JUNE 2023

Using Population-Based Outcome Measures to Assess the Impact of Telehealth Expansion on Medicare Beneficiaries' Access to Care and Quality of Care

A report by American Institutes for Research® for the Medicare Payment Advisory Commission

The views expressed in this report are those of the authors. No endorsement by MedPAC is intended or should be inferred.

MEDPAC
Medicare Payment Advisory Commission

425 I Street, NW | Suite 701 | Washington, DC 20001 | (202) 220-3700 | www.medpac.gov

Using Population-Based Outcome Measures to Assess the Impact of Telehealth Expansion on Medicare Beneficiaries' Access to Care and Quality of Care

FINAL REPORT

Prepared for:
Ledia Tabor

Medicare Payment Advisory Commission
425 I Street NW, Suite 701
Washington, DC 20001

Prepared by:
Morteza Saharkhiz, PhD
Tanvi Rao, PhD
Sara Parker Lue, PhD
Sara Borelli, PhD
Karin Johnson, PhD
Guido Cataife, PhD

American Institutes for Research®
1400 Crystal Drive, 10th Floor
Arlington, VA 22202-3289

Advancing Evidence.
Improving Lives.

Acronyms

ACS	Ambulatory care sensitive
ADI	Area Deprivation Index
AHRFs	Area Health Resource Files
APM	Alternative Payment Model
APRN	Advanced practice registered nurse
BPCI	Bundled Payments for Care Improvement
CCA	Consolidated Appropriations Act
CCW	Chronic Conditions Warehouse
CME	Common Medicare Environment
CMS	Centers for Medicare & Medicaid Services
DID	Difference-in-differences
ED	Emergency department
E/M	Evaluation and management
ESRD	End-stage renal disease
FFS	Fee-for-service
FIPS	Federal Information Processing System
GAFs	Geographic Adjustment Factors
GAO	Government Accountability Office
GPCIs	Geographic Practice Cost Indexes
HCC	Hierarchical condition category
HCPCS	Healthcare Common Procedure Coding System
HSA	Hospital Service Area
IPPS	Inpatient Prospective Payment System
MA	Medicare Advantage
MDM	Master Data Management
MedPAC	Medicare Payment Advisory Commission
NGACO	Next Generation Accountable Care Organization
PA	Physician assistant
PFS	Physician Fee Schedule
PHE	Public Health Emergency
POS	Place of Service
PSW	Propensity score weighting
RAS	Risk Adjustment System
SAF	Standard analytic file
SAIPE	Small Area Income and Poverty Estimates
UIC	Urban Influence Code
ZCTA	ZIP Code Tabulation Area

Contents

Exhibits

Executive Summary

Background. In June 2023, the Medicare Payment Advisory Commission (MedPAC) will present a mandated report to the Congress on telehealth including the impact of expanded telehealth coverage on access to care and quality of care, which could ultimately help inform the question of whether and how Medicare should permanently cover telehealth expansion. The objective of this study is to use the data available to date to assess the feasibility of using population-based measures to estimate the association between telehealth use and quality of care and access to care when both telehealth and in-person visits are available to fee-for-service (FFS) Medicare beneficiaries.

Methodology. We used Medicare FFS administrative data to compare population-based outcomes across geographic areas with different levels of telehealth service use. We examined population-based measures that capture

- quality of care, including ambulatory care sensitive (ACS) hospitalizations and emergency department (ED) visits per 1,000 FFS Medicare beneficiaries;
- access to care, including clinician encounters per FFS Medicare beneficiary and a breakdown of clinician encounters by provider type; and
- cost of care, including total cost of care for Part A and Part B services per FFS Medicare beneficiary and a breakdown of cost by service type.

The geographic unit for the study is Hospital Service Areas (HSA). The study compares the second half of 2021 (treatment period) with the second half of 2019 (baseline period). The second half of 2021 was selected as the treatment period of analysis because it was less affected by the COVID-19 pandemic compared with the early stages of the pandemic. This approach helps produce estimates that represent, as much as possible, the effects of telehealth in a relatively more “normal” period, given the data available at the time of conducting the analyses.

We created three levels of intensity for the usage of telehealth services by ranking HSAs based on the number of telehealth visits per 1,000 beneficiaries in the second half of 2021. We assigned the bottom third of HSAs to the Low telehealth intensity group, the middle third of HSAs to the Medium group, and the top third of HSAs to the High group.

We compared outcomes in Medium and High telehealth intensity HSAs to Low telehealth intensity HSAs using a difference-in-differences (DID) approach. The DID approach controls for factors that remain constant over time within HSAs (e.g., HSA urbanicity). However, factors that change between the baseline and treatment period could confound the association between telehealth intensity and population-based outcomes. For example, sociodemographic characteristics of FFS beneficiaries may change as the share of beneficiaries enrolled in FFS Medicare changes. Therefore, we controlled for several time-varying covariates including the sociodemographic characteristics of FFS enrolled beneficiaries, average risk scores, and new and cumulative COVID-19 cases per 10,000 people.

Key Findings. Our key findings concerning the association between telehealth intensity and population-based outcomes are as follows:

- *Telehealth and quality*: The High telehealth intensity HSAs were associated with an increase in ACS hospitalizations relative to the Low telehealth intensity HSAs (with the magnitude of increase being 6.90% of the baseline rate). For context, Low telehealth intensity HSAs (the comparison group) had a sharp decline in ACS hospitalizations (32% decline in 2021 relative to baseline). Thus, ACS hospitalizations also decreased on average among High telehealth intensity HSAs but at a slower rate than in the Low telehealth intensity HSAs. The High telehealth intensity group was not associated with a change in ACS ED visits. Additionally, our analysis did not show an association between telehealth intensity and ACS hospitalizations or ED visits when comparing areas of Medium telehealth intensity with areas of Low telehealth intensity.
- *Telehealth and access.* High telehealth intensity HSAs were associated with an increase in the overall number of clinician encounters per beneficiary relative to the Low telehealth intensity HSAs (with the magnitude of increase being 2.67% of the baseline rate). Although there was a relative increase in all types of clinician encounters, the largest increases came from encounters with hospitalists (7.11% increase) followed by encounters with advanced practice registered nurses (APRNs) and physician assistants (PAs) (6.60% increase). There was no statistically significant difference between the Medium and Low telehealth intensity groups in total clinician encounters per beneficiary.
- *Telehealth and cost.* High telehealth intensity HSAs were associated with an increase in the total cost of care per beneficiary relative to the Low telehealth intensity HSAs (with the magnitude of increase being 2.47% of the baseline rate). The High telehealth intensity group was also associated with an increase in cost of care per beneficiary for the following service types: inpatient, skilled nursing facility, physician, and durable medical equipment. These increases in cost are consistent with the increases in ACS hospitalizations and clinician encounters experienced by these HSAs. Our analysis did not show an association between telehealth intensity and total cost of care per beneficiary when comparing areas of Medium telehealth intensity with areas of Low telehealth intensity.

Key findings from these analyses are robust to various specification changes, including the use of propensity score weighting (PSW) to assign weights to HSAs in the DID model so that the HSAs are more similar in terms of baseline sociodemographic and geographic characteristics.

Discussion. The association between High telehealth intensity and a smaller reduction (increase relative to the Low telehealth intensity group) in ACS hospitalizations may seem counterintuitive because we also found relative increases in all types of clinician encounters in the High telehealth intensity group, and ACS hospitalizations are generally viewed as hospital admissions that could have been prevented with timely ambulatory care. A potential interpretation could be that the increased ACS hospitalizations may have been driven by an overall increase in health care utilization in High telehealth intensity HSAs relative to the Low telehealth intensity HSAs, which increased both telehealth and other health care services (including ACS hospitalizations). However, our findings remained robust to controlling for in-person utilization, which was defined as the number of non-telehealth visits with clinicians per FFS beneficiary. This suggests that an overall increase in utilization in the High telehealth intensity areas is not the likely explanation for our findings. Also, these findings do not seem to be driven by the Low telehealth intensity areas being a poor comparison group for areas with High and Medium telehealth intensity because the three groups had similar outcome trends in the baseline period.

A more plausible interpretation of these findings relates to one of the limitations of this study, namely, the use of the second half of 2021 to approximate post-pandemic outcomes. This time period overlapped with the continuing public health emergency (PHE) and the surge in COVID-19 cases due to the Delta variant of COVID-19, which peaked in early September 2021. This time period also includes the beginning of the surge in cases due to the Omicron variant, which began in December 2021. Therefore, pandemic-related care disruptions and the resulting changes in population behaviors likely confounded our analyses. Nationally, there were steep declines in ACS hospitalizations and ED visits between the second half of 2019 and the second half of 2021. These declines were directly related to the COVID-19 pandemic, which disrupted the health care system and fundamentally altered the behavior of the population, the way individuals access the health care system, and their likelihood of acquiring ACS conditions (e.g., see Becker et al., 2022). Although these trends were observed at the national level, differences in the timing and implementation of local regulations on masking, school closures, screening for illnesses, vaccine uptake, and local transmission of other viruses (Chow et al., 2023), as well as differential rates of rebound of health care systems across the country all put different geographic areas at very different points in terms of their recovery from the PHE in 2021. These differences directly affected our analyses. For example, systems that already have enhanced telehealth capabilities in place may also be more likely to adapt to pandemic restrictions and challenges relatively quickly (e.g., Whaley et al., 2022). This may translate into higher ACS hospitalizations compared with areas that were unable to navigate the complexities associated with the pandemic and experienced more drastic interruptions in care.

Conclusions. Our study suggests that using a population-based approach to inform Medicare's decision to permanently expand telehealth is feasible. The main obstacle observed (the confounding effects of the pandemic throughout the second half of 2021) will fade over time as the outcome trends return to normal and the differential effects of the pandemic across HSAs tend to disappear. This can be confirmed by repeating the current analysis with post-pandemic data from a time period in which there is less pandemic-related volatility, such as the second half of 2022. Other options to mitigate the bias arising from the pandemic include using alternative outcomes such as non-respiratory ACS hospitalizations (e.g., diabetes-related hospitalizations), which were much less affected by the ebb and flow of the pandemic's care disruptions.

1. Background

Telehealth includes health care services delivered through a range of online, video, telephone, and other communication methods. Historically, traditional Medicare has been limited by statute to paying only for telehealth services under the Physician Fee Schedule (PFS) when they are provided to beneficiaries who received the service at a clinician's office or certain health care facilities (known as "originating sites") located in a rural area, with some exceptions. However, to maintain access to care and help limit community spread of COVID-19 during the PHE, Medicare temporarily expanded coverage for telehealth under the PFS to all Medicare beneficiaries regardless of their location, including telehealth visits provided to patients at home (CMS, 2022). During the PHE, many providers and beneficiaries embraced telehealth (ASPE, 2020; GAO, 2022). In a 2022 report, the U.S. Government Accountability Office (GAO) found that telehealth use under Medicare increased tenfold from about 5 million services in April–December 2019 to more than 53 million services during the same months in 2020. For these reasons, Congress and the Centers for Medicare and Medicaid Services (CMS) are considering the possibility of making the PHE expansions permanent.

In the March 2021 report to the Congress, MedPAC presented a policy option that policy makers continue to cover telehealth services with the potential for clinical benefit for a limited time (1 or 2 years) after the end of the COVID-19 PHE. The motivation was to allow time to gather evidence on the effects of telehealth services (including audio-only) on access to care, quality of care, and cost outcomes, which could ultimately inform the question of whether Medicare should permanently cover telehealth expansion (MedPAC, 2021).

In the Consolidated Appropriations Act (CAA), 2023, the Congress extended Medicare's telehealth expansions through December 31, 2024. In the CAA, 2022, the Congress mandated that the Commission submit a report by June 2023 on the use of telehealth services in Medicare during the PHE, the impact of expanded telehealth coverage on access to care and quality of care, Medicare payment policy for telehealth services under the PFS and the payment systems for federally qualified health centers and rural health clinics, and alternative approaches to paying for telehealth services.

The objective of this study is to use the data available to date to assess the feasibility of using population-based outcomes to estimate the association between telehealth use and population-based outcomes for access to care and quality of care when both telehealth and in-person visits are available to FFS Medicare beneficiaries. In this context, "population-based outcomes approach" means a study of various measures of health or health-related events (e.g., access to care and quality of care) for all individuals in a geographic region, based on administrative data, including claims (Kindig, 2015).

2. Methodology

To study how telehealth affected population-based outcomes, we used population-based measures that MedPAC previously used to analyze quality of care (ACS hospitalizations and ED visits per 1,000 FFS Medicare beneficiaries) and access to care (all clinician encounters, composed of in-person and telehealth encounters, per FFS Medicare beneficiary and a breakdown of clinician encounters by provider type). We also analyzed corresponding costs (total cost of care for Part A and Part B services per FFS Medicare beneficiary and a breakdown of cost by service type). The geographic unit for the study is the HSA. The study period is the second half of 2019 (baseline period) and the second half of 2021 (treatment period). This approach provides two symmetrical periods that are as close to normal times (i.e., less affected by the COVID-19 pandemic) as possible, which helps mimic a long-term scenario.

The main methodological challenge for this study was that the independent variable, telehealth intensity was not randomly assigned across HSAs. Instead, telehealth intensity was heavily correlated with sociodemographic characteristics and other variables that confound health outcomes. For example, MedPAC's analysis of 2021 FFS Medicare claims found that beneficiaries who are younger, qualify for Medicare because of end-stage renal disease (ESRD) or disability, have lower income, and live in urban areas use a higher number of telehealth services on average (MedPAC, 2023). This nonrandom assignment has the potential to bias the estimates and yield unreliable findings.

To address this challenge, we conducted a quasi-experimental DID analysis comparing health outcomes for areas with different levels of telehealth intensity. A DID approach controls for all factors that remain constant over time within the geographic regions under study.[1] A DID approach does not account for changes between the baseline and treatment period (i.e., compositional change within the HSAs included in the study). However, this problem has limited impact on our study given the short period of time being analyzed. This limitation could, however, have a larger impact on future studies that MedPAC may conduct with data from extended time periods.[2]

The following sections describe the period of study, health care markets, study measures, covariates, and the empirical strategy.

2.1. Period of Study

The baseline period is the second semester of 2019 (defined as July–December 2019, before the PHE and the expansion of telehealth) and the treatment period is the second semester of 2021 (defined as July–December 2021, after COVID-19 vaccines were widely available to Medicare beneficiaries and the expansion of telehealth). Experiences during the early months of the pandemic may not be appropriate to use when studying changes in population-based outcomes. We used data from the second semester of 2021 because they were the most recent data available

[1] This is less of a concern for the two quality measures because they are risk adjusted. However, the access measure and the cost measure are not risk adjusted.

[2] We also implemented a methodology that supplements the DID model with propensity score weighting to create a comparison group that better resembles the treatment groups based on baseline sociodemographic and geographic characteristics. Although we did not find propensity score weighting to substantially alter the DID results in this study, this methodology may be implemented by MedPAC in future years.

at the time of this study. In addition, the symmetry of these periods alleviates concerns about seasonality in the data.[3]

2.2. Health Care Markets

We used HSAs to represent health care markets. The Dartmouth Atlas of Health Care defines HSAs as local health care markets that satisfy most of the residents' health care needs, including hospitalizations (Dartmouth Atlas Project, 2022a). There are 3,436 HSAs in the United States, and most contain only one hospital. Given the purpose behind their construction and the granularity that they allow, the HSA is the geographic level we chose for the calculation of the outcome measures.

An alternative market area that we considered for this study, hospital referral regions, are geographically larger; there are 306 of these regions in the United States. Given their size, hospital referral regions may mask important variations in outcomes within an already populous geographic area. A second alternative was to use MedPAC market areas, which are derived from core-based statistical areas from the Office of Management and Budget. However, MedPAC market areas were also deemed to be too large; there are about 1,200 in the United States (MedPAC, 2019).

2.3. Study Measures

Exhibit 1 provides an overview of the outcomes and the treatment variable that we used in the analysis. We discuss each of these variables in detail below.

Exhibit 1. Outcomes and Treatment Variables Used in the Analysis

Variable type	Variable name	Specification	Notes
Treatment	Telehealth intensity	Groupings based on the number of telehealth visits per 1,000 FFS Medicare beneficiaries: Low (<33rd percentile), Medium (33rd to 66th percentile), and High (>66th percentile)	See Exhibit A–1 in Appendix A for telehealth codes.
Outcome: Quality	ACS hospitalizations rate (risk adjusted)	Number of hospitalizations and observation stays with specified acute and chronic ACS conditions per 1,000 FFS Medicare beneficiaries	MedPAC-modified AHRQ PQIs[a,b]
	ACS ED visit rate (risk adjusted)	Number of ED visits with specified acute and chronic ACS conditions per 1,000 FFS Medicare beneficiaries	
Outcome: Access	Number of clinician encounters	Number of clinician encounters, including in-person and telehealth encounters, per FFS Medicare beneficiary	Previously specified and used by MedPAC[c]
	Number of clinician encounters by provider type	A breakdown of clinician encounters per FFS Medicare beneficiary by the following provider types: primary care physicians, specialists, APRNs and PAs, other practitioners, and hospitalists	

[3] In addition, we used data from the two semesters of 2018 and the first semester of 2019 to check the parallel trends assumption underlying the DID.

Variable type	Variable name	Specification	Notes
Outcome: Cost	Total cost of care	Sum of Medicare payments, beneficiary cost sharing, and primary payer payments for Part A and Part B services per FFS Medicare beneficiary	See CCW Technical guidance[d]
	Total cost of care by service type	A breakdown of cost per FFS Medicare beneficiary by the following service types: inpatient, outpatient, skilled nursing facility, home health, hospice, physician, and durable medical equipment	

Note. ACS = ambulatory care sensitive; AHRQ = Agency for Healthcare Research and Quality; APRN = advanced practice registered nurse; CCW = Chronic Conditions Warehouse; ED = emergency department; FFS = fee-for-service; HSA = Hospital Service Area; PA = physician assistant; PQI = Prevention Quality Indicator. Primary care physicians include physicians from family medicine, internal medicine, pediatric medicine, and geriatric medicine. Other practitioners include clinicians such as physical therapists, psychologists, social workers, and podiatrists.
[a] Agency for Healthcare Research and Quality. (2022, July). *Prevention quality indicators technical specifications*. https://qualityindicators.ahrq.gov/measures/PQI_TechSpec
[b] Feng, Z., Silver, B., Segelman, M., Jones, M., Ingber, M. J., Beadles, C., & Pickett, R. (2019, August). *Developing risk-adjusted avoidable hospitalizations and emergency department visits quality measures*. https://www.medpac.gov/wp-content/uploads/import_data/scrape_files/docs/default-source/contractor-reports/august2019_riskadjusted_ah_av_measures_contractor_sec.pdf
[c] Medicare Payment Advisory Commission. (2022, March). *Report to the Congress: Medicare payment policy*. https://www.medpac.gov/wp-content/uploads/2022/03/Mar22_MedPAC_ReportToCongress_SEC.pdf.
[d] Chronic Conditions Warehouse. (2022, September). *Getting Started with CMS Medicare Administrative Research Files*. https://www2.ccwdata.org/documents/10280/19002248/ccw-technical-guidance-getting-started-with-cms-medicare-administrative-research-files.pdf

2.3.1. Treatment: Telehealth

The telehealth intensity measure is based on utilization in the second half of 2021 (the treatment period) instead of the baseline period because telehealth was restricted prior to the COVID-19 PHE. We identified telehealth visits using the physician and outpatient claims Standard Analytic Files (SAFs). We used Healthcare Common Procedure Coding System (HCPCS) codes for telehealth-eligible services published by CMS (CMS, 2022), together with Place of Service (POS) and HCPCS modifier codes 95, GT, GQ, and G0, which are necessary to define telehealth use for codes that are not specific to telehealth. We also used codes that are specific to the CMS Innovation Center waivers for the Bundled Payments for Care Improvement (BPCI) Advanced and the Next Generation Accountable Care Organization (NGACO) models to fully capture telehealth use within existing value-based care initiatives. In addition, we included codes for remote services, virtual or e-visit check-ins, and telephone evaluation and management (E/M). We excluded codes for originating site telehealth services and interprofessional internet consultation services because they are not patient-facing services. Exhibit A–1 in Appendix A provides more details.

We considered two options for measuring telehealth intensity in a market area: (a) the proportion of PFS claims that are provided by telehealth and (b) the number of PFS claims provided by telehealth per 1,000 beneficiaries. We chose the second option, the rate, because the first option conflates the effects of variation in total (i.e., telehealth and non-telehealth) clinician visits with the effect of variation in telehealth visits. To illustrate this point, consider the following example:

- HSA A has 250 telehealth visits and 500 in-person visits.
- HSA B has 500 telehealth visits and 1,000 in-person visits (double amounts but identical proportion).

Further, assume the two HSAs have the same number of FFS beneficiaries (e.g., 1,000). The proportion of PFS claims that are telehealth are identical (i.e., 0.5) for both HSAs, but the rate measures are different (HSA A = 250/1,000; HSA B = 500/1,000). Because the proportion measure has identical value for the two HSAs, it fails to reflect the effects brought by the much higher number of telehealth visits in HSA B. Instead, the rate measure reflects the higher number of visits in HSA B, making it a more suitable measure to study the association between telehealth and population-based outcomes.

When calculating the numerator of the measure, we attributed telehealth encounters to HSAs using the location of the beneficiary based on the Common Medicare Environment (CME) custom enrollment file. In the denominator, we included all beneficiaries who had FFS Parts A and B for the entire semester. After calculating the number of telehealth visits per 1,000 FFS Medicare beneficiaries for each HSA during the treatment period, we assigned each HSA one of the three telehealth intensity groups of Low, Medium, or High based on their ranking in terms of the number of telehealth visits per 1,000 beneficiaries. We assigned the bottom third of HSAs to the Low group, the middle third of HSAs to the Medium group, and the top third of HSAs to the High group. We discuss how we used the telehealth intensity groups in **2.5. Empirical Strategy**.

Ideally, we would have created a sub-measure and conducted an analysis focusing on how audio-only encounters relate to the study outcomes of interest. However, only 3 of 86 telehealth-eligible codes reliably indicate audio-only encounters with a physician or other qualified health professional because of PHE flexibilities (MedPAC, 2022). Therefore, we did not conduct a regression-based DID analysis focused on audio-only encounters.

2.3.2. Outcome: Quality

We studied two quality measures: risk-adjusted ACS hospitalizations and ED visit rates (Feng et al., 2019). MedPAC developed these two claims-based outcome measures to compare quality of care within and across different populations due to the adverse impact on beneficiaries and high cost of these events. We used MedPAC's preexisting SAS codes and specifications to calculate both quality measures.

Two categories of ACS conditions are included in the measures: chronic (e.g., diabetes, asthma, hypertension) and acute (e.g., bacterial pneumonia, cellulitis). Conceptually, an ACS hospitalization or ED visit refers to hospital use that could have been prevented with timely, appropriate, high-quality care. For example, if a diabetic patient's primary care physician and specialists effectively control the condition and they have a system in place to allow for urgent visits, then the patient may be able to avoid a visit to the ED for a diabetic crisis.

2.3.3. Outcome: Access

We studied clinician encounters per FFS Medicare beneficiary and examined their breakdown by the following provider types: primary care physicians, specialists, APRNs and PAs, other practitioners, and hospitalists.[4] Primary care physicians include clinicians in family medicine, internal medicine, pediatric medicine, and geriatric medicine. Other practitioners include clinicians such as physical therapists, psychologists, social workers, and podiatrists. We used MedPAC's preexisting SAS codes and specifications and calculated these measures using the physician claims SAFs.

The number of clinician encounters per beneficiary offers a direct measure of health care access. Encounters are a measure of entry into the health care system. Entry can be a first step toward timely use of services. However, it is an aggregate measure that may mask important differences by provider type. As an example, before the pandemic, from 2015 to 2019, while the number of primary care physician encounters per beneficiary fell by 2.5% per year, encounters with APRNs and PAs per beneficiary rose by 11.2% per year (MedPAC, 2022). The breakdown by provider type allows for a more nuanced examination of how the expansion of telehealth affects different parts of the health care system. For example, while telehealth expansion may increase access to routine preventive visits to primary care clinicians, it may have a limited impact on patients' access to certain specialists (e.g., proceduralists).

2.3.4. Outcome: Cost

We studied the total cost of care for Part A and Part B services per FFS Medicare beneficiary. Following Chronic Conditions Warehouse (CCW) technical guidance, total cost of care includes Medicare payments, beneficiary cost sharing, and primary payer payments. To Medicare payments, we also added back advanced payments that CMS makes to Alternative Payment Model (APM) participants that are recouped through claim payments to those providers (e.g., population-based payments in the NGACO model) (CCW, 2022).

In addition to total cost of care for Part A and Part B services per FFS Medicare beneficiary, we studied costs by the following service types: inpatient, outpatient, skilled nursing facility, home health, hospice, physicians, and durable medical equipment. This allows us to better understand any source of change in total costs because of telehealth use. One possibility is that if telehealth positively affects access to primary care, then over time we could expect lower utilization of more expensive sources of care, and hence a decrease in certain costs such as inpatient costs. We calculated the cost for each service type using its respective claims data SAFs.

2.4. Covariates

Exhibit 2 presents the full list of beneficiary and market characteristic variables used as covariates in the DID model. **2.4.1. HSA Medicare Population Characteristics** details the rationale and data sources for the beneficiary characteristics and **2.4.2. HSA Market Characteristics** describes the market characteristics. Exhibits A-2 and A-3 in Appendix A describe the data sources in detail.

We identified potential covariates that, based on the literature, could affect both the study outcomes and telehealth intensity and could possibly change between the baseline period and the treatment period (Bose et al., 2022; Eberly et al., 2020). In general, we controlled for variables that were found, in descriptive analysis, to correlate with both the outcome variables and telehealth intensity and that varied over time between the baseline and treatment periods (see **3.1.2. HSA Characteristics** for related descriptive statistics). Certain variables, mainly HSA

[4] Each encounter is identified as a unique combination of beneficiary ID, claim ID, and National Provider Identifier.

sex and racial/ethnic group composition, showed very little variation between the two time periods, but we still controlled for them, following the literature. Adding such variables does not bias our estimates because variables that are mostly constant over time do not have explanatory power in a DID model.[5]

Exhibit 2. Covariates Used in DID

Covariates
HSA Medicare population characteristics
Share of Medicare beneficiaries enrolled in FFS
Shares of FFS beneficiaries under age 65, 65–74, 75–84, and 85+
Share of FFS male/female/unknown sex beneficiaries
Shares of FFS White/Black/Hispanic/Asian/other/unknown race beneficiaries
Share of FFS beneficiaries fully/partially eligible for Medicaid
Average HCC risk score and its square for FFS Medicare beneficiaries
Share of FFS beneficiaries attributed to APMs
Average ADI for FFS Medicare beneficiaries
HSA market characteristics
Population size
New and cumulative COVID-19 cases per 10,000 people

Note. ADI = Area Deprivation Index; APM = alternative payment model; DID = difference-in-differences; FFS = fee-for-service; HCC = hierarchical condition category; HSA = Hospital Service Area.

2.4.1. HSA Medicare Population Characteristics

In this section, we describe population characteristics of Medicare beneficiaries listed in Exhibit 2.

Share of Medicare beneficiaries enrolled in FFS: Beneficiaries self-select FFS or Medicare Advantage (MA) plan enrollment. Enrollment in MA plans increased significantly during our study period; therefore, the composition of sample beneficiaries in an HSA could have changed in a way that correlates with both telehealth use and population-based outcomes. For instance, if MA plans attract disproportionately younger Medicare beneficiaries in an HSA, then the Medicare FFS population included in our study could show lower telehealth use and a change in population-based outcomes for the same reason. Thus, we used the share of beneficiaries enrolled in FFS Medicare as a DID covariate. We calculated this variable from the CME custom enrollment file.

Share of FFS beneficiaries by age, sex, race, and Medicaid eligibility: The demographic characteristics of an area affect both the outcomes and the treatment. As an example, younger Medicare beneficiaries use telehealth more frequently and may have better health outcomes than

[5] Some additional variables were considered for inclusion in the DID model but were not included either because they did not have any variation across the baseline and treatment periods or because only baseline data were available. These variables were used to generate propensity score weights and are described in detail in Appendix C.

older beneficiaries. Thus, we controlled for the share of Medicare FFS beneficiaries who were under age 65, 65–74, 75–84, and 85 and older; the share of FFS Medicare beneficiaries who were male, female, or unknown sex; the share of FFS Medicare beneficiaries who were White, Black, Hispanic, Asian, other, or unknown race; and the share of FFS Medicare beneficiaries who were eligible for Medicaid. We calculated these variables using the CME custom enrollment file. We used the Research Triangle Institute race code for determining race and ethnicity because it improves the coding accuracy for Hispanic beneficiaries (Eicheldinger & Bonito, 2008).

Average hierarchical condition category risk score and its square for FFS Medicare beneficiaries: The health status and disease severity of the underlying population in an area affect both the outcomes and the treatment. Thus, we controlled for average hierarchical condition category (HCC) risk score.[6] In addition, the distribution of HCC risk scores at the beneficiary level is right skewed (i.e., a small number of beneficiaries have high HCC risk scores). These beneficiaries may drive both the aggregate health care quality and access outcomes. Hence, we also used the average of HCC risk score squared to capture the disproportionate effects that beneficiaries with high HCC risk scores may have on the outcomes. We calculated these variables using the CME custom enrollment file and Risk Adjustment System (RAS) data.[7]

Share of FFS beneficiaries attributed to APMs: APMs are motivated to lower health care costs, and they may also incentivize the use of telehealth services, which could confound the effects of telehealth (Samson et al., 2021). Hence, we controlled for the share of FFS Medicare beneficiaries attributed to APMs. To calculate this variable, we determined whether each FFS beneficiary was attributed to an APM by linking the CME custom enrollment file with the Master Data Management (MDM) beneficiary extract.

Average Area Deprivation Index for FFS Medicare beneficiaries: We controlled for the Area Deprivation Index (ADI) as a proxy for social determinants of health, which may affect telehealth use and health outcomes (Kind & Buckingham, 2018). We obtained the ADI from the University of Wisconsin School of Medicine and Public Health and assigned each FFS beneficiary an ADI based on their 9-digit ZIP Code.

2.4.2. HSA Market Characteristics

The variables discussed in this section were available at the county level. To aggregate these characteristics to the HSA level, we created a crosswalk between counties and HSAs and estimated the population for the part of each county that overlaps with an HSA. This crosswalk allowed us to calculate a weight for each county that overlapped with an HSA that was proportional to the population of the county that resides within the HSA. We constructed this crosswalk by combining (a) the crosswalk between ZIP Codes and HSAs from the Dartmouth

[6] Risk-adjusted ACS hospitalizations and ED visit rates already adjust for comorbidities, among other factors. However, HCC risk scores are also included in the DID model to ensure a uniform methodology across outcomes. This partial redundancy in the adjustment of the regressions for some of the outcomes does not have substantial implications for the estimated associations.

[7] Note that risk scores in a given year are based on chronic conditions in the prior year, which could lead to the possibility of risk scores being affected by the treatment itself and thus not be an appropriate control variable in the regression. This, however, seems of little concern in our setup as 2019 risk scores are based on 2018 chronic conditions (thus unrelated to the treatment), and 2021 risk scores are based on 2020 chronic conditions, which are also plausibly unaffected by the way we define treatment (i.e., telehealth intensity in 2021). Controlling for risk scores could become a concern if the study were repeated over a longer period. In that context, lagged risk scores could be used to mitigate the issue.

Atlas (2022b) and (b) the crosswalk between ZIP Code Tabulation Areas and counties from the U.S. Census Bureau (2022). Under the assumption that the population of the county is roughly homogeneous, the created measure is a good proxy for HSA market characteristics. As summarized in Exhibit 2, we controlled for the following market characteristics.

Population size: We controlled for the population size of an HSA because providers in larger HSAs could be more likely to adopt telehealth, and HSA size can also affect population health outcomes through various channels. For instance, larger markets have favorable impacts on provider profitability, which in turn could affect quality of care and health outcomes (Kaufman et al., 2016). We obtained population data from the Census Bureau (2021).

New and cumulative COVID-19 cases per 10,000 people: We created proxies for the prevalence of COVID-19 in each HSA during the second half of 2021 and used them to control for the pandemic's effect on telehealth use and health outcomes. We controlled for both the number of newly confirmed COVID-19 cases during the second semester of 2021 and the cumulative COVID-19 cases up until the second semester of 2021 per 10,000 people. We obtained the number of COVID-19 cases from the *New York Times* database (2021).

2.5. Empirical Strategy

We conducted a regression-based DID analysis to identify how telehealth intensity affected the study outcomes. The DID analysis identifies the effect of the telehealth intensity by comparing the average change in an outcome for HSAs with Medium or High telehealth intensity between the second semester of 2019 and the second semester of 2021 with the average change in that outcome for HSAs with Low telehealth intensity during the same period. The regression-based DID approach automatically controls for any baseline differences in outcome levels between the three groups and for any time-invariant characteristic affecting the outcomes at the HSA level. Time-varying confounders are not automatically adjusted; to adjust for these confounders, we included covariates in the model (described in **2.4. COVARIATES**). We implemented this approach by estimating the following equation for each outcome using data from the second semester of 2019 and 2021:

$$Y_{it} = \beta_0 + \beta_1 M_i + \beta_2 H_i + \beta_3 Post_t + \boldsymbol{\beta_4} M_i \times Post_t + \boldsymbol{\beta_5} H_i \times Post_t + X_{it}\Gamma + \varepsilon_{it} \quad (1)$$

where Y_{it} is an outcome of interest for HSA i at time t (e.g., ACS hospitalizations). M_i and H_i are indicators for HSAs that have Medium and High telehealth intensity, respectively. $Post_t$ is an indicator for the treatment period (i.e., second semester of 2021). X_{it} is a matrix containing time-varying confounders, and ε_{it} is the error term. β_4 and β_5 are the coefficients of interest and provide the effect of Medium and High telehealth intensity on the outcome, relative to the Low telehealth intensity, after controlling for change in time-varying confounding factors. β_3 provides context for the coefficients of interest; it shows the average change in the outcome between the baseline and treatment period for Low telehealth intensity HSAs. Following standard practice, we estimated heteroskedastic robust standard errors and clustered them at the HSA level.

The key assumption of DID models for producing reliable results is that the trajectory of the outcomes (after controlling for covariates) would have been identical for the Low, Medium, and High telehealth intensity groups had the telehealth expansion not happened. This is usually referred to as the "parallel trends assumption." Even though this assumption is not directly testable, we can approximate it by testing for the existence of differences in trends in the outcomes between the three groups during the baseline period. We test this assumption by estimating the following DID equation using data from the first and second semesters of 2018 and 2019:

$$Y_{it} = \alpha_0 + \alpha_1 M_i + \alpha_2 H_i + \alpha_3 Pre_{t1} + \alpha_4 Pre_{t2} + \alpha_5 Pre_{t3} + \boldsymbol{\alpha}_6 M_i \times Pre_{t1} \\ + \boldsymbol{\alpha}_7 M_i \times Pre_{t2} + \boldsymbol{\alpha}_8 M_i \times Pre_{t3} + \boldsymbol{\alpha}_9 H_i \times Pre_{t1} + \boldsymbol{\alpha}_{10} H_i \times Pre_{t2} \\ + \boldsymbol{\alpha}_{11} H_i \times Pre_{t3} + X_{it}\Delta + \epsilon_{it} \quad (2)$$

where Pre_{t1} is an indicator for the first semester of 2018, Pre_{t2} is an indicator for the second semester of 2018, and Pre_{t3} is an indicator for the first semester of 2019 (the second semester of 2019 is omitted from the regression and serves as the reference time period). The coefficients $\boldsymbol{\alpha}_6$ - $\boldsymbol{\alpha}_{11}$ measure the existence of "pre-trends." For instance, the $\boldsymbol{\alpha}_6$ coefficient measures the average change in the outcome between the first semester of 2018 and the second semester of 2019 for the Medium telehealth intensity group relative to the change in the Low telehealth intensity group over the same time period. We tested the parallel trends assumption separately for HSAs that have Medium and High telehealth intensity. The parallel trends test passes (or more correctly, we fail to reject it) for HSAs with Medium telehealth intensity if α_6, α_7, and α_8 are jointly zero. If all three coefficients are statistically equal to zero, then it means that the average outcome for Medium and Low intensity HSAs followed a similar path in 2018 and 2019. Otherwise, the data would suggest that the Medium and Low intensity groups behaved differently in the baseline period. Similarly, when testing for parallel trends between High and Low telehealth intensity HSAs, we test whether the α_9, α_{10}, and α_{11} parameters are jointly zero.

3. Results

This chapter has three sections. First, in **3.1. Descriptive Statistics**, we describe the characteristics of FFS Medicare beneficiaries included in our analysis and show that they are consistent with other studies of FFS beneficiaries. We also illustrate some of the systematic differences in demographic and market characteristics between the Low, Medium, and High telehealth intensity HSAs. Next, in **3.2. Main Findings**, we present the DID estimates for the impact of Medium or High telehealth intensity relative to the Low telehealth intensity, on quality, access, and cost outcomes. We also discuss the validity of the DID parallel trends assumption in our study. Lastly, in **3.3. Robustness Checks**, we present additional analyses that add more confidence in our main findings.

3.1. Descriptive Statistics

3.1.1. Beneficiary Characteristics

We examined the characteristics of our samples of beneficiaries and found them to be generally consistent with prior analyses of the FFS Medicare population (MedPAC, 2022). The characteristics of the samples were also generally consistent across the time periods. In the second semester of 2019, the sample included 31.6 million FFS Medicare beneficiaries (i.e., those who were alive and entitled to FFS Parts A and B for the full 6 months). In the second semester of 2021, the sample included 2.3 million fewer FFS Medicare beneficiaries because of an increase in MA enrollment. The average age of the beneficiaries in the samples was 71.2 and 71.7 years in the second semesters of 2019 and 2021, respectively. In both samples, most of these beneficiaries were female (around 55%), non-Hispanic White (around 80%), and living in an urban area (around 79%). Thirteen percent had full dual eligibility for Medicaid and Medicare for 6 months in the second semester of both 2019 and 2021. Four percent had partial dual eligibility for 6 months in the second semester of 2019, which declined to 3% in the second semester of 2021. The share of beneficiaries attributed to an APM for at least 1 month increased from 38% in the second semester of 2019 to about 40% in the second semester of 2021. Average ADI decreased from 49 in the second semester of 2019 to 47 in the second semester of 2021. Exhibit 3 provides a summary of these findings.

Exhibit 3. Beneficiary Characteristics

	2019	2021
Number of beneficiaries alive and with FFS Part A and B for 6 months	31,625,338	29,335,407
Average age	71.2	71.7
Percentage female	54.6	54.7
Percentage non-Hispanic White	79.6	80.3
Percentage Black	8.8	7.8
Percentage Hispanic	5.5	5.3
Percentage Asian	2.8	2.9
Percentage with full Medicaid eligibility for 6 months	12.6	12.5
Percentage with partial Medicaid eligibility for 6 months	3.9	3.4

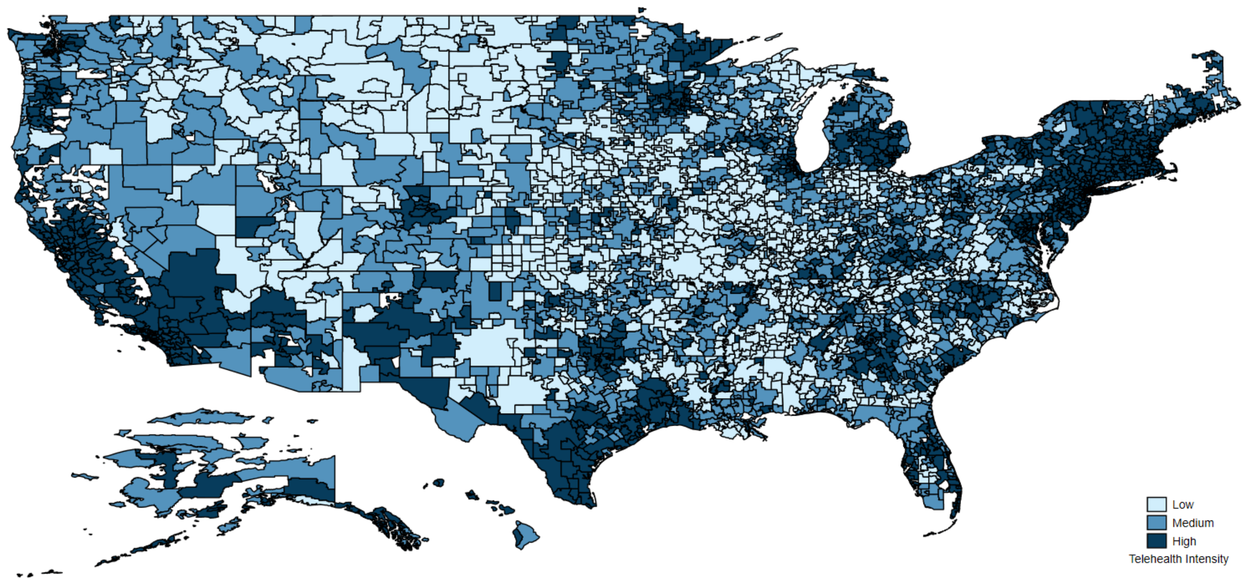
	2019	2021
Percentage attributed to an APM for at least 1 month	38.1	39.6
Percentage living in urban areas	78.4	79.0
Average ADI	48.9	47.3

Note. ADI = Area Deprivation Index; APM = Alternative Payment Model; FFS = fee-for-service. The statistics pertain to the second half of the year.

3.1.2. HSA Characteristics

Exhibit 4 presents a heatmap of telehealth intensity at the HSA level. As described in **2.3.1. Treatment: Telehealth**, we assigned HSAs to Low, Medium, or High telehealth intensity groups according to the terciles of telehealth visits per 1,000 FFS Medicare beneficiaries in the second semester of 2021. Telehealth visits per 1,000 FFS Medicare beneficiaries averaged 174, 311, and 679 visits in Low, Medium, and High telehealth intensity groups, respectively.

Exhibit 4. Low, Medium, and High Telehealth Intensity HSAs



Note. Low, Medium, and High denote telehealth intensity levels in the second semester of 2021.

We then calculated averages for the various characteristics at baseline (before the telehealth expansion) by telehealth intensity (Exhibit 5) and tested whether the averages were statistically different between the three groups in order to empirically select covariates for the DID model (see **2.4. Covariates**).[8] We found that HSAs having Medium or High telehealth intensity in 2021 were different at baseline in many characteristics from the HSAs with Low telehealth intensity. In particular:

[8] At the time of this study, some market characteristics were not available for the second semester of 2021; as a result, we cannot control for them in the DID regressions. We return to this point in robustness checks.

- *Age and gender composition:* The three groups were similar in terms of average age (around 71 years) and gender composition (around 54% were female).
- *Racial composition and income:* While the Low and Medium groups were similar in terms of racial composition and income, the High group was more diverse and included a higher percentage of low-income beneficiaries (using eligibility for Medicaid as a proxy for income). The average percentage of FFS Medicare beneficiaries who were non-Hispanic White decreased from around 88% in the Low and Medium groups to around 78% in the High group. The percentage of FFS Medicare beneficiaries with full Medicaid eligibility increased from around 11% in the Low and Medium groups to around 15% in the High group. This finding is consistent with MedPAC's prior findings that beneficiaries with lower incomes use telehealth services more frequently (MedPAC, 2023).
- *Urban residency and ADI:* There were substantial differences in the percentage of FFS Medicare beneficiaries living in urban areas between the three groups. The average percentage was around 24%, 41%, and 77% for the Low, Medium, and High groups, respectively. This is consistent with MedPAC's prior findings that beneficiaries living in urban areas use telehealth more frequently (MedPAC, 2023). Average ADI showed a similar pattern, but in the opposite direction. The average ADI was around 73 for the Low group, decreasing to 66 for the Medium group, and further decreasing to 45 for the High group.
- *Market characteristics:* The Low group had more hospital beds but fewer primary care physicians per 10,000 people than the Medium group, which in turn showed the same pattern relative to the High group. On average, the Low group had 37 hospital beds per 10,000 people, while the Medium group had 29 and the High group had 26. On average, the Low group had 10 primary care physicians per 10,000 people compared with 11 physicians in the Medium group and 15 in the High group.

Exhibit 5. Average Characteristics by Telehealth Intensity at Baseline

	Low	Medium	High
Average age	71.3	70.7	71.1
Percentage female	53.7	53.4	54.5
Percentage non-Hispanic White	88.2	87.0	77.5
Percentage Black	5.4	5.9	7.1
Percentage Hispanic	2.2	2.8	8.1
Percentage Asian	0.3	0.6	3.6
Percentage with full Medicaid eligibility for 6 months	11.0	11.8	15.2
Percentage with partial Medicaid eligibility for 6 months	5.1	5.1	3.8
Percentage attributed to an APM for at least 1 month	33.3	34.3	36.9
Percentage living in urban areas	24.4	41.2	76.8
Average ADI	73.0	66.2	44.9
Hospital beds per 10,000 people	37.3	29.2	26.1
Primary care physicians per 10,000 people	10.5	11.5	15.0

Note. ADI = Area Deprivation Index; APM = Alternative Payment Model. Low, Medium, and High denote telehealth intensity groups in the second semester of 2021. All statistics are an average over HSAs and pertain to the second semester of 2019.

3.2. Main Findings

3.2.1. Impact Estimates

Exhibit 6 presents a summary of the study findings. All impact estimates presented are significant at the 90% level or above. We describe these findings in detail in the sections that follow. The implications and possible interpretations of these results are discussed in **4. DISCUSSION**.

Exhibit 6. Impact of Telehealth Intensity on Quality, Access, and Cost

	Medium relative to Low	High relative to Low
ACS hospitalizations per 1,000 beneficiaries		↗
ACS ED visits per 1,000 beneficiaries		
Clinician encounters per beneficiary		↗
Primary care physicians		↗
Specialists	↘	↗
APRNs/PAs	↗	↗
Other practitioners		↗
Hospitalists		↗
Total cost of care per beneficiary		↗
Inpatient		↗
Outpatient		↘
Skilled nursing facility		↗
Home health	↘	↘
Hospice		
Physician		↗
Durable medical equipment	↗	↗

Note. ACS = ambulatory care sensitive; APRNs = advanced practice registered nurses; PA = physician assistant. Blue (or ↗) denotes a statistically significant increase. Red (or ↘) denotes a statistically significant decrease. Blank denotes a statistically insignificant result. Impact estimates show the change between the second semester of 2019 and the second semester of 2021 relative to Low telehealth intensity.

Quality Outcomes: HSAs with High telehealth intensity were associated with an increase in ACS hospitalizations, but there was no effect on ACS ED visits. Also, there were no effects on ACS hospitalizations or ED visits for Medium telehealth intensity.

Exhibit 7 presents the findings. For context, the overall trend across HSAs in the Low telehealth intensity group was a sharp decline of 8.14 ACS hospitalizations per 1,000 beneficiaries (32.05% of the baseline rate)[9] and a decline of 12.12 ACS ED visits per 1,000 beneficiaries per semester

[9] The percentages report the estimate as a share of the average for that outcome in the relevant telehealth group in the second semester of 2019. This is to provide a sense of magnitude.

(26.22% of the baseline rate) between the second semester of 2019 and the second semester of 2021.

Medium telehealth intensity: The analysis did not show a statistically significant effect for ACS hospitalizations or ED visits.

High telehealth intensity: The High group was associated with an increase of 1.63 ACS hospitalizations per 1,000 beneficiaries per semester, which represents a 6.90% increase. Given the decline of 8.14 ACS hospitalizations in the Low group, the impact estimate implies that ACS hospitalizations decreased on average among High telehealth intensity HSAs, but at a slower rate than in the Low telehealth intensity HSAs. There was no statistically significant effect for ACS ED visits.

Exhibit 7. Change in ACS Events per 1,000 Beneficiaries per Semester

	Low	Medium relative to Low	High relative to Low
Hospitalizations	-8.14 (-32.05%) ***	0.41 (1.70%)	1.63 (6.90%) ***
ED visits	-12.12 (-26.22%) ***	-0.58 (-1.29%)	0.10 (0.27%)

Note. ACS = ambulatory care sensitive; ED = emergency department. Low, Medium, and High denote telehealth intensity. Estimates show the change between the second semester of 2019 and the second semester of 2021. The denominator for the percentages is that group's average in the second semester of 2019. ***, **, and * denote statistical significance at 1, 5, and 10% level, respectively.

Access Outcomes: HSAs with High telehealth intensity were associated with an increase in clinician encounters per beneficiary, but there was no effect for HSAs with Medium telehealth intensity.

Exhibit 8 presents the findings. For context, the overall trend across HSAs in the Low group was a decline of 0.23 clinician encounters per beneficiary per semester between the second semester of 2019 and the second semester of 2021, which is 2.72% of the baseline rate.

Medium telehealth intensity: There was no statistically significant difference between the Medium and Low groups in total clinician encounters per beneficiary per semester. However, there was a 0.04 increase in encounters with APRNs and PAs (a 2.77% increase relative to the base rate).[10] This was offset by a 0.03 decrease in encounters with specialists (a 0.62% decrease relative to the base rate), which is why there is no statistically significant effect for encounters with all types of clinicians. The analysis showed no difference between the Medium and Low groups for encounters with primary care physicians, hospitalists, or other practitioners.

High telehealth intensity: The High group was associated with an increase of 0.30 clinician encounters per beneficiary per semester, which is 2.67% of the baseline rate. While the analysis shows that the largest increases in the level of clinician encounters occurs among specialists, for whom encounters increased by 0.09 per beneficiary, this is the smallest change relative to the

[10] Note that because there was an overall decline in APRN/PA visits between the second semester of 2019 and the second semester of 2021, this increase represents a slower rate of decline for the Medium intensity group.

baseline rate (1.49%). The largest increases for the High telehealth intensity HSAs come from encounters with hospitalists, which increased by 0.03 (7.11% of the baseline rate). Similarly, encounters with APRNs/PAs increased by 0.07 visits (6.60% of the baseline rate) and encounters with other practitioners increased by 0.07 visits (4.00% of the baseline rate). Encounters with primary care physicians increased by 0.04 per beneficiary per semester during the study period (1.95% of the baseline rate).

Exhibit 8. Change in Clinician Encounters per Beneficiary per Semester

	Low	Medium relative to Low	High relative to Low
All clinicians	-0.23 (-2.72%) *	0.02 (0.22%)	0.30 (2.67%) ***
Primary care physicians	-0.07 (-6.00%)	0.00 (0.14%)	0.04 (1.95%) **
Specialists	-0.03 (-0.61%)	-0.03 (-0.62%) *	0.09 (1.49%) ***
APRNs/PAs	0.10 (8.21%) **	0.04 (2.77%) ***	0.07 (6.60%) ***
Other practitioners	-0.25 (-18.45%) ***	0.01 (0.94%)	0.07 (4.00%) ***
Hospitalists	0.02 (7.06%) *	0.00 (0.57%)	0.03 (7.11%) ***

Note. APRNs = advanced practice registered nurses; PAs = physician assistants. Low, Medium, and High denote telehealth intensity. Estimates show the change between the second semester of 2019 and the second semester of 2021. The denominator for the percentages is that group's average in the second semester of 2019. ***, **, and * denote statistical significance at 1, 5, and 10 % level, respectively.

Cost Outcomes: HSAs with High telehealth intensity were associated with an increase in the total cost of care per beneficiary, but there was no effect for HSAs with Medium telehealth intensity.

Exhibit 9 presents the findings. For context, the overall trend across HSAs in the Low group was an increase in the average total cost of care per beneficiary per semester of $424.54 between the second semester of 2019 and the second semester of 2021 (6.92% of the baseline rate).

Medium telehealth intensity: There was no statistically significant difference in the total cost of care per beneficiary per semester, but there were two significant changes to specific categories of costs that approximately offset one another. Home health costs decreased relative to the Low telehealth intensity HSAs by $7.44 (3.75% of the baseline) while durable medical equipment costs increased by $3.35 (1.87% of the baseline). Although these estimates do not exactly offset one another, the resulting difference is not statistically different from zero. The analysis shows no statistically significant difference between the Medium and Low groups for the cost of inpatient care, outpatient care, skilled nursing facility care, hospice care, or physician services.

High telehealth intensity: The High group was associated with an overall $164.99 increase in the total cost of care per beneficiary per semester, an increase of 2.47% of the baseline cost. The largest increase relative to the baseline averages comes from skilled nursing facility care, which increased by $45.14 in the study period (9.51% of the baseline average). Costs for durable medical equipment increased by $9.68 (6.12% of the baseline average), costs for physician services increased by $100.54 (4.89% of the baseline average), and costs for inpatient care increased by $63.63 (3.22% of the baseline average). In contrast, costs for home health decreased by $20.47 (7.60% of the baseline average) and costs for outpatient care decreased by $31.93 (1.98% of the baseline average). The analysis showed no statistically significant effect for hospice care.

Exhibit 9. Change in Total Cost of Care per Beneficiary per Semester (in Dollars)

	Low	Medium relative to Low	High relative to Low
All claim types	424.54 (6.92%) ***	18.08 (0.30%)	164.99 (2.47%) ***
Inpatient	21.00 (1.29%)	5.66 (0.33%)	63.63 (3.22%) ***
Outpatient	138.16 (6.83%) **	-22.07 (-1.19%)	-31.93 (-1.98%) *
Skilled nursing facility	24.75 (4.69%)	25.90 (5.94%)	45.14 (9.51%) **
Home health	101.75 (57.52%) ***	-7.44 (-3.75%) **	-20.47 (-7.60%) ***
Hospice	24.53 (22.38%) **	0.63 (0.55%)	-1.60 (-1.27%)
Physician	128.55 (8.67%) ***	12.05 (0.74%)	100.54 (4.89%) ***
Durable medical equipment	-14.21 (-7.52%)	3.35 (1.87%) *	9.68 (6.12%) ***

Note. Low, Medium, and High denote telehealth intensity. Estimates show the change between the second semester of 2019 and the second semester of 2021. The denominator for the percentages is that group's average in the second semester of 2019. ***, **, and * denote statistical significance at 1, 5, and 10% level, respectively.

3.2.2. The Parallel Trends Assumption

In general, the DID estimates capture associations between telehealth intensity and the outcomes. A necessary condition to interpret these associations as causal (i.e., that the difference in outcomes between the Low, Medium, and High telehealth intensity groups is due to the difference in the usage of telehealth services) is that the Low telehealth intensity group provides a valid counterfactual for the outcomes in the Medium or High groups in the absence of telehealth expansion. This assumption is not testable, but we have more confidence in its validity if the outcomes for the Low, Medium, and High telehealth intensity HSAs moved in parallel (i.e., had similar patterns) before the expansion of telehealth. This is usually referred to as the parallel trends assumption test.

We checked whether there is a statistically significant difference in outcomes between the Low and Medium groups and between the Low and High groups in 2018 and 2019 after adjusting for the differences in the covariates (see 2.5. Empirical Strategy). We found that for Medium telehealth intensity HSAs, the parallel trends test passes for three of the four main outcomes (ACS hospitalizations, ACS ED visits, and total cost of care) and fails for the clinician encounters outcome. In addition, for High telehealth intensity HSAs, the parallel trends test passes for two of the four main outcomes (ACS ED visits and total cost of care) and fails for the other main outcomes (ACS hospitalizations and clinician encounters). However, a visual inspection of the trends during the same period shows that the trends are largely similar. For instance, looking at a graph of the trends for clinician encounters, we see that HSAs in the Low, Medium, or High groups did have very similar trends before the expansion of telehealth even though the trends were statistically different (Exhibit B-4.1 in Appendix B). Thus, we conclude that the violations of the parallel trends assumption detected by the formal tests are primarily driven by the high precision of our estimates. The small magnitude of the differences suggests that these violations are not a major concern.

3.3. Robustness Checks

3.3.1. Controlling for Geographic Adjustment Factors

Medicare payments change from year to year because of changes in the geographic adjustment factors (GAFs). Because the GAFs vary across HSAs and because they enter the payments multiplicatively (as opposed to additively), a DID strategy cannot perfectly adjust for them, and it is better to make the cost comparable across HSAs before estimating the model. CMS has developed standardized payment amounts for this purpose. Standardized payment amounts are hypothetical Medicare payments calculated as if claims were priced based on the national amounts without adjusting for GAFs or including other factors that make a cross-sectional comparison invalid, such as payments for indirect medical education (CMS, 2020). However, because of lack of access to complete data on standardized payment amounts, we used the actual paid amount. Therefore, it is possible that results are driven by the differences in the GAFs as opposed to differences in resource use between areas of varying telehealth intensities.

As a robustness check, we constructed a measure of the hospital wage index from the Inpatient Prospective Payment System (IPPS) and measures of geographic practice cost indexes (physician work, practice expense, and malpractice insurance) from the PFS by HSA and semester. We included these four measures as covariates in the DID regressions for all outcomes, but their inclusion did not affect the findings qualitatively; there were only modest changes in impact estimates. For example, after controlling for GAFs, the estimate for the impact of High telehealth intensity on the total cost of care, relative to Low telehealth intensity, decreased only by $4.25 per beneficiary per semester, from $164.99 to $160.74, which is a negligible amount. Exhibit 10 presents the impact estimates after adjusting for the differences in GAFs for the main outcomes.

Exhibit 10. Effect of Controlling for GAFs on the Impact Estimates

Outcome	Unadjusted impact	Adjusted impact
ACS hospitalizations per 1,000 beneficiaries per semester	1.63 (6.90%) ***	1.34 (5.69%) ***
ACS ED visits per 1,000 beneficiaries per semester	0.10 (0.27%)	0.19 (0.54%)
Clinician encounters per beneficiary per semester	0.30 (2.67%) ***	0.19 (1.65%) ***
Total cost of care per beneficiary per semester (in dollars)	164.99 (2.47%) ***	160.74 (2.41%) ***

Note. ACS = ambulatory care sensitive; ED = emergency department; GAFs = geographic adjustment factors. Impact estimates are for the High telehealth intensity group and they are relative to the Low telehealth intensity group. The denominator for the percentages is the baseline average. ***, **, and * denote statistical significance at 1, 5, and 10% level, respectively.

3.3.2. Controlling for In-Person Utilization

Our main specification does not control for in-person utilization. This is because controlling for the changes in in-person utilization would have prevented us from capturing the effects that telehealth use may have on outcomes through the follow-up in-person visits. Indeed, more telehealth use may lead to more in-person use via downstream effects (Bavafa et al., 2018), which in turn may cause better outcomes. However, not controlling for in-person utilization could cause an omitted variable bias in our findings because the effect of telehealth usage may be confounded with the effect of general health care utilization. Therefore, as a robustness check, we controlled for in-person utilization in a separate analysis.

We constructed a measure of in-person utilization as the number of non-telehealth visits per FFS Medicare beneficiary by HSA and semester and included it as a covariate in the DID regressions for quality and cost outcomes. Exhibit 11 shows that the impact estimates do not change substantially after adjusting for the differences in in-person utilization, reducing concerns that the findings are driven by the omitted variable bias.[11]

Exhibit 11. Effect of Controlling for In-Person Utilization on the Impact Estimates

Outcome	Unadjusted impact	Adjusted impact
ACS hospitalizations per 1,000 beneficiaries per semester	1.63 (6.90%) ***	1.60 (6.80%) ***
ACS ED visits per 1,000 beneficiaries per semester	0.10 (0.27%)	0.24 (0.67%)
Total cost of care per beneficiary per semester (in dollars)	164.99 (2.47%) ***	163.96 (2.46%) ***

Note. ACS = ambulatory care sensitive; ED = emergency department. Impact estimates are for the High telehealth intensity group and they are relative to the Low telehealth intensity group. The denominator for the percentages is the baseline average. ***, **, and * denote statistical significance at 1, 5, and 10% level, respectively.

3.3.3. Excluding Small HSAs

MedPAC's prior work on ACS hospitalizations and ED visit rates has shown that measures calculated using a denominator with fewer than 500 beneficiaries have low reliability. Thus, as a robustness check, we excluded small HSAs from the sample and repeated the analysis. Small HSAs are defined as HSAs that have fewer than 500 FFS beneficiaries during the study period. This definition resulted in the exclusion of 162 HSAs out of 3,436 HSAs that are in the main sample. Exhibit 12 presents the findings for the main outcomes. As the exhibit shows, excluding small HSAs has a negligible effect on the impact estimates.

[11] Because in-person utilization is affected by the telehealth intensity, this approach cannot definitively rule out that a general increase in utilization is driving the results. A more rigorous approach to rule out the possibility that such confounding bias is affecting our estimates would require finding an exogenous source of telehealth intensity. To partially compensate for this limitation, we also tried a different definition for the telehealth intensity—the ratio of telehealth to in-person visits—and found similar results.

Exhibit 12. Effect of Excluding Small HSAs on the Impact Estimates

Outcome	Unadjusted impact	Adjusted impact
ACS hospitalizations per 1,000 beneficiaries per semester	1.63 (6.90%) ***	1.62 (6.90%) ***
ACS ED visits per 1,000 beneficiaries per semester	0.10 (0.27%)	-0.01 (-0.03%)
Clinician encounters per beneficiary per semester	0.30 (2.67%) ***	0.34 (2.95%) ***
Total cost of care per beneficiary per semester (in dollars)	164.99 (2.47%) ***	130.75 (1.96%) ***

Note. ACS = ambulatory care sensitive; ED = emergency department; HSAs = Hospital Service Areas. Impact estimates are for the High telehealth intensity group and they are relative to the Low telehealth intensity group. The denominator for the percentages is the baseline average. ***, **, and * denote statistical significance at 1, 5, and 10% level, respectively.

3.3.4. Propensity Score Weighting

A DID methodology is robust to differences that do not change over time and accounts for differences that do change over time through the covariates included in the model (see **2.5. Empirical Strategy**). However, because the three telehealth intensity groups are different in many observed characteristics, they may also be different in unobserved characteristics. For example, at the time of this study, market characteristics data were unavailable for 2021, and, as a result, we could not include them in the model. Propensity score weighting (PSW) can reduce these concerns by making Low, Medium, and High groups more comparable on observed characteristics; presumably, when the groups are similar on observed characteristics, they are similar on unobserved characteristics as well.

To this aim, we used multinomial logistic regression to estimate the propensity scores for each HSA (i.e., the probability of that HSA having Low, Medium, or High telehealth intensity) based on their observed characteristics during the second semester of 2019. Then, we weighted each HSA in the DID regressions by the inverse of their propensity score for their observed telehealth intensity; this resulted in propensity score-weighted DID impacts. Statistically speaking, these estimates are calculated such that HSAs that are similar to HSAs in their respective group will be weighted less, while HSAs that are similar to HSAs in other groups will be weighted more. This increases comparability and reduces the potential impact of confounders on the results (see **Appendix C. Methodology for Propensity Score Weighted DID** for details).

We noted in **3.1.2. HSA Characteristics** that there were statistically significant differences between Low, Medium, and High telehealth intensity groups for many characteristics. After applying the propensity score weights, the differences were statistically insignificant for the majority of characteristics. The enhanced comparability between the three groups also resulted in the satisfaction of the parallel trends test for many more outcomes. After applying the propensity score weights to the DID regressions, the impact estimates did not change qualitatively for the main outcomes with the exception of ACS ED visits. Investigating this further, we found that the increase in ED visits disappeared if we excluded the small HSAs. In addition, we know from MedPAC's prior work that the ACS ED visits measure is unreliable when HSAs with a small number of beneficiaries are included in the analysis. Thus, propensity score weighted DID estimates concur with the main findings, which increases our confidence in them. Exhibit 13 presents a comparison of the main outcomes.

Exhibit 13. Effect of Propensity Score Weighting on the Impact Estimates

Outcome	Unadjusted impact	Adjusted impact
ACS hospitalizations per 1,000 beneficiaries per semester	1.63 (6.90%) ***	1.44 (6.12%)
ACS ED visits per 1,000 beneficiaries per semester	0.10 (0.27%)	8.90 (24.70%) **
Clinician encounters per beneficiary per semester	0.30 (2.67%) ***	0.24 (2.17%) **
Total cost of care per beneficiary per semester (in dollars)	164.99 (2.47%) ***	260.18 (3.90%) **

Note. ACS = ambulatory care sensitive; ED = emergency department. Impact estimates are for the High telehealth intensity group and they are relative to the Low telehealth intensity group. The denominator for the percentages is the baseline average. ***, **, and * denote statistical significance at 1, 5, and 10% level, respectively.

3.3.5. Separating Rural and Urban HSAs

One of the biggest differences between the three telehealth intensity groups is the level of urbanicity. As described earlier, the average percentage of beneficiaries living in an urban area was 24%, 41%, and 77% for the Low, Medium, and High groups, respectively. Given the higher urbanicity of the High telehealth intensity group, it may be concerning that the differences in outcomes are due to differences in urbanicity levels, rather than due to differences in telehealth usage. Therefore, as a robustness check, we separated HSAs into rural and urban subsamples and repeated the DID analysis. If the differences in outcomes were due to differences in urbanicity, then we would expect telehealth usage to have no impact on the outcome measures within the urban or rural subsample.

We split the sample by the median share of beneficiaries living in urban areas in the second semester of 2019 (15.7%). HSAs above the median were classified as urban, and HSAs below the median were classified as rural.[12] Exhibit 14 presents a cross-tabulation of HSAs by telehealth intensity and urbanicity. As expected, most HSAs in the urban subsample are in the High telehealth intensity group while most HSAs in the rural subsample are in the Low telehealth intensity group.

Exhibit 14. Cross-Tabulation of HSAs by Telehealth Intensity and Urbanicity

	Urban	Rural
Low	305	829
Medium	514	655
High	899	234
Total	1,718	1,718

[12] Choosing other values (e.g., the mean [47.4%]), as the threshold for urbanicity does not affect the analysis because most HSAs are completely urban or rural.

Exhibit 15 presents the findings for the main outcomes. The impact estimates for the High group based on the urban and rural subsamples show the same pattern as in the full sample, although the impact estimates for the rural subsample are smaller. For ACS hospitalizations, the result is no longer significant for the rural subsample, but this is likely due to a smaller sample size. In addition, while the High group in the rural subsample does not show an effect on the total cost of care for all claim types combined, it shows an increase in the total cost of care for physician services, which is consistent with the increase in clinician encounters per beneficiary. Thus, overall, the findings indicate that limiting the sample to only urban or rural HSAs does not eliminate the association between High telehealth intensity and outcomes. [13]

Exhibit 15. Effect of Splitting the Sample by Urbanicity on the Impact Estimates

Outcome	Full sample	Urban sample	Rural Sample
ACS hospitalizations per 1,000 beneficiaries per semester	1.63 (6.90%) ***	1.30 (5.61%) ***	0.92 (3.65%)
ACS ED visits per 1,000 beneficiaries per semester	0.10 (0.27%)	-0.08 (-0.25%)	-1.00 (-2.02%)
Clinician encounters per beneficiary per semester	0.30 (2.67%) ***	0.36 (3.06%) ***	0.19 (2.09%) **
Total cost of care per beneficiary per semester (in dollars)	164.99 (2.47%) ***	211.69 (3.13%) ***	2.31 (0.04%)

Note. ACS = ambulatory care sensitive; ED = emergency department. Impact estimates are for the High telehealth intensity group and they are relative to the Low telehealth intensity group. The denominator for the percentages is the baseline average. ***, **, and * denote statistical significance at 1, 5, and 10% level, respectively.

[13] A potential explanation for the smaller impact estimates for rural HSAs is that telehealth usage is lower in rural areas. For reference, in the High telehealth intensity group, the median (mean) level of telehealth services per 1,000 beneficiaries was 507 (568) in the rural sample and 651 (708) in the urban subsample.

4. Discussion

We used Medicare administrative data to estimate the association between telehealth use and population-based outcomes for quality of care (as measured by risk-adjusted ACS hospitalizations and ED visits per 1,000 beneficiaries), access to care (as measured by clinician encounters per beneficiary), and cost of care (as measured by total cost of care per beneficiary), when both telehealth and in-person visits were available to FFS Medicare beneficiaries. We identified the effect of telehealth usage by comparing the average change in an outcome for HSAs with Medium or High telehealth intensity between the second half of 2019 and the second half of 2021 with the average change in that outcome for HSAs with Low telehealth intensity during the same period. We discuss our findings in the following subsections.

4.1 Interpretation of the Results

Telehealth and quality. There was no significant association between telehealth intensity and ACS hospitalizations and ED visits for the areas with Medium telehealth intensity compared to areas with Low telehealth intensity. In areas with High telehealth intensity, however, we found an increase in ACS hospitalizations relative to the Low telehealth intensity areas. Because of the large decrease in ACS hospitalizations overall, this amounts to a slower decline in ACS hospitalizations in the High group than in the Low group. At the same time, we found no effect on ACS ED visits. The slower decline in ACS hospitalizations may seem counterintuitive because we also found relative increases in all types of clinician encounters in the High group, including primary care (discussed next), and because ACS hospitalizations are generally viewed as hospital admissions that could have been prevented with timely ambulatory care. Importantly, the smaller reduction (increase relative to the Low group) in ACS hospitalizations persists even when we control for the rate of in-person clinician services per beneficiary in the area.

Telehealth and access. There was no association between telehealth intensity and total clinician encounters for areas with Medium telehealth intensity. In contrast, the areas with High telehealth intensity had increases in clinician encounters both overall and for each provider type relative to the Low telehealth intensity areas—namely, an increase of 0.30 encounters per beneficiary per semester overall, or 2.67% of the 2019 average. Although our measure of encounters included both telehealth and in-person encounters, the relative size of this increase is in line with the increase in telehealth-only primary care services reported by MedPAC (MedPAC, 2023). The three types of clinicians with the largest relative increases in the number of encounters were hospitalists (7.1%), APRNs/PAs (6.6%), and other practitioners (4.0%). The other practitioners category includes both clinical psychologists and clinical social workers, two groups that MedPAC identified as providing the highest rates of telehealth services per clinician (MedPAC, 2023). The finding on APRNs/PAs is likely to have been influenced by the PHE, as there were 28 states with waived or suspended licensure or practice agreements for APRNs that at least partially overlapped with our study period in the second half of 2021 (NCSBN, 2022). Many of these waivers reduced the restrictions on the types of care APRNs could provide and the amount of direct supervision that was needed, so the number of encounters that APRNs and PAs were permitted to supervise was at an elevated level, both in person and via telehealth. Additionally, some encounters with APRNs and PAs may be attributed to physicians via “incident to” billing, so these figures likely underestimate the magnitude of encounters with APRNs and PAs.

Telehealth and costs. Areas with Medium telehealth intensity had no significant change in the overall average cost of care, which is consistent with our finding that this group did not have increased clinician encounters during the period. The High telehealth intensity areas experienced increases in the total cost of care approximately in line with the increase in encounters as well: clinician encounters increased 2.67%, while average costs of physician services (after adjusting for GAFs) increased by 3.85% relative to Low telehealth intensity areas. In addition, inpatient costs increased in line with the increases in ACS hospitalizations in these areas. An unexpected result was that both groups experienced a decline in average home health spending relative to Low telehealth intensity areas. This is likely to be a consequence of a phenomenon that is correlated with but not caused by telehealth use that we were unable to completely control for, such as shortages of home health aides that varied by locale and were exacerbated by the PHE (Graham, 2022).

The interpretation of these findings, particularly the association between High telehealth intensity and increased ACS hospitalizations, requires a detailed discussion that includes certain methodological considerations.[14] One potential interpretation is that the increased ACS hospitalizations in the High group may have been driven by a general increase in health care utilization in High telehealth intensity HSAs relative to the Low group, which increased both telehealth and other health care services (including ACS hospitalizations). However, as we mentioned above, the findings persist even when we control for the rate of in-person visits with clinicians. This suggests that a general increase in utilization in the area is not the likely explanation for this finding.

A more plausible interpretation focuses on one of the limitations of this study, namely, the use of the second half of 2021 to approximate post-pandemic outcomes. This time period overlapped with the surge in cases due to the Delta variant of COVID-19, which peaked in early September 2021, and also includes the beginning of the surge in cases due to the Omicron variant, which began in December 2021.[15] In addition, the PHE continued to influence patient behavior in 2021 through delayed needed medical care (NPR, 2021). This could have played a role in the observed increase in ACS hospitalizations for the High telehealth intensity HSAs. For instance, clinicians may identify issues during telehealth encounters which in turn may cause beneficiaries to seek inpatient care rather than further delaying it, leading to a relative increase in ACS hospitalizations for the High telehealth intensity HSAs.

Between the second half of 2019 and the second half of 2021, ACS hospitalizations dropped 32.0% and ACS ED visits dropped 26.2% across all HSAs. These declines are directly related to the COVID-19 pandemic, which disrupted the health care system and fundamentally altered the behavior of the population, the way individuals accessed the health care system, and their likelihood of acquiring ACS conditions (NPR, 2021; Becker, et al. 2022). Although these trends were observed at the national level, differences in the timing and implementation of local regulations on masking, school closures, screening for illnesses, vaccine uptake, and local transmission of other viruses can all affect the degree and timing of these phenomena across communities (Chow et al., 2023). Furthermore, different geographic areas have health care systems

[14] Other studies have found a similar association; for example, see https://www.ncbi.nlm.nih.gov/pmc/articles/PMC8972029/

[15] The exact date of the peak of the Delta variant depends on the metric used; using a 7-day rolling average of new cases per day, the peak occurred on September 1, 2021, with 166,105 cases per day. The Omicron surge began in December 2021, with average daily cases steadily increasing, reaching 384,714 on December 31, 2021.

with different degrees of resilience and different abilities to rebound from the pandemic. All these factors suggest that by the second half of 2021, different geographic areas were at different points in terms of their recovery from the PHE in ways that directly affected our outcomes.

These differences across HSAs are also likely to be correlated with the level of telehealth intensity. For example, systems that have enhanced telehealth capabilities already in place may also be more likely to adapt to pandemic restrictions and challenges relatively quickly (e.g., Whaley et al., 2022). These differences—the timing of the COVID-19 case surges, implementation of and compliance with mask and social distancing mandates, the speed of health care system responses across HSAs, and so forth—are very likely to have an impact on both the telehealth intensity in an area and outcomes such as ACS hospitalizations and clinician encounters. The controls included in our analysis, such as COVID-19 cases, are unlikely to sufficiently adjust for all these confounding factors. For instance, a surge in cases in one area may lead to hospitals in other HSAs reaching maximum capacity due to hospital transfers even if there is no surge in COVID-19 cases in those HSAs (see Ladyzhets & Dutton, 2022; NPR, 2022; and Wermus, 2021 for examples). Scarce health care capacity could lead to delayed primary care (which is likely to result in increased ACS hospitalizations) by some beneficiaries and increased telehealth use by others. Over time, however, we would expect that the large transient effects of the PHE that vary by community would diminish, and patients and health care systems will converge on a "new normal." We discuss some additional PHE-related trends and other possible phenomena that could be influencing these results—such as reverse causality, where increased hospitalizations lead to increased follow-up care via telehealth—in the next section.

4.2 Options for Future Work

In our study, we compared HSAs with Medium or High telehealth intensity to HSAs with Low telehealth intensity. There are two ways to evaluate the validity of this comparison. The first approach consists of testing for differences between the Low, Medium, and High telehealth intensity groups in terms of observable characteristics that are likely to affect our outcomes of interest. The second approach consists of assessing whether the trends in outcomes were similar between the three groups before the expansion of telehealth benefits. We have presented evidence, through the main findings and the robustness checks, that our model performs well in both of these validation approaches. **This suggests that the econometric model used in this study is adequate to conduct an estimation of the causal effects of telehealth on quality, access, and cost outcomes.** However, as we discuss below, we believe a number of additional refinements to the period of study and the outcome measures are required before interpreting these results as causal.

Alternative period of study. As discussed above, we found ample evidence that the PHE continued to influence the way Medicare beneficiaries sought and received health care throughout our period of study. Thus, the first and most simple refinement is to repeat the analysis with data from a time period in which there is less pandemic-related volatility in the outcomes of interest, especially ACS hospitalizations and ED visits. **Therefore, an option for future studies is to use a more recent period of study, such as the second half of 2022**.

Alternative outcome: ACS hospitalizations less likely to be affected by PHE. As noted above, one way that the pandemic has influenced our outcomes is in terms of substantial reductions in ACS hospitalizations. Other studies confirm this observation and provide a useful nuance. For example, a recent study shows that respiratory-related ACS hospitalizations have declined substantially, starting early in the pandemic (around March 2020) and continued decreasing until the end of that study (March 2021), but other ACS hospitalizations, particularly those related to diabetes, were much less affected (Becker et al., 2022). Other evidence, such as the substantially reduced incidence of influenza in the 2021–2022 flu season, also supports the intuitive finding that PHE-related public health measures reduced the incidence of respiratory conditions in particular (CDC, 2022). This suggests that we may be able to more accurately approximate trends by narrowing our quality measure to consider outcomes that have suffered less pronounced effects from the pandemic. **Therefore, an option for future studies is to study respiratory and non-respiratory ACS hospitalizations separately**. This could be an inclusive outcome (e.g., all ACS hospitalizations, excluding pneumonia, chronic obstructive pulmonary disease, or asthma) or more specific outcomes corresponding to non-respiratory ACS conditions, such as diabetes, hypertension, heart failure, and/or urinary tract infection. Similarly, because intensive care unit (ICU) admissions are associated with the most advanced stages of a condition, it constitutes an outcome of relevance for this study and may be less influenced by the pandemic relative to hospitalizations. As such, **an option for future studies is to include ICU admissions among ACS hospitalizations.** By minimizing the impact of COVID-19-related changes that could differ across HSAs, we would expect the finding of increased ACS hospitalizations for the High telehealth HSAs to be substantially reduced or possibly even reversed.

Alternative outcome: Non-ACS hospitalizations. If the association between high telehealth intensity and increased ACS hospitalizations is, as we conjecture, because of the differential impact of the PHE across communities by the second half of 2021, then we would likely observe a difference in the estimates for ACS versus non-ACS hospitalizations. That is, if the PHE is biasing the results, we would expect an increase in both ACS and non-ACS hospitalizations for these HSAs. If telehealth utilization improves access to and quality of ambulatory care, then we would expect that ACS hospitalizations would increase at a slower rate than non-ACS hospitalizations for these HSAs. The difference between the two would convey some information on the beneficial effects of increased telehealth intensity on quality of care. **Therefore, an option for future studies is to conduct additional tests to explore the differential effect of telehealth on ACS hospitalizations relative to other types of hospitalizations.**

Alternative outcome: Timing of telehealth and outcomes. Lastly, medically complex patients are more likely to use telehealth (Dixit et al., 2022; Hatef et al., 2022). If these patients are more likely to be hospitalized because of their medical needs and are receiving follow-up care via telehealth, the higher rate of hospitalizations in an area could cause an increase in telehealth usage rather than vice versa. This scenario could be addressed by identifying the timing of telehealth usage relative to the outcomes of interest. For instance, we could measure the association between telehealth intensity in 2021 and ACS hospitalizations in 2022. **An option for future studies is to assess the timing of telehealth encounters relative to the outcomes of interest to understand the mechanism underlying any association between telehealth and hospitalizations or other utilization measures.**

5. Conclusions

We found that increased telehealth usage was associated with improved access to care (as measured by clinician encounters) and increases in the total cost of care, but only for the High telehealth intensity HSAs, and not for the areas with Medium telehealth intensity. The growth in costs was commensurate with the growth in access. In addition, we found that increased telehealth usage was associated with an increase in ACS hospitalizations, but only for High telehealth intensity areas. Finally, we did not find an association between telehealth usage and ACS ED visits for Medium or High telehealth intensity areas.

Overall, our study suggests that using population-based outcomes to inform Medicare's decision on the permanent expansion of telehealth is feasible. The main concern we have encountered in this respect relates to the finding of an increase in ACS hospitalizations in High telehealth intensity areas; this counterintuitive result is likely driven by the fact that, given the timing of the study, we had to use the second half of 2021 to represent "normal times"—in other words, to mimic the post-COVID-19 PHE period. As more recent data become available, we expect the volatility in the data associated with the ebb and flow of the pandemic's care disruptions to be greatly reduced, diminishing the risk of confounding and improving the reliability of the estimates. In addition, the findings presented here yield valuable lessons that would prove useful for future studies focused on the impacts of telehealth. These include recommending the use of alternative outcome measures that are less likely to be affected by the pandemic, such as non-respiratory ACS hospitalizations or ACS ICU stays, which could also help reduce confounding. Finally, alternative outcome measures—such as the rate of ACS hospitalizations relative to other types of hospitalizations or the timing of telehealth encounters relative to hospitalizations—could also help produce a more reliable and nuanced understanding of the impact of telehealth on the quality of care.

References

Assistant Secretary for Planning and Evaluation. (2020). *Medicare beneficiary use of telehealth visits: Early data from the start of the COVID-19 pandemic*. https://aspe.hhs.gov/sites/default/files/private/pdf/263866/hp-issue-brief-medicare-telehealth.pdf

Bavafa, H., Hitt, L. M., & Terwiesch, C. (2018). The impact of e-visits on visit frequencies and patient health: Evidence from primary care. *Management Science, 64*(12), 5461–5480. https://doi.org/10.1287/mnsc.2017.2900

Becker, N. V., Karmakar, M., Tipirneni, R., & Ayanian, J. Z. (2022). Trends in hospitalizations in ambulatory care-sensitive conditions during the COVID-19 pandemic. *JAMA New Open, 5*(3). https://jamanetwork.com/journals/jamanetworkopen/fullarticle/2790207

Bose, B., Dun, C., Zhang, G. Q., Walsh, C., Makary, M. A., & Hicks, C. W. (2022). Medicare beneficiaries in disadvantaged neighborhoods increased telemedicine use during the COVID-19 pandemic. *Health Affairs, 41*(5), 635–642. https://www.healthaffairs.org/doi/10.1377/hlthaff.2021.01706

Centers for Disease Control and Prevention. (2022). *2021–2022 estimated influenza burden*. https://www.cdc.gov/flu/about/burden/2021-2022.htm

Centers for Medicare & Medicaid Services. (2022, October). *COVID-19 emergency declaration blanket waivers for health care providers*. https://www.cms.gov/files/document/covid-19-emergency-declaration-waivers.pdf

Centers for Medicare & Medicaid Services. (2022, June 17). *List of telehealth services*. https://www.cms.gov/Medicare/Medicare-General-Information/Telehealth/Telehealth-Codes

Centers for Medicare & Medicaid Services. (2020, May). *CMS standardization methodology for allowed amount—v.10.* https://resdac.org/sites/datadocumentation.resdac.org/files/CMS%20Part%20A%20and%20Part%20B%20Price%20%28Payment%29%20Standardization_Detailed%20Methods_2.pdf

Chow, E. J., Uyeki, T. M., & Chu, H. Y. (2023). The effects of the COVID-19 pandemic on community respiratory virus activity. *Nature Reviews Microbiology, 21,* 195–210. https://www.nature.com/articles/s41579-022-00807-9

Chronic Conditions Warehouse. (2022). *Getting started with CMS Medicare administrative research files*. https://www2.ccwdata.org/documents/10280/19002248/ccw-technical-guidance-getting-started-with-cms-medicare-administrative-research-files.pdf

Dartmouth Atlas Project. (2022a). *FAQ*. https://www.dartmouthatlas.org/faq/

Dartmouth Atlas Project. (2022b). *Supplemental data: Crosswalks.* https://data.dartmouthatlas.org/supplemental/#crosswalks

Daw, J. R., & Hatfield, L. A. (2018). Matching and regression to the mean in difference-in-differences analysis. *Health Services Research, 53,* 4138–4156. https://doi.org/10.1111/1475-6773.12993

Dixit, R. A., Ratwani, R. M., Bishop, J. A., Schulman, K., Sharp, C., Palakanis, K., & Booker, E. (2022). The impact of expanded telehealth availability on primary care utilization. *Npj Digital Medicine, 5*(141). https://doi.org/10.1038/s41746-022-00685-8

Eberly, L. A., Kallan, M. J., & Julien, H. M. (2020). Patient characteristics associated with telemedicine access for primary and specialty ambulatory care during the COVID-19 pandemic. *JAMA Network Open, 3*(12). https://jamanetwork.com/journals/jamanetworkopen/fullarticle/2774488

Eicheldinger, C., & Bonito, A. (2008). More accurate racial and ethnic codes for Medicare administrative data. *Health Care Financing Review, 29*(3), 27–42. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4195038/

Federal Communications Commission. (2022, June 21). *Form 477 county data on internet access services*. https://www.fcc.gov/form-477-county-data-internet-access-services

Feng, Z., Silver, B., Segelman, M., Jones, M., Ingber, M. J., Beadles, C., & Pickett, R. (2019). *Developing risk-adjusted avoidable hospitalizations and emergency department visits quality measures*. Prepared for the Medicare Payment Advisory Commission. https://www.medpac.gov/wp-content/uploads/import_data/scrape_files/docs/default-source/contractor-reports/august2019_riskadjusted_ah_av_measures_contractor_sec.pdf

Graham, J. (2022, February 3). *Pandemic-fueled shortages of home health aides strand patients without care*. CNN. https://www.cnn.com/2022/02/03/health/home-health-care-aide-shortage-khn-partner-wellness/index.html

Hatef, E., Lans, D., & Bandeian, S. (2022). Outcomes of in-person and telehealth ambulatory encounters during COVID-19 within a large commercially insured cohort. *JAMA Network Open, 5*(4). https://jamanetwork.com/journals/jamanetworkopen/article-abstract/2791531

Health Resources & Services Administration. (n.d.). *Data downloads.* https://data.hrsa.gov/data/download

Karimi, M., Lee, E. C., Couture, S. J., Gonzales, A. B., Grigorescu, V., Smith, S. R., De Lew, N., & Sommers, B. D. (2022, February 1). *National survey trends in telehealth use in 2021: Disparities in utilization and audio vs. video services* (Research Report No. HP-2022-04). U.S. Department of Health and Human Services, Office of the Assistant Secretary for Planning and Evaluation. https://www.aspe.hhs.gov/reports/hps-analysis-telehealth-use-2021

Kaufman, B. G., Thomas, S. R., Randolph, R. K., Perry, J. R., Thompson, K. W., Holmes, G. M., & Pink, G. H. (2016). The rising rate of rural hospital closures. *The Journal of Rural Health, 32*(1), 35–43. https://pubmed.ncbi.nlm.nih.gov/26171848/

Kind A. J. H., & Buckingham W. (2018). Making neighborhood disadvantage metrics accessible −The Neighborhood Atlas. *New England Journal of Medicine, 378*(26), 2456–2458. https://www.nejm.org/doi/10.1056/NEJMp1802313

Kindig, D. (2015, April 6). What are we talking about when we talk about population health? *Health Affairs Blog*.

https://www.healthaffairs.org/do/10.1377/forefront.20150406.046151/full/

Ladyzhets, B., & Dutton, A. (2022, March 16). Idaho COVID-19 surge drove patient transfers, strained out-of-state hospitals, new data shows. *Idaho Capital Sun*. https://idahocapitalsun.com/2022/03/16/idaho-covid-19-surge-drove-patient-transfers-strained-out-of-state-hospitals-new-data-shows/

Medicare Payment Advisory Commission. (2023, January). *Mandated report on telehealth: Updates on telehealth use and beneficiary and clinician experiences*. MedPAC. https://www.medpac.gov/wp-content/uploads/2023/01/MedPAC-Telehealth-Jan-2023.pdf

Medicare Payment Advisory Commission. (2021, March). *Report to the Congress: Medicare Payment Policy*. https://www.medpac.gov/document/march-2021-report-to-the-congress-medicare-payment-policy/

Medicare Payment Advisory Commission. (2022, March). *Report to the Congress: Medicare payment policy*. https://www.medpac.gov/wp-content/uploads/2022/03/Mar22_MedPAC_ReportToCongress_SEC.pdf

Medicare Payment Advisory Commission. (2021, June). *Report to the Congress: Medicare and the health care delivery system*. https://www.medpac.gov/wp-content/uploads/import_data/scrape_files/docs/default-source/reports/jun21_medpac_report_to_congress_sec.pdf

Medicare Payment Advisory Commission. (2019, June). *Report to the Congress: Medicare and the health care delivery system*. https://www.medpac.gov/wp-content/uploads/import_data/scrape_files/docs/default-source/reports/jun19_medpac_reporttocongress_sec.pdf

Medicare Payment Advisory Commission. (2022, July). *July 2022 data book: Health care spending and the Medicare Program. Chapter 2*. https://www.medpac.gov/wp-content/uploads/2022/07/July2022_MedPAC_DataBook_Sec2_SEC.pdf

NCSBN. (2022). *State response to COVID-19 (APRNs) as of May 2, 2022 – 12:00 PM (CST).* https://www.ncsbn.org/public-files/APRNState_COVID-19_Response.pdf

New York Times. (2021). *Coronavirus (COVID-19) data in the United States*. https://github.com/nytimes/covid-19-data

NPR. (2021, October). *Household experiences in America during the Delta variant outbreak*. https://media.npr.org/assets/img/2021/10/20/National%20Report_Oct2021-FINAL.pdf

Samson, L., Tarazi, W., Turrini, G., & Sheingold, S. (2021, December). *Medicare beneficiaries' use of telehealth services in 2020: Trends by beneficiary characteristics and location* (Issue Brief No. HP-2021-27). U.S. Department of Health and Human Services, Office of the Assistant Secretary for Planning and Evaluation. https://aspe.hhs.gov/sites/default/files/documents/a1d5d810fe3433e18b192be42dbf2351/medicare-telehealth-report.pdf

Stone, W. (2022, February 10). In rural America, patients are waiting for care – sometimes with deadly consequences. *NPR*. https://www.npr.org/sections/health-shots/2022/02/10/1078134622/in-rural-america-patients-are-waiting-for-care-sometimes-with-deadly-consequence

United States Census Bureau. (n.d.). *Small area income and poverty estimates (SAIPE) interactive tool*. https://www.census.gov/data-tools/demo/saipe

United States Census Bureau. (2022, May 16). *ZIP Code tabulation areas (ZCTAs)*. https://www.census.gov/programs-surveys/geography/guidance/geo-areas/zctas.html

United States Census Bureau. (2021, December 17). *Population and housing unit estimates tables*. https://www.census.gov/programs-surveys/popest/data/tables.html

United States Department of Agriculture. (2019, October 24n.d.). *Urban influence codes*. https://www.ers.usda.gov/data-products/urban-influence-codes/

United States Government Accountability Office. (2022, September). *MEDICARE telehealth: Actions needed to strengthen oversight and help providers educate patients on privacy and security risks*. https://www.gao.gov/assets/gao-22-104454.pdf

University of Wisconsin School of Medicine and Public Health. (n.d.) 2015 Area Deprivation Index v2.0. Center for Health Disparity Research. https://www.neighborhoodatlas.medicine.wisc.edu

Wermus, K. (2021, December 20). Patients marooned in emergency rooms for a week because COVID surge prevents transfers. *Newsweek*. https://www.newsweek.com/patients-marooned-emergency-rooms-week-because-covid-surge-prevents-transfers-1661391

Whaley, C. M., Ito, Y., Kolstad, J. T., Cowling, D. W., & Handel, B. (2022) The health plan environment in California contributed to differential use of telehealth during the COVID-19 pandemic. *Health Affairs, 41*(12). https://doi.org/10.1377/hlthaff.2022.00464

Appendix A. Detailed Data Sources

Exhibit A–1.16 Telehealth Codes

Category	HCPCS Codes	Notes
Medicare-approved telehealth services	Codes available from https://www.cms.gov/Medicare/Medicare-General-Information/Telehealth/Telehealth-Codes (updated 6/17/2022; downloaded 8/9/2022)	Claims must also meet one or more of the following requirements: HCPCS modifier code = GQ or 95 or GT or G0. GQ is for asynchronous services in Alaska or Hawaii. 95 is for services provided after 3/1/2020. GT applies to distant site services billed under Critical Access Hospital method II on institutional claims.[a] G0 is for claims for telehealth services that are furnished on or after January 1, 2019, for purposes of diagnosis, evaluation, or treatment of symptoms of an acute stroke.[b] For carrier claims, PLCSRVC (place of service) = 02, which was required to be used for distant site services provided before 3/1/2020 and may also be used after that date.
	99201	HCPCS 99201 was on the CMS list of telehealth services for 2019 and 2020 but does not appear on the most recent list because the code was eliminated after 2020. Therefore, include HCPCS 99201 in 2020 and earlier years if claims meet any of the requirements in the previous row.
Remote services	99453, 99454, or 99457	These codes are specifically designed to allow for remote monitoring and fewer in-person visits.[c]
	Codes specific to Innovation Center Models NGACO: G9481–G9489; G0438–G0439[d] BPCI Advanced: G9978–G9986[e]	POS 12 (beneficiary's home) should be used when the beneficiary's place of residence was the originating site (applicable to all NGACO telehealth billing codes G9481–G9489). Annual Wellness Visits (G0438–G0439) are the exception, in that they are billed with POS 02 when the beneficiary's place of residence was the originating site.
Virtual/e-visit check-ins	Communication technology–based services: G2012, G2010, G2250, or G2251 E-visits: 99421–99423 or G2061–G2063	G2250 & G2251 are for clinicians who cannot bill E/M services, so they should be excluded from analyses focused specifically on primary care providers versus care teams. 99421–99423 and G2061–G2063 were created in 2020, so they should not appear in the 2019 file.

Category	HCPCS Codes	Notes
Telephone E/M codes ("audio only")	99441–99443 for evaluation and management visits provided by telephone[f] or 98966–98968 for telephone assessment and management services by qualified nonphysicians[g]	CMS began paying for these codes on 3/1/2020, so there should be zero records with these codes in the 2019 file.

Note. BPCI = Bundled Payments for Care Improvement; CMS = Centers for Medicare & Medicaid Services; E/M = Evaluation and Management; HCPCS = Healthcare Common Procedure Coding System; NGACO = Next Gen Accountable Care Organization. "Nonphysicians" means anyone who can't bill E/M services (i.e., in this usage, "nonphysicians" does not include physician assistants or nurse practitioners).

[a] Centers for Medicare & Medicaid Services. (2017, December 4). *Elimination of the GT modifier for telehealth services*. https://www.cms.gov/Outreach-and-Education/Medicare-Learning-Network-MLN/MLNMattersArticles/downloads/MM10152.pdf

[b] Centers for Medicare & Medicaid Services. (2018, November 27). *New modifier for expanding the use of telehealth for individuals with stroke*. https://www.cms.gov/Outreach-and-Education/Medicare-Learning-Network-MLN/MLNMattersArticles/Downloads/MM10883.pdf

[c] Association of American Medical Colleges (AAMC). (2021). *2021 Medicare coverage of remote physiologic monitoring (RPM)*. https://www.aamc.org/media/55306/download

[d] Centers for Medicare & Medicaid Services. (2021, May). *Next generation ACO model telehealth expansion waiver*. https://innovation.cms.gov/files/x/nextgenaco-telehealthwaiver.pdf

[e] InnoviHealth. (n.d.). Coronavirus and telehealth cheatsheet. https://www.findacode.com/medical-code-sets/covid19-card.pdf

[f] Telehealth.HHS.gov. (n.d.). Billing and coding Medicare Fee-for-Service claims. https://telehealth.hhs.gov/providers/billing-and-reimbursement/billing-and-coding-medicare-fee-for-service-claims/

[g] Centers for Medicare & Medicaid Services. (2020, December 3). *2021 annual update to the therapy code list.* https://www.cms.gov/files/document/mm12126.pdf.

Exhibit A–2. Data Sources for HSA Medicare Population Characteristics

Item	Description
A	**Attribution to HSAs:** Items #B to #O are calculated for each HSA, year (2018, 2019, and 2021), and semester (January to June, July to December). All variables are from CME custom enrollment files unless otherwise stated. Beneficiaries are attributed to HSAs using the first valid monthly ZIP Code for that year and semester (invalid ZIP Codes begin with 99999 or 00000). ZIP Code to HSA crosswalk is obtained from https://data.dartmouthatlas.org/supplemental/#crosswalks, the 2019 version.
B	**Medicare beneficiaries:** Records in CME enrollment file that have (a) Part A and B enrollment during the entire semester and (b) a death date after the first day of that semester. Part A and B enrollment is assessed using the part of the column MEDICARE_ENR_[YY] applicable to the semester. The first six characters apply to the first semester and the last six characters apply to the second semester. Values of C, E, F, H, J, K, L, M, N, P, or Q are counted as Part A and B enrollment. Death date is obtained from the column BENE_DEATH_DT.
C	**FFS Medicare beneficiaries:** Records in #B that have FFS Part A and B enrollment during the entire semester. FFS enrollment is assessed using the part of the column MEDICARE_ENR_[YY] applicable to the semester. Values of E or M are counted as FFS Part A and B enrollment.
D	**Share of Medicare beneficiaries enrolled in FFS Medicare:** Count of records in #C divided by the count of records in #B.
E	**Shares of FFS Medicare beneficiaries ages 64-, 65–74, 75–84, and 85+:** Beneficiaries in #C are assigned to one of four groups (64-, 65–74, 75–84, 85+) based on their age; then the counts for each group are divided by the number of records in #C. Age is determined using the column BENE_BIRTH_DT as of the first day of the semester.
F	**Share of FFS Medicare beneficiaries with male/female/unknown gender:** Analogous to #G. Sex is determined using the column SEX.
G	**Shares of FFS Medicare beneficiaries with White/Black/Hispanic/Asian/other/unknown race:** Analogous to #G. Race/ethnicity is determined using the column RTI_RACE_CD.
H	**Share of FFS Medicare beneficiaries fully/partially eligible for Medicaid:** Beneficiaries in #C are counted as fully/partially eligible for Medicaid if they have at least 1 month of full/partial eligibility for Medicaid during the semester; then the count is divided by the number of records in #C. Medicaid eligibility is assessed using the part of the column DUAL_STUS_20[YY] applicable for the semester. Values of 02, 04, or 08 are counted as full dual eligibility. Values of 01, 03, 05, or 06 are counted as partial dual eligibility. Note that a beneficiary may be counted as a full dual and a partial dual for the same semester.
I	**Share of FFS Medicare beneficiaries attributed to APMs:** Count of beneficiaries in #C that have at least 1 month of APM attribution during the semester divided by the count of records in #C. Attribution to APMs is assessed by linking CME custom enrollment files and cleaned MDM Beneficiary extract on BID. MDM is cleaned by selecting records with beneficiary category code of F or blank, keeping one record per beneficiary, and counting the number of months that [Beneficiary Alignment Effective Date] and [Beneficiary Alignment End Date] overlap with the semester. Partial overlap will be counted as a full month.
J	**Share of FFS Medicare beneficiaries attributed to ACOs:** Identical to #K, with an additional condition that only MDM records that have one of the ACO program IDs are selected before cleaning the MDM. ACO program IDs include 08 (SSP), 18 (CEC), 21 (NGACO), 53 (Vermont All-Payer ACO Model), or 63 (DC).

Item	Description
K	**Share of FFS Medicare beneficiaries having ESRD:** Count of beneficiaries in #C who have ESRD for at least 1 month during the semester divided by the count of beneficiaries in #C. Having ESRD is assessed using the part of the column ESRD_[YY] applicable to the semester. Value of 1 is counted as having ESRD.
L	**Average HCC risk scores for FFS Medicare beneficiaries:** Beneficiaries in #C are assigned a risk score by linking CME custom enrollment files with MedPAC cleaned risk scores on BID. Then, the average risk score is calculated.
M	**Average HCC risk scores squared for FFS Medicare beneficiaries:** Analogous to #N. HCC Risk scores will be squared before taking the average.
N	**Share of FFS Medicare beneficiaries residing in urban, rural micropolitan, rural adjacent, and rural non-adjacent areas:** Beneficiaries in #C are assigned to one of four groups (urban, rural micropolitan, rural adjacent, rural nonadjacent) based on their first valid Social Security Administration (SSA) county code for that semester (invalid SSA county codes begin with 99999 or 00000) and the 2013 Urban Influence Codes (UICs); then the counts for each group are divided by the number of records in #C. Urban areas are defined as UICs 1 and 2. Rural micropolitan areas are defined as UICs 3, 5, and 8. Rural adjacent areas are defined as UICs 4, 6, and 7. Rural nonadjacent areas are defined as UICs 9,10, 11, and 12. UICs, based on Federal Information Processing System (FIPS) county codes, are obtained from https://www.ers.usda.gov/data-products/urban-influence-codes. A crosswalk between 2013 SSA county codes and 2013 FIPS county codes is obtained from https://data.nber.org/ssa-fips-state-county-crosswalk/2013/ssa_fips_state_county2013.dta.
O	**Average ADI for FFS Medicare beneficiaries:** Beneficiaries in #C are assigned an ADI based on their first valid 9-digit ZIP Code for that semester; then, the average ADI is calculated. 2020 ADI is obtained from https://www.neighborhoodatlas.medicine.wisc.edu/.

Exhibit A–3. Data Sources for HSA Market Characteristics

Item	Description
A	**Attribution to HSAs:** Items #B to #G are calculated for each county, year (2018, 2019, 2021), and semester (January to June, July to December). Then, for each year and semester, county-level statistics are converted to the HSA level using county-level weights proportional to the population of the county residing in that HSA. A list of HSAs and their overlapping counties along with the population living in the intersection of HSA and county pairs is created by combining 2019 ZIP Code to HSA crosswalk obtained from https://data.dartmouthatlas.org/supplemental/#crosswalks and 2010 ZIP Code Tabulation Area (ZCTA) to county relationship files obtained from https://www2.census.gov/geo/docs/maps-data/data/rel/zcta_county_rel_10.txt. County code changes between 2010 and 2021 are accounted for based on https://www.census.gov/programs-surveys/geography/technical-documentation/county-changes.html.
B	**Population:** The Census Bureau provides population estimates for each county on an annual basis at https://www.census.gov/programs-surveys/popest/data/tables.html. Population estimates for 2018 and 2019 are obtained from https://www2.census.gov/programs-surveys/popest/datasets/2010-2019/counties/totals/co-est2019-alldata.csv, variables [POPESTIMATE2018] and [POPESTIMATE2019]. Population estimates for 2021 are obtained from https://www2.census.gov/programs-surveys/popest/datasets/2020-2021/counties/totals/co-est2021-alldata.csv, the variable [POPESTIMATE2021]. Values for each semester of a year are set to the value of that year.
C	**Number of COVID-19 cases/deaths per 10,000 people:** The *New York Times* provides the cumulative count of cases and deaths for each county on a daily basis at https://github.com/nytimes/covid-19-data. The counts for 2021 are obtained from https://github.com/nytimes/covid-19-data/blob/master/us-counties-2021.csv. The data for a few records that are at the city level are evenly divided between their overlapping counties based on https://github.com/nytimes/covid-19-data#geographic-exceptions. New cases and deaths for a semester are calculated by subtracting the cumulative count as of the last day of that semester from the cumulative count as of the last day of the prior semester. For a few counties for which new cases or deaths are less than zero, the values are set to zero. The counts are divided by the county-level population described in item #B and multiplied by 10,000. New cases and deaths for both semesters of 2018 and 2019 are set to zero.
D	**Number of hospital beds and the number of primary care physicians per 10,000 residents:** The Health Resources and Services Administration (HRSA) publishes the Area Health Resource Files (AHRFs) for each county on an annual basis at https://data.hrsa.gov/data/download. The number of hospital beds for 2018 is obtained from https://data.hrsa.gov//DataDownload/AHRF/AHRF_2019-2020_SAS.zip, variable name [F0892118]. The number of hospital beds for 2019 is obtained from https://data.hrsa.gov//DataDownload/AHRF/AHRF_2020-2021_SAS.zip, variable name [F0892119]. The number of primary care physicians for 2018 and 2019 is obtained from https://data.hrsa.gov//DataDownload/AHRF/AHRF_2020-2021_SAS.zip by adding the variables [F14675YY], [F14677YY], and [F14679YY], where YY equals 18 for 2018 and 19 for 2019. The numbers are divided by the county-level population described in item #B and multiplied by 10,000. Values for each semester of a year are set to the value of that year. The values for 2021 are not available for this study.
E	**Poverty rate and median household income:** The Census Bureau provides estimates for poverty rate and median household income for each county on an annual basis at https://www.census.gov/programs-surveys/saipe/data/datasets.html. The estimates for 2018 are obtained from https://www2.census.gov/programs-surveys/saipe/datasets/2018/2018-state-and-county/est18all.xls, variables [Poverty%, All Ages] and [Median Household Income], respectively. The estimates for 2019 are obtained from https://www2.census.gov/programs-

Item	Description
	surveys/saipe/datasets/2019/2019-state-and-county/est19all.xls, variables [Poverty%, All Ages] and [Median Household Income], respectively. Values for each semester of a year are set to the value of that year. The estimates for 2021 are not available for this study. Note that weighted county-level median household income is not the same as the median household income for the HSA.
F	**Share of residential housing units with downstream internet speed of 0.2+/10+/25+ Mbps:** The Federal Communications Commission provides estimates of internet availability for each county as of the last day of June and the last day of December of each year at https://www.fcc.gov/form-477-county-data-internet-access-services. Share of residential housing units with downstream internet speed of 0.2+/10+/20+ Mbps for the two semesters of 2018 and 2019 are obtained from https://www.fcc.gov/sites/default/files/form_477_county-level_data.xlsx, variable names [Tier_1], [Tier_2], and [Tier_3], respectively. These variables are categorical, taking one of the six possible categories (0%, 0%–20%, 20%–40%, 40%–60%, 60%–80%, and 80%–100%). The categories are converted to cardinal values using the mid-point of each category. The value for each semester is set to the value as of the last day of the prior semester. For example, the value for the first semester of 2018 is set to internet availability as of Dec. 31, 2017. The estimates for 2021 are not available for this study.
G	**Average IPPS hospital wage index:** The CMS publishes the IPPS hospital wage index at the core-based statistical area (CBSA) level along with county to CBSA crosswalk for each year at https://www.cms.gov/Medicare/Medicare-Fee-for-Service-Payment/AcuteInpatientPPS. The values for 2018 are obtained from https://www.cms.gov/Medicare/Medicare-Fee-for-Service-Payment/AcuteInpatientPPS/Acute-Inpatient-Files-for-Download-Items/FY2018-Final-Rule-Correction-Notice-Files. The values for 2019 are obtained from https://www.cms.gov/medicaremedicare-fee-service-paymentacuteinpatientppsacute-inpatient-files-download/files-fy-2019-final-rule-and-correction-notice. The values for 2021 are obtained from https://www.cms.gov/medicare/acute-inpatient-pps/fy-2021-ipps-final-rule-home-page. Values for each semester of a year are set to the value of that year.
H	**Average PFS geographic practice cost indexes (GPCIs):** The CMS publishes GPCIs at MAC-Locality level for each year at https://www.cms.gov/Medicare/Medicare-Fee-for-Service-Payment/PhysicianFeeSched/PFS-Relative-Value-Files. The GPCIs for 2018, 2019, and 2021 are obtained from RVU18D, RVU19C, and RVU21D. In addition, the CMS publishes ZIP Code to MAC-Locality crosswalks for each year at https://www.cms.gov/medicare/medicare-fee-for-service-payment/feeschedulegeninfo. These ZIP Code to MAC-Locality crosswalks are used along with the ZIP Code to HSA crosswalk to determine GPCIs for each ZIP Code in an HSA. Then, average GPCIs for each HSA are calculated. Values for each semester of a year are set to the value of that year.

Appendix B. DID Visualizations

The attached PDF file is composed of two sections: (a) summary visualizations presenting the findings for all outcomes in one graph, and (b) detailed visualizations for each outcome separately, where we present a time series of that outcome between the first semester of 2018 and the second semester of 2021, impact estimates, and parallel trends tests. This file includes bookmarks for easy navigation.

DID-Viz-v6.0.pdf

Appendix C. Methodology for Propensity Score Weighted DID

In this section, we describe methodological details related to estimating propensity score weighted (PSW) DID. In **C.1. Comparison Group Construction**, we describe how PSW improves the balance between Low, Medium, and High telehealth intensity HSAs at baseline. Subsequently, we describe the variables we used to generate propensity score weights and DID estimation using PSW.

C.1. Comparison Group Construction

PSW improves the balance between Low, Medium, and High telehealth intensity HSAs by giving more weight to HSAs that are similar to the HSAs in the other groups. Exhibit C–1 provides the intuition using a simple example. Imagine that there are two types of HSAs (differing covariates), and for simplicity, assume we have only Low and High telehealth intensities. Type A HSAs have a one out of four chance of having High telehealth intensity, while Type B HSAs have a two out of three chance. Suppose that there are 12 HSAs of each type. We would expect three Type A HSAs and eight Type B HSAs to have High telehealth intensity, with the remaining nine Type A HSAs and four Type B HSAs having Low telehealth intensity (Exhibit C–1, Panel A). If we were to estimate the DID equation using this sample, then we would compare the outcomes of a treatment group that includes mostly Type B HSAs with the outcomes of a comparison group that includes mostly Type A HSAs (Exhibit C–1, Panel B). If Type A and B HSAs were different in terms of their unobservable characteristics, then this approach would result in biased estimates. Inverse probability weighting can correct this imbalance by decreasing the influence of overrepresented observations (Type A in the comparison group and Type B in the treatment group) and increasing the influence of underrepresented observations (Type A in the treatment group and Type B in the comparison group). In this example, the nine Type A HSAs in the comparison group would have a weight of 1/0.75, and the three Type A HSAs in the treatment group would have a weight of 1/0.25, which creates a perfect balance for Type A HSAs between the treatment and a comparison group because $9 \times [1/(1-0.25)\,] = 3 \times (1/0.25)$ (Panel C). The intuition is similar for Type B HSAs.

Exhibit C–1. PSW Improves the Balance Between Treatment and Control HSAs

Variable	Type A HSAs	Type B HSAs
Panel A: Setup		
Number of HSAs	12	12
Probability of having Low telehealth intensity	3/4	1/3
Probability of having High telehealth intensity	1/4	2/3
Expected number of HSAs having Low telehealth intensity	12 × 3/4 = 9	12 × 1/3 = 4
Expected number of HSAs having High telehealth intensity	12 × 1/4 = 3	12 × 2/3 = 8
Panel B: Composition of the sample without weighting		
Number of HSAs in the comparison group	9	4
Number of HSAs in the treatment group	3	8

Variable	Type A HSAs	Type B HSAs
Panel C: Composition of the sample with inverse probability weighting		
Weighted number of HSAs in the comparison group	9 × 4/3 = 12	4 × 3/1 = 12
Weighted number of HSAs in the treatment group	3 × 4/1 = 12	8 × 3/2 = 12

We estimated the propensity scores for each HSA (i.e., the probability of an HSA having Low, Medium, and High telehealth intensity) based on its observable characteristics during the baseline period using a multinomial logistic regression (the sum of the three probabilities is 1 for each HSA). Following standard guidelines, we estimated the propensity scores based on characteristics that correlate highly with both the outcomes and the treatment (see **C.2. Propensity Score Weighting Variables**). We assessed the estimated propensity scores and their associated weights by (a) checking the distribution of propensity scores for overlap between Low, Medium, and High telehealth intensities; (b) checking the balance of covariates between the three groups (i.e., difference in means) in the baseline period with and without using the inverse propensity score weights; and (c) checking the balance of covariates within blocks of propensity scores (e.g., quantiles).

C.2. Propensity Score Weighting Variables

Exhibit C–2 shows the full list of variables used for estimating the propensity scores and as covariates in the DID model. Only the baseline values are used for estimating the propensity scores whereas the DID model uses both the baseline period and the treatment period values.

The rationale behind controlling for various covariates in the DID model is discussed in **2.4. Covariates**. All of these variables, with the exception of COVID-19 cases that had no variation during the baseline period, were also used for PSW. However, some variables were used for PSW but not as DID covariates because there was either no variation over time in the variable (e.g., share of FFS beneficiaries in urban areas) or the data were limited to the baseline period (e.g., number of hospital beds per 10,000 residents). Below, we describe the rationale behind using these variables for PSW. Additionally, we included baseline telehealth use as a variable for PSW. Even though telehealth use prior to the PHE was very limited, baseline use may help predict telehealth adoption after the PHE.

Share of FFS beneficiaries in urban areas. During the PHE, Medicare beneficiaries residing in rural areas used telehealth with less intensity than beneficiaries residing in urban areas (MedPAC, 2021). Thus, we controlled for the share of FFS Medicare beneficiaries residing in urban and rural areas. Following MedPAC's prior work, we used Urban Influence Codes (UICs) from the U.S. Department of Agriculture (2019) to assign each beneficiary to either urban, rural micropolitan, rural-adjacent, or rural nonadjacent areas based on their county. Urban counties are defined as those counties that contain an urban cluster of 50,000 or more people. Rural micropolitan counties contain a cluster of 10,000 to 50,000 people. Rural-adjacent counties are adjacent to urban areas and without a city of at least 10,000 people. Rural nonadjacent counties are not adjacent to an urban area and do not have a city with at least 10,000 people.

Number of hospital beds and primary care physicians. We used two standard measures for health care supply-side factors, which could affect health care quality and access: (a) the number of hospital beds per 10,000 people and (b) the number of primary care physicians per 10,000 people. The Area Health Resource Files (AHRFs) contain the number of hospital beds and the number of primary care physicians by year and county (Health Resources & Services Administration, n.d.).

Median household income and poverty rate. We used the baseline median household income and poverty rates as a proxy for economic conditions, which may influence access to and quality of health care. The annual county-level median household income and poverty rates are available in the Small Area Income and Poverty Estimates (SAIPE) from the Census Bureau (n.d.).

Share of residential housing units with 10+ Mbps internet service. We used the share of residential housing units having access to 10+ Mbps internet service, which is conventionally considered high-speed internet, as a proxy for internet access. The Federal Communications Commission (2022) publishes these data semiannually for each county.

Exhibit C–2. Covariates Used in PSW and DID

Covariate	Analysis
HSA Medicare Population Characteristics	
Share of Medicare beneficiaries enrolled in FFS	PSW/DID
Shares of FFS beneficiaries under age 65, 65–74, 75–84, and 85+	PSW/DID
Share of FFS male/female/unknown sex beneficiaries	PSW/DID
Shares of FFS White/Black/Hispanic/Asian/other/unknown race beneficiaries	PSW/DID
Share of FFS beneficiaries fully/partially eligible for Medicaid	PSW/DID
Average HCC risk score and its square for FFS Medicare beneficiaries	PSW/DID
Share of FFS beneficiaries attributed to APMs	PSW/DID
Average ADI for FFS Medicare beneficiaries	PSW/DID
Share of FFS beneficiaries in urban areas	PSW
Telehealth visits per 1,000 FFS Medicare beneficiaries	PSW
HSA Market Characteristics	
Population size	PSW/DID
New and cumulative COVID-19 cases per 10,000 people	DID
Number of hospital beds per 10,000 people	PSW
Number of primary care physicians per 10,000 people	PSW
Median household income	PSW
Poverty rate	PSW
Share of residential housing units with 10+ Mbps internet service	PSW

Note. ADI = Area Deprivation Index; APM = alternative payment model; DID = difference-in-differences; FFS = fee-for-service; HCC = hierarchical condition category; HAS = Hospital Service Area; Mbps = Megabits per second; PSW = propensity score weighting.

C.3. DID Estimation With PSW

To obtain the propensity score weighted DID estimates, we estimated Equation (1) noted in **2.5. Empirical Strategy** while weighting each HSA by the inverse of the propensity score for its observed telehealth intensity. To test for the parallel trends assumption, we estimated Equation (2) noted in the same section, again weighting each HSA by the inverse of the propensity score for its observed telehealth intensity.

A potential limitation of PSW relates to the possibility of obtaining weights that are extreme in value. This leads to HSAs that, given their weights, become overly influential in the effect estimates, making the findings hinge on the inclusion or exclusion of a very small number of HSAs. To investigate this problem, we analyzed the prevalence of extreme PSW weights. The prevalence of extreme weights was found to be small (2.9%).

About the American Institutes for Research®

Established in 1946, the American Institutes for Research® (AIR®) is a nonpartisan, not-for-profit organization that conducts behavioral and social science research and delivers technical assistance both domestically and internationally in the areas of education, health, and the workforce. AIR's work is driven by its mission to generate and use rigorous evidence that contributes to a better, more equitable world. With headquarters in Arlington, Virginia, AIR has offices across the U.S. and abroad. For more information, visit AIR.ORG.

AIR® Headquarters
1400 Crystal Drive, 10th Floor
Arlington, VA 22202-3289
+1.202.403.5000 | AIR.ORG