AJPM FOCUS
INCLUSIVITY IN PEOPLE, METHODS, AND OUTCOMES
RESEARCH ARTICLE

Effect of Telemedicine Use on Medical Spending and Health Care Utilization: A Machine Learning Approach

Ayesha Jamal, PhD

Introduction: This study analyzes the effect of telemedicine use on healthcare utilization and medical spending for patients with chronic mental illness.

Methods: Using the IBM MarketScan Research database from 2009 to 2018, this study examined the timing of users' first telemedicine use and identified similar periods for non-users by using random forest and random forest proximity matching. A difference-in-differences approach, which tests whether there are differences in the study outcomes before and after the actual/predicted first use among the treated group (users) compared with the control group (non-users), was then used to assess the impact of telemedicine. Analyses were done in 2021.

Results: Comparing users with non-users after matching suggested that telemedicine use both increases the number of overall outpatient visits (0.461; 95% CI=0.280, 0.642; $p<0.001$) related to psychotherapy and evaluation and management services, and decreases the number of in-person visits (0.280; 95% CI= −0.446, −0.114; $p=0.001$) for patients with chronic mental health diagnoses. Total medical spending was not significantly affected. Additionally, no evidence was found of telemedicine use being associated with an increased probability of an emergency department visit or hospitalization.

Conclusions: The study findings suggest that telemedicine use is associated with an increase in outpatient care utilization for patients with chronic mental health diagnoses. No substantive changes in medical spending, the probability of an emergency department visit, or the probability of hospitalization were noted. Results provide insights into the effect of telemedicine use on spending and healthcare utilization for patients with chronic mental illness. These findings may inform research to guide future telemedicine policies and interventions.

INTRODUCTION

The mental health crisis is severe in the U.S., where 1 in 5 adults experience mental illness each year and 1 in 25 adults live with a serious mental illness.[1] Compared with the general population, people with mental illness have greater physical health morbidity and mortality, and those with serious mental illness are likely to die 10–25 years

From the Arthur J. Bauernfeind College of Business, Department of Economics and Finance, Murray State University Murray, Kentucky

Address correspondence to: Ayesha Jamal, PhD, Department of Economics and Finance, Murray State University, 307H Business Building, Murray KY 42071. E-mail: ajamal1@murraystate.edu.
2773-0654/$36.00
https://doi.org/10.1016/j.focus.2023.100127

earlier, on average.[2–4] Access issues and unmet patient needs have further exacerbated the severity of the crisis. More than 150 million people in the U.S. live in a mental health professional shortage area, and nearly one fifth of adults with mental illness report an unmet need for mental health treatment.[5,6]

Telemedicine has emerged as a promising solution to bridge at least some of the gaps in access. It offers certain advantages to the patient; for example, virtual visits tend to be cheaper than in-person visits, reduce travel-related expenses, allow remote access to distant locations, and reduce anxiety and stigma associated with in-person visits.[7,8] Additionally, in terms of treatment method, compared with other illnesses, it is easier to substitute in-person care with virtual care for patients with chronic mental illness. Providers mostly use a combination of medication and therapy to treat patients with chronic mental illness, both of which can be prescribed and delivered remotely. Telemedicine offers potential cost savings to the overall healthcare system as well, in that it can shift patients from more expensive settings such as emergency departments (EDs) to less costly office visits, which patients with mental illness tend to underutilize.[9–11]

Because of several state and federal policy changes, which temporarily removed barriers to access during the pandemic, telemedicine utilization increased overall.[12,13] It is no surprise that when telemedicine use peaked at the beginning of the coronavirus disease 2019 (COVID-19) pandemic, between March and August of 2020, 40% of all telemedicine visits were for mental health and substance use disorders.[14] However, these policy changes are not yet permanent, as there are concerns regarding the potentially negative consequences. These concerns include that telemedicine utilization may encourage excessive use and spending or provide subpar quality of care.[15,16]

Numerous randomized clinical trials have demonstrated that telemedicine and in-person visits are comparable qualitatively in treating patients with mental illness.[17–19] However, outside of clinical trials, evidence for patients with mental illness is limited, with research focused primarily on Medicare and Medicaid beneficiaries.[20,21]

This study expands on existing literature by examining the privately insured patients with chronic mental illness in a retrospective study. This study uses private insurance claims data to understand when patients start using telemedicine, and it uses a machine learning method combined with a difference-in-differences analysis to assess the impact of telemedicine on spending and healthcare use. This study aims to test whether telemedicine use is associated with increased engagement (measured by the total number of psychotherapy and evaluation and management visits), increased total medical or pharmaceutical spending, and reduced quality of care (specifically, an increase in the probability of an ED visit or hospitalization) for patients with mental illness.

METHODS

Study Sample

This study used the IBM MarketScan Research database from 2009 to 2018. The database contains inpatient, outpatient, and prescription claims records for millions of employees and their dependents. The claims are granular, containing information on patient diagnoses, location of services, procedure codes, amount billed, amount paid by insurance, and amount paid by patient. Patients are uniquely identified by an identification number that is consistent across years, and claims and allows for patients to be followed over time.

The sample was constructed by starting with the universe of enrollees who either had a telemedicine visit, or had an in-person visit for a chronic mental illness. Chronic mental illness was identified using the Agency for Healthcare Research and Quality's chronic condition indicators with ICD-9 and ICD-10 diagnostic codes. This study defined a visit as a patient-provider interaction in an outpatient setting for evaluation and management purposes or psychotherapy (based on the most used services for telemedicine). This study further categorized visits as either in-person or telemedicine using procedure codes. For the exact list of procedure codes, please see Appendix A1 (available online). All telemedicine users were included in the sample, but the number of non-users was restricted to 100,000 randomly selected enrollees.

Claim-level information was aggregated for each patient at a quarterly level. The data were further restricted to quarters during which patients were fully enrolled and for which pharmaceutical claims information was available.

Measures

Separate dichotomous outcome variables were created to indicate any ED use and any hospitalization during a quarter. Next, outcome variables related to visits and spending were calculated on a quarterly basis. Outpatient visits included: (1) total visits, (2) in-person visits, and (3) telemedicine visits. Spending variables included: (1) total medical spending on inpatient and outpatient services, (2) total pharmaceutical/drug spending, and (3) total out-of-pocket spending on medical and pharmaceutical services. The first 2 spending variables included total spending by or on behalf of the enrollee, excluding cash payments and premiums, whereas the last spending variable included enrollee spending in the form of copayments, co-insurance, and deductibles. All the monetary values were converted to real values using the consumer price index for medical care. Natural logarithms of all spending variables were used for estimation.

Additional outcome variables, which were used to assess the validity of the research design, included health shocks and an indicator for nonpreventable ED visits. Health shocks represent a sudden deterioration in health that is arguably exogenous to telemedicine use by the virtue of being unexpected or nonpreventable. Existing literature has used various variables as measures of health shocks, including, but not limited to, serious illness, injury, BMI, and hospital admissions.[22–26]

In this study health shocks were identified as nonpreventable hospitalizations, nonpreventable ED visits, and injury-related ED visits. A measure for health shocks is constructed by taking a natural logarithm of total spending on nonpreventable hospitalizations, nonpreventable ED visits, and injury-related ED visits. Agency for Healthcare Research and Quality's prevention quality indicators were used to identify nonpreventable hospitalizations, nonpreventable and injury-related ED visits from the claims database using the New York University's ED algorithm.[27,28]

Time was normalized for each user such that time, t, was measured with respect to the quarter of first observed telemedicine use (e.g., $t = 0$ is the quarter of first use, $t = -1$ is a quarter before, and so on). Throughout the remainder of the text, $t < 0$ is referred to as the preperiod and $t \geq 0$ is referred to as the postperiod.

Statistical Analysis

The estimation procedure could be broken down into 2 steps. In the first step, treatment and control groups were constructed. In the second step, a difference-in-differences method was used for estimation and inference.

The treatment group included telemedicine users who had information available for the past and the next 4 quarters relative to their quarter of first telemedicine use. The control group was constructed by finding observations of non-users who were most like the quarter of first telemedicine use of users (i.e., $t = 0$) based on preperiod dynamics of all outcome variables, patient characteristics, patients' chronic mental health diagnoses, and contemporaneous health shocks.

To find a control group, the data set was restructured such that each quarterly observation included the past 4 periods' shocks, spending, and utilization variables, along with current shocks, diagnoses, and patient characteristics. Current values of outcomes of interest were excluded. Then, machine learning was used to find a placebo first period for non-users that was compared with the first period of telemedicine use (i.e., observations similar to $t = 0$) for users in terms of lagged and some current values.

Four machine learning methods (classification and regression trees, random forest, logistic regressions, and neural network) were used to predict the probability of an observation containing the first telemedicine use, for both users and non-users, and the algorithm with the best performance, based on the accuracy of prediction, was selected. This yielded random forest as the primary prediction approach, with 76.5% accuracy.[29,30] The random forest model identified time periods where patients were likely to start using telemedicine based on the past dynamics of all variables, and current health shocks, patient characteristics, and mental health diagnosis. A subsample was created of all the observations that were predicted to have first telemedicine use, for both users and non-users, and was used for matching.

A valuable by-product of random forest is a matrix of proximity scores, which measures the similarity between any 2 observations based on multiple dimensions. The proximity scores allow for matching using high-dimensional data, where the model automatically selects the important variables, variable interactions, and values.[31] The method has been successfully applied in several studies in other areas.[32,33]

The scores range from 0 to 1, with 1 being a perfect match. Then, the highest proximity score was used for each $t = 0$ observation to uniquely match it with placebo first-use observation, $t = 0$, of non-users. Using placebo first-use observations, the time for non-users was normalized in the same way as for telemedicine users. The resulting group of non-users with placebo first-use served as the control group. This study focused on a 9-quarter window, with 4 quarters before and 4 quarters after the first use or placebo first-use. All matched pairs were kept with proximity scores ≥0.35, which was the eightieth percentile of the proximity scores. The matching procedure is described in detail in Appendix A2 (available online). For details on each algorithm's performance and tuning parameters see Appendix Table 1 (available online).

Once a panel data of matched pairs of users and non-users with actual or placebo first-use quarters was constructed, a difference-in-differences method was used to estimate the effect of telemedicine on outcomes.[34] This analysis tested whether there were differences in study outcomes before and after the event, the actual/predicted first quarter among treated group (users) compared with the control group (non-users). A set of covariates (age, dummy for metropolitan statistical area [MSA], insurance type, a set of year dummies, contemporaneous shocks [excluded for secondary outcomes], and patient fixed effects) was included to improve the balance between the 2 groups.

R version 4.1.1. was used for machine learning and random forest proximity matching, and Stata version BE 17.0 was used for difference-in-differences estimation. Analyses were conducted in 2021−2022.

RESULTS

The initial sample included 2.9 million quarters belonging to 37,928 telemedicine users and 100,000 non-users. Telemedicine utilization was very low with just >2% of observations with any telemedicine use. Highest utilization was by patients with a diagnosis related to depression, followed by anxiety, post-traumatic stress disorder, and attention deficit/hyperactivity disorder (Appendix Table 2, available online).

Summary statistics for users and non-users are provided in Appendix Table 3 (available online). For users, statistics were reported by the period (before/after) of first telemedicine use. Mean values for all spending and utilization variables were higher for users than non-users, regardless of the period relative to first telemedicine use, with the only exception being out-of-pocket spending in the before period.

Timing of users' first telemedicine use is analyzed in Appendix Figure 1 (available online). Plots of healthcare utilization and spending variables over time revealed an upward trend, with a sharp increase around the time of first telemedicine use, followed by a sharp decline.

After matching, 51,606 observations belonging to 5,734 patients were identified. Half of the patients were users and half were non-users, and each patient could be tracked for 9 quarters.

The quality of matching was assessed by comparing users and non-users using 2 separate methods: (1) by visually inspecting preperiod time trends of various spending and utilization variables and (2) by statistically testing preperiod differences in means at $p < 0.05$.

Figure 1 shows time trends of various spending and utilization variables. Time (t) on the horizontal axis represents quarters relative to first telemedicine use (or placebo first-use). Mean values are plotted separately for users and non-users. Plots revealed a similar trajectory for users and non-users in the preperiod, which suggested the 2 groups were comparable.

Preperiod summary statistics are provided in Table 1. Columns 1 and 2 show mean values for non-users and users, respectively. The p-values associated with the differences in means between the 2 groups are included in the third column. Summary statistics for variables related to healthcare use, shocks, and spending confirm the visual impression conveyed in Figure 1. None of these differences was significant, except for the logarithm of out-of-pocket spending, which was slightly lower for users (p=0.027).

There were slight differences in patient characteristics. Users consisted of 4 percentage points (p<0.001) more females and had 3 percentage points (p<0.001) greater number of people residing in MSAs. For both groups, anxiety was the most common diagnosis, followed by depression. A total of 58% of users and 56% of non-users had at least 1 anxiety-related diagnosis, and approximately 45% of patients in each group had a depression-related diagnosis.

Results are provided from difference-in-differences specification, after matching, in Table 2. The regression included an indicator for post-period (1 if t>=0, and 0 if t<0), an indicator for treatment group (1 for users and 0 for non-users), and an interaction term between the 2 indicators. The coefficient on the interaction term was the estimate of interest, which measured the pre- and post-difference in outcome for users compared with that for non-users. This study included covariates and provided results separately, with and without t=0 included in the post-period.

The estimates for overall visits and telemedicine visits were positive and statistically significant across both specifications, but they were larger with t=0 included in the post-period. This study found that telemedicine users had 0.461 greater number of overall visits (95% CI=0.280,0.642; p<0.001) despite a 0.280 reduction in in-person visits (95% CI= −0.446, −0.114; p=0.001). Estimates for the probability of hospitalization or any ED visit were small and insignificant.

This study did not find significant differences in medical spending and drug spending. Out-of-pocket spending was 10.9% lower when t=0 was included in the post-period (95% CI= −0.197, −0.021; p=0.015), and it remained negative but became insignificant when t=0 was excluded.

For reference, event-study estimates are provided for users with the control group excluded (Appendix Table 4, available online). All spending and utilization estimates were large and significant when the control group was excluded.

For visual clarity of dynamic patterns, a difference-in-differences was estimated by replacing the post-indicator with time dummies. The coefficients on the interactions between time dummies and the treatment indicator are reported with their 95% CIs in Figure 2. The figure confirms that users and non-users had similar trends in the preperiod, as coefficients for all variables were insignificant and close to 0. The post-period trends were also not

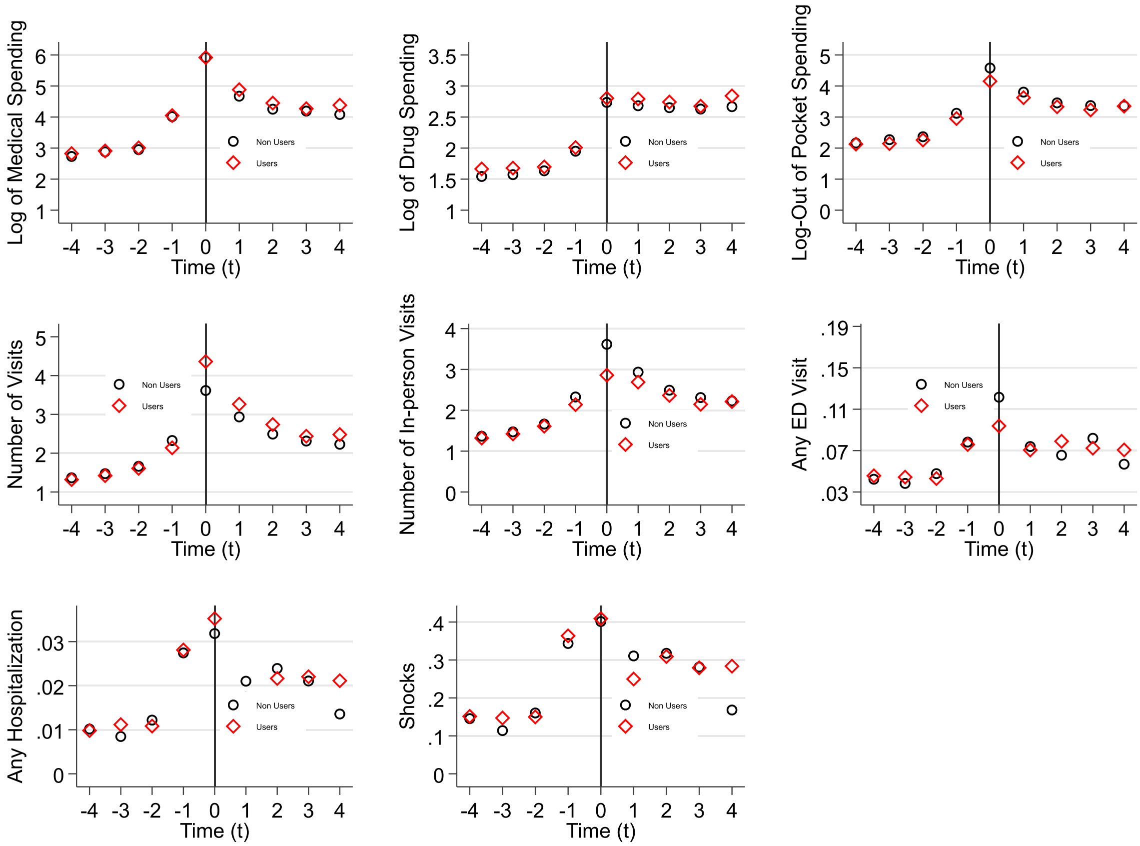

Figure 1. Variable trends by treatment status.
Note: The figure shows mean values for utilization and spending variables over time for users and non-users. Time (t) on the horizontal axis represents quarters relative to the first telemedicine use (or placebo first use). Variables on the y-axis include log of medical spending, log of drug spending, log of out-of-pocket spending, number of visits, number of in-person visits, any ED visit, any hospitalization, and shocks.

differentiable for medical spending, drug spending, shocks, and the probability of hospitalization.

Figure 2 also shows that just as patients started using telemedicine, the number of user visits became clearly and meaningfully greater than that of non-users. Simultaneously the number of in-person visits fell, along with the probability of an ED visit, and out-of-pocket spending. These effects seemed to dissipate over time, although the increased telemedicine use was still present 4 quarters out. There was some evidence of patients substituting ED care with telemedicine, which likely reduced their out-of-pocket spending, though the pattern did not hold after t=0.

A key assumption to ensure internal validity of the difference-in-differences model was that in the absence of telemedicine use, the difference between users and non-users would have been constant.[34,35] Although the assumption could not be statistically tested, it was supported by the similarity of preperiod trends shown in Figure 2.

Validity of the research design was further assessed by conducting falsification tests using outcomes that were not supposed to be affected by telemedicine use. Health shocks and the probability of nonpreventable ED visits were used as alternative placebo outcomes. Estimates for both variables were small and insignificant, suggesting that the 2 groups were similar in terms of exogenous health changes in the pre- and post-periods.

DISCUSSION

Descriptive analyses revealed empirical challenges, in that patients who start using telemedicine do so following large spikes in medical spending and health shocks, on average, and hence they are quite different from non-users. Given an upward pretrend in all outcomes of interest, any comparison between users and non-users that does not account for variable dynamics will likely yield biased estimates.[36,37] To address these challenges, a longitudinal sample of telemedicine users and similar

Table 1. Preperiod Summary Statistics

Variables	Non-users	Users	*p*-value
Patient characteristics			
Age (years)	27.95	27.70	0.208
Female (%)	47	51	**<0.001**
MSA (%)	88	91	**<0.001**
Anxiety (%)	56	58	**<0.001**
Depression (%)	45	46	0.158
PTSD (%)	28	30	**<0.001**
ADHD (%)	19	21	**0.001**
Health status			
Shocks	0.20	0.21	0.466
Nonpreventable ED visits (%)	1	1	0.848
Health care use			
Claims	6.78	6.54	0.298
Visits	1.70	1.62	0.094
In-person visits	1.70	1.62	0.094
Any ED visit (%)	5	5	0.837
Any hospitalization (%)	1	1	0.787
Spending			
OOP drug spending ($)	24.63	24.29	0.767
Total drug spending ($)	172.52	193.77	0.060
OOP medical spending ($)	121.20	119.38	0.758
Total medical spending ($)	997.78	1,116.09	0.328
Log drug spending	1.67	1.76	0.130
Log medical spending	3.14	3.19	0.198
Log OOP	2.48	2.37	**0.001**
Number of observations	11,468	11,468	
Number of patients	2,867	2,867	

Note: This table includes preperiod mean values for the variables listed. Unit of observation is patient-quarter. *p*-values for the differences between users and non-users are reported in the last column. Boldface indicates statistical significance ($p<0.05$).
ADHD, attention deficit/hyperactivity disorder; ED, emergency department; MSA, Metropolitan Statistical Area; OOP, out-of-pocket spending; PTSD, post-traumatic stress disorder.

non-users was constructed using machine learning, and then difference-in-differences analysis was used to assess the impact of telemedicine on the outcomes.

Using difference-in-differences analysis, this study found that telemedicine use is associated with an increase in the overall number of outpatient visits, despite a reduction in in-person visits. The effect of telemedicine use on the number of visits for privately insured patients in this study parallels the trends reported in other studies, which have looked at Medicaid and Medicare patients and found an association between higher telemedicine availability and an increase in visit rates.[20,21] This could be beneficial for people with chronic mental health conditions, who, despite higher morbidity, have fewer routine checkup visits, experience delayed diagnoses, and have higher treatment dropout rates than the general population.[9–11]

Interestingly, increased use of outpatient care is not accompanied by an increase in medical spending. One possible explanation could be that telemedicine visits during the studied time period were cheaper than in-person visits.[7] As patients substitute cheaper virtual visits in the place of more expensive in-person visits, despite an overall increase in outpatient care, overall medical costs could remain somewhat similar.

This study is the first, to the best of the author's knowledge, that examined the timing of users' telemedicine use and showed that once variable dynamics are accounted for, in addition to patient characteristics, the differences in medical spending between users and non-users are no longer significant. Previous studies have examined the relationship without controlling for preperiod trends. These studies have found telemedicine users as having considerably higher costs than non-users, but they focused on different health conditions and subgroups.[15,38]

In addition, this study did not find significant differences in the probabilities of ED visits or hospitalization

Table 2. Difference-in-Differences Estimates

Outcomes	With t=0				Without t=0			
	Coeff	95% CI		*p*-value	Coeff	95% CI		*p*-value
Health status								
Shocks	−0.008	−0.072	0.056	0.775	−0.008	−0.067	0.052	0.795
Any nonpreventable ED visit	−0.000	−0.004	0.003	0.865	0.001	−0.003	0.005	0.536
Health care use								
Visits	0.461	0.354	0.568	**<0.001**	0.261	0.146	0.376	**<0.001**
In-person visits	−0.280	−0.383	−0.176	**<0.001**	−0.132	−0.245	−0.019	**0.022**
Telemedicine visits	0.741	0.714	0.768	**<0.001**	0.393	0.369	0.417	**<0.001**
Any hospitalization	0.000	−0.002	0.003	0.781	−0.001	−0.004	0.002	0.492
Any ED visit	−0.005	−0.013	0.003	0.249	0.004	−0.005	0.013	0.381
Spending								
Log (medical spending)	0.027	−0.073	0.127	0.596	0.035	−0.078	0.148	0.547
Log (drug spending)	0.018	−0.041	0.077	0.541	0.016	−0.050	0.081	0.642
Log (out of pocket)	−0.124	−0.201	−0.048	**0.001**	−0.058	−0.143	0.027	0.182
Number of observations	51,606	51,606	51,606	51,606	45,872	45,872	45,872	45,872
Number of patients	5,734	5,734	5,734	5,734	5,734	5,734	5,734	5,734

Note: This table shows the regression coefficients from difference-in-differences. Each row represents an outcome variable of interest. All regressions include patient fixed effects, year fixed effects and the following patient level controls: age, dummy for metropolitan statistical area (MSA), insurance type, and contemporaneous health shocks (excluded for health status outcomes). All spending variables are in logarithms of real dollars. Estimates are provided separately with and without including t=0 in the post period, where t=0 indicates the quarter of first telemedicine use (or placebo first use). Boldface indicates statistical significance (p<0.05).
Coeff, coefficient; ED, emergency department.

between users and non-users. These results suggest that the quality of care delivered via telemedicine is on par with that of in-person visits. These findings need to be considered in the context of earlier work demonstrating the limitations of telemedicine use, which found that patients with telemedicine follow-up visits after ED-discharge were more likely to return to the ED within 30 days and had greater hospital utilization compared with patients with in-person follow-up.[39] However, these results are consistent with previous literature that looked at mental health specifically in an RCT setting and found no difference in outcomes between telemedicine interventions and in-person interventions.[40,41]

Limitations

This study has several limitations, as it focuses on the pre−COVID-19 period. Major policy changes, many impermanent, have allowed for increased telehealth utilization since the start of the COVID-19 pandemic.[42] For example, the U.S. Centers for Medicare and Medicaid Services and major insurers adopted temporary coverage policies providing telehealth payment parity.[43]

As the longevity of these policies continues to be debated, several factors need to be considered. This study's finding that telemedicine use is not associated with increased medical costs, despite increased utilization, needs to be considered in the context that telemedicine visits were cheaper than in-person visits in the pre-COVID period.

In the future, with telehealth payment parity, medical spending may rise with increased utilization, especially if patient cost sharing for telemedicine visits remains lower than for in-person visits. Also, any advantages to patients with chronic mental health conditions, who tend to underutilize health care, may disappear if telemedicine expansion policies are not accompanied by policies that address mental health provider shortages.

This study focused on patients with chronic mental illnesses who are privately insured. This study did not cover patients with government insurance and those who are uninsured. The results could differ for patients with different coverage. Separately, despite controlling for selection bias using matching and difference-in-differences, it is possible that some unobserved characteristics among users were not accounted for and influenced the patterns observed.

Finally, the study uses the probability of ED visits and of hospitalizations to measure the quality of care over a narrow 4-quarter window after treatment. Results could differ over a longer period and may be different if other proxies for quality of care are used. Future studies are needed to determine the generalizability of these findings.

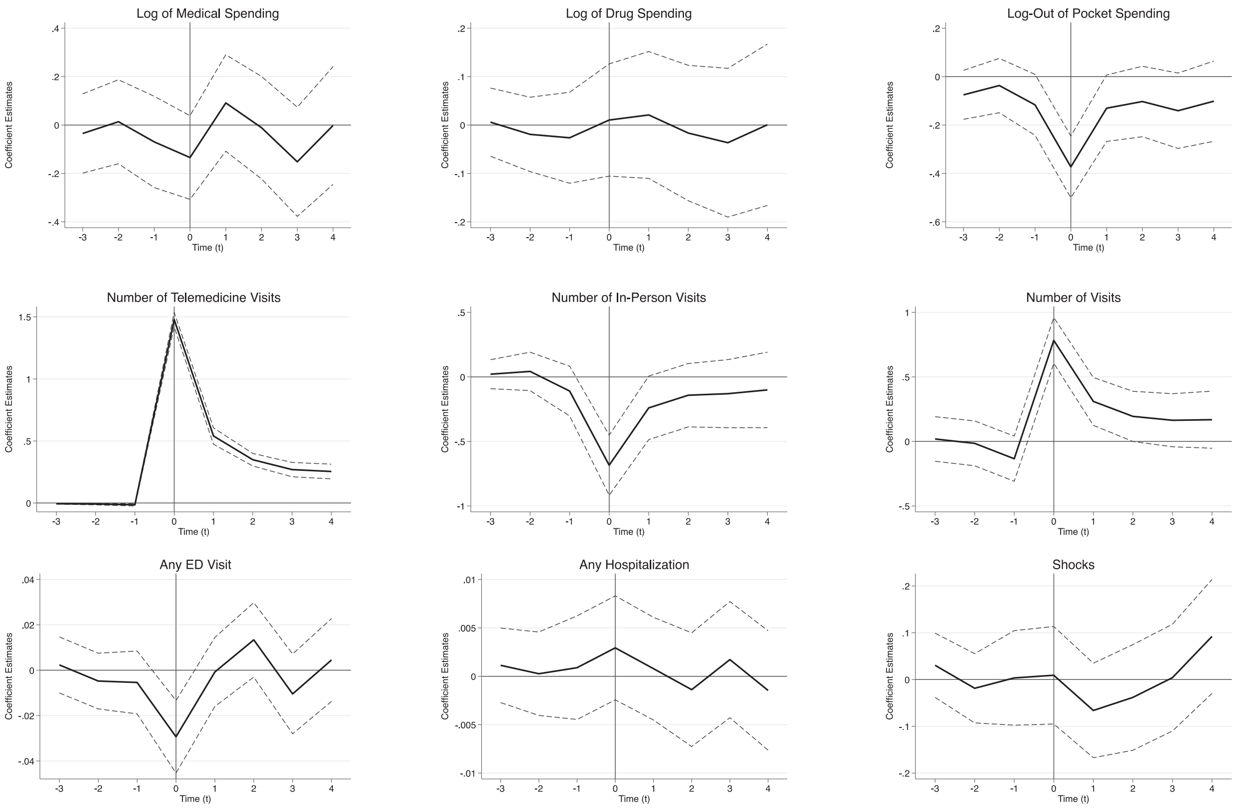

Figure 2. Event studies for outcome variables.
Note: The solid black line shows how the difference between users and non-users changes over time relative to the baseline difference (in $t= -4$). Outcomes include log of medical spending, log of drug spending, log of out-of-pocket spending, number of telemedicine visits, number of in-person visits, number of visits, any ED visits, any hospitalization, and shocks. All spending variables are in logarithms of real dollars. All regressions include patient fixed effects, year fixed effects and patient level controls. The dotted back lines represent 95% CI.

CONCLUSIONS

This study identifies that telemedicine leads to higher outpatient care utilization without significantly affecting hospitalizations, ED visits, and spending. The results highlight the importance of telemedicine utilization, in improving outpatient care utilization, for patients with chronic mental health conditions, without affecting the overall medical spending.

ACKNOWLEDGMENTS

The author thanks the Department of Pharmacy Systems, Outcomes and Policy and the Center for Pharmacoepidemiology and Pharmacoeconomic Research at University of Illinois at Chicago for access to the IBM MarketScan Research database. The interpretation and use of this database are solely the author's responsibility, and all mistakes are their own. The author acknowledges and thanks the reviewers and journal editors for their review and helpful comments.

Declaration of interest: none.

CREDIT AUTHOR STATEMENT

Ayesha Jamal: Conceptualization, Methodology, Software, Validation, Formal analysis, Investigation, Resources, Data curation, Writing — review & editing, Visualization.

SUPPLEMENTARY MATERIALS

Supplementary material associated with this article can be found in the online version at doi:10.1016/j.focus.2023.100127.

REFERENCES

1. Ahrnsbrak R, Bose J, Hedden SL, Lipari RN, Park-Lee E. Key substance use and mental health indicators in the United States: results from the 2016 national survey on drug use and health; Published 2017. https://www.samhsa.gov/data/sites/default/files/NSDUH-FFR1-2016/NSDUH-FFR1-2016.pdf. Accessed December 14, 2022.
2. Quirk H, Crank H, Harrop D, Hock E, Copeland R. Understanding the experience of initiating community-based physical activity and social support by people with serious mental illness: a systematic review using a meta-ethnographic approach. *Syst Rev.* 2017;6(1):214. https://doi.org/10.1186/s13643-017-0596-2.

3. DE HERT M, CORRELL CU, BOBES J, et al. Physical illness in patients with severe mental disorders. I. Prevalence, impact of medications and disparities in health care. *World Psychiatry.* 2011;10(1):52–77. https://doi.org/10.1002/j.2051-5545.2011.tb00014.x.
4. Fiorillo A, Sartorius N. Mortality gap and physical comorbidity of people with severe mental disorders: the public health scandal. *Ann Gen Psychiatry.* 2021;20(1):52. https://doi.org/10.1186/s12991-021-00374-y.
5. Mental health care health professional shortage areas (HPSAs). *Kaiser Family Foundation website.* 2022 Published https://www.kff.org/other/state-indicator/mental-health-care-health-professional-shortage-areas-hpsas/?currentTimeframe=0&sortModel=%7B%22colId%22:%22Location%22,%22sort%22:%22asc%22%7D.
6. Stacy W, AAMC. A growing psychiatrist shortage and an enormous demand for mental health services; Published 2022. https://www.aamc.org/news/growing-psychiatrist-shortage-enormous-demand-mental-health-services. Accessed December 14, 2022.
7. Cuellar A, Pomeroy JML, Burla S, Jena AB. Outpatient care among users and nonusers of direct-to-patient telehealth: observational study. *J Med Internet Res.* 2022;24(6):e37574. https://doi.org/10.2196/37574.
8. Chen PV, Helm A, Caloudas SG, et al. Evidence of phone vs video-conferencing for mental health treatments: a review of the literature. *Curr Psychiatry Rep.* 2022;24(10):529–539. https://doi.org/10.1007/s11920-022-01359-8.
9. Olfson M, Mojtabai R, Sampson NA, et al. Dropout from outpatient mental health care in the United States. *Psychiatr Serv.* 2009;60(7):898–907. https://doi.org/10.1176/ps.2009.60.7.898.
10. Heiberg IH, Nesvåg R, Balteskard L, et al. Diagnostic tests and treatment procedures performed prior to cardiovascular death in individuals with severe mental illness. *Acta Psychiatr Scand.* 2020;141(5):439–451. https://doi.org/10.1111/acps.13157.
11. Dixon LB, Holoshitz Y, Nossel I. Treatment engagement of individuals experiencing mental illness: review and update. *World Psychiatry.* 2016;15(1):13–20. https://doi.org/10.1002/wps.20306.
12. Shaver J. The state of telehealth before and after the COVID-19 pandemic. *Prim Care.* 2022;49(4):517–530. https://doi.org/10.1016/j.pop.2022.04.002.
13. Suran M. Increased use of Medicare telehealth during the pandemic. *JAMA.* 2022;327(4):313. https://doi.org/10.1001/jama.2021.23332.
14. Rae M, Amin K, Cox C, Panchal N, Miller BF. Telehealth has played an outsized role meeting mental health needs during the COVID-19 pandemic. *Kaiser Family Foundation website.* 2022 https://www.kff.org/coronavirus-covid-19/issue-brief/telehealth-has-played-an-outsized-role-meeting-mental-health-needs-during-the-covid-19-pandemic/.
15. Ashwood JS, Mehrotra A, Cowling D, Uscher-Pines L. Direct-to-consumer telehealth may increase access to care but does not decrease spending. *Health Aff (Millwood).* 2017;36(3):485–491. https://doi.org/10.1377/hlthaff.2016.1130.
16. Resneck JS, Abrouk M, Steuer M, et al. Choice, transparency, coordination, and quality among direct-to-consumer telemedicine Websites and apps treating skin disease. *JAMA Dermatol.* 2016;152(7):768–775. https://doi.org/10.1001/jamadermatol.2016.1774.
17. García-Lizana F, Muñoz-Mayorga I. What about telepsychiatry? A systematic review. *Prim Care Companion J Clin Psychiatry.* 2010;12(2).. https://doi.org/10.4088/PCC.09m00831whi.
18. Chakrabarti S. Usefulness of telepsychiatry: A critical evaluation of videoconferencing-based approaches. *World J Psychiatry.* 2015;5(3):286–304. https://doi.org/10.5498/wjp.v5.i3.286.
19. Hilty DM, Ferrer DC, Parish MB, Johnston B, Callahan EJ, Yellowlees PM. The effectiveness of Telemental health: A 2013 review. *Telemed J E Health.* 2013;19(6):444–454 [Review]. https://doi.org/10.1089/tmj.2013.0075.
20. Wang B, Huskamp HA, Rose S, et al. Association between telemedicine use in nonmetropolitan counties and quality of care received by Medicare beneficiaries with serious mental illness. *JAMA Netw Open.* 2022;5(6):e2218730. https://doi.org/10.1001/jamanetworkopen.2022.18730.
21. Cole MB, Lee EK, Davoust M, Carey K, Kim JH. Comparison of visit rates before vs after telehealth expansion among patients with mental health diagnoses treated at federally qualified health centers. *JAMA Netw Open.* 2022;5(11):e2242059. https://doi.org/10.1001/jamanetworkopen.2022.42059.
22. García-Gómez P, Kippersluis H van, O'Donnell O, Doorslaer E van. Long-term and spillover effects of health shocks on employment and income. *J Hum Resour.* 2013;48(4):873–909. https://doi.org/10.1353/jhr.2013.0031.
23. Morudu P, Kollamparambil U. Health shocks, medical insurance and household vulnerability: evidence from South Africa. *PLOS ONE.* 2020;15(2):e0228034. https://doi.org/10.1371/journal.pone.0228034.
24. Mohanan M. Causal effects of health shocks on consumption and debt: quasi-experimental evidence from bus accident injuries. *Rev Econ Stat.* 2013;95(2):673–681. https://doi.org/10.1162/REST_a_00262.
25. Wagstaff A. The economic consequences of health shocks: evidence from Vietnam. *J Health Econ.* 2007;26(1):82–100. https://doi.org/10.1016/j.jhealeco.2006.07.001.
26. Fadlon I, Nielsen TH. Family labor supply responses to severe health shocks: evidence from Danish administrative records. *Am Econ J Appl Econ.* 2021;13(3):1–30. https://doi.org/10.1257/app.20170604.
27. Faculty and research. New York University website. https://wagner.nyu.edu/faculty/billings/nyued-background. Accessed December 14, 2022.
28. Prevention quality indicators overview. *Agency for Healthcare Research and Quality* [website]. https://qualityindicators.ahrq.gov/measures/pqi_resources. Accessed October 11, 2022.
29. Breiman L. *Mach Learn.* 2001;45(1):5–32. https://doi.org/10.1023/A:1010933404324.
30. James G, Witten D, Hastie T, Tibshirani R. *Tree-based methods.* An Introduction to Statistical Learning. Springer, 2013303–335 103. https://doi.org/10.1007/978-1-4614-7138-7_8.
31. Zhao P, Su X, Ge T, Fan J. Propensity score and proximity matching using random forest. *Contemp Clin Trials.* 2016;47:85–92. https://doi.org/10.1016/j.cct.2015.12.012.
32. Pierola A, Epifanio I, Alemany S. An ensemble of ordered logistic regression and random forest for child garment size matching. *Comput Ind Eng.* 2016;101:455–465. https://doi.org/10.1016/j.cie.2016.10.013.
33. Nielsen MS, Rohde V. A surrogate model for estimating extreme tower loads on wind turbines based on random forest proximities. *J Appl Stat.* 2022;49(2):485–497. https://doi.org/10.1080/02664763.2020.1815675.
34. Dimick JB, Ryan AM. Methods for evaluating changes in health care policy: the difference-in-differences approach. *JAMA.* 2014;312(22):2401–2402. https://doi.org/10.1001/jama.2014.16153.
35. Angrist JD, Pischke JS. *MostlyHarmless Econometrics: an Empiricist's Companion.* 1st ed. Princeton University Press, 2008.
36. Ashenfelter O. Estimating the effect of training programs on earnings. *Rev Econ Stat.* 1978;60(1):47. https://doi.org/10.2307/1924332.
37. Heckman JJ, Smith JA. The pre-programme earnings dip and the determinants of participation in a social programme. Implications for simple programme evaluation strategies. *Econ J.* 1999;109(457):313–348. https://doi.org/10.1111/1468-0297.00451.
38. Weiner JP, Bandeian S, Hatef E, Lans D, Liu A, Lemke KW. In-person and telehealth ambulatory contacts and costs in a Large US insured cohort before and during the COVID-19 pandemic. *JAMA Netw Open.* 2021;4(3):e212618. https://doi.org/10.1001/jamanetworkopen.2021.2618.
39. Shah VV, Villaflores CW, Chuong LH, et al. Association between in-person vs telehealth follow-up and rates of repeated hospital visits among patients seen in the emergency department. *JAMA Netw Open.* 2022;5(10):e2237783. https://doi.org/10.1001/jamanetworkopen.2022.37783.
40. Drago A, Winding TN, Antypa N. Videoconferencing in psychiatry, a meta-analysis of assessment and treatment. *Eur Psychiatry.* 2016;36:29–37. https://doi.org/10.1016/j.eurpsy.2016.03.007.

41. Bashshur RL, Shannon GW, Bashshur N, Yellowlees PM. The empirical evidence for telemedicine interventions in mental disorders. *Telemed J E Health*. 2016;22(2):87–113. https://doi.org/10.1089/tmj.2015.0206.
42. Telehealth policy changes after the COVID-19 public health emergency; 2023. *Last Updated*. https://telehealth.hhs.gov/providers/policy-changes-during-the-covid-19-public-health-emergency/policy-changes-after-the-covid-19-public-health-emergency. Accessed March 31, 2023.
43. Morse S. Telehealth payment parity only good through 2023; Published 2023. https://www.healthcarefinancenews.com/news/telehealth-payment-parity-only-good-through-2023. Accessed March 31, 2023.