the control and treatment groups by only using the date of records. The Y2Y-DID, as introduced in (Bandiera et al. 2005), considers similar periods but in different years (or other time frames) to distinguish between treatment and control groups. This approach controls for time trends and seasonality. The Y2Y-DID is, in fact, a combination of the interventions and the placebo test.

In our study, the intervention is the pandemic outbreak. Considering March 14, 2020 as the real pandemic outbreak, we specify March 14, 2019 as the placebo pandemic outbreak. Our DID analysis uses the following four groups of patient records in which the records from 2019/2020 are considered as the control/treatment group:

- Control group before treatment: patient records from January 1, 2019 to March 14, 2019;
- Control group after treatment: patient records from March 14, 2019 to December 31, 2019;
- Treatment group before treatment: patient records from January 1, 2020 to March 14, 2020;
- Treatment group after treatment: patient records from March 14, 2020 to December 31, 2020.

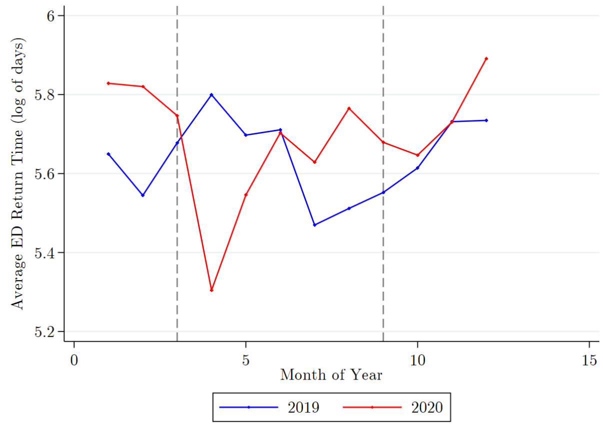
5.1.3. DID Diagnostics. As discussed above, the DID has two prominent assumptions: no anticipation and parallel trends. We explained why our methodology satisfies the no anticipation assumption, earlier in Section 5.1.1. However, we acknowledge that because our data is observational (not experimental), no anticipation is not a design assumption which means it is not testable and cannot be assured by randomization. In this section, we elaborate on PT and the additional steps to take when it is violated.

Usually, visualizing the outcome trends of the control and treatment groups over time can illustrate whether PT exists in a data sample. Figure 3 depicts the trend of the ED return time (log-transformed) in 2019 and 2020 at month level (Part (a)) and week level (Part (b)). We observe the outcome trends in periods before 14 March in both 2019 and 2020 are almost similar and they vary after 14 March. We may conclude that our sample satisfies the PT assumption.

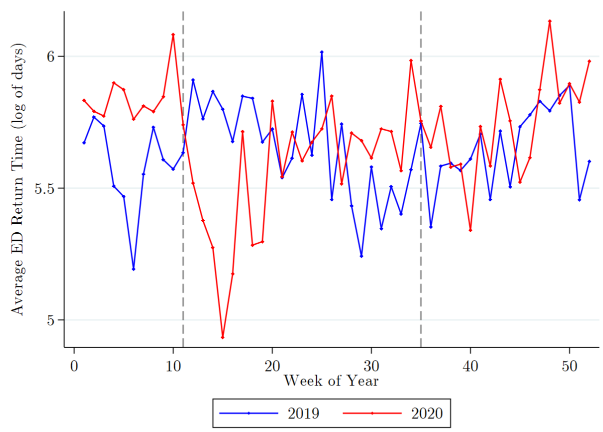
Most of the DID literature focuses on panel data, which contains records of specific individuals across the time (for example, years). On the contrary, repeated cross-sections, which contain samples of somehow different individuals across the time, are investigated to some limited extent. However, almost all of the methods and frameworks developed for panel data apply to repeated cross-sections. Although the individuals may vary throughout the years, they are selected randomly from the population. This is why repeated cross-sections can be treated as panel data.

The compositional shift in repeated cross-sections violates the PT assumption. Compositional shift refers to the change in the distribution of individuals in different periods that is correlated with the treatment assignment. In other words, variations in units' characteristics can change the composition of the sample. This change can be accounted for by incorporating time-varying covariates into the DID analysis. Then, the PT assumption can be validated in sub-samples. This is called conditional PT.

Figure 3 **Average ED Return Time of Psychiatric Patients in Years 2019-2020.**



(a) Month

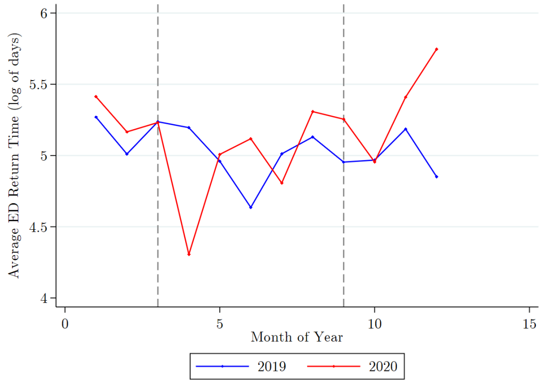


(b) Week

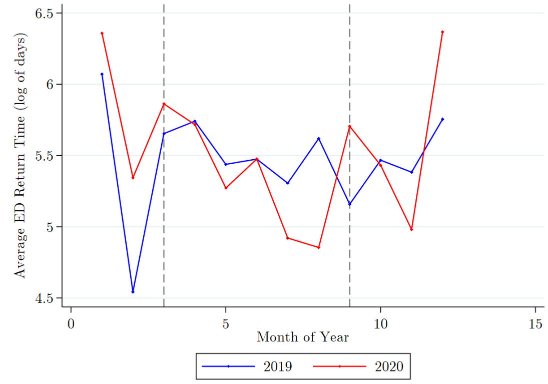
Later in Section 6.3, Figure 5 shows that the existence of compositional shift, for example in the distribution of diagnosis classes, is unlikely in our data. However, we delve into this assumption by considering the trends in sub-samples. As an example, Figure 4 illustrates the outcome trends for the sub-sample of male patients with “F20” diagnosis class in SpecialPsych hospital (Parts (a) and (c)) and GeneralPsych (Parts (b) and (d)).

We observe the trends are pretty similar between the control and treatment groups in periods before March 14. We conclude that the PT exists in the subsamples, i.e., the conditional PT assumption is satisfied. The existence of Conditional PT requires the DID framework to contain covariates. This means that if we use a two-way-fixed-effect (TWFE) model, we should include the control variables in the regression. Although the Operations literature did not show a meticulous investigation of the models proposed for performing DID analysis with covariates, we acknowledge that adding covariates to the model opens doors for other types of biases.

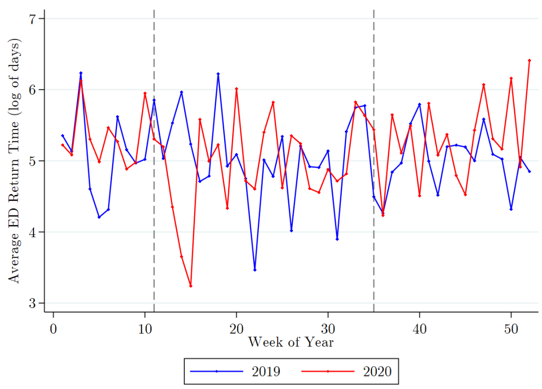
Figure 4 **Average ED Return Time of Male Patients with "F20" ICD-10 Codes in Years 2019-2020.**



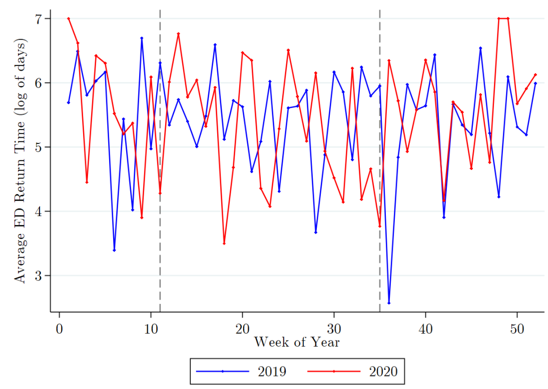
(a) SpecialPsych-Month



(b) GeneralPsych-Month



(c) SpecialPsych-Week



(d) GeneralPsych-Week

Incorporating time-varying covariates into the TWFE specification makes the DID estimates biased as they are correlated with the unit fixed effects (Arkhangelsky & Imbens 2023). An initial solution can be using Propensity Score (P-score) matching to create a balanced sample in terms of the covariate. P-scores are the probability of being in the treatment/control group given the covariate value ($\Pr(D = 1|X)$). Besides, two enhanced methods are developed by building upon the P-score method. Abadie (2005) developed the Inverse Probability Weighting (IPW) method by using the P-scores from the covariates to weight the ATT estimates over individuals. This method only requires two assumptions: conditional PT and common support. The latter assumption means the categories of individuals in terms of covariate values exist in both control and treatment groups.

In case we cannot ensure the correctness of P-scores, Sant'Anna & Zhao (2020) develops an improved method by controlling for covariates two times, once in calculating the P-scores and once in calculating outcomes adjusted by regressions (it is called Outcome Regression (OR) method introduced by Heckman et al. (1997)). Sant'Anna & Zhao (2020) names their proposed method Doubly Robust (DR) DID. We use both these methods as robustness checks in Section 8.6.

These other considerations for obtaining unbiased DID estimates are homogeneity across units (individuals) and across time. Homogeneous treatment effect across individuals means there is no heterogeneity between individuals. Including patient characteristics as controls in the regression model resolves the heterogeneity of individuals. Homogeneous treatment effect across time is referred to as constant treatment effect. If the treatment effect varies over time, then we have dynamic/carry-over effects. However, Arkhangelsky & Imbens (2023) declares that the constant treatment effect assumption can be ignored in the DID setting with block assignment. In block assignment, a subsample of units is treated after a common starting date. The following matrix visualizes the scheme of treatment assignment over individuals (rows) across periods (columns):

$$D = \begin{pmatrix} 0 & 0 & 0 & 0 & \cdots & 0 & 0 \\ 0 & 0 & 0 & 0 & \cdots & 0 & 0 \\ 0 & 0 & 0 & 0 & \cdots & 0 & 0 \\ 0 & 0 & 0 & 1 & \cdots & 1 & 1 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & 1 & \cdots & 1 & 1 \\ 0 & 0 & 0 & 1 & \cdots & 1 & 1 \end{pmatrix}$$

Our setting resembles the block assignment design because the treatment is the pandemic outbreak after which all individuals in the data are considered treated. Thus, we do not have the dynamic treatment effect concern even though the pandemic affects individuals differently across the months after March 2020.

5.2. Admission Evaluation from Triage Notes

Triage notes represent a significant yet often underutilized source of clinically relevant condition information. While typically aligned with triage severity scores, these notes may offer a more nuanced perspective that could help predict the likelihood of hospital admission for patients presenting to the ED with psychiatric disorders. The lack of attention to this resource is primarily due to the complexities inherent in handling unstructured text data. However, advancements in machine learning (ML) and natural language processing (NLP) techniques present an opportunity to tap into this potential. Leveraging NLP techniques could reveal patterns and indicators within triage notes that may enhance the predictive accuracy of admission likelihood, extending beyond the capabilities of traditional triage scores and demographics. In this section, we acquire the expert evaluation from the psychiatrist of our research team to perform a professional evaluation on a randomly selected sample of triage notes from our SpecialPsych data set. The aim is to obtain patients' admission likelihoods through their triage notes.

We have carefully curated a representative subset, consisting of 2703 unique patient records, from the data set SpecialPsych. Comprising 20% of the total data, this subset was selected using a random but stratified approach, ensuring balanced representation across diverse pandemic periods and hospitalization statuses. This stratified approach aims to maintain the holistic representativeness of the entire data set within our chosen sample. The choice of a 20% subset serves two primary purposes. First, it provides sufficient data for an in-depth expert evaluation. Second, it prevents overburdening the expert with excessive information, thereby preserving the integrity and quality of the evaluation process.

We provided the psychiatrist with the 2,703 patient records to acquire his evaluation. We shuffled the prepared data set to eliminate any potential order effects that might influence the training. The psychiatrist's assessment focused on each patient's potential need for hospitalization, as inferred from their triage notes. We also supplied the expert with two additional variables, the patients' ages and backgrounds, to facilitate a more contextual and informed evaluation. To uphold the independence of the expert's judgment, we consciously avoided intervening in this evaluation process and did not provide any more information like the visit dates.

In collaboration with the expert, we have converted his qualitative assessments into quantitative labels for further analysis and modeling. Thus, the expert's assessment is encoded into numerical labels: 1 for patients who should have been admitted, 0 for those who should have been discharged, and 0.5 for uncertain cases. Using this expert-labeled data set, we trained a machine-learning model to label the remaining 80% of the data set. We chose a set of BERT models for this purpose. Specifically, we fine-tuned the pre-trained models Bio Clinical BERT (Alsentzer et al. 2019) and PubMedBERT (Gu et al. 2020). We concluded PubMedBERT performed better so we chose this model to predict the admission likelihood of patients. This methodology yielded an AUC (Area Under the Curve) of approximately 75% in predicting patients' admission likelihood.

The text data undergoes preprocessing steps such as tokenization, conversion to lowercase, removal of stop words, and lemmatization, which are standard preprocessing steps in NLP tasks. This results in clean, concise, and meaningful text ready for subsequent analysis. The preprocessed notes are then tokenized and transformed into embeddings using the pre-trained PubMedBERT model. These embeddings, which are numerical representations of the notes, enable the ML model to understand and learn from the text data. Finally, a transformer model is built for sequence classification. This model is trained with the preprocessed text and additional features such as age, and the target variable is the expert's evaluation. The model uses a cross-entropy loss function suitable for multiclass classification tasks.

The model is designed as a multimodal transformer, merging the capabilities of the BERT model with a fully connected layer that processes additional numerical features. The BERT model generates embeddings from the triage notes, and these embeddings are combined with additional numerical features. This combined feature set is then passed through fully connected layers to output the final prediction. We calculate several evaluation metrics to monitor the model's performance, including accuracy, F1-score, precision, recall, and AUC. These metrics provide a comprehensive understanding of the model's performance, balancing the aspects of identifying correct and incorrect predictions, managing false positives and false negatives, and distinguishing between different classes. Table 2 shows the values of these different performance measures in training and fine-tuning the PubMedBERT model and predicting unseen data labels. We observe that the algorithm obtained 75% AUC in the training data set and 57% in the test data set.

Table 2 **Performance Metrics of Training the PubMedBERT on the Subsample Evaluated by the Psychiatrists.**

Data Set	Training Loss	Validation Loss	Accuracy	F1	Precision	Recall	AUC
Training	0.7706	0.8315	0.6580	0.4741	0.4389	0.5158	0.7448
Test		1.1925	0.6118	0.1518	0.1224	0.2000	0.5671

We label the remaining 80% patient records of the SpecialPsych by using the fine-tuned trained PubMedBERT algorithm. Furthermore, we use this algorithm on the GeneralPsych by using the chief complaints of patients. We note that the purpose here is to obtain a somewhat good prediction of admission likelihood. That is, we do not strive for the best possible prediction; instead, we aim to obtain a variable for estimating admission likelihood that is not contaminated with intervention or other confounding factors in our causal analysis.

6. Empirical Model

In this section, we outline the empirical model formulated to examine the influence of COVID-19 and government response policies on patient outcomes. The outcome of interest is the ED return time which is the interval between consecutive ED visits by patients. We also consider a mediator for the impact of the pandemic on the ED return time of psychiatric patients. The mediator is admission, or in other words, hospitalization decision.

6.1. Base Model

We use the DID framework to create our empirical model. DID is a statistical technique used in econometrics and epidemiology to estimate the effect of a specific intervention or treatment by accounting for time trends and potential biases across groups. DID methodology enables us to capture the confounding impacts of time trends and cross-sectional differences, thereby providing more accurate attribution of observed changes to the treatment under study.

We begin our analysis with a linear probability model that regresses ED return time on the Pandemic outbreak. Our base model is shown in Equation (1). Although this model is sometimes called DID in Operations literature, we acknowledge that this model essentially is a difference-in-means. Later in Section 6.2, we illustrate how this model turns into a DID.

$$\log(ED\_ReturnTime)_{ijt} = \beta_0 + \delta Pandemic_t + \mu_i + \nu_j + y_t + \gamma_t + \varepsilon_{ijt}, \tag{1}$$

where i indexes patients, j indexes hospitals (GeneralPsych or SpecialPsych), and t indexes time (regarding the specific ED visit). Although there is no overlap between hospital patients, as each psychiatric patient in Quebec can only be served by hospitals in her corresponding catchment area, we use the indexes i and j to differentiate between hospital records. That said, we introduce two levels of fixed effects: μ_i represents patient-specific fixed effects, and ν_j denotes hospital-specific fixed effects. Moreover, we introduce the year-month fixed effects (y_t for years and γ_t for months) to capture any seasonality and time-variant patterns. Finally, ε_{ijt} are the error terms. We cluster standard errors at the hospital-patient level to account for the potential intra-group correlation within hospital-patient groups. The outcome variable $\log(ED\_ReturnTime)$ is the log-transformed ED return time.

The coefficient δ and the variable $Pandemic_t$ denote two different specifications. In the first specification, $Pandemic_t$ is a single binary variable differentiating the periods before 14 March ($Pandemic_t = 0$) and after it ($Pandemic_t = 1$); δ is then a single coefficient estimating the DID estimator of the pandemic impact. In the second specification, $Pandemic_t$ is an array of three binary variables $Wave1_t$, $Wave2_t$, and $Wave3_t$ that take values as follows:

- $Wave1_t$ is 1 if t is between 14 March (inclusive) and 1 September, and 0 otherwise;
- $Wave2_t$ is 1 if t is between 1 September (inclusive) and 25 December, and 0 otherwise;
- $Wave3_t$ is 1 if t is after 25 December (inclusive), and 0 otherwise.

In this specification, δ is an array of δ_1, δ_2, and δ_3 that estimates the DID coefficients of the pandemic waves, accordingly. The first specification captures the pandemic outbreak effects, and the second specification captures the effects of the pandemic's waves. We contemplate the starting dates of these waves by the recorded events in the course of the pandemic in Quebec[6]. As a robustness check, we change the wave starting dates in Section 8.5.

The concept of "intervention" is not directly clear in the context of our research. While COVID-19 impacts everyone regardless of their mental health status, identifying an unaffected control group for the DID framework appears challenging. Consequently, the suitability of a DID design for the problem addressed in this paper comes into question. Alternative empirical approaches, such as interrupted time series, might be more qualified. Nonetheless, we have devised strategies to navigate this methodological challenge, targeting various assumptions of the DID framework. These will be elaborated on in the next sub-section.

[6] `https://en.wikipedia.org/wiki/Timeline_of_the_COVID-19_pandemic_in_Quebec`

6.2. Year-to-Year DID

We use the Y2Y-DID to solve the issue of having no unaffected control group. Cao et al. (2022) is the closest work to this study in the Operations literature that studied the impact of the pandemic on gig workers' behaviors. They followed Bandiera et al. (2005) who used this approach for a similar problem in a different context. We follow these works to build up the most appropriate empirical model as follows in Equation (2). This regression is a TWFE model that estimates the DID estimator of the impact of the pandemic on the ED return time:

$$\begin{aligned} \log(ED\_ReturnTime)_{ijt} = \beta_0 + \rho_t + \tau_t + \delta Pandemic\_placebo_t \times Y2020_t \\ + \mu_i + \nu_j + \gamma_t + \varepsilon_{ijt}, \end{aligned} \tag{2}$$

in which, $Pandemic\_placebo_t$ is 1 if t is after March 14 (regardless of the year) and 0 otherwise. ρ_t is the same as $Pandemic\_placebo_t$ to account for the fixed effect of intervention groups. $Y2020_t$ is a binary variable taking 1 if the t is in 2020 and 0 otherwise. τ_t is exactly similar and denotes the fixed effect of years.

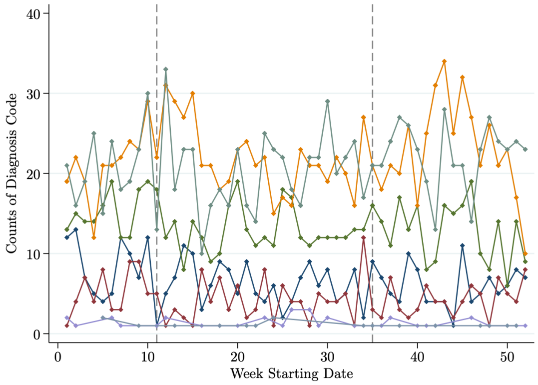
Note that in this model, we consider only the data from 2019 and 2020, i.e., we discard the data of ED visits in 2018 and 2021 because we do not have data of counterpart ED visits to compare with these visits. To this end, we ignore the pandemic's third wave which commenced on 25 December 2020 due to a small sample size. We consider the last week of December 2020 as part of the second wave. However, we perform a robustness check by specifying the last week of December 2020 as the pandemic's third wave in Section 8.3.

6.3. Patient characteristics

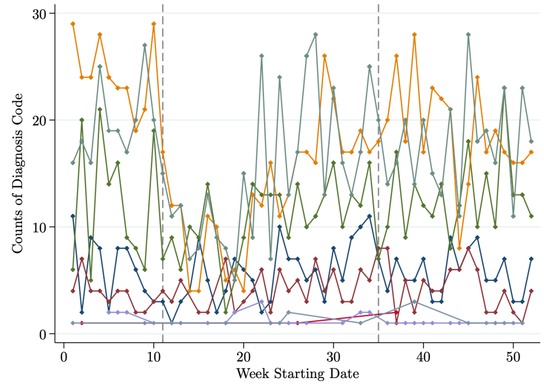
Note that our primary reference for Y2Y-DiD, Bandiera et al. (2005), studied panel data, i.e., they studied the same individuals across the years. The panel data has the advantage of studying a similar population of patients in different time stamps. This brings an identical distribution of patients across the periods. On the other hand, Cao et al. (2022) studied cross-sectional data, although they did not explicitly mention it. The cross-sectional data studies different sets of individuals across the years. Our study resembles more to (Cao et al. 2022) as the set of patients visiting the EDs after the pandemic outbreak is not completely the same set of patients seeking acute care in the EDs before the pandemic.

A critical part of employing DID in repeated cross-sections is to ensure that there is no compositional shift. As explained in Section 5.1.3, compositional shift causes the DID estimates to be biased by violating the parallel trends assumption. To this end, we should either ensure there is no compositional shift or show a conditional parallel trend exists. In this study, no compositional shift means the patient characteristics distributions remain the same across periods. To this end,

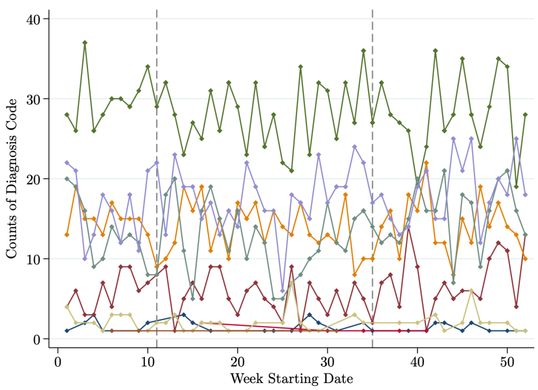
Figure 5 Weekly ED Visit Counts of Psychiatric Patients with Different Diagnoses in Years 2019-2020.



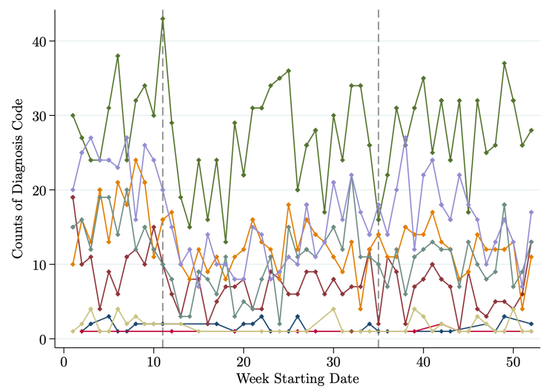
(a) GeneralPsych-2019



(b) GeneralPsych-2020



(c) SpecialPsych-2019



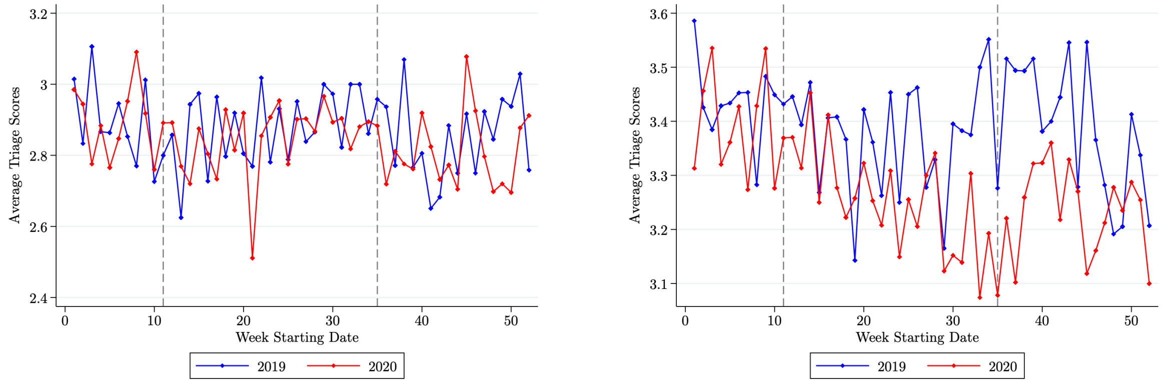
(d) SpecialPsych-2020

we should introduce patient controls to the base model given in Equation (1) to account for any discrepancies among the individuals in the sample.

These controls consist of patients' age, sex, marital status, severity of condition, and diagnosis codes. The diagnosis code classifies the patient's condition based on ICD-10-CM Codes. The triage score reflects the severity of the patient's condition based on the Canadian Triage and Acuity Scale (CTAS), with values ranging from 1 (Resuscitation) to 5 (Non-urgent). The gender indicates the patient's sex (male or female). Marital status encapsulates the patient's marital status, including categories such as single, married, divorced, and others.

Figure 5 shows the counts of daily ED visits averaged over weeks for each diagnosis class representing a category of mental illnesses. We observe that the frequency of ED visits within a certain type of disorder does not have a meaningful change over time, relative to other types of disorders. Figure 6 illustrates the average triage scores over time in 2019 and 2020. The trends in the GeneralPsych data are similar, but they diverge in the SpecialPsych sample. This may indicate that the composition of patients partially shifted during the pandemic period.

Figure 6 Average Triage Scores of Psychiatric Patients in Years 2019-2020 .



(a) GeneralPsych (b) SpecialPsych

We employ these control variables in our base regression model to control for specific patient characteristics. Moreover, we incorporate the patient's age and its squared term in the model. Note that they control for time-varying patterns in the population under study. The matrix variable $\boldsymbol{X}$ in the following equation represents the controls and $\boldsymbol{\theta}$ contains the coefficients.

$$\log(ED\_ReturnTime)_{ijt} = \beta_0 + \rho_t + \tau_t + \delta Pandemic\_placebo_t \times Y2020_t + \mu_i + \nu_j + \gamma_t + \boldsymbol{\theta X}_{ijt} + \varepsilon_{ijt}, \tag{3}$$

6.4. Mediation of Hospitalization

In the base model, we capture the impact of the pandemic on the ED return time. However, a fair question is how hospitalization could intervene in this relationship. A hospitalized patient receives the highest levels of acute care and treatment for the time she spends in the hospital. A discharged (not admitted) patient, on the other hand, should seek community care in case her problem is not solved completely in her ED visit. Thus, upon departing the ED, a psychiatric patient will pursue community care, or in case of emergency, seek another round of ED care.

The pandemic affects both the admission likelihood of psychiatric patients (Fasshauer et al. 2021a) and the provision of community care and social workers' services (Nicholas et al. 2023). To this end, we contemplate that the pandemic might impact the patient outcomes of both hospitalized and discharged patients, but in different ways. We account for these multifaceted impacts by introducing the admission variable as the mediator of the relationship between the COVID-19 pandemic and psychiatric patients' ED return time. The following set of equations represents this mediation framework.

$$\log(ED\_ReturnTime)_{ijt} = \beta_0^e + \rho_t^e + \tau_t^e + \delta^e Pandemic\_placebo_t \times Y2020_t + \alpha^e Admitted_{ijt} + \mu_i^e + \nu_j^e + \gamma_t^e + \theta^e \boldsymbol{X}_{ijt} + \varepsilon_{ijt}^e, \tag{4}$$

$$
\begin{aligned}
Admitted_{ijt} = {} & \beta_0^a + \rho_t^a + \tau_t^a + \delta^a Pandemic\_placebo_t \times Y2020_t \\
& + \mu_i^a + \nu_j^a + \gamma_t^a + \theta^a \boldsymbol{X}_{ijt} + \varepsilon_{ijt}^a .
\end{aligned} \tag{5}
$$

The coefficients $\beta_0, \rho, \tau, \delta, \mu, \nu, \gamma, \theta$ and the error term ε in Equation 4 are distinguished by the superscript e from those in Equation 5 which are superscripted by a. The variable *Admitted* is a binary indicator of whether a patient was hospitalized ($Admitted = 1$) or discharged ($Admitted = 0$) following their ED visit.

We choose the linear probability model (LPM) for estimating the admission likelihood for ease of coefficient interpretation. Wooldridge (2010) justifies this choice and claims LPM performs quite well in problems where the objective is to obtain partial effects. Wooldridge (2010) continues that using LPM becomes even more justified if most of the covariates are categorical variables with a limited number of values. This is the case in our setting in which all the explanatory variables are discrete except patients' age.

6.5. Instrumented DiD

We include hospital and patient fixed effects to account for time-invariant heterogeneity in the sample. We also incorporate patient characteristics to ensure conditional parallel trends assumption is satisfied. Moreover, we introduce time fixed effects to capture time-varying heterogeneity and seasonality across the years. We take all these measures to reduce the likelihood of biased estimates due to issues of the observational sample in which the sample is not randomly designed.

Another prevalent problem in a sample-based causal framework is endogeneity. To account for potential endogeneity problems in our model, we further extend our DID setup and incorporate an Instrumental Variable (IV) approach. The variable *Admitted*, which represents whether a patient was admitted or discharged, might correlate with unobserved factors influencing the ED return time. This correlation could bias our estimates of the effects of interest. We use the admission evaluations from Section 5.2 as an IV for the variable *Admitted* to mitigate this potential endogeneity.

The variable *ExpertEval* contains our expert's evaluation of whether the patient should have been hospitalized ($ExpertEval = 1$) or not ($ExpertEval = 0$), representing the expert opinion of the appropriate course of action. The expert sat in a quiet place in his office and made his decision on whether a patient in the SpecialPsych data set should be hospitalized or not. The ED doctors who make these decisions experience several pressures and restrictions, such as patient overcrowding in the ED, lack of stretchers, extensive ED boarding times for admission, and lack of beds in inpatient wards to accommodate patients waiting in the ED. Our expert did not experience any of these strains in his decision-making process. In other words, his evaluations are free of any operational and system-related unobserved factors. Moreover, the notes evaluated by the expert were randomly chosen from the SpecialPsych data set, and the expert was unaware whether a case he examined belonged to the pre-pandemic or the pandemic period. Therefore, $ExpertEval$ is presumed to be correlated with the admission decision but uncorrelated with the error term in the ED return time equation. By adopting this IV approach, the DID coefficients in equations (4-5) can be estimated by the conventional two-stage least squares (2SLS) (Ye et al. 2023). The following set of equations aims to provide unbiased estimates of the effect of admission and the pandemic on the ED return time by accounting for potential endogeneity.

$$\begin{aligned} \log(ED\_ReturnTime)_{ijt} &= \beta_0^e + \rho_t^e + \tau_t^e + \delta^e Pandemic\_placebo_t \times Y2020_t \\ &\quad + \alpha^e \widehat{Admitted}_{ijt} + \mu_i^e + \nu_j^e + \gamma_t^e + \theta^e \boldsymbol{X}_{ijt} + \varepsilon_{ijt}^e, \qquad (6) \\ Admitted_{ijt} &= \beta_0^a + \rho_t^a + \tau_t^a + \delta^a Pandemic\_placebo_t \times Y2020_t \\ &\quad + \alpha^a ExpertEval_{ijt} + \mu_i^a + \nu_j^a + \gamma_t^a + \theta^a \boldsymbol{X}_{ijt} + \varepsilon_{ijt}^a. \qquad (7) \end{aligned}$$

Note that we have the expert evaluation only for the SpecialPsych data set as it contains detailed triage notes. Unfortunately, we could not obtain the triage notes of patients' ED visits for the GeneralPsych data set. Therefore, we run the above instrumented Y2Y-DID model only for the SpecialPsych sample.

7. Results

This section presents the results of the models described in Section 6. We start with the base model and then illustrate the results of the extended models.

7.1. Primary Results

Table 3 shows the results of the base model given in Equation (1) using the whole three years of data. The first row in columns (1)-(3) illustrates that the pandemic reduces the ED return time. Columns (4)-(6) presents similar results for the pandemic waves. The first and second waves clearly decrease the ED return time, while the third wave has no significant effect on the patient outcome. The results remain the same either with patient fixed effects and controls or without them. Columns (2) and (5) show the results with patient fixed effects, and columns (3) and (6) show the results with patient characteristics controls. All the effects are significant at $p-value = 0.001$ level.

The above results show a difference in means rather than a difference in differences. As explained in Section 6.2, the Y2Y-DID model given in Equation (2) provides us with a DID analysis. This model utilizes patient records from the placebo periods in 2019 as the control group. Table 4 shows the results of this Y2Y-DID model. The coefficient of "Outbreak×Y2020" is the DID estimator of the effect of the pandemic on psychiatric patients' ED return times. The first row of columns (1)-(3) shows this DID estimator is negative, meaning the pandemic reduced the ED return time of psychiatric patients. Columns (4)-(6) illustrate the DID estimators of the pandemic waves.

Table 3 The Effects of the Pandemic on Patients' ED Return Time for the Whole Three Years of Data.

	Dependent Variable: ED Return Time					
	(1)	(2)	(3)	(4)	(5)	(6)
Pandemic Period	-0.360^{***}	-0.509^{***}	-0.503^{***}			
	(0.085)	(0.112)	(0.112)			
Pandemic Wave 1				-0.372^{***}	-0.476^{***}	-0.471^{***}
				(0.089)	(0.119)	(0.118)
Pandemic Wave 2				-0.353^{***}	-0.612^{***}	-0.606^{***}
				(0.103)	(0.132)	(0.132)
Pandemic Wave 3				-0.072	-0.110	-0.095
				(0.215)	(0.366)	(0.357)
Constant	5.272^{***}	4.153^{***}	3.757	5.283^{***}	4.117^{***}	3.663
	(0.094)	(0.126)	(2.721)	(0.098)	(0.128)	(2.723)
Observations	23131	23131	23131	23131	23131	23131
Patient FE	No	Yes	Yes	No	Yes	Yes
Controls	No	No	Yes	No	No	Yes

Standard errors are in parentheses.
FE denotes fixed effects. All regressions have hospital FE, year FE, and month FE.
$^{+}$ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

The coefficient of "Wave 1×Y2020" is the DID estimator of the effect of the pandemic's first wave on psychiatric patients' ED return times. The coefficient of the second wave has a similar interpretation. Both the first and second waves of the pandemic decrease the ED return time, however, the magnitude and significance of the second wave is lower than the first wave. Note that the fixed effect of Y2020 is significantly positive. This means the pandemic increases the ED return times when comparing visits in 2020 to visits in 2019. We intentionally include this effect in the result table to depict the complicated mechanisms behind the effects of the pandemic on the psychiatric patients' usage of acute ED care.

7.2. Hospitalization Mediation Effect

The mediation role of hospitalization in the relationship between the pandemic outbreak/waves and ED return time is one of the possible mechanisms that may explain why psychiatric patients return to the ED faster amid the pandemic outbreak. Table 5 illustrates the results of the mediation analysis formulated in equations (4-5). The top segment of the table presents results from regressing ED return time and the bottom segment shows results of regressing admission likelihood.

In the top segment of Table 5, columns (1) and (4) show that without patient fixed effects and controls, the impact of admission on ED return time is insignificant. However, the other columns illustrate that hospitalization increases the ED return time significantly when including patient fixed effects and controls. This suggests that admission has a positive impact on patient outcomes (reducing ED return time). On the other hand, the bottom segment of the table depicts that the pandemic outbreak/waves have no significant effects on admission after including patient fixed effects and controls. This may suggest that hospitalization does not mediate the relationship

Table 4 The Effects of the Pandemic on Patients' ED Return Time in the Y2Y-DID across 2019-2020.

	Dependent Variable: ED Return Time					
	(1)	(2)	(3)	(4)	(5)	(6)
Outbreak×Y2020	-0.267^{**}	-0.488^{***}	-0.499^{***}			
	(0.102)	(0.141)	(0.141)			
Wave 1×Y2020				-0.312^{**}	-0.532^{***}	-0.539^{***}
				(0.109)	(0.151)	(0.151)
Wave 2×Y2020				-0.210^{+}	-0.411^{**}	-0.428^{**}
				(0.115)	(0.154)	(0.156)
Y2020	0.290^{***}	1.306^{***}	1.210^{***}	0.290^{***}	1.299^{***}	1.205^{***}
	(0.086)	(0.127)	(0.157)	(0.086)	(0.126)	(0.157)
Constant	5.353^{***}	4.503^{***}	3.144	5.353^{***}	4.506^{***}	3.142
	(0.088)	(0.115)	(3.675)	(0.088)	(0.115)	(3.679)
Observations	15491	15491	15491	15491	15491	15491
Patient FE	No	Yes	Yes	No	Yes	Yes
Controls	No	No	Yes	No	No	Yes

Standard errors are in parentheses.
FE denotes fixed effects. All regressions have hospital FE and month FE.
Outbreak refers to the pandemic outbreak, and waves refer to the pandemic waves.
Y2020 is the fixed effect of the Year 2020.
$^{+}$ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

between the pandemic events and ED return time. However, we should be cautious in considering these results as they may suffer inaccuracy and biasedness.

A legitimate concern about including admission in our empirical model is the existence of unobserved variables affecting both admission likelihood and ED return time. As discussed in Section 6.5, we use the expert opinion on patients' admissions as the instrumental variable correcting for the potential endogeneity. Table 6 shows the results from the model given in equations (6-7).

Similarly to Table 5, the top segment of Table 6 presents the results from regressing ED return time, and the bottom segment shows the results of regressing admission likelihood. From the bottom segment of the table, it is evident that the expert's opinion on a patient's admission is highly correlated with the patient's actual admission, as its coefficients in all the specifications are positive and significant at the 0.001 level. Interestingly, columns (2) and (3) show that the pandemic outbreak reduces the admission likelihood significantly. Also, the first and second waves of the pandemic decrease the admission likelihood from columns (5)-(6).

The top segment of Table 6 reveals that hospitalization increases the ED return time in specifications with the patient fixed effects and controls. Thus, we conclude that hospitalization mediates the impact of the pandemic on ED return time. In other words, psychiatric patients' ED return times are shorter during the pandemic, possibly because their admission likelihood becomes lower amid the pandemic outbreak.

It's important to point out that these findings are derived solely from the SpecialPsych dataset, as the GeneralPsych dataset does not include triage notes. Consequently, this limitation results

Table 5 **The Effects of the Pandemic and Hospitalization on Patients' ED Return Time in 2019-2020.**

	(1)	(2)	(3)	(4)	(5)	(6)
	Dependent Variable: ED Return Time					
Admitted	-0.022	0.614^{***}	0.628^{***}	-0.023	0.614^{***}	0.628^{***}
	(0.044)	(0.073)	(0.075)	(0.044)	(0.073)	(0.075)
Outbreak×Y2020	-0.265^{**}	-0.485^{***}	-0.494^{***}			
	(0.102)	(0.141)	(0.142)			
Wave 1×Y2020				-0.310^{**}	-0.523^{***}	-0.530^{***}
				(0.109)	(0.151)	(0.151)
Wave 2×Y2020				-0.207^{+}	-0.418^{**}	-0.431^{**}
				(0.115)	(0.155)	(0.158)
Y2020	0.288^{***}	1.358^{***}	1.234^{***}	0.288^{***}	1.352^{***}	1.263^{***}
	(0.086)	(0.128)	(0.152)	(0.086)	(0.127)	(0.157)
Constant	5.363^{***}	4.297^{***}	2.865	5.364^{***}	4.300^{***}	2.863
	(0.094)	(0.119)	(3.677)	(0.094)	(0.119)	(3.681)
	Dependent Variable: Admission Likelihood					
Outbreak×Y2020	0.102^{***}	-0.004	-0.007			
	(0.019)	(0.034)	(0.032)			
Wave 1×Y2020				0.092^{***}	-0.013	-0.014
				(0.020)	(0.036)	(0.034)
Wave 2×Y2020				0.115^{***}	0.012	0.005
				(0.021)	(0.038)	(0.036)
Y2020	-0.104^{***}	-0.085^{**}	-0.091^{**}	-0.104^{***}	-0.086^{**}	-0.092^{**}
	(0.016)	(0.029)	(0.033)	(0.016)	(0.029)	(0.033)
Constant	0.449^{***}	0.336^{***}	0.444	0.449^{***}	0.337^{***}	0.444
	(0.016)	(0.028)	(0.749)	(0.016)	(0.028)	(0.749)
Observations	15491	15491	15491	15491	15491	15491
Patient FE	No	Yes	Yes	No	Yes	Yes
Controls	No	No	Yes	No	No	Yes

Standard errors are in parentheses.
FE denotes fixed effects. All regressions have hospital FE and month FE.
Outbreak refers to the pandemic outbreak, and waves refer to the pandemic waves.
Y2020 is the fixed effect of the Year 2020.
$^{+}$ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

in about a halved sample size, diminishing the statistical power of our analysis. This reduction in sample size could account for the relatively constrained effects of the pandemic outbreak and pandemic waves on ED return time observed in the upper segment of Table 6, especially when compared to the corresponding values in Table 5.

8. Robustness Checks

The empirical model that we have developed in this paper has a year-to-year difference-in-differences design. This setting encompasses several assumptions and specification choices. In this section, we investigate the robustness of our results regarding these assumptions.

Table 6 **The Effects of the Pandemic and Hospitalization on Patients' ED Return Time in the Y2Y-DID across 2019-2020, Corrected for Endogeneity in Admission.**

	(1)	(2)	(3)	(4)	(5)	(6)
	Dependent Variable: ED Return Time					
Admitted	−2.354***	0.956[a]	1.427+	−2.357***	0.955[a]	1.427+
	(0.233)	(0.604)	(0.803)	(0.234)	(0.606)	(0.803)
Outbreak×Y2020	−0.064	−0.282[b]	−0.225			
	(0.169)	(0.190)	(0.207)			
Wave 1×Y2020				−0.113	−0.285	−0.229
				(0.173)	(0.206)	(0.222)
Wave 2×Y2020				−0.001	−0.275	−0.218
				(0.193)	(0.199)	(0.216)
Y2020	−0.021	1.174***	1.137***	−0.021	1.174***	1.136***
	(0.136)	(0.152)	(0.197)	(0.136)	(0.152)	(0.197)
Constant	6.312***	3.920***	2.454	6.313***	3.921***	2.456
	(0.137)	(0.258)	(4.574)	(0.137)	(0.259)	(4.575)
	Dependent Variable: Admission Likelihood					
Admitted (Expert)	0.316***	0.144***	0.110***	0.316***	0.144***	0.110***
	(0.012)	(0.019)	(0.019)	(0.012)	(0.019)	(0.019)
Outbreak×Y2020	0.020	−0.131**	−0.127***			
	(0.026)	(0.042)	(0.039)			
Wave 1×Y2020				0.005	−0.150***	−0.142***
				(0.028)	(0.044)	(0.041)
Wave 2×Y2020				0.039	−0.097*	−0.100**
				(0.030)	(0.048)	(0.045)
Y2020	−0.075***	−0.005	−0.019	−0.075***	−0.008	−0.021
	(0.022)	(0.037)	(0.041)	(0.022)	(0.037)	(0.040)
Constant	0.228***	0.282***	0.505	0.228***	0.284***	0.516
	(0.020)	(0.036)	(0.861)	(0.020)	(0.036)	(0.859)
Observations	8716	8716	8716	8716	8716	8716
Patient FE	No	Yes	Yes	No	Yes	Yes
Controls	No	No	Yes	No	No	Yes

[a] p-value=0.11. [b] p-value=0.14.
Standard errors are in parentheses.
FE denotes fixed effects. All regressions have hospital FE and month FE.
Outbreak refers to the pandemic outbreak, and waves refer to the pandemic waves.
Y2020 is the fixed effect of the Year 2020.
$^{+}$ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

8.1. Week Fixed Effects

In the original models, we account for time-varying trends by including the year and month fixed effects. We choose months over weeks to avoid over-specification and/or introducing confusing fluctuations to the models. However, we check if our models are robust concerning the week fixed effects. The check reveals that our original results stay intact by replacing the month fixed effect with the week fixed effect. The result tables are not shown for the sake of paper length.

8.2. The COVID-19 Policies

We have delineated the effects of COVID-19 on psychiatric patient outcomes by referencing the pandemic's outbreak date and the start dates of its waves. This method contrasts records from before and after the pandemic to capture the shifts in patient outcome trends. This approach captures the discrepancy between different periods by ruling out time-related effects and patterns. The rationale is no deviation should have existed between the compared periods had the pandemic not occurred. Although this approach provides some insights, it does not consider the policies and measures taken at the state/regional levels, which could affect patient outcomes and health systems performance.

As a robustness check, we redefine the treatment group by replacing the pandemic outbreak date with pandemic severity measures. For this purpose, we utilize the Oxford COVID-19 Government Response Tracker (OxCGRT) dataset. The OxCGRT provides us with a plethora of indexes and measures related to the pandemic severity and the intensity of governments' policies. Among them, we choose the Stringency Index (SI) which shows the strictness of the government of Quebec's policies in response to the pandemic events. Also, we consider Government Response Index (GRI), Containment and Health Index (CHI), and Economic Support Index (ESI) to control for several types of the government's responses. Moreover, we use the cumulative numbers of confirmed infected cases and confirmed deaths due to COVID-19 to account for the pandemic severity on a continuous scale. The calculation details of these measures are given in the appendix. The full details are available in Hale et al. (June 2023).

Table 7 shows the impacts of the policy indexes and pandemic severity measures on the ED return time. The first (second) column contains the results of regressing ED return time (admission likelihood). We observe that higher Stringency Index results in lower ED return time, while, as the number of confirmed infected cases rise, the ED return time increases. A possible reason could be that rising number of cases make the patients afraid of getting infected, so they avoid visiting the ED, however, as the government responds to this rise by enacting stricter policies (so, higher Stringency Index), the patient's concern is lifted. That said, we conclude that our original results are robust as the pandemic reduces the ED return time of psychiatric patients.

8.3. Pandemic Wave 3 in Y2Y-DID

As explained in Section 6.2, we exclude Wave 3 from our analysis in the Y2Y-DUD, as this wave commences in the last week of December 2020. However, to ensure the robustness of our results, we include this wave in the analysis. As all the results remain the same, we conclude they are robust. For the sake of paper length, we do not provide the result tables.

Table 7 The Effects of the Government's Responses to the Pandemic on Patients' ED Return Time for the Whole Three Years of Data .

	Dependent Variable:	
	ED Return Time	Admission Likelihood
Admitted	0.583^{***}	
	(0.061)	
Stringency Index	-0.038^{*}	0.002
	(0.016)	(0.003)
Confirmed Cases	0.000005^{***}	0.0000001
	(0.0000008)	(0.0000001)
Confirmed Deaths	-0.0002^{***}	0.000007
	(0.00004)	(0.000009)
Constant	-8.440^{***}	1.401^{***}
	(2.075)	(0.460)
Observations	23131	23131

Standard errors are in parentheses.
FE denotes fixed effects. All regressions have hospital, patient, and month FE.
All regressions have controls for patient characteristics.
$^{+}$ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

8.4. Ward LOS

We propose hospitalization as the mediator of the impact of the COVID-19 pandemic on the ED return time. Note that hospitalization (or admission) is a binary variable. Here, we replace it with the Ward LOS as a robustness check. Ward LOS is 0 for discharged patients (Admitted=0) and is a positive value for hospitalized patients (Admitted=1). So, it gives us a version of hospitalization on a continuous scale. We re-run all the original models with Ward LOS replacing Admission. All the results remain the same, ensuring our model specifications are robust.

8.5. Outbreak and Wave Dates

We determine the pandemic outbreak and wave dates according to the official declarations of emergency states in Quebec. However, we acknowledge that the pandemic starts impacting the health system and patients on a slightly different timeline because usually, the governments' responses take place a little after the pandemic event occurs. That said, we deviate the pandemic outbreak and wave dates by two weeks earlier and two weeks later than the declared dates. The results remain the same, with the impact of the pandemic outbreak and waves on the ED return time becoming significant in the instrumented Y2Y-DID model. Note that these effects are insignificant in Table 6. We conclude that our results are completely robust in this respect. Tables of results are not shown for the sake of the paper's length.

8.6. TWFE Model with Covariates

As discussed in Section 5.1.3, we observe conditional PT exists in our sample. So, we develop a TWFE model with patient characteristics as covariates. However, incorporating time-varying covariates may introduce bias into our estimates. Here, we check the robustness of our results by employing two methods developed to resolve this issue. The first method is the Inverse Probability Weighting (IPW) (Abadie 2005) and the second method is Doubly Robust DID (DR-DID) (Sant'Anna & Zhao 2020). We employ both methods in our analysis and observe the results are similar to our original results in terms of direction (coefficients' signs) and significance. We conclude that our original TWEF model is not biased by including time-varying covariates.

8.7. Non-psychiatric Visits

Our analysis consists of data from two different hospitals: SpecialPsych and GeneralPsych. SpecialPsych is from a psychiatry hospital, so it contains only psychiatry visits. GeneralPsych is from a general hospital that serves a variety of patients with different illnesses. In this data set, we have considered only psychiatric patients who visited the ED for mental disorders. However, some of those patients returned to the ED for non-psychiatric visits. We did not count those visits in our original study. That said, the original results of this study are conservative estimates of the actual patient outcomes. Nevertheless, as a robustness check, we consider all the psychiatric and non-psychiatric returns of psychiatric patients in the GeneralPsych data set. The new results show no deviation from the original results. We conclude our results are robust regarding counting the non-psychiatric visits.

8.8. ED Return Time of Non-returning Patients

As we do not have information on non-returning patients, the ED return time is, in fact, a right-censored variable in our empirical model. In other words, the ED return time of a non-returning patient is missing. To solve this issue, we have considered non-returning patients as if they returned in 1097 days (3 years plus one day). In other words, we consider the log of ED return time of non-returning patients to be 7, which is the log of ≈ 1097.

We choose this value because the time horizon of our data set is three years. In this setting, all patients returned to the ED after a maximum of three years from their first visit. We believe this assumption does not attenuate the validity of our results as we have log-transformed the ED return time that minimizes the effect of large values by normalizing the distribution of the ED return time. However, as a robustness check, we check other limits like 900 days, 1500 days, and 1800 days. The original results remain unchanged in all these checks.

8.9. Placebo Tests

We have used a Y2Y-DID to identify the pandemic effect on ED return time. In this model, we have used placebo event dates to create control groups, as explained in Section 6.2. Moreover, we have controlled for month fixed effect to capture any month-level seasonalities. However, to make our model's identification more reliable, we follow Bandiera et al. (2005) by introducing a 2019 placebo outbreak date as a robustness check. The placebo outbreak occurs on 14 March 2019. We modify our model's identification by incorporating this placebo outbreak date along with the real pandemic outbreak date. The results show that the placebo outbreak in 2019 has no significant effect on the ED return time, while the actual outbreak in 2020 decreases the ED return time significantly. This placebo test ensures that our DID identification is robust.

9. Discussion

This research studies the dynamics of healthcare delivery during the COVID-19 pandemic, with a specific focus on ED return times and the mediation role of hospital admission. The results highlight the reduction in the ED return time amid the pandemic outbreak, suggesting that psychiatric patients generally returned to the ED quicker during the pandemic compared to the pre-pandemic period. Furthermore, the analysis reveals how this relationship evolved throughout the pandemic waves. On the other hand, hospitalization of psychiatric patients mediates the impact of the pandemic on ED return time. While hospitalization strongly increases the ED return time, the pandemic reduces the admission likelihood, suggesting that the pandemic made the health systems admit fewer patients in psychiatry wards. This results in shorter ED return times highlighting a possible mechanism for how the pandemic reduces psychiatric patients' ED return times.

By employing hospital, patient, year, and month fixed effects, as well as including patient characteristics controls and clustering standard errors at the hospital-patient level, we ensure the robustness of the results of our empirical models. Moreover, using an instrumental variable approach allows us to infer better causal relationships between the pandemic, admission likelihood, and ED return time. We utilize expert evaluations as a strong instrument for patient admission likelihood. Also, we perform robustness checks to investigate whether the results are vulnerable to changes in the specification assumptions. The checks show the robustness of our results.

We acknowledge the limitations of our data for generalizing the findings of this paper to other contexts and populations. Our data consists of only two hospitals in Montreal, Canada, which may not represent the whole country or other countries with different health systems, cultures, or epidemiological situations. We also focus on a specific group of psychiatric patients who seek acute care in emergency departments, which may not reflect the experiences and outcomes of other psychiatric patients who receive care in different settings or who do not seek care at all. Therefore, we caution against the applicability of these findings to other settings and populations.

We warn about the ethical implications of this paper for the mental health care of psychiatric patients during the pandemic, as we do not measure the actual mental health status or quality of life of the patients, nor the reasons why they sought or avoided care in EDs. We acknowledge that the pandemic and government policies could have negative effects on the mental health of psychiatric patients. They may have faced barriers or disincentives to access or receive care in EDs, such as fear of infection, stigma, discrimination, or lack of resources. That said, we do not draw normative conclusions from our findings to acknowledge the potential harms and benefits of the pandemic and the government policies for the mental health and well-being of psychiatric patients.

10. Conclusion

EDs are integral to delivering acute care, especially for patients experiencing psychiatric emergencies. However, EDs frequently struggle with high patient volumes, overcrowding, and extended lengths of stay, all of which can impede prompt and efficient care. To meet the escalating demand for ED care and alleviate pressure on hospital resources, several operational facets of EDs can be strategically optimized. Key among these are refining the triage processes, embracing telemedicine or telepsychiatry services for psychiatric consultations, augmenting communication and coordination among healthcare providers, and improving patient flow management. By concentrating on these areas, hospitals stand a better chance at fortifying their capacity to handle patient surges, subsequently improving the overall quality of care extended to patients with psychiatric disorders.

This study underscores the complexity and adaptability of healthcare systems in response to a major global crisis. The dynamics of patient admission and ED return times reveal how hospital procedures and patient severity interact in nuanced ways amidst an evolving pandemic. These findings have critical implications for healthcare management and policy, particularly in preparing for future public health emergencies. Future research should delve deeper into understanding these relationships and the underlying mechanisms to further refine strategies for optimal healthcare delivery during crises.

References

Abadie, A. (2005). Semiparametric difference-in-differences estimators. *The review of economic studies*, *72*(1), 1–19.

Alsaeed, D., Al-Ozairi, A., Alsarraf, H., Albarrak, F., & Al-Ozairi, E. (2023). Are we ready for the next pandemic? lessons learned from healthcare professionals' perspectives during the covid-19 pandemic. *Frontiers in Public Health*, *11*, 1048283.

Alsentzer, E., Murphy, J. R., Boag, W., Weng, W.-H., Jin, D., Naumann, T., & McDermott, M. (2019). Publicly available clinical bert embeddings. *arXiv preprint arXiv:1904.03323*.

Arkhangelsky, D., & Imbens, G. (2023). Causal models for longitudinal and panel data: A survey. Tech. rep., National Bureau of Economic Research.

Ashenfelter, O. C., & Card, D. (1984). Using the longitudinal structure of earnings to estimate the effect of training programs.

Bandiera, O., Barankay, I., & Rasul, I. (2005). Social preferences and the response to incentives: Evidence from personnel data. *The Quarterly Journal of Economics*, *120*(3), 917–962.

Breton, C., Han, J. Y., Sim, P., & Tabbara, M.-D. (2020). Covid-19 canadian provinces measures dataset. *Center of Excellence on the Canadian Federation*. Available at github.com/charlesbreton/COVID19-Canada-Provinces.

Brooks, S. K., Webster, R. K., Smith, L. E., Woodland, L., Wessely, S., Greenberg, N., & Rubin, G. J. (2020). The psychological impact of quarantine and how to reduce it: rapid review of the evidence. *The lancet*, *395*(10227), 912–920.

Bueno-Notivol, J., Gracia-García, P., Olaya, B., Lasheras, I., López-Antón, R., & Santabárbara, J. (2021). Prevalence of depression during the covid-19 outbreak: A meta-analysis of community-based studies. *International journal of clinical and health psychology*, *21*(1), 100–196.

Caniglia, E. C., & Murray, E. J. (2020). Difference-in-difference in the time of cholera: a gentle introduction for epidemiologists. *Current epidemiology reports*, *7*, 203–211.

Cao, X., Zhang, D., & Huang, L. (2022). The impact of the covid-19 pandemic on the behavior of online gig workers. *Manufacturing & Service Operations Management*, *24*(5), 2611–2628.

Carpiniello, B., Tusconi, M., Zanalda, E., Di Sciascio, G., & Di Giannantonio, M. (2020). Psychiatry during the covid-19 pandemic: a survey on mental health departments in italy. *BMC psychiatry*, *20*, 1–12.

Chen, W., Argon, N. T., Bohrmann, T., Linthicum, B., Lopiano, K., Mehrotra, A., Travers, D., & Ziya, S. (2022). Using hospital admission predictions at triage for improving patient length of stay in emergency departments. *Operations Research*.

Croskerry, P. (2012). Perspectives on diagnostic failure and patient safety. *Healthc Q*, *15*(Spec No), 50–56.

Cunnigham, S. (2021). Callaway and Sant'anna DD estimator. `https://causalinf.substack.com/p/callaway-and-santanna-dd-estimator`. Accessed: 2023-12-20.

Currie, J., Kleven, H., & Zwiers, E. (2020). Technology and big data are changing economics: Mining text to track methods. In *AEA Papers and Proceedings*, vol. 110, (pp. 42–48). American Economic Association 2014 Broadway, Suite 305, Nashville, TN 37203.

Deshpande, S., Marla, L., Scheller-Wolf, A., & Singh, S. P. (2022). Capacity management in a pandemic with endogenous patient choices and flows. *arXiv preprint arXiv:2207.05016*.

Duden, G. S., Gersdorf, S., & Stengler, K. (2022). Global impact of the covid-19 pandemic on mental health services: A systematic review. *Journal of Psychiatric Research*, *154*, 354–377.

Durand, A.-C., Gentile, S., Devictor, B., Palazzolo, S., Vignally, P., Gerbeaux, P., & Sambuc, R. (2011). Ed patients: how nonurgent are they? systematic review of the emergency medicine literature. *The American journal of emergency medicine*, *29*(3), 333–345.

Fasshauer, J. M., Bollmann, A., Hohenstein, S., Hindricks, G., Meier-Hellmann, A., Kuhlen, R., Broocks, A., Schomerus, G., & Stengler, K. (2021a). Emergency hospital admissions for psychiatric disorders in a german-wide hospital network during the covid-19 outbreak. *Social psychiatry and psychiatric epidemiology*, *56*, 1469–1475.

Fasshauer, J. M., Bollmann, A., Hohenstein, S., Mouratis, K., Hindricks, G., Meier-Hellmann, A., Kuhlen, R., Broocks, A., Schomerus, G., & Stengler, K. (2021b). Impact of covid-19 pandemic on involuntary and urgent inpatient admissions for psychiatric disorders in a german-wide hospital network. *Journal of psychiatric research*, *142*, 140–143.

Feizi, A., Tucker, A., & Baker, W. (2022). Impact of emergency department workload on triage behavior. *Available at SSRN 4071155*.

Gaida, F. (2023). *Factors facilitating or hindering the effectiveness of emergency departments in their responses to the needs of major users of these mental health services*. Ph.D. thesis, University of Montreal.

Goldenberg, M. N., & Parwani, V. (2021). Psychiatric emergency department volume during covid-19 pandemic. *The American Journal of Emergency Medicine*, *41*, 233.

Gu, Y., Tinn, R., Cheng, H., Lucas, M., Usuyama, N., Liu, X., Naumann, T., Gao, J., & Poon, H. (2020). Domain-specific language model pretraining for biomedical natural language processing.

Hale, T., Angrist, N., Goldszmidt, R., Kira, B., Petherick, A., Phillips, T., Webster, S., Cameron-Blake, E., Hallas, L., Majumdar, S., & Tatlow, H. (2021). A global panel database of pandemic policies (oxford covid-19 government response tracker). *Nature human behaviour*, *5*(4), 529–538.
URL `https://doi.org/10.1038/s41562-021-01079-8`

Hale, T., Petherick, A., Phillips, T., Anania, J., de Mello, B. A., Angrist, N., Barnes, R., Boby, T., Cameron-Blake, E., Cavalieri, A., Di Folco, M., Edwards, B., Ellen, L., Elms, J., Furst, R., Ribeiro, L. G., Green, K., Goldszmidt, R., Hallas, L., Kamenkovich, N., Kira, B., Laping, S., Luciano, M., Majumdar, S., Oliveira, T. M., Nagesh, A., Radhika Pott, Ren, L., Sampaio, J., Tatlow, H., Torness, W., Wade, A., Webster, S., Wood, A., Zha, H., Zhang, Y., & Vaccaro, A. (June 2023). Variation in government responses to covid-19. *Blavatnik School of Government Working Paper*, *Version 15.*.
URL `www.bsg.ox.ac.uk/covidtracker`

Heckman, J. J., Ichimura, H., & Todd, P. E. (1997). Matching as an econometric evaluation estimator: Evidence from evaluating a job training programme. *The review of economic studies*, *64*(4), 605–654.

Hejazian, H., Kucukyazici, B., Nasiry, J., Verter, V., & Frank, D. (2023). The impact of hospital and patient characteristics on psychiatry readmissions. *Available at SSRN 4438262*.

Hollander, J. E., & Carr, B. G. (2020). Virtually perfect? telemedicine for covid-19. *New England Journal of Medicine*, *382*(18), 1679–1681.

Hong, W. S., Haimovich, A. D., & Taylor, R. A. (2018). Predicting hospital admission at emergency department triage using machine learning. *PloS one*, *13*(7), e0201016.

Hoot, N. R., & Aronsky, D. (2008). Systematic review of emergency department crowding: causes, effects, and solutions. *Annals of emergency medicine*, *52*(2), 126–136.

Huang, J.-A., Tsai, W.-C., Chen, Y.-C., Hu, W.-H., & Yang, D.-Y. (2003). Factors associated with frequent use of emergency services in a medical center. *Journal-Formosan Medical Association*, *102*(4), 222–228.

Johnson, S., Dalton-Locke, C., Vera San Juan, N., Foye, U., Oram, S., Papamichail, A., Landau, S., Rowan Olive, R., Jeynes, T., Shah, P., et al. (2021). Impact on mental health care and on mental health service users of the covid-19 pandemic: a mixed methods survey of uk mental health care staff. *Social psychiatry and psychiatric epidemiology*, *56*, 25–37.

Jurcik, T., Jarvis, G. E., Zeleskov Doric, J., Krasavtseva, Y., Yaltonskaya, A., Ogiwara, K., Sasaki, J., Dubois, S., & Grigoryan, K. (2021). Adapting mental health services to the covid-19 pandemic: reflections from professionals in four countries. *Counselling Psychology Quarterly*, *34*(3-4), 649–675.

Koekkoek, B., van Meijel, B., & Hutschemaekers, G. (2006). " difficult patients" in mental health care: a review. *Psychiatric Services*, *57*(6), 795–802.

Krishnamoorthy, Y., Nagarajan, R., Saya, G. K., & Menon, V. (2020). Prevalence of psychological morbidities among general population, healthcare workers and covid-19 patients amidst the covid-19 pandemic: A systematic review and meta-analysis. *Psychiatry research*, *293*, 113382.

Kupcova, I., Danisovic, L., Klein, M., & Harsanyi, S. (2023). Effects of the covid-19 pandemic on mental health, anxiety, and depression. *BMC psychology*, *11*(1), 1–7.

McGuire, A. B., Flanagan, M. E., Kukla, M., Rollins, A. L., Myers, L. J., Bass, E., Garabrant, J. M., & Salyers, M. P. (2021). Inpatient mental healthcare before and during the covid-19 pandemic. In *Healthcare*, vol. 9, (p. 1613). MDPI.

Nair, S., Kannan, P., Mehta, K., Raju, A., Mathew, J., & Ramachandran, P. (2021). The covid-19 pandemic and its impact on mental health services: the provider perspective. *Journal of Public Health*, *43*(Supplement_2), ii51–ii56.

Nicholas, D. B., Samson, P., Hilsen, L., & McFarlane, J. (2023). Examining the covid-19 pandemic and its impact on social work in health care. *Journal of Social Work*, *23*(2), 334–349.

Ornell, F., Borelli, W. V., Benzano, D., Schuch, J. B., Moura, H. F., Sordi, A. O., Kessler, F. H. P., Scherer, J. N., & von Diemen, L. (2021). The next pandemic: impact of covid-19 in mental healthcare assistance in a nationwide epidemiological study. *The Lancet Regional Health–Americas*, *4*.

Parker, C. A., Liu, N., Wu, S. X., Shen, Y., Lam, S. S. W., & Ong, M. E. H. (2019). Predicting hospital admission at the emergency department triage: A novel prediction model. *The American journal of emergency medicine*, *37*(8), 1498–1504.

Phillips, C. (2020). Relationships between workload perception, burnout, and intent to leave among medical–surgical nurses. *JBI Evidence Implementation*, *18*(2), 265–273.

Roncero, C., García-Ullán, L., de la Iglesia-Larrad, J. I., Martín, C., Andrés, P., Ojeda, A., González-Parra, D., Pérez, J., Fombellida, C., Álvarez-Navares, A., et al. (2020). The response of the mental health network of the salamanca area to the covid-19 pandemic: The role of the telemedicine. *Psychiatry research*, *291*, 113252.

Roy, D., Tripathy, S., Kar, S. K., Sharma, N., Verma, S. K., & Kaushal, V. (2020). Study of knowledge, attitude, anxiety & perceived mental healthcare need in indian population during covid-19 pandemic. *Asian journal of psychiatry*, *51*, 102083.

Sabbatini, A. K., Kocher, K. E., Basu, A., & Hsia, R. Y. (2016). In-hospital outcomes and costs among patients hospitalized during a return visit to the emergency department. *Jama*, *315*(7), 663–671.

Saghafian, S., Hopp, W. J., Van Oyen, M. P., Desmond, J. S., & Kronick, S. L. (2012). Patient streaming as a mechanism for improving responsiveness in emergency departments. *Operations Research*, *60*(5), 1080–1097.

Saghafian, S., Hopp, W. J., Van Oyen, M. P., Desmond, J. S., & Kronick, S. L. (2014). Complexity-augmented triage: A tool for improving patient safety and operational efficiency. *Manufacturing & Service Operations Management*, *16*(3), 329–345.

Sant'Anna, P. H., & Zhao, J. (2020). Doubly robust difference-in-differences estimators. *Journal of Econometrics*, *219*(1), 101–122.

Shaw, H., Robertson, S., & Ranceva, N. (2021). What was the impact of a global pandemic (covid-19) lockdown period on experiences within an eating disorder service? a service evaluation of the views of patients, parents/carers and staff. *Journal of Eating Disorders*, *9*, 1–11.

Shields, M., Tonmyr, L., Gonzalez, A., Weeks, M., Park, S.-B., Robert, A.-M., Blair, D.-L., & MacMillan, H. L. (2021). Symptoms of major depressive disorder during the covid-19 pandemic: results from a representative sample of the canadian population. *Health Promotion & Chronic Disease Prevention in Canada: Research, Policy & Practice*, *41*(11).

Shobassy, A., Nordsletten, A., Ali, A., Bozada, K., Malas, N., & Hong, V. (2022). Effects of the covid-19 pandemic in a psychiatric emergency service: utilization patterns and patient perceptions. *The American Journal of Emergency Medicine*, *52*, 241.

Sibeoni, J., Manolios, E., Costa-Drolon, E., Meunier, J.-P., Verneuil, L., & Revah-Levy, A. (2021). Clinical practice during the covid-19 pandemic: a qualitative study among child and adolescent psychiatrists across the world. *Child and adolescent psychiatry and mental health*, *15*, 1–15.

Singh, S., Lin, Y.-L., Nattinger, A. B., Kuo, Y.-F., & Goodwin, J. S. (2015). Variation in readmission rates by emergency departments and emergency department providers caring for patients after discharge. *Journal of hospital medicine*, *10*(11), 705–710.

Vaičekauskaitė, R., Babarskienė, J., & Grubliauskienė, J. (2021). Mental health services during the pandemic: the experiences of lithuanian psychologists. In *SOCIETY. INTEGRATION. EDUCATION. Proceedings of the International Scientific Conference*, vol. 7, (pp. 217–229).

Warner, L. S. H., Galarraga, J. E., Litvak, O., Davis, S., Granovsky, M., & Pines, J. M. (2018). The impact of hospital and patient factors on the emergency department decision to admit. *The Journal of emergency medicine*, *54*(2), 249–257.

Weibelzahl, S., Reiter, J., & Duden, G. (2021). Depression and anxiety in healthcare professionals during the covid-19 pandemic. *Epidemiology & Infection*, *149*, e46.

Wooldridge, J. M. (2010). *Econometric analysis of cross section and panel data*. MIT press.

Wowak, K. D., Lalor, J. P., Somanchi, S., & Angst, C. M. (2023). Business analytics in healthcare: Past, present, and future trends. *Manufacturing & Service Operations Management*.

Xiong, J., Lipsitz, O., Nasri, F., Lui, L. M., Gill, H., Phan, L., Chen-Li, D., Iacobucci, M., Ho, R., Majeed, A., et al. (2020). Impact of covid-19 pandemic on mental health in the general population: A systematic review. *Journal of affective disorders*, *277*, 55–64.

Yalçın, M., Baş, A., Bilici, R., Özdemir, Y. Ö., Beştepe, E. E., Kurnaz, S., Güneş, M., Kurt, V., & Koyu, E. (2021). Psychiatric emergency visit trends and characteristics in a mental health epicenter in istanbul during covid-19 lockdown. *Social psychiatry and psychiatric epidemiology*, *56*, 2299–2310.

Yang, H., Bin, P., & He, A. J. (2020). Opinions from the epicenter: an online survey of university students in wuhan amidst the covid-19 outbreak1. *Journal of Chinese Governance*, *5*(2), 234–248.

Ye, T., Ertefaie, A., Flory, J., Hennessy, S., & Small, D. S. (2023). Instrumented difference-in-differences. *Biometrics*, *79*(2), 569–581.

Online Appendix to "The Weakest Goes to the Wall: The Impact of the Covid-19 Pandemic on Psychiatric Patients"

Appendix EC.1: COVID-19 Policy Indexes

We explain the construct and calculations of the COVID-19 policy indices and measures. This section is an exact copy of the working paper Hale et al. (June 2023) that is written by the creators of the OxCGRT data set. The OxCGRT calculates several indices to give an overall impression of government activity: the Government Response Index, the Containment and Health Index, the Stringency Index, and the Economic Support Index. These indexes are different combinations of policies, as shown in Table EC.1.

Table EC.1 COVID-19 Indices Breakdown.

Policies	Government Response Index	Containment and Health Index	Stringency Index	Economic Support Index
C1	×	×	×	
C2	×	×	×	
C3	×	×	×	
C4	×	×	×	
C5	×	×	×	
C6	×	×	×	
C7	×	×	×	
C8	×	×	×	
E1	×			×
E2	×			×
H1	×	×	×	
H2	×	×		
H3	×	×		
H6	×	×		
H7	×	×		
H8	×	×		

The policies in Table EC.1 are from three categories of policies: containment and closure policies (C), economic policies (E), and health system policies (H). The details of each policy are as follows.

EC.1.1. C - Containment and Closure Policies

- C1 - School Closing (was in place since 10 March 2020 until the end of horizon):

—0 - no measures,

—1 - recommend closing, or all schools open with alterations resulting in significant differences compared to non-Covid-19 operations,

—2 - require closing (only some levels or categories, e.g. just high school or just public schools)

—3 - require closing all levels.

- C2 - Workplace closing (was in place since 11 March 2020 until the end of horizon):

—0 - no measures,

—1 - recommend closing (or recommend work from home), or all businesses open with alterations resulting in significant differences compared to non-Covid-19 operation,

—2 - require closing (or work from home) for some sectors or categories of workers,

—3 - require closing (or work from home) for all-but-essential workplaces (e.g. grocery stores, doctors).

- C3 - Cancel public events (was in place since 11 March 2020 until the end of horizon):

—0 - No measures,

—1 - Recommend canceling,

—2 - Require canceling.

- C4: Restrictions on gatherings (was in place since 12 March 2020 until the end of horizon):

—0 - no restrictions,

—1 - restrictions on very large gatherings (the limit is above 1000 people),

—2 - restrictions on gatherings between 101-1000 people,

—3 - restrictions on gatherings between 11-100 people,

—4 - restrictions on gatherings of 10 people or less.

- C5 - Close public transport (was not in place at all, so we can simply ignore this indicator):

—0 - No measures,

—1 - Recommend closing (or significantly reduce volume/route/means of transport available),

—2 - Require closing (or prohibit most citizens from using it).

- C6 - Stay at home requirements (was in place since 16 March 2020 until the end of horizon):

—0 - no measures,

—1 - recommend not leaving house,

—2 - require not leaving house with exceptions for daily exercise, grocery shopping, and 'essential' trips,

—3 - require not leaving house with minimal exceptions (e.g. allowed to leave once a week, or only one person can leave at a time, etc.)

- C7 - Restrictions on internal movement (was in place since 19 March 2020 until the end of horizon, but vaccinated people got exemption in the last few months, i.e., after May 2021):

—0 - No measures,

—1 - Recommend not to travel between regions/cities,

—2 – internal movement restrictions in place.

- C8 -International travel controls (was in place since 11 March 2020 until the end of horizon):

—0 - no restrictions,

—1 - screening arrivals,

—2 - quarantine arrivals from some or all regions,

—3 - ban arrivals from some regions,

—4 - ban on all regions or total border closure.

EC.1.2. E - Economic Policies

- E1 - Income support (for households) (was in place since 15 March 2020 until the end of horizon):

—0 - no income support,

—1 - the government is replacing less than 50% of lost salary (or if a flat sum, it is less than 50% median salary),

—2 - the government is replacing 50% or more of lost salary (or if a flat sum, it is greater than 50% median salary).

- E2 - Debt/contract relief (for households) (was in place since 17 March 2020 until the end of horizon):

—0 - no debt/contract relief,

—1 - narrow relief, specific to one kind of contract,

—2 - broad debt/contract relief.

- E3 - Fiscal measures (was in place since 19 March 2020 until the end of horizon):

—Numerical variable,

—Record monetary value in USD of fiscal stimuli, includes any spending or tax cuts NOT included in E4, H4 or H5,

—0 - no new spending that day.

- E4 - Giving international support (was in place since 7 July 2021 until the end of horizon, i.e., for two months):

—Numerical variable,

—Record monetary value in USD,

—0 - no new spending that day.

EC.1.3. H - Health System Policies

- H1 - Public information campaigns (was in place since 20 January 2020 until the end of horizon):

—0 - no Covid-19 public information campaign,

—1 - public officials urging caution about Covid-19,

—2- coordinated public information campaign (e.g. across traditional and social media).

- H2 - Testing policy (was in place since 9 March 2020 until the end of horizon):

—0 - no testing policy,

—1 - only those who both (a) have symptoms AND (b) meet specific criteria (e.g. key workers, admitted to hospital, came into contact with a known case, returned from overseas),

—2 - testing of anyone showing COVID-19 symptoms,

—3 - open public testing (e.g. "drive through" testing available to asymptomatic people).

- H3 - Contact tracing (was in place since 5 October 2020 until the end of horizon):

—0 - no contact tracing,

—1 - limited contact tracing; not done for all cases,

—2 - comprehensive contact tracing; done for all identified cases.

- H4 - Emergency investment in healthcare (was in place since 30 March 2020 until the end of horizon):

—Numerical variable,

—Record monetary value in USD,

—0 - no new spending that day.

- H5 - Investment in vaccines (only on March 21, however, note that it was only from the government of Quebec):

—Numerical variable,

—Record monetary value in USD,

—0 - no new spending that day.

- H6 - Facial Coverings (was in place since 20 April 2020 until the end of horizon):

—0 - No policy,

—1 - Recommended,

—2 - Required in some specified shared/public spaces outside the home with other people present or some situations when social distancing is not possible,

—3 - Required in all shared/public spaces outside the home with other people present or all situations when social distancing is not possible,

—4 - Required outside the home at all times regardless of location or presence of other people.

- H7 - Vaccination policy (was in place since 14 December 2020 until the end of horizon):

—0 - No availability,

—1 - Availability for ONE of following: key workers/ clinically vulnerable groups (non elderly) / elderly groups,

—2 - Availability for TWO of following: key workers/ clinically vulnerable groups (non elderly) / elderly groups,

—3 - Availability for ALL of following: key workers/ clinically vulnerable groups (non elderly) / elderly groups,

—4 - Availability for all three plus partial additional availability (select broad groups/ages),

— 5 - Universal availability.

- H8 - Protection of elderly people (was in place since 12 March 2020 until the end of horizon):

— 0 - no measures,

— 1 - Recommended isolation, hygiene, and visitor restriction measures in LTCF[7]s and/or elderly people to stay at home,

— 2 - Narrow restrictions for isolation, hygiene in LTCFs, some limitations on external visitors and/or restrictions protecting elderly people at home,

— 3 - Extensive restrictions for isolation and hygiene in LTCFs, all non-essential external visitors prohibited, and/or all elderly people required to stay at home and not leave the home with minimal exceptions, and receive no external visitors.

EC.1.4. V - Vaccination Policies

- V1 - Vaccine prioritisation (was in place since 7 December 2020 until the end of horizon):

— 0 – no plan,

— 1 – a prioritised plan is in place,

— 2 – universal/general eligibility; no prioritisation between groups.

- V2A - Vaccine eligibility/availability (was in place since 14 December 2020 until the end of horizon):

— 0 – no categories are receiving vaccines,

— 1 – vaccines are available to some categories,

— 2 – vaccines are available to anyone over the age of 16 years,

— 3 – vaccines are available to anyone over the age of 16 years PLUS one or both of 5-15 years and 0-4 years.

- V3 - Vaccine financial support (was in place since 14 December 2020 until the end of horizon):

— 0 – no availability,

— 1 – full cost to the individual for all categories identified in V2,

— 2 – full cost to the individual for some categories identified in V2, some subsidy for other categories,

— 3 - partial funding by the government for all of the categories identified in V2,

— 4 – partial funding by the government for some categories identified in V2, full funding for other categories,

— 5 – all categories fully funded by the government.

All of the above indices use ordinal indicators where policies are ranked on a simple numerical scale. In order to aggregate these indicators into an index, they first calculate a sub-index score

[7] Long Term Care Facility

that normalizes each of the indicators. Some indicators (C1-C7, E1, H1, H6, H7, and H8) have an additional binary flag variable that can be either 0 or 1. The codebook above has details about each indicator and what the different values represent. Because different indicators (j) have different maximum values (N_j) in their ordinal scales, and only some have flag variables, each sub-index score must be calculated separately. The different indicators that contribute to the indices are:

Table EC.2 Policy Indicators Contributing to the COVID-19 Indices

Indicator	Max value (N_j)	Flag? (F_j)
C1	3 (0, 1, 2, 3)	Yes=1
C2	3 (0, 1, 2, 3)	Yes=1
C3	2 (0, 1, 2)	Yes=1
C4	4 (0, 1, 2, 3, 4)	Yes=1
C5	2 (0, 1, 2)	Yes=1
C6	3 (0, 1, 2, 3)	Yes=1
C7	2 (0, 1, 2)	Yes=1
C8	4 (0, 1, 2, 3, 4)	No=0
E1	2 (0, 1, 2)	Yes=1
E2	2 (0, 1, 2)	No=0
H1	2 (0, 1, 2)	Yes=1
H2	3 (0, 1, 2, 3)	No=0
H3	2 (0, 1, 2)	No=0
H6	4 (0, 1, 2, 3, 4)	Yes=1
H7	5 (0, 1, 2, 3, 4, 5)	Yes=1
H8	3 (0, 1, 2, 3)	Yes=1

Each sub-index score (I) for any given indicator (j) on any given day (t), is calculated by the function described in equation 1 based on the following parameters:

- the maximum value of the indicator (N_j)
- whether that indicator has a flag ($F_j = 1$ if the indicator has a flag variable, or 0 if the indicator does not have a flag variable)
- the recorded policy value on the ordinal scale ($v_{j,t}$)
- the recorded binary flag for that indicator, if that indicator has a flag ($f_{j,t}$)

$$I_{j,t} = 100\frac{v_{j,t} - 0.5(F_j - f_{j,t})}{N_j}$$

This normalizes the different ordinal scales to produce a sub-index score between 0 and 100, where each full point on the ordinal scale is equally spaced. For indicators that have a flag variable, if this flag is recorded as 0 (i.e. if the policy is geographically targeted or for E1 if the support only applies to informal sector workers) then this is treated as a half-step between ordinal values.

Note that the database only contains flag values if the indicator has a non-zero value. If a government has no policy for a given indicator (i.e. the indicator equals zero) then the corresponding flag is blank/null in the database. For the purposes of calculating the index, this is equivalent to a sub-index score of zero. In other words, $I_{j,t} = 0$ if $v_{j,t} = 0$.

For a given jurisdiction, their non-vaccinated and vaccinated indices are simple averages of the individual component indicators for each group. Then, the index is calculated as follows, where k is the number of component indicators in an index, and I_j is the sub-index score for an individual indicator.

$$index = \frac{1}{k}\sum_{j=1}^{k} I_j$$