JUNE 2024

Updated Analysis: 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 the 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

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

Final Report

Tanvi Rao, PhD, Angshuman Gooptu, PhD, Karin Johnson, PhD,
Guido Cataife, PhD, Steven Susana-Castillo, MPH

April 2024

American Institutes for Research® | AIR.ORG

AIR®
Advancing Evidence.
Improving Lives.

This report was prepared for Ledia Tabor
Medicare Payment Advisory Commission
425 I Street NW, Suite 701
Washington, DC 20001

Acronyms

ACS	Ambulatory care sensitive
ADI	Area Deprivation Index
APM	Alternative Payment Model
APRN	Advanced practice registered nurse
BPCI	Bundled Payments for Care Improvement
CAH	Critical access hospitals
CAA	Consolidated Appropriations Act
CCW	Chronic Conditions Warehouse
CME	Common Medicare Environment
CMS	Centers for Medicare & Medicaid Services
DID	Difference-in-differences
ED	Emergency department
FFS	Fee-for-service
FIPS	Federal Information Processing System
FQHC	Federally Qualified Health Centers
GAFs	Geographic Adjustment Factors
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
PTT	Parallel trends test
REH	Rural Emergency Hospitals
RHC	Rural Health Clinics
SAF	Standard analytic file
UIC	Urban Influence Code
ZCTA	ZIP Code Tabulation Area

Contents

Exhibits

Executive Summary

Background

The objective of this study is to update the previous analysis that the American Institutes for Research® (AIR®) conducted to inform the Medicare Payment Advisory Commission's (MedPAC) June 2023 report to Congress (AIR, 2023). The previous report used population-based measures to describe the association between telehealth use and quality, access, and cost when both telehealth and in-person visits are available to fee-for-service (FFS) Medicare beneficiaries. In the previous study, results were likely susceptible to the confounding effects of COVID-19, as there were documented surges in COVID-19 prior to and during the treatment period (July to December 2021) (Truelove et al., 2022). This study uses more recent available data from a time when there were fewer extreme surges of COVID-19 cases during the treatment period (July to December 2022) (Ahmad et al., 2022). Findings from this study can be more relevant and informative when assessing the impact of telehealth utilization in a post-pandemic environment. Also, this study looks at the effects of telehealth separately for urban and rural beneficiaries and differentiates behavioral and non-behavioral telehealth use. The findings of this study could inform policymakers' discussions with regards to whether current telehealth flexibilities, which are temporarily in place through the end of December 2024, should be made permanent.

Methodology

We used a difference-in-differences (DID) approach to compare changes in population-based outcomes across areas with different levels of telehealth service use. Using Medicare FFS administrative data, we examined population-based measures that capture

- quality of care, including risk-adjusted 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 semester of 2022 (treatment period) with the second semesters of 2018 and 2019 (baseline period). The second semester of 2022 was selected as the treatment period because it was the most recently available data at the time of the study, and it could produce estimates that represent, as much as possible, the effects of telehealth in a period relatively free of the influence of the COVID-19 pandemic.

To measure the effects of telehealth usage for behavioral telehealth services separately from non-behavioral telehealth services, each HSA was assigned two treatment level rankings: a Low, Medium, or High ranking for behavioral telehealth visits per 1,000 beneficiaries in the treatment period, and a Low, Medium, or High ranking for non-behavioral telehealth visits per 1,000 beneficiaries in the treatment period. These rankings are not mutually exclusive. For example, the same HSA can be in the High group for both behavioral telehealth and non-behavioral telehealth services.

We classified HSAs as urban or rural and within each grouping used a DID approach to compare outcomes in (1) HSAs with Medium and High non-behavioral telehealth intensity to HSAs with Low non-behavioral telehealth intensity and (2) HSAs with Medium and High behavioral telehealth intensity to HSAs with Low behavioral telehealth intensity. The identifying assumption of the DID model is that the trajectory of the outcomes would have been identical for the two treatment groups (HSAs with Medium and High telehealth intensity) and the control group (HSAs with Low telehealth intensity), absent differences in telehealth use on account of the telehealth expansion. The DID approach controls for factors that remain constant over time within HSAs. However, factors that differentially change from the baseline to the treatment period between groups can 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 versus Medicare Advantage (MA) 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.

In addition, we implemented three robustness checks to test whether the main impact estimates are robust to (1) controlling for differences in geographic adjustment factors across HSAs, (2) controlling for in-person utilization, and (3) excluding small HSAs with less than 500 beneficiaries. In the key findings section below, we report results of the robustness checks if they contradict a finding from the main specification.

Key Findings

We found that for both urban and rural HSAs, those in the Medium and Low telehealth intensity groups were considerably more comparable at baseline than were HSAs in the High and Low telehealth intensity groups. This was evidenced by looking at the magnitude of differences in baseline characteristics across Low, Medium, and High telehealth intensity HSAs and by looking at the frequency with which the parallel trends test (PTT) passed when outcome

trends were compared across Medium and Low groups versus High and Low groups.[1] Therefore, in this study, drawing conclusions regarding the effects of telehealth are more reliable when Medium and Low groups are compared than when High and Low groups are compared.

Another noteworthy point to aid in the interpretation of findings is that we found a strong overlap in HSAs classified as Low, Medium, and High telehealth intensity when grouped according to utilization of non-behavioral telehealth and behavioral telehealth services. This was especially true for urban HSAs, where the correlation coefficient between the two treatments was 0.59, which indicates a strong correlation. Overall, this indicates that it is challenging to interpret the effects of non-behavioral telehealth and behavioral telehealth services as distinct treatments, as each treatment may be reflective of telehealth utilization as a whole in the area.

In the main body of the report, we discuss impact estimates for HSAs with Medium and High telehealth intensity. However, the overall conclusions of this study discussed below are based only on impact estimates from HSAs with Medium telehealth intensity and on outcomes that passed the PTT among HSAs with Medium telehealth intensity. As discussed above, the impact estimates for Medium versus Low telehealth intensity can reasonably be interpreted as reflecting reliable associations, whereas the impact estimates for High versus Low telehealth intensity should not be interpreted as reflecting reliable associations given the high proportion of outcomes that did not pass the PTT.

- ***Telehealth and quality.*** **We do not observe an association between telehealth intensity and quality outcomes (ACS hospitalizations and ACS ED visits).** When comparing the Medium and Low groups for urban HSAs, there is no association between non-behavioral telehealth intensity and ACS hospitalizations and ACS ED visits.[2] For rural HSAs, there is no association between non-behavioral telehealth intensity and ACS ED visits.
- ***Telehealth and access.*** **In some of our analyses, we find evidence that higher telehealth intensity is associated with fewer clinician encounters.** For both non-behavioral and behavioral telehealth urban HSAs, HSAs with Medium telehealth intensity are associated with a decrease in clinician encounters per beneficiary when compared to the Low group.

[1] In the Medium versus Low analysis, the PTT passed for 75 percent of the main outcomes for urban non-behavioral telehealth intensity HSAs, 50 percent of the main outcomes for rural non-behavioral telehealth intensity HSAs, 100 percent of the main outcomes for urban behavioral telehealth intensity HSAs, and 100 percent of outcomes for rural behavioral telehealth intensity HSAs. In the High versus Low analysis, the PTT passed for 25 percent of outcomes for urban non-behavioral telehealth intensity HSAs, 25 percent of outcomes for rural non-behavioral telehealth intensity HSAs, 100 percent of outcomes for urban behavioral telehealth intensity HSAs, and 50 percent of outcomes for rural behavioral telehealth intensity HSAs.

[2] Recall that we do not estimate quality outcomes for behavioral telehealth intensity. It is also important to note that the PTT did not pass for one quality outcome: ACS hospitalizations in rural HSAs.

The magnitude of the association is small relative to baseline clinician encounters, ranging from a decline of 1.18 percent (0.13 fewer clinician encounters per beneficiary) to a decline of 1.60 percent (0.17 fewer clinician encounters per beneficiary) for urban and rural HSAs, respectively. Rural HSAs with Medium behavioral telehealth intensity are also associated with a decrease in overall clinician encounters per beneficiary once small HSAs are excluded—a decline of 0.79 percent of the baseline rate (0.07 fewer clinician encounters per beneficiary).

- ***Telehealth and costs.*** **Although the results related to telehealth and costs are not sufficiently conclusive, we find some evidence that higher telehealth intensity is associated with a decrease in the total cost of care.** In urban HSAs, those with Medium non-behavioral telehealth intensity are not associated with the total cost of care per beneficiary outcome. However, rural HSAs with Medium non-behavioral telehealth intensity are associated with a decrease in the total cost of care per beneficiary, but this decrease is not robust to the exclusion of small HSAs. In urban and rural HSAs, those with Medium behavioral telehealth intensity are associated with a decrease in the total cost of care per beneficiary. The magnitude of the association is small relative to baseline total cost of care, ranging from a decrease of 2.18 percent (132.45 fewer dollars per beneficiary) to a decline of 1.82 percent (108.70 fewer dollars per beneficiary) in urban and rural HSAs respectively. In urban HSAs the decrease in the total cost of care is also accompanied by a decrease in physician costs. However, other cost components, such as skilled nursing facility and hospice costs, also show a decrease, and these cost components can be less reliably tied to the intensity of behavioral telehealth treatment.

Discussion. Focusing on the findings for HSAs with Medium telehealth intensity, we observe that greater telehealth intensity is not associated with the quality outcomes and is associated in some analyses with a small decrease in clinician encounters and the total cost of care. However, given the heterogeneity in the findings, caution should be exercised in strongly interpreting any given study result as a causal estimate of the impact of telehealth.

The decrease in clinician encounters observed may relate to a transitory effect of the COVID-19 pandemic. The second half of 2022, which is the treatment period of this study, was a period when the pandemic had largely subsided in the majority of the United States. A relative decline in clinician encounters in HSAs with Medium telehealth intensity compared with HSAs with Low telehealth intensity between the baseline period of 2018 and 2019 and the post-period of 2022, could reflect less pent-up demand for care in HSAs with Medium telehealth intensity, as these areas were able to see providers more regularly during the pandemic and had less of a need to catch-up with postponed or canceled care.

Findings from the current analysis do differ from findings reported by AIR (2023). Our previous study found that increased telehealth usage was associated with slightly higher clinician encounters, increases in the total cost of care, and an increase in ACS hospitalizations. The difference in the post-periods across the two studies could help explain this difference. The post-period of the previous study (second half of 2021) was confounded by the COVID-19 pandemic, but during the post-period of the current study (second half of 2022), the majority of the United States population was in a location with low COVID-19 community transmission levels. For instance, previous differences across HSAs, in terms of prevalence of COVID-19 or their response preparedness, very likely had an impact on both the telehealth intensity in an area and outcomes, such as ACS hospitalizations, making it difficult to cleanly interpret the observed increases in utilization. The estimated relationships in the current study may reflect a mitigation of the confounding previously observed.

The main limitation encountered in this analysis was the lack of baseline comparability between the High and Low telehealth intensity groups and, to a lesser extent, between the Medium and Low telehealth intensity groups. To address this, MedPAC could consider propensity score weighting. The selection of more comparable groups at baseline, in terms of observable characteristics, may more likely yield parallel baseline trends in study outcomes between these group, but PSW improved parallel trends only marginally in the previous iteration of this study.

Another challenge encountered was a strong overlap in HSAs classified as Low, Medium, and High telehealth intensity when grouped according to utilization of non-behavioral telehealth and behavioral telehealth services, which makes it challenging to interpret non-behavioral telehealth and behavioral telehealth as distinct treatments. In this report, we recommend some refinements to the regression analysis with additional controls and outcome variables to better identify the separate effects of behavioral and non-behavioral telehealth intensity.

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 such services are provided to beneficiaries who receive the service at a clinician's office or certain health care facility (known as the "originating site") located in a rural area, with some exceptions. However, to maintain access to care and help limit community spread of COVID-19 during the public health emergency (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 (Centers for Medicare & Medicaid Services [CMS], 2022b). During the PHE, many providers and beneficiaries embraced telehealth (Office of the Assistant Secretary for Planning and Evaluation, 2020; U.S. Government Accountability Office, 2022; MedPAC, 2023b). MedPAC's June 2023 report to Congress detailed that between 2019 and 2020, the number of FFS beneficiaries who received at least one telehealth service paid under the PFS accelerated rapidly from 239,000 to 14.2 million (40 percent of Part B FFS beneficiaries) and then declined in 2021 to 9.7 million (29 percent of Part B FFS beneficiaries).

Congress and CMS are considering the possibility of making the PHE flexibilities permanent. In its March 2021 report to Congress, MedPAC presented a policy option that policymakers continue to cover telehealth services with 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).

As the PHE continued, some important permanent and temporary changes were made to Medicare's telehealth policy (U.S. Department of Health and Human Services, 2023):

- ***Permanent changes.*** The Consolidated Appropriations Act (CAA) of 2021 and the CAA of 2022 legislated that Medicare would permanently pay for telehealth behavioral services received in a patient's home. However, some requirements were imposed requiring that an in-person, face-to-face, non-telehealth service takes place within 6 months of the telehealth behavioral health service and that an in-person visit takes place annually thereafter (U.S. Department of Health and Human Services, 2022).

  Other permanent changes included allowing Federally Qualified Health Centers (FQHCs) and Rural Health Clinics (RHCs) to serve as distant site providers for behavioral telehealth services, removing geographic restrictions for originating sites for behavioral telehealth services, allowing behavioral telehealth services to be delivered using audio-only

communication platforms, and (effective beginning in calendar year 2023) designating Rural Emergency Hospitals (REHs) as eligible originating sites for telehealth.

- ***Temporary changes applicable through December 31, 2024.*** The CAA of 2023 (effective December 2022) extended many of Medicare's telehealth expansions through December 31, 2024. The temporary expansions included continuing to pay for telehealth services received by Medicare patients in their home, allowing FQHCs and RHCs to serve as distant site providers for non-behavioral telehealth services; removing geographic restrictions for originating sites for non-behavioral telehealth services, allowing some non-behavioral telehealth services to be delivered using audio-only communication platforms, and allowing telehealth services to be provided by all eligible Medicare providers.

  Additionally, the requirement for an in-person visit within 6 months of an initial behavioral telehealth service (and annually thereafter) was waived until December 31, 2024.

In the CAA of 2022, Congress mandated that MedPAC 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 FQHCs and RHCs, and alternative approaches to paying for telehealth services. In response to this mandate, MedPAC's June 2023 report to Congress included a chapter on telehealth. One section of the report described the analysis we conducted using population-based measures to describe the association between telehealth use and quality, access, and cost when both telehealth and in-person visits were available to FFS Medicare beneficiaries. Findings of this analysis suggested that during the pandemic, greater telehealth use was associated with little change in measured quality, slightly improved access to care for some beneficiaries, and slightly increased costs to the Medicare program (MedPAC, 2023b; AIR, 2023). However, the findings from the previous study should not be interpreted causally because of the confounding effects of COVID-19 and other variables that we could not measure, which could affect both the use of telehealth and patient outcomes.

The objective of this study is to update the previous analysis that AIR conducted to inform MedPAC's June 2023 report to Congress (AIR, 2023). In addition, this study aims to address some limitations of the previous work. To address the confounding effects of COVID-19 in the previous study, the treatment period for this study uses more recent data from a time when there were fewer extreme surges of COVID-19. This study also looks at the effects of telehealth separately for urban and rural beneficiaries and differentiates between behavioral and non-behavioral telehealth use, to allow for additional granularity.

The findings of this study can inform policymakers' discussions with regards to making permanent (or not) current flexibilities that are temporarily in place through the end of December 2024.

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 prior AIR report and MedPAC's 2023 report to Congress included some discussion of conceptually how telehealth can impact these outcomes. The geographic unit for the study is the HSA. The study period is the second half of 2018 and 2019 (baseline period) and the second half of 2022 (treatment period). This approach provides an analysis period that is 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 or disability, have lower income, and live in urban areas use a higher number of telehealth services on average (MedPAC, 2023b). 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. A DID approach does not account for differential changes between treatment and control groups (i.e., compositional change within the HSAs included in the study) from the baseline to the treatment period.

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

2.1. Period of Study

For the DID analysis, the baseline period covers a period before the PHE and the expansion of telehealth and is the second semester of 2018 and 2019 (defined as July–December of each year), and the treatment period is the second semester of 2022 (defined as July–December 2022, after COVID-19 vaccines were widely available to Medicare beneficiaries and the expansion of

telehealth). We analyze the same portion of the year during the baseline and treatment periods to help alleviate concerns about seasonality in the data. 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 2022 because there were no major COVID surges in this period and because these were the most recent data available at the time of this study. To study outcome trends, we used all semesters of data between 2018 and 2022; and to test parallel trends at baseline, we used all semesters of 2018 and 2019.

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, 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 variables 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 for behavioral telehealth and non-behavioral telehealth services	Groupings based on the number of telehealth visits per 1,000 FFS Medicare beneficiaries: Low (<33rd percentile), Medium (33rd–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]

Variable type	Variable name	Specification	Notes
	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 (including hospitalists), APRNs and PAs, and other practitioners	
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, carrier (hereafter referred to as "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. The programs used to calculate the ACS measures have been updated since the 2019 report cited here. 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*. Medicare Payment Advisory Commission (MedPAC). 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 2022 (the treatment period). We identified telehealth encounters from the outpatient and carrier Standard Analytic Files (SAFs). We used Healthcare Common Procedure Coding System (HCPCS)

codes for telehealth-eligible services published by CMS (2022b), together with Place of Service (POS), HCPCS modifier codes, and revenue center codes detailed in Appendix A, Exhibit A-1, which are necessary to define telehealth use before the pandemic and after the PHE. We also used codes that are specific to the CMS Innovation Center model telehealth waivers to fully capture telehealth use within existing value-based care initiatives. Additionally, we included codes for virtual or e-visit check-ins and telephone evaluation and management codes. We excluded codes for remote monitoring, originating site telehealth services and interprofessional internet consultation services because those are not patient facing services.

As noted in our original report, we considered two options for measuring telehealth intensity: (1) proportion of services that are provided by telehealth and (2) rate of services provided by telehealth per 1,000 beneficiaries (AIR, 2023). 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.

We separately analyzed behavioral telehealth (identified by specific HCPCS codes, Current Procedural Terminology codes, and revenue center codes) and non-behavioral telehealth services. Because telehealth flexibilities for behavioral health services have been permanently extended, it is important for policymakers to understand their potential benefits separately from non-permanently extended (non-behavioral health) services.

When calculating the numerator of the measures, we attributed behavioral and non-behavioral telehealth encounters to HSAs using location of the beneficiary's residence on the claim. In the denominator, we included all beneficiaries who were alive and who had Part A and B coverage for the entire 6 months of the semester. After calculating the number of behavioral health and non-behavioral health telehealth visits per 1,000 beneficiaries for each HSA during the baseline and treatment periods, we assigned each HSA one of the three treatment levels of Low, Medium, or High based on their ranking of the number of behavioral and non-behavioral telehealth visits per 1,000 beneficiaries in the treatment period. We assigned the bottom third of HSAs to the Low level, the middle third of HSAs to the Medium level, and the top third of HSAs to the High level. Since we are considering telehealth visits for behavioral services separately from non-behavioral services, each HSA had two treatment level rankings—a Low, Medium, or High ranking for behavioral telehealth visits per 1,000 beneficiaries in the treatment period; and a Low, Medium, or High ranking for non-behavioral telehealth visits per 1,000 beneficiaries in the treatment period. Note, however, that these rankings are not mutually exclusive: for instance, the same HSA can be in the High group for both behavioral telehealth and non-behavioral telehealth services.

We discuss how we will use the treatment levels in the **EMPIRICAL STRATEGY** section.

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 pre-existing 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 the number of clinician in-person and telehealth encounters per FFS Medicare beneficiary and their breakdown by the following provider types: primary care physicians, physicians from other specialties, advanced practice registered nurses (APRNs) and physician assistant (PAs), and other practitioners. Primary care specialties 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.

The number of clinician encounters per beneficiary offers a direct measure of health care access. However, it is an aggregate measure that may mask important differences by specialty and type of provider. For example, before the pandemic, from 2015 to 2019, while the number of primary care physician encounters per beneficiary fell by 2.5 percent per year, encounters with APRNs and PAs per beneficiary rose by 11.2 percent per year (MedPAC, 2022a). The breakdown by provider type allows a more nuanced examination of how the expansion of telehealth affects different parts of the health care process. For example, while telehealth expansion may increase the access to routine preventive visits to primary care physicians, APRNs, and PAs, it may limit patients' access to physicians from specialties other than primary care.

We used MedPAC's pre-existing programming SAS codes and specifications to calculate these measures. Most encounters were captured in the Carrier SAF, where each encounter is identified as a unique combination of beneficiary ID, claim ID, and National Provider Identifier. Since FQHCs, RHCs, and Critical Access Hospitals (CAHs) method II encounters are reimbursed under special payment rates for providing telehealth services, we also considered the number of clinician

encounters originating from FQHCs, RHCs, and CAH method II billings[3] (U.S. Department of Health and Human Services, 2023b). Outpatient SAF files provide information about facility type and type of service that allowed for identification of encounters with FQHCs and RHCs (defined here as unique claim IDs). Similarly, the outpatient SAF file was used to identify CAH method II encounters using information on the provider number, the facility type, and relevant revenue centers. Here too, encounters are defined as unique claims IDs. Exhibit A-2 in Appendix A encloses the relevant codes.

To compute each access measure at the HSA level, we counted the number of encounters in an HSA and divided it by the number of FFS Medicare beneficiaries in that HSA.

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 Next Generation Accountable Care Organization [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-3 and A-4 in Appendix A describe the data sources in detail.

[3] "CAH method II billing" refers to situations in which clinicians reassign their billing rights to a CAH and cannot bill for services under the standard Medicare PFS: https://www.cms.gov/outreach-and-education/medicare-learning-network-mln/mlnproducts/downloads/critaccesshospfctsht.pdf

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, on account of the changing population composition of FFS Medicare beneficiaries, between the baseline and treatment periods.

Exhibit 2. Covariates Used in DID

Covariates
HSA Medicare population characteristics
Share of Medicare beneficiaries enrolled in FFS
Share of FFS beneficiaries under age 65, 65–74, 75–84, and 85+
Share of FFS male/female/unknown sex beneficiaries
Share 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 the rationale behind controlling for population characteristics of Medicare beneficiaries listed in Exhibit 2.

Share of Medicare beneficiaries enrolled in FFS: Beneficiaries self-select FFS or 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 Common Medicare Environment (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. For example, younger Medicare beneficiaries use telehealth more frequently and may have better health outcomes than 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 scores.[4] 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 data.

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

[4] 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.

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 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 U.S. 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 2022 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 2022 and the cumulative COVID-19 cases up until the second semester of 2022 per 10,000 people. We obtained the number of COVID-19 cases from the *New York Times* database.

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 telehealth intensity by comparing the average change in an outcome for HSAs with Medium or High telehealth intensity between the second semesters of 2018 and 2019 and the second semester of 2022 with the average change in that outcome for HSAs with Low telehealth intensity during the same period.[5] The regression-based DID approach automatically controls for any baseline difference in outcome levels between the treatment and comparison groups, as well as for any time-invariant characteristic affecting the outcomes at the HSA level. Time-varying confounders are not

[5] We exclude the first semesters of 2018 and 2019 to avoid introducing potential seasonality issues in our estimates.

automatically adjusted; to adjust for these confounders, we included covariates in the model (described in **Covariates**).

The regression analysis was conducted for four separate subgroups:

- urban[6] HSAs for telehealth utilization related to non-behavioral health,
- urban HSAs for telehealth utilization related to behavioral health,
- rural HSAs for telehealth utilization related to non-behavioral health, and
- rural HSAs for telehealth utilization related to behavioral health.

Since the two quality outcomes—ACS hospitalizations and ED visits—are not conceptually related to behavioral health, we do not estimate these outcomes when analyzing behavioral telehealth. All other outcomes are estimated for each of the four subgroups. Exhibit 3 lists the treatment and comparison groups for each analysis.

Exhibit 3. Treatment and Comparison Groups

Type of HSA and Telehealth Category	Treatment	Comparison
Urban HSAs		
Non-behavioral telehealth	• High telehealth intensity for non-behavioral health services • Medium telehealth intensity for non-behavioral health services	Low telehealth intensity for non-behavioral health services
Behavioral telehealth	• High telehealth intensity for behavioral health services • Medium telehealth intensity for behavioral health services	Low telehealth intensity for behavioral health services
Rural HSAs		
Non-behavioral telehealth	• High telehealth intensity for non-behavioral health services • Medium telehealth intensity for non-behavioral health services	Low telehealth intensity for non-behavioral health services
Behavioral telehealth	• High telehealth intensity for behavioral health services • Medium telehealth intensity for behavioral health services	Low telehealth intensity for behavioral health services

[6] We split the sample by the median share of beneficiaries living in urban areas in the second semester of 2019. HSAs above the median were classified as urban, and HSAs below the median were classified as rural. Choosing other values (e.g., the mean) as the threshold for urbanicity does not affect the analysis because beneficiaries in most HSAs live in the same type of area.

2.5.1. DID Estimation Equation

We estimated the following equation for each outcome using data from the second semesters of 2018, 2019, and 2022:

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

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 2022). 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 HSAs with Low telehealth intensity. Following standard practice, we estimated heteroskedastic robust standard errors and clustered them at the HSA level.

2.5.2. Parallel Trends Test

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[7]:

$$\begin{aligned} Y_{it} = \alpha_0 + \alpha_1 M_i + \alpha_2 H_i + \alpha_3 Pre_{t1} + \alpha_4 Pre_{t2} + \alpha_5 Pre_{t3} + \boldsymbol{\alpha_6} M_i \times Pre_{t1} \\ + \boldsymbol{\alpha_7} M_i \times Pre_{t2} + \boldsymbol{\alpha_8} M_i \times Pre_{t3} + \boldsymbol{\alpha_9} H_i \times Pre_{t1} + \boldsymbol{\alpha_{10}} H_i \times Pre_{t2} \\ + \boldsymbol{\alpha_{11}} H_i \times Pre_{t3} + X_{it}\Delta + \epsilon_{it} \quad (2) \end{aligned}$$

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

[7] We used both semesters of a year to test for parallel trends, even though the DID analysis only uses the second semester of a year, to allow for sufficient granularity in trend analysis, given that we extracted claims only for a limited number of baseline years.

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 PTT passes (or more precisely, 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 HSAs with Medium and Low intensity 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 HSAs with High and Low telehealth intensity, 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 illustrate some of the systematic differences in demographic and market characteristics between the Low, Medium, and High telehealth intensity HSAs. We also show trends in study outcomes over the full 2018–2022 period. Next, in **3.2. Main Findings: Non-Behavioral Telehealth**, we present, for the non-behavioral telehealth treatment, results of the PTT, testing whether baseline outcome trends are parallel when comparing Medium and Low, as well as High and Low, telehealth intensity HSAs (a key assumption for the validity of the DID estimates); the DID estimates for the impact of Medium or High telehealth intensity, relative to Low telehealth intensity, on quality, access, and cost outcomes; and results of robustness checks. Section **3.3. Main Findings: Behavioral Telehealth** is analogous to Section 3.2 and discusses the same set of analyses for the behavioral telehealth treatment.

3.1. Descriptive Statistics

3.1.1. Beneficiary Characteristics

Exhibit 4 summarizes the characteristics of our sample of beneficiaries, which were generally consistent with prior analyses of the FFS Medicare population (MedPAC, 2022b). The characteristics of the samples were also generally consistent across the time periods. In the second semesters of 2018 and 2019, the sample included 32.0 and 31.1 million FFS Medicare beneficiaries, respectively. In the second semester of 2022, the sample included 3.8 million fewer FFS Medicare beneficiaries relative to 2018, consistent with an increase in MA enrollment over these years. The average ages of the beneficiaries in the samples were 71.0, 71.3, and 72.0 years in the second semesters of 2018, 2019 and 2022, respectively. In all samples, most of these beneficiaries were female (around 55 percent), non-Hispanic White (around 80 percent), and living in an urban area (around 79 percent). The percentage with full dual eligibility for Medicaid and Medicare for 6 months declined from 13.2 percent to 12.4 percent between the second semesters of 2018 and 2022; and 4.2 percent had partial dual eligibility for 6 months in the second semester of 2018, which declined to 3 percent in the second semester of 2022. The share of beneficiaries attributed to an APM for at least 1 month increased from 35.4 percent in the second semester of 2018 to 46.9 percent in the second semester of 2022. Average ADI decreased from 51.2 in the second semester of 2018 to 46.5 in the second semester of 2022.

Exhibit 4. Beneficiary Characteristics

	2018	2019	2022
Number of beneficiaries alive and with FFS Part A and B for 6 months	31,883,969	31,147,616	28,248,932
Average age	71.0	71.3	72.0
Percentage female	54.8	54.7	54.7
Percentage non-Hispanic White	79.4	79.6	80.6
Percentage Black	9.0	8.7	7.3
Percentage Hispanic	5.8	5.7	5.3
Percentage Asian	2.7	2.7	3.0
Percentage with full Medicaid eligibility for 6 months	13.2	12.6	12.4
Percentage with partial Medicaid eligibility for 6 months	4.2	3.9	3.0
Percentage attributed to an APM for at least 1 month	35.4	39.0	46.9
Percentage living in urban areas	78.3	78.3	79.3
Average ADI	51.2	50.8	46.5

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

We calculated averages for the various characteristics at baseline (before the telehealth expansion) by non-behavioral telehealth intensity (Exhibit 5a) and behavioral telehealth intensity (Exhibit 5b). We also conducted t-tests to examine how HSAs having Medium or High telehealth intensity in 2022 differed at baseline from HSAs with Low telehealth intensity. In general, both Medium and High telehealth intensity HSAs were statistically significantly different than HSAs with Low telehealth intensity for most characteristics examined, but in terms of magnitude, baseline differences were much larger when comparing High versus Low telehealth intensity HSAs than when comparing Medium versus Low telehealth intensity HSAs. Moreover, characteristics associated with telehealth uptake differed across urban and rural HSAs.

- Age and gender composition:
  - Among urban HSAs, the High group was, on average, slightly older and slightly more likely to be female compared to the Low and Medium groups (as seen across both Exhibits 5a and 5b).

  - Among rural HSAs, interestingly, the pattern was reversed. The High group was, on average, younger and less likely to be female.

- Racial composition and income:
  - Among urban HSAs, the Low and Medium groups were more similar in terms of racial composition and income, but the High group had a lower percentage of non-Hispanic White beneficiaries and included a higher percentage of low-income beneficiaries (using eligibility for Medicaid as a proxy for income) compared to both Low and Medium groups. For instance, for the non-behavioral telehealth treatment (Exhibit 5a), the average percentage of FFS Medicare beneficiaries who were non-Hispanic White decreased from around 88.8 percent in the Low group to around 85.8 percent in the Medium group and to around 74.8 percent in the High group. The percentage of FFS Medicare beneficiaries with full Medicaid eligibility increased from around 10.6 percent in the Low group to 11.1 percent in the Medium group and to around 15.9 percent in the High group. This finding is consistent with MedPAC's prior findings that beneficiaries with lower incomes use telehealth services more frequently (MedPAC, 2023a).
  - Among rural HSAs, characteristics associated with telehealth uptake were more mixed. For the non-behavioral telehealth treatment, the pattern was similar to urban HSAs. However, for the behavioral telehealth treatment, the High and Medium groups had a greater percentage of non-Hispanic White beneficiaries compared to the Low group.

- ADI:
  - Among urban HSAs, the High group had substantially lower average ADI compared to the Medium and Low groups (as seen across both Exhibits 5a and 5b). For instance, for the non-behavioral telehealth treatment, average ADI decreased from around 66.9 percent in the Low group to around 62.4 percent in the Medium group and to around 38.3 percent in the High group.
  - Among rural HSAs, average ADI followed a similar pattern across Low, Medium, and High HSAs, but the difference in ADI between the groups was smaller.

- Population Size:
  - Among urban HSAs, the High group was much more populous than the Medium and Low groups, as seen in both Exhibits 5a and 5b.
  - Among rural HSAs, population size followed a similar pattern across Low, Medium, and High groups, but the difference in population size between the groups was small.

Exhibit 4a. Average Baseline Characteristics by Non-Behavioral Telehealth Intensity

	Low	Medium	High	Mean Difference Medium vs. Low	Mean Difference High vs. Low
Urban HSAs					
Average age	70.85	70.71	71.38	-0.14**	0.53***
Percentage female	54.23	54.16	54.88	-0.07	0.65***
Percentage non-Hispanic White	88.82	85.78	74.79	-3.04***	-14.03***
Percentage Black	6.13	7.67	7.44	1.54***	1.31***
Percentage Hispanic	1.60	2.91	9.93	1.32***	8.34***
Percentage Asian	0.54	0.81	4.51	0.28***	3.98***
Percentage with full Medicaid eligibility for 6 months	10.62	11.13	15.86	0.51**	5.24***
Percentage with partial Medicaid eligibility for 6 months	4.88	4.46	3.05	-0.41***	-1.83***
Percentage attributed to an APM for at least 1 month	35.94	38.26	36.97	2.31**	1.03
Average ADI	66.87	62.39	38.26	-4.48***	-28.60***
Population Size (in 10,000 people)	11.67	29.15	140.14	17.49***	128.48***
Rural HSAs					
Average age	71.35	70.72	70.48	-0.63***	-0.86***
Percentage female	53.73	53.14	52.46	-0.58***	-1.27***
Percentage non-Hispanic White	88.12	88.34	83.24	0.22	-4.88***
Percentage Black	5.85	4.93	4.04	-0.92**	-1.81***
Percentage Hispanic	1.78	2.76	6.90	0.98***	5.11***
Percentage Asian	0.28	0.37	1.02	0.08***	0.74***
Percentage with full Medicaid eligibility for 6 months	11.58	13.11	14.49	1.53***	2.91***
Percentage with partial Medicaid eligibility for 6 months	5.63	5.90	6.12	0.27	0.49**
Percentage attributed to an APM for at least 1 month	29.13	28.54	25.42	-0.59	-3.71***
Average ADI	76.59	73.60	70.74	-2.99***	-5.85***
Population Size (in 10,000 people)	2.43	3.31	3.88	0.88***	1.45***

Note. ADI = Area Deprivation Index; APM = Alternative Payment Model. Low, Medium, and High denote non-behavioral telehealth intensity groups in the second semester of 2022. All statistics are an average among HSAs

and are average values for the second semester of 2018 and 2019. *** and ** denote statistical significance at 1 and 5 percent levels, respectively.

Exhibit 5b. Average Baseline Characteristics by Behavioral Telehealth Intensity

	Low	Medium	High	Mean Difference Medium vs. Low	Mean Difference High vs. Low
Urban HSAs					
Average age	70.75	70.80	71.32	0.05	0.57***
Percentage female	53.90	54.11	54.99	0.21**	1.09***
Percentage non-Hispanic White	83.05	84.54	78.30	1.49	-4.75***
Percentage Black	8.06	6.62	7.39	-1.44***	-0.67
Percentage Hispanic	5.09	4.91	7.02	-0.18	1.93***
Percentage Asian	0.50	1.42	3.90	0.92***	3.40***
Percentage with full Medicaid eligibility for 6 months	12.36	11.37	14.83	-0.99**	2.47***
Percentage with partial Medicaid eligibility for 6 months	5.49	4.33	3.05	-1.16***	-2.44***
Percentage attributed to an APM for at least 1 month	35.60	35.68	38.63	0.08	3.03***
Average ADI	70.91	60.51	39.95	-10.41***	-30.96***
Population Size (in 10,000 people)	13.61	37.15	126.47	23.54***	112.86***
Rural HSAs					
Average age	71.08	71.02	70.45	-0.06	-0.63***
Percentage female	53.43	53.26	52.89	-0.16**	-0.54***
Percentage non-Hispanic White	84.67	89.70	90.56	5.03***	5.88***
Percentage Black	6.91	3.98	2.33	-2.94***	-4.58***
Percentage Hispanic	3.54	2.45	2.86	-1.08***	-0.68
Percentage Asian	0.34	0.45	0.80	0.12**	0.46***
Percentage with full Medicaid eligibility for 6 months	12.54	11.98	14.64	-0.56**	2.09***
Percentage with partial Medicaid eligibility for 6 months	6.34	5.24	5.44	-1.10***	-0.90***
Percentage attributed to an APM for at least 1 month	26.78	29.48	30.24	2.70***	3.46***
Average ADI	77.79	73.05	66.64	-4.74***	-11.16***
Population Size (in 10,000 people)	2.39	3.18	4.66	0.79***	2.27***

Note. ADI = Area Deprivation Index; APM = Alternative Payment Model. Low, Medium, and High denote behavioral telehealth intensity groups in the second semester of 2022. All statistics are an average over HSAs and are average values for the second semester of 2018 and 2019. *** and ** denote statistical significance at 1 and 5 percent levels, respectively.

3.1.3. Outcome Trends

Exhibits 6–9 plot unadjusted trends between 2018 and 2022 for the four main study outcomes—risk-adjusted ACS hospitalizations per 1,000 beneficiaries, risk-adjusted ED visits per 1,000 beneficiaries, total clinician encounters per beneficiary, and the total cost of care per beneficiary—by telehealth intensity group. As previously discussed, the two quality outcomes—ACS hospitalizations (Exhibit 6) and ACS ED visits (Exhibit 7)—are not conceptually related to behavioral health; therefore, we plot these outcomes only for the non-behavioral health treatment.

Among urban HSAs, risk-adjusted ACS hospitalizations in the Low and Medium groups followed nearly an identical and overlapping trend up until 2021, semester 2. During this time ACS hospitalizations in the High group were lower. After 2021, semester 2, ACS hospitalizations in the High and Medium groups increased relative to the Low group. Among rural HSAs, risk-adjusted hospitalizations tended to be slightly higher in the Low telehealth group than in the Medium and High groups, both before and after the PHE and telehealth expansion (Exhibit 6).

Among urban HSAs, risk-adjusted ACS ED visits were highest in the Low group, followed by the Medium group, and then the High group, which had the lowest ACS ED visit rates. This general pattern remained the same both before and after the PHE and telehealth expansion, but ED visits overall declined. Among rural HSAs, the pattern was reversed; risk-adjusted ACS ED visits were highest in the High group, followed by the Medium group, and then the Low group. ACS ED visits in the Medium group declined slightly relative to the Low group after the PHE and telehealth expansion (Exhibit 7).

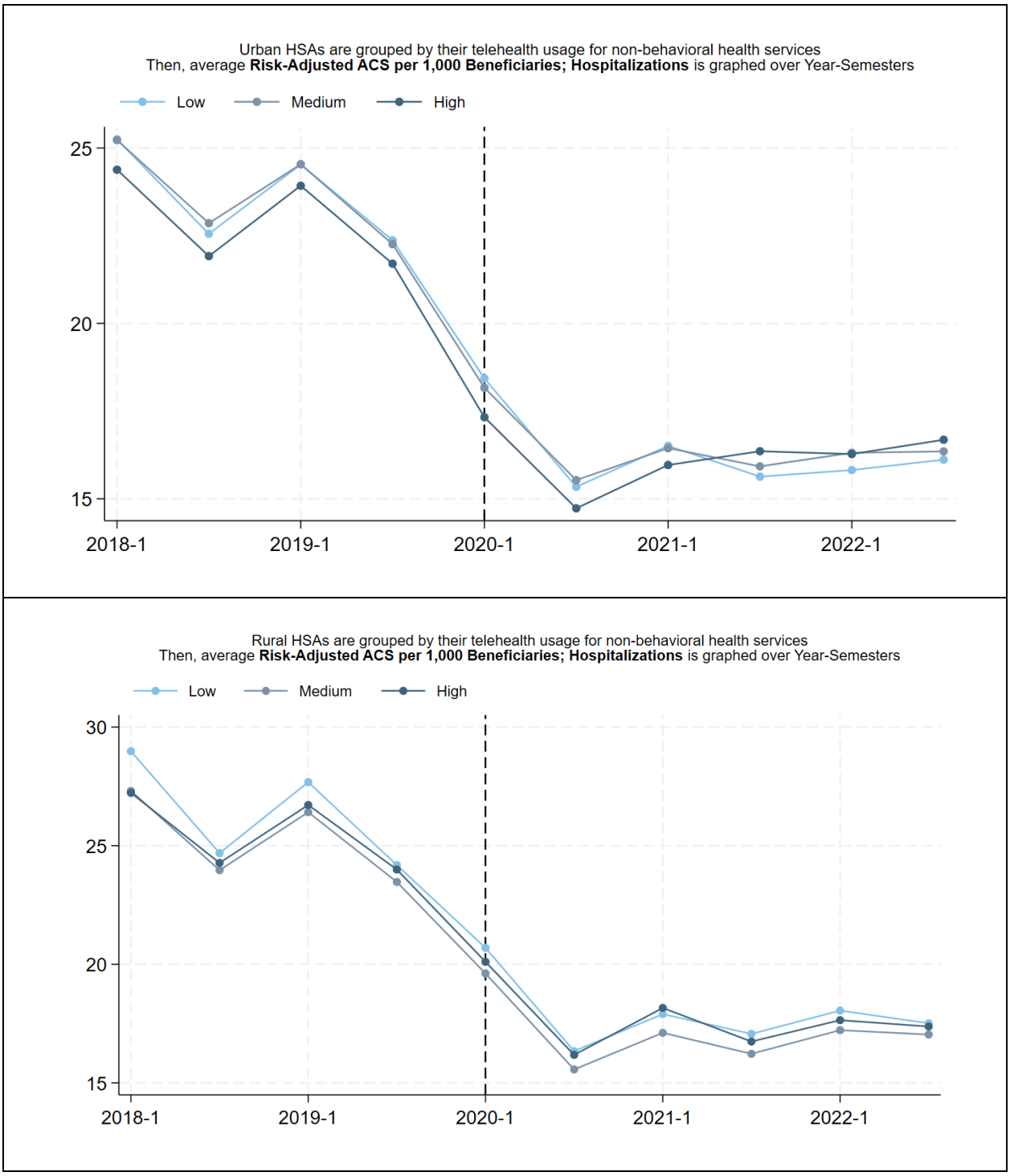
Among urban HSAs, clinician encounters per beneficiary were highest among the High telehealth intensity group, followed by the Medium group, and then the Low group. This general pattern remained the same both before and after the PHE and telehealth expansion, but clinician encounters declined overall. Among rural HSAs, the differences in clinician encounters between the three groups were smaller throughout the study period compared to urban HSAs. Similar to urban HSAs, clinician encounters declined overall after the PHE and telehealth expansion (Exhibit 8).

Among urban HSAs, the total cost of care per beneficiary was substantially higher in the High group than in the Medium and Low groups both before and after the PHE and telehealth expansion. Among rural HSAs, the trends were more mixed. For instance, for the non-

behavioral telehealth treatment (upper right quadrant), the Low and High groups had nearly overlapping trend lines for the total cost of care per beneficiary both before and after the PHE and telehealth expansion. In contrast, for the behavioral telehealth treatment (lower right quadrant), the Low group had substantially higher total cost of care per beneficiary compared to the Medium and High groups (Exhibit 9).

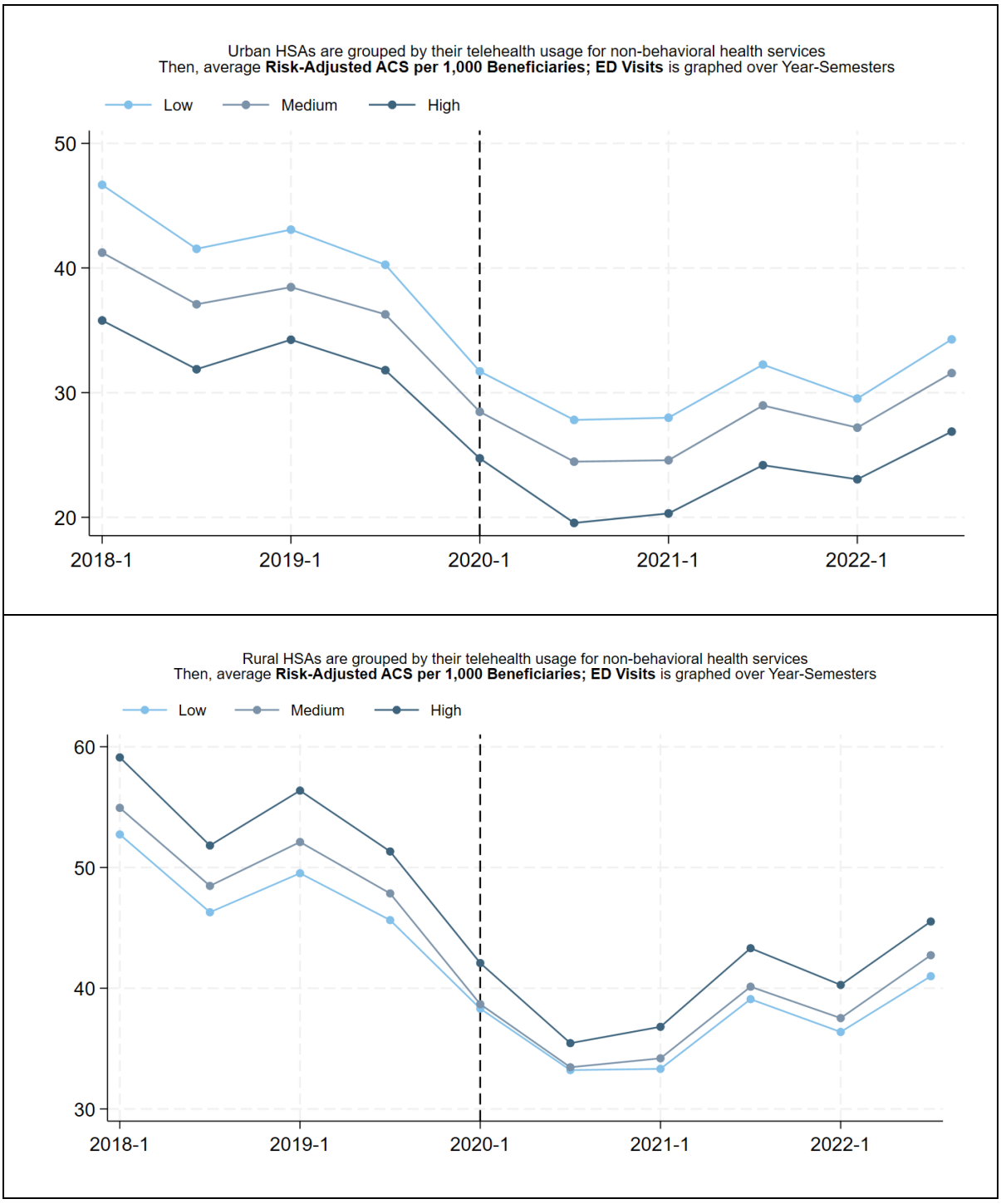
Appendix B contains outcome trends for the full set of outcomes analyzed in this study, including encounters by provider type and cost of care by service type.

Exhibit 6. Average Risk-Adjusted ACS Hospitalizations per 1,000 Beneficiaries, 2018–2022



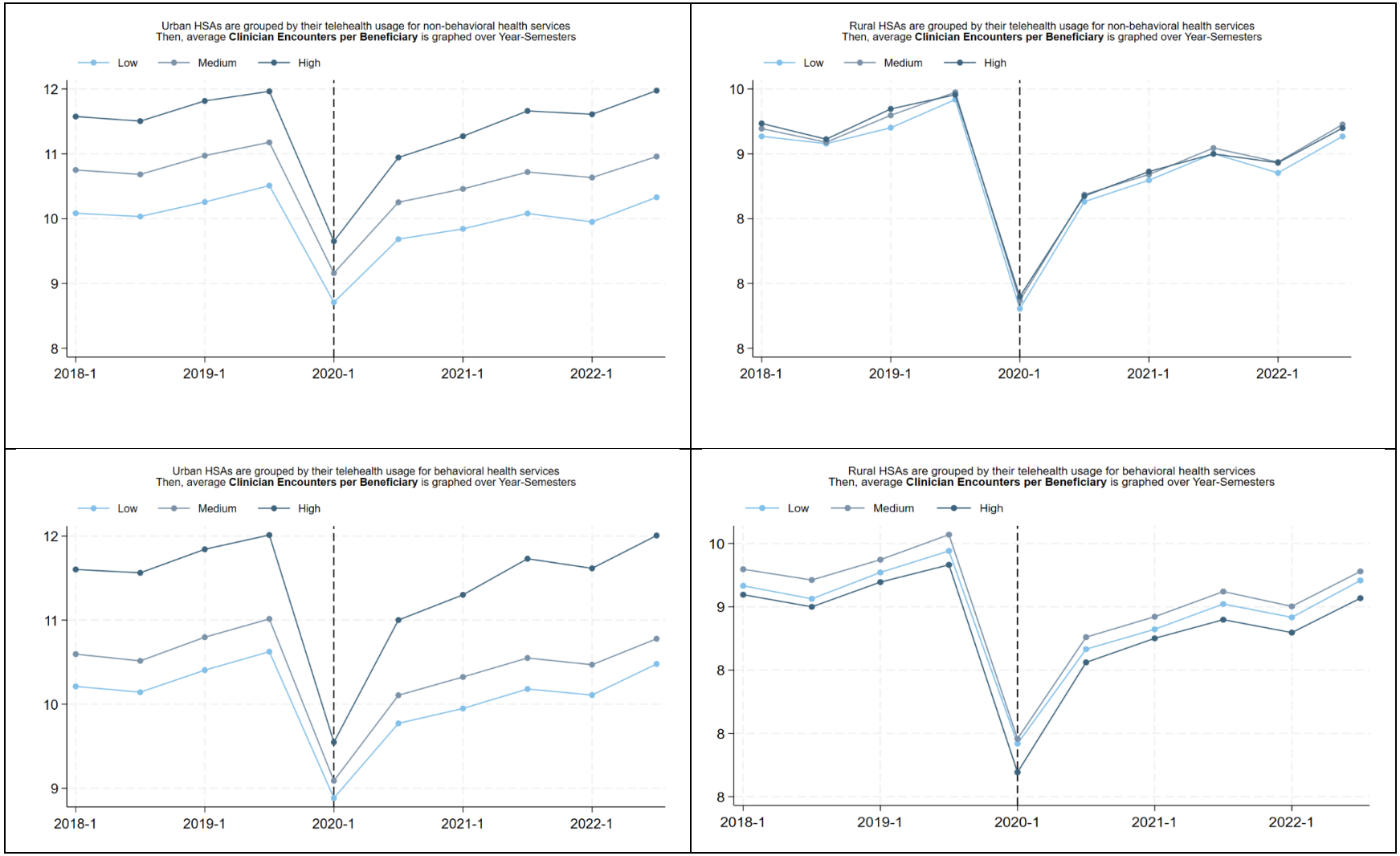
Note. Low, Medium, and High denote telehealth intensity levels in the second semester of 2022.

Exhibit 7. Average Risk-Adjusted ACS ED Visits per 1,000 Beneficiaries, 2018–2022



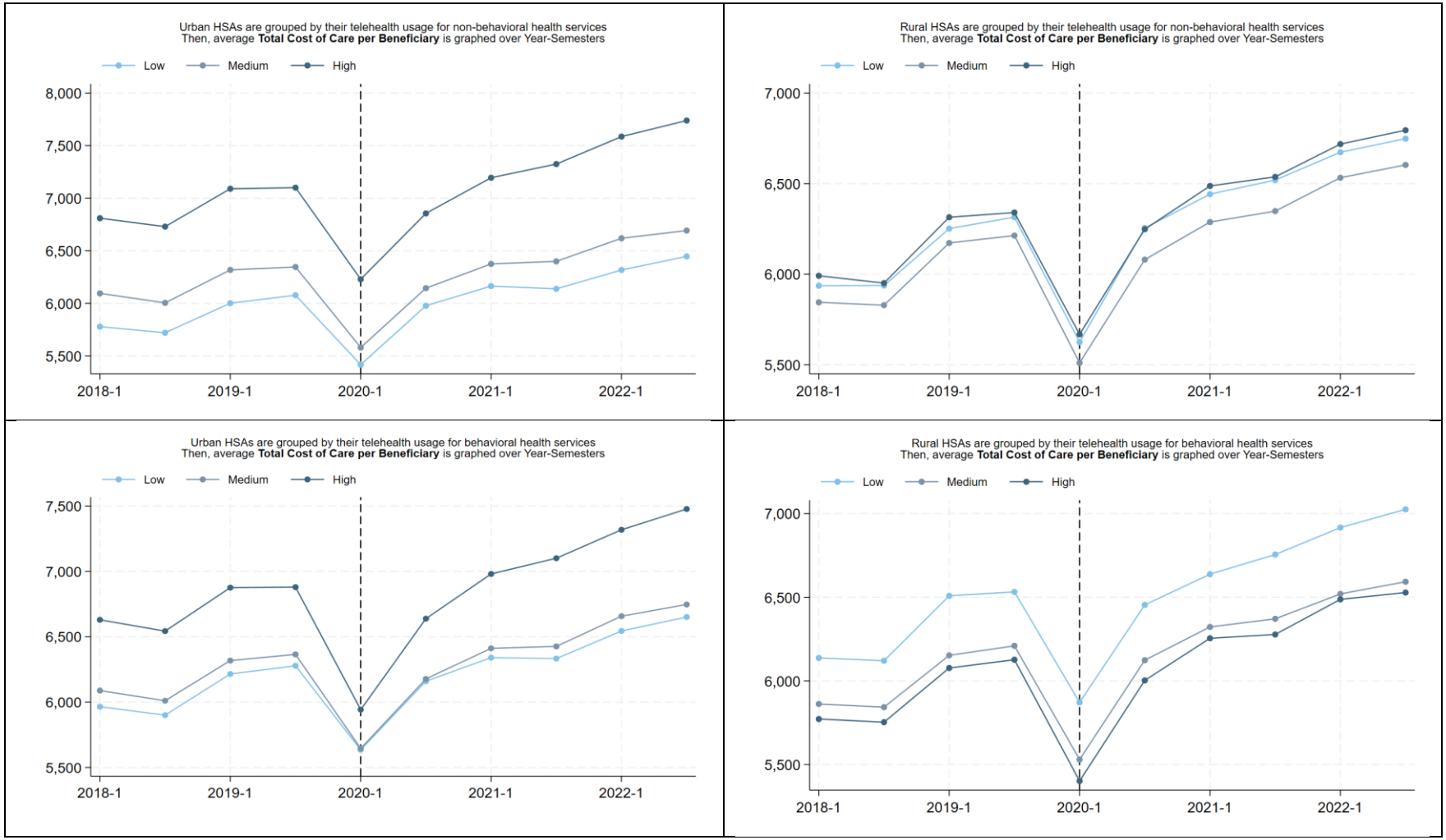
Note. Low, Medium, and High denote telehealth intensity levels in the second semester of 2022.

Exhibit 8. Average Clinician Encounters per Beneficiary, 2018–2022



Note. Low, Medium, and High denote telehealth intensity levels in the second semester of 2022.

Exhibit 9. Average Total Cost of Care per Beneficiary, 2018–2022



Note. Low, Medium, and High denote telehealth intensity levels in the second semester of 2022.

3.2. Main Findings: Non-Behavioral Telehealth

We explore the impact of non-behavioral telehealth utilization on quality, access, and cost outcomes separately from the impact of behavioral telehealth utilization on access and cost outcomes. This section presents DID estimates of the association between High versus Low non-behavioral telehealth intensity HSAs and Medium versus Low non-behavioral telehealth intensity HSAs. The next section, Section 3.3, presents DID estimates of the association between High versus Low behavioral telehealth intensity HSAs and Medium versus Low behavioral telehealth intensity HSAs. Since we anticipate potentially different effects in rural and urban areas, the impacts of behavioral and non-behavioral telehealth utilization are estimated separately for urban and rural HSAs.

In addition to presenting the impact estimates in both Sections 3.2 (non-behavioral telehealth) and 3.3 (behavioral telehealth), we discuss whether the estimates for telehealth impact on outcomes are valid based on PTTs and a series of robustness checks. The robustness checks tested whether the main impact estimates are robust to (1) differences in geographic adjustment factors across HSAs, (2) controlling for in-person utilization, and (3) excluding small HSAs with less than 500 beneficiaries.

3.2.1. The Parallel Trends Assumption

The main necessary condition for DID estimates to reflect causal relationships or associations that are not due to confounding factors 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 there is 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.

We checked whether there is a statistically significant difference in outcomes between the Medium and Low non-behavioral telehealth intensity groups and between the High and Low non-behavioral telehealth intensity groups during a period prior to the pandemic, from the first semester of 2018 through the second semester of 2019 (PTT period). For parallel trends, differences in outcomes between Medium and Low or High and Low groups for the first three semesters (first and second semesters of 2018 and first semester of 2019) cannot be significantly different relative to the second semester of 2019 in a joint hypothesis test. We used a multivariate regression analysis to adjust for the differences in covariates that are included in the DID models (see **2.5. EMPIRICAL STRATEGY**). To pass parallel trends, the p-value of the joint hypothesis test must not be statistically significant at the 1 percent, 5 percent, or 10 percent significance levels (i.e., should have p-value > 0.10). Therefore, our approach for testing

parallel trends counts even small baseline differences (with low statistical significance at the 10 percent significance level) between groups as indicative of PTT failure.

Exhibit 10 summarizes parallel trends comparing Medium and High to Low non-behavioral telehealth intensity HSAs by urban and rural HSA status.

Urban HSAs: For urban HSAs with Medium non-behavioral telehealth intensity, three of the four main outcomes (ACS hospitalizations, ACS ED visits, and total clinician encounters) passed the PTT and one outcome (total cost of care) failed the PTT. For urban HSAs with High non-behavioral telehealth intensity, one of the four main outcomes (ACS hospitalizations) passed the PTT and the remaining other main outcomes (ACS ED visits, total clinician encounters, and total cost of care) failed the PTT. A visual inspection of the trends for urban HSAs during the PTT period shows that the trends for ACS hospitalizations (Exhibit 6; top panel) are largely similar for High, Medium, and Low non-behavioral telehealth intensity groups. However, slight differences in trends are observed for ACS ED visits (Exhibit 7; top panel), clinician encounters (Exhibit 8; second quadrant), and total cost of care (Exhibit 9; second quadrant) for HSAs in the High non-behavioral telehealth intensity group compared to those in the Low group.[8]

Across all 15 outcomes evaluated for the impact of non-behavioral telehealth utilization, outcomes for urban HSAs with Medium non-behavioral telehealth intensity passed the PTT more frequently (80 percent; 12 out of 15 outcomes) than did outcomes for urban HSAs with High non-behavioral telehealth intensity (40 percent; 6 out of 15 outcomes).

Rural HSAs: For rural HSAs with Medium non-behavioral telehealth intensity, two of the four main outcomes (ACS ED visits and total cost of care) passed the PTT and the other two main outcomes (ACS hospitalizations and clinician encounters) failed the PTT. For rural HSAs with High non-behavioral telehealth intensity, one of the four main outcomes (total cost of care) passed the PTT and the other three main outcomes (ACS hospitalizations, ACS ED visits, and clinician encounters) failed the PTT. A visual inspection of the trends for ACS hospitalizations (Exhibit 6; bottom panel), ACS ED visits (Exhibit 7; bottom panel), and clinician encounters (Exhibit 8; first quadrant) generally supports the PTT findings.

Across all 15 outcomes evaluated for the impact of non-behavioral telehealth utilization, outcomes for rural HSAs with Medium non-behavioral telehealth intensity passed the PTT more frequently (67 percent; 10 out of 15 outcomes) than did outcomes for rural HSAs with High non-behavioral telehealth intensity (33 percent; 5 out of 15 outcomes).

[8] While the visual outcome trends during the baseline period usually reinforces PTT findings, sometimes visual discrepancies may not always result in a statistically significant difference in the PTT. Since the PTT accounts for overlapping confidence intervals and allows us to control for covariates, we favor the statistical test (as opposed to visual inspection) to determine whether trends for outcomes were parallel prior the pandemic.

When describing the DID impact of telehealth utilization on quality, access, and cost outcomes, we also need to account for the presence or lack of parallel trends between the Medium and Low groups or between the High and Low groups. We refer to a statistically significant impact of Medium or High non-behavioral telehealth utilization HSAs as an *association* if it is statistically significant and the PTT compared to Low non-behavioral telehealth intensity HSAs passed. For an estimated impact that is not statistically significant but the PTT passed, we refer to the impact as *no association*. Finally, for a statistically significant or not statistically significant impact of Medium or High groups where the PTT compared to the Low group failed, we refer to the result as *association likely due to confounding factors* or *no association potentially due to confounding factors*, respectively.

Therefore, for results that are:

- Statistically significant and the PTT passed, we will note that there is an *association*.
- Statistically significant and the PTT failed, we will note that there is an *association potentially due to confounding factors*.
- Not statistically significant and the PTT passed, we will note that there is *no association*.
- Not statistically significant and the PTT failed, we will note that there is *no association potentially due to confounding factors*.

Exhibit 10. Parallel Trends Test for Quality, Access, and Cost, for Non-Behavioral Telehealth Intensity

	Urban HSAs		Rural HSAs	
	Medium relative to Low	High relative to Low	Medium relative to Low	High relative to Low
ACS hospitalizations per 1,000 beneficiaries	PTT Pass	PTT Pass		
ACS ED visits per 1,000 beneficiaries	PTT Pass		PTT Pass	
Clinician encounters per beneficiary	PTT Pass			
Primary care physicians	PTT Pass			
Specialists (including Hospitalists)	PTT Pass			
APRNs/PAs	PTT Pass		PTT Pass	
Other practitioners				

	Urban HSAs		Rural HSAs	
	Medium relative to Low	High relative to Low	Medium relative to Low	High relative to Low
Total cost of care per beneficiary			PTT Pass	PTT Pass
Inpatient	PTT Pass	PTT Pass	PTT Pass	PTT Pass
Outpatient	PTT Pass		PTT Pass	
Skilled nursing facility	PTT Pass	PTT Pass	PTT Pass	PTT Pass
Home health	PTT Pass	PTT Pass	PTT Pass	PTT Pass
Hospice	PTT Pass	PTT Pass	PTT Pass	PTT Pass
Physician	PTT Pass		PTT Pass	
Durable medical equipment		PTT Pass	PTT Pass	

Note. ACS = ambulatory care sensitive; APRNs = advanced practice registered nurses; PA = physician assistant. Green (or PTT Pass) denotes that the PTT passed, as the differences between Medium or High non-behavioral telehealth intensity HSAs and Low intensity HSAs were not statistically significant (p-value > 0.10). Blank denotes that the PTT failed, as the differences between Medium or High non-behavioral telehealth intensity HSAs and Low intensity HSAs were statistically significant (p-value ≤ 0.10). To statistically test for parallel trends, a joint hypothesis test was conducted to ensure differences in outcomes between High and Low groups or Medium and Low groups over three semesters (2018 semester 1, 2018 semester 2, and 2019 semester 1) were similar and not statistically significant from the second semester in 2019.

3.2.2. Impact Estimates

Exhibit 11 presents a summary of the impact estimates related to the impact of non-behavioral telehealth intensity on quality, access, and costs for urban HSAs separately from rural HSAs. 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 11. Impact of Non-Behavioral Telehealth Intensity on Quality, Access, and Cost

	Urban HSAs		Rural HSAs	
	Medium relative to Low	High relative to Low	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 (including Hospitalists)				
APRNs/PAs		↘	↗	
Other practitioners	↘		↘	↘
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-shaded cells with a blue arrow pointing up (or ↗) denote a statistically significant increase, PTT passed. White-shaded cells with a blue arrow pointing up (or ↗) denote a statistically significant increase, PTT failed. Orange-shaded cells with a red arrow pointing down (or ↘) denote a statistically significant decrease, PTT passed. White-shaded cells with a red arrow pointing down (or ↘) denote a statistically significant decrease, PTT failed. Grey-shaded cells denote no association, PTT passed. Blank (white) cells denote no association, PTT failed. We only consider DID results significant if the DID coefficient has a p-value ≤ 0.05. Impact estimates show the change between the baseline period (average of the second semesters of 2018 and 2019) and the treatment period (second semester of 2022) relative to Low non-behavioral telehealth intensity.

<u>Preview of Findings for Quality Outcomes in Urban HSAs</u>: There is no association with ACS hospitalizations or ED visits for urban HSAs with Medium non-behavioral telehealth intensity. Urban HSAs with High non-behavioral telehealth intensity are associated, potentially due to confounding factors, with an increase in ACS ED visits.

Exhibit 12 presents these findings. For context, the overall trend across urban HSAs in the Low non-behavioral telehealth intensity group is a sharp decline of 4.05 ACS hospitalizations per 1,000 beneficiaries per semester (18.07 percent of the baseline rate)[9] and a decline of 7.36 ACS ED visits per 1,000 beneficiaries per semester (18.09 percent of the baseline rate) between the baseline period (average of the second semesters of 2018 and 2019) and the second semester of 2022.

Urban Medium non-behavioral telehealth intensity: Urban HSAs with Medium non-behavioral telehealth intensity are not associated with ACS hospitalizations or ED visits compared to urban HSAs with Low non-behavioral telehealth intensity.

Urban High non-behavioral telehealth intensity: Urban HSAs with High non-behavioral telehealth intensity are not associated with ACS hospitalizations. The High non-behavioral HSAs are associated, potentially due to confounding factors, with an increase in ACS ED visits.

<u>Preview of Findings for Quality Outcomes in Rural HSAs</u>: Rural HSAs with Medium non-behavioral telehealth intensity are not associated with ACS hospitalizations. Rural HSAs with Medium non-behavioral telehealth intensity are not associated, potentially due to confounding factors, with ED visits. Rural HSAs with High non-behavioral telehealth intensity are not associated, potentially due to confounding factors, with ACS hospitalizations and ED visits.

Exhibit 12 also presents the findings for rural HSAs. For context, the overall trend across rural HSAs in the Low non-behavioral telehealth intensity group is a sharp decline of 9.64 ACS hospitalizations per 1,000 beneficiaries per semester (39.48 percent of the baseline rate) and a decline of 9.00 ACS ED visits per 1,000 beneficiaries per semester (19.29 percent of the baseline rate) between the baseline period and the second semester of 2022.

Rural Medium non-behavioral telehealth intensity: Rural HSAs with Medium non-behavioral telehealth intensity are not associated, potentially due to confounding factors, with ACS hospitalizations. In addition, the Medium group is not associated with ED visits.

Rural High non-behavioral telehealth intensity: Rural HSAs with High non-behavioral telehealth intensity are not associated, potentially due to confounding factors, with ACS hospitalizations and ED visits.

[9] The percentages report the estimate as a share of the average for that outcome in the relevant telehealth group for the second semesters of 2018 and 2019. This is to provide a sense of magnitude.

Exhibit 12. Change in ACS Events per 1,000 Beneficiaries per Semester for Non-Behavioral Telehealth HSAs

	Low	Medium relative to Low	High relative to Low
Urban HSAs			
ACS hospitalizations per 1,000 beneficiaries per semester	-4.05 (-18.07%) ***	-0.01 (-0.06%)	0.54 (2.45%)
ACS ED visits per 1,000 beneficiaries per semester	-7.36 (-18.09%) ***	0.62 (1.56%)	1.72 (5.28%) **
Rural HSAs			
ACS hospitalizations per 1,000 beneficiaries per semester	-9.64 (-39.48%) ***	0.51 (2.17%)	-0.01 (-0.06%)
ACS ED visits per 1,000 beneficiaries per semester	-9.00 (-19.29%) ***	-0.75 (-1.53%)	-0.94 (-1.80%)

Note. Estimates presented in the exhibit are coefficients from DID regressions along with the standard error and statistical significance. ACS = ambulatory care sensitive; ED = emergency department. Low, Medium, and High denote telehealth intensity. Estimates show the change between the baseline period (average of second semesters of 2018 and 2019) and the second semester of 2022. The denominator for the percentages is that group's average in the baseline period. *** and ** denote statistical significance at 1 and 5 percent levels, respectively.

<u>Preview of Findings for Access Outcomes in Urban HSAs</u>: Urban HSAs with Medium non-behavioral telehealth intensity are associated with a decrease in clinician encounters per beneficiary. Urban HSAs with High non-behavioral telehealth intensity are not associated, potentially due to confounding factors, with clinician encounters.

Exhibit 13 presents these findings. For context, the overall trend between the baseline period and the second semester of 2022 across urban HSAs in the Low non-behavioral group is a decline (not statistically significant) of 0.33 clinician encounters per beneficiary per semester (3.20 percent of the baseline rate).

Urban Medium non-behavioral telehealth intensity: Urban HSAs with Medium non-behavioral telehealth intensity are associated with a decrease of 0.13 clinician encounters per beneficiary per semester, which is 1.18 percent of the baseline rate. In terms of clinician types, the Medium group is associated with a 0.04 per beneficiary per semester decrease in encounters with primary care physicians (a 1.96 percent decrease relative to the baseline rate). The Medium group is not associated with encounters with specialists (including hospitalists) and APRN/PAs. The Medium group is associated, potentially due to confounding factors, with a decrease in other practitioner encounters.

Urban High non-behavioral telehealth intensity: Urban HSAs with High non-behavioral telehealth intensity are not associated, potentially due to confounding factors, with clinician encounters. In terms of specific clinician types, the High group is not associated, potentially due to confounding factors, with encounters with primary care physicians, specialists (including hospitalists), and other practitioners. The High group is associated, potentially due to confounding factors, with a decrease in encounters with APRNs/PAs.

Preview of Findings for Access Outcomes in Rural HSAs: Rural HSAs with High and Medium non-behavioral telehealth intensity are not associated, potentially due to confounding factors, with clinician encounters.

Exhibit 13 presents these findings for rural HSAs. For context, the overall trend across HSAs in the Low group is a decline of 0.35 clinician encounters per beneficiary per semester between the baseline period and the second semester of 2022, which is 3.79 percent of the baseline rate.

Rural Medium non-behavioral telehealth intensity: Rural HSAs with Medium non-behavioral telehealth intensity are not associated, potentially due to confounding factors, with clinician encounters. However, such HSAs with Medium intensity are associated with a 0.04 increase in encounters with APRNs and PAs (a 2.97 percent increase relative to the base rate). For specific clinician types, the Medium group is not associated, potentially due to confounding factors, with primary care physicians and specialists (including hospitalists). Finally, the Medium group is associated, potentially due to confounding factors, with a decrease in other practitioner encounters.

Rural High non-behavioral telehealth intensity: Rural HSAs with High non-behavioral telehealth intensity are not associated, potentially due to confounding factors, with clinician encounters and particularly, encounters with specialists (including hospitalists) and APRNs/PAs. The High group is associated, potentially due to confounding factors, with an increase in primary care physician and a decrease in other practitioner encounters.

Exhibit 13. Change in Clinician Encounters per Beneficiary per Semester for Non-Behavioral Telehealth HSAs

	Low	Medium relative to Low	High relative to Low
Urban HSAs			
All clinicians	-0.33 (-3.20%)	-0.13 (-1.18%) **	-0.03 (-0.24%)
Primary care physicians	-0.14 (-8.28%) **	-0.04 (-1.96%) **	0.02 (1.19%)
Specialists (including hospitalists)	-0.28 (-5.09%) **	-0.04 (-0.71%)	0.04 (0.60%)

	Low	Medium relative to Low	High relative to Low
APRNs/PAs	0.22 (15.69%) **	-0.01 (-0.96%)	-0.07 (-6.65%) ***
Other practitioners	-0.09 (-6.12%)	-0.04 (-2.52%) **	-0.02 (-0.86%)
Rural HSAs			
All clinicians	-0.35 (-3.79%) **	-0.03 (-0.35%)	0.03 (0.35%)
Primary care physicians	-0.25 (-15.80%) ***	-0.01 (-0.43%)	0.06 (3.35%) **
Specialists (including hospitalists)	-0.12 (-2.52%)	-0.03 (-0.64%)	-0.03 (-0.53%)
APRNs/PAs	-0.03 (-1.78%)	0.04 (2.97%) **	0.03 (2.43%)
Other practitioners	-0.05 (-3.54%)	-0.03 (-2.03%) **	-0.03 (-2.59%) **

Note. Estimates presented in the exhibit are coefficients from DID regressions along with the standard error and statistical significance. APRNs = advanced practice registered nurses; PAs = physician assistants. Low, Medium, and High denote non-behavioral telehealth intensity. Estimates show the change between the baseline period (average of second semesters of 2018 and 2019) and the second semester of 2022. The denominator for the percentages is that group's average in the baseline period. *** and ** denote statistical significance at 1 and 5 percent levels, respectively

Preview of Findings for Cost Outcomes in Urban HSAs: Urban HSAs with Medium telehealth intensity have no association, potentially due to confounding factors, with total cost of care per beneficiary. Urban HSAs with High non-behavioral telehealth intensity are associated, potentially due to confounding factors, with an increase in the total cost of care per beneficiary.

Exhibit 14 presents these findings. For context, the overall trend across urban HSAs in the Low group is an increase in the average total cost of care per beneficiary per semester of $776.02 between the baseline period and the second semester of 2022 (13.24 percent of the baseline rate).

Urban Medium non-behavioral telehealth intensity: Urban HSAs with Medium non-behavioral telehealth intensity are not associated, potentially due to confounding factors, with the total cost of care per beneficiary per semester. However, the Medium group is associated with a decrease in home health costs by $13.39 per beneficiary per semester (5.99 percent of the baseline), The Medium group is not associated with the cost of care for inpatient, outpatient, skilled nursing facilities; hospice; and physician service types. Finally, the Medium group is not associated, potentially due to confounding factors, with durable medical equipment costs.

Urban High non-behavioral telehealth intensity: Urban HSAs with High non-behavioral telehealth intensity are associated, potentially due to confounding factors, with an increase in the total cost of care per beneficiary per semester. For costs based on claim types, the High group is associated with an increase in inpatient costs by $68.24 per beneficiary per semester (3.40 percent of the baseline average), skilled nursing facility costs by $37.40 per beneficiary per semester (7.83 percent of the baseline average), and hospice costs by $7.87 per beneficiary per semester (6.01 percent of the baseline average). The High group is associated, potentially due to confounding factors, with an increase in physician costs. The High group is not associated with home health and durable medical equipment costs. Finally, the High group is not associated, potentially due to confounding factors, with outpatient costs.

Preview of Findings for Cost Outcomes in rural HSAs: Rural HSAs with Medium non-behavioral telehealth intensity are associated with a decrease in total cost of care per beneficiary. Rural HSAs with High non-behavioral telehealth intensity are not associated with the total cost of care per beneficiary.

Exhibit 14 presents these findings for rural HSAs. For context, the overall trend across rural HSAs in the Low group is an increase in the average total cost of care per beneficiary per semester by $988.33 between the baseline period and the second semester of 2022 (16.21 percent of the baseline rate).

Rural Medium non-behavioral telehealth intensity: Rural HSAs with Medium non-behavioral telehealth intensity are associated with a decrease in the total cost of care per beneficiary per semester by $87.34 (1.46 percent of the baseline rate). The decrease is primarily driven by a $60.47 decrease in outpatient cost (3.09 percent of the baseline). The Medium group is not associated with inpatient, skilled nursing facilities, home health, hospice, physician, and durable medical equipment costs.

Rural High non-behavioral telehealth intensity: Rural HSAs with High non-behavioral telehealth intensity are not associated with total cost of care. However, the High group is associated with a $73.92 increase in inpatient costs (4.14 percent of the baseline). The High group is also associated, potentially due to confounding factors, with a decrease in outpatient and an increase in durable medical equipment costs. The High group is not associated with skilled nursing facility, home health, and hospice costs. Finally, the High group is not associated, potentially due to confounding factors, with physician costs.

Exhibit 14. Change in Total Cost of Care per Beneficiary per Semester (in Dollars) for Non-Behavioral Telehealth HSAs

	Low	Medium relative to Low	High relative to Low
Urban HSAs			
All claim types	776.02 (13.24%) ***	-68.24 (-1.13%)	144.73 (2.16%) ***
Inpatient	206.48 (12.68%) ***	20.93 (1.19%)	68.24 (3.40%) ***
Outpatient	189.19 (11.47%) **	-26.07 (-1.66%)	-26.38 (-1.78%)
Skilled nursing facility	105.87 (25.88%) **	-18.65 (-4.91%)	37.40 (7.83%) **
Home health	39.65 (19.78%) **	-13.39 (-5.99%) ***	-2.37 (-0.82%)
Hospice	-28.30 (-21.96%) **	-1.66 (-1.35%)	7.87 (6.01%) **
Physician	236.91 (14.22%) ***	-23.64 (-1.30%)	64.02 (2.95%) ***
Durable medical equipment	26.22 (14.71%) **	-5.77 (-3.41%)	-4.06 (-2.64%)
Rural HSAs			
All claim types	988.33 (16.21%) ***	-87.34 (-1.46%) **	59.18 (0.94%)
Inpatient	6.17 (0.38%)	-4.03 (-0.24%)	73.92 (4.14%) **
Outpatient	657.36 (31.42%) ***	-60.47 (-3.09%) **	-76.09 (-3.85%) **
Skilled nursing facility	166.60 (29.93%) **	1.91 (0.40%)	23.11 (4.49%)
Home health	44.48 (27.34%) **	-6.01 (-3.28%)	-2.67 (-1.20%)
Hospice	43.77 (45.68%) ***	-3.40 (-3.32%)	-0.97 (-0.99%)
Physician	26.78 (1.93%)	-19.61 (-1.37%)	28.73 (1.90%)
Durable medical equipment	43.17 (21.89%) ***	4.27 (2.22%)	13.15 (7.07%) **

Note. Estimates presented in the exhibit are coefficients from DID regressions along with the standard error and statistical significance. Low, Medium, and High denote non-behavioral telehealth intensity. Estimates show the change between the baseline period (average of second semesters of 2018 and 2019) and the second semester of 2022. The denominator for the percentages is that group's average in the baseline period. *** and ** denote statistical significance at 1 and 5 percent levels, respectively.

3.2.3 Robustness Checks

3.2.3.1. 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, 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 the price of claims was 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 (GPCIs) (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. In Exhibit 15, we present the DID results unadjusted for the GAFs and the DID results adjusted for GAFs for HSAs with Medium non-behavioral telehealth intensity and High non-behavioral telehealth intensity compared to HSAs with Low non-behavioral telehealth intensity.

Urban HSAs: For urban HSAs with Medium and High non-behavioral telehealth intensity, adjusting for GAFs in the DID regressions for the quality, overall access, and total cost outcomes does not substantially change impact estimates in terms of statistical significance and interpretation.

Rural HSAs: For rural HSAs with Medium and High non-behavioral telehealth intensity, adjusting for GAFs in the DID regressions for the quality, overall access, and total cost outcomes does not substantially change impact estimates in terms of statistical significance and interpretation.

Exhibit 15. Effect of Controlling for GAFs on the Impact Estimates for Non-Behavioral Telehealth HSAs

Outcome	Unadjusted impact (Medium vs. Low)	Adjusted Impact (Medium vs. Low)	Unadjusted impact (High vs. Low)	Adjusted Impact (High vs. Low)
Urban HSAs				
ACS hospitalizations per 1,000 beneficiaries per semester	-0.01 (-0.06%)	0.01 (0.06%)	0.54 (2.45%)	0.59 (2.67%)
ACS ED visits per 1,000 beneficiaries per semester	0.62 (1.56%)	0.65 (1.64%)	1.72 (5.28%) **	1.81 (5.54%) **

Outcome	Unadjusted impact (Medium vs. Low)	Adjusted Impact (Medium vs. Low)	Unadjusted impact (High vs. Low)	Adjusted Impact (High vs. Low)
Clinician encounters per beneficiary per semester	-0.13 (-1.18%) **	-0.12 (-1.10%) **	-0.03 (-0.24%)	-0.02 (-0.13%)
Total cost of care per beneficiary per semester (in dollars)	-68.24 (-1.13%)	-56.30 (-0.93%)	144.73 (2.16%) ***	148.39 (2.21%) ***
Rural HSAs				
ACS hospitalizations per 1,000 beneficiaries per semester	0.51 (2.17%)	0.57 (2.42%)	-0.01 (-0.06%)	-0.05 (-0.19%)
ACS ED visits per 1,000 beneficiaries per semester	-0.75 (-1.53%)	-0.76 (-1.55%)	-0.94 (-1.80%)	-0.94 (-1.79%)
Clinician encounters per beneficiary per semester	-0.03 (-0.35%)	-0.02 (-0.17%)	0.03 (0.35%)	0.05 (0.48%)
Total cost of care per beneficiary per semester (in dollars)	-87.34 (-1.46%) **	-101.64 (-1.70%) **	59.18 (0.94%)	39.36 (0.63%)

Note. Estimates presented in the exhibit are coefficients from DID regressions along with the standard error and statistical significance. ACS = ambulatory care sensitive; ED = emergency department; GAFs = geographic adjustment factors. Impact estimates are provided separately for the High and Medium non-behavioral telehealth intensity groups, and are relative to the Low non-behavioral telehealth intensity group. The denominator for the percentages is the baseline average. *** and ** denote statistical significance at 1 and 5 percent levels, respectively.

3.2.3.2. Controlling for In-Person Utilization

Our main specification does not control for in-person utilization of healthcare services. 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. In Exhibit 16, we present the DID results unadjusted for in-person utilization and the DID results adjusted for in-person utilization for HSAs with Medium and High non-behavioral telehealth intensity compared to HSAs with Low non-behavioral telehealth intensity.

Urban HSAs: For urban HSAs with Medium non-behavioral telehealth intensity, adjusting for in-person utilization in the DID regressions for the quality and total cost outcomes does not substantially change impact estimates in terms of interpretation of the association.

However, for urban HSAs with High telehealth intensity, adjusting for in-person utilization changes the association and interpretation for both quality outcomes: ACS hospitalizations and ACS ED visits. For ACS hospitalizations per beneficiary per semester, while the High group is not associated in the unadjusted DID estimates, when we adjust for in-person utilization it is associated with an increase of 0.79 per beneficiary per semester (3.60 percent of the baseline average). For ACS ED visits per beneficiary per semester, the High group is associated with an increase, potentially due to confounding factors in the unadjusted DID estimates, but the association no longer exists when we adjust for in-person utilization. Therefore, the association of the High group with the quality outcomes is susceptible to omitted variable bias from excluding in-person utilization.

Rural HSAs: Exhibit 16 shows that after adjusting for in-person utilization, the impact estimates on quality and total cost of care outcomes for HSAs with Medium and High non-behavioral telehealth intensity do not change in terms of interpretation.

Exhibit 16. Effect of Controlling for In-Person Utilization on the Impact Estimates for Non-Behavioral Telehealth HSAs

Outcome	Unadjusted impact (Medium vs. Low)	Adjusted Impact (Medium vs. Low)	Unadjusted impact (High vs. Low)	Adjusted Impact (High vs. Low)
Urban HSAs				
ACS hospitalizations per 1,000 beneficiaries per semester	-0.01 (-0.06%)	0.12 (0.53%)	0.54 (2.45%)	0.79 (3.60%) **
ACS ED visits per 1,000 beneficiaries per semester	0.62 (1.56%)	0.05 (0.14%)	1.72 (5.28%) **	0.66 (2.01%)
Total cost of care per beneficiary per semester (in dollars)	-68.24 (-1.13%)	-45.55 (-0.75%)	144.73 (2.16%) ***	187.47 (2.79%) ***
Rural HSAs				
ACS hospitalizations per 1,000 beneficiaries per semester	0.51 (2.17%)	0.58 (2.50%)	-0.01 (-0.06%)	0.07 (0.27%)
ACS ED visits per 1,000 beneficiaries per semester	-0.75 (-1.53%)	-0.98 (-2.00%)	-0.94 (-1.80%)	-1.18 (-2.26%)
Total cost of care per beneficiary per semester (in dollars)	-87.34 (-1.46%) **	-88.11 (-1.47%)**	59.18 (0.94%)	58.36 (0.93%)

Note. Estimates presented in the exhibit are coefficients from DID regressions along with the standard error and statistical significance. ACS = ambulatory care sensitive; ED = emergency department. Impact estimates are provided separately for the High and Medium non-behavioral telehealth intensity groups, and are relative to the Low non-behavioral telehealth intensity group. The denominator for the percentages is the baseline average. *** and ** denote statistical significance at 1 and 5 percent levels, respectively

3.2.3.3. 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 181 HSAs out of 3,436 HSAs that are in the main sample.

As a final robustness check, we excluded small HSAs from the sample because the outcome rates are considered less reliable due to the small denominator. In Exhibit 17, we present the DID results for the full sample and the DID results restricted to the sample excluding small HSAs for HSAs with Medium and High non-behavioral telehealth intensity compared to the Low group.

Urban HSAs: For urban HSAs with Medium and High non-behavioral telehealth intensity, excluding small HSAs in the DID regressions for quality, overall access, and total cost outcomes does not substantially change impact estimates in terms of interpretation of the association.

Rural HSAs: Exhibit 17 also presents these results for rural HSAs. The impact estimates of rural HSAs with Medium non-behavioral telehealth intensity on ACS hospitalizations, ACS ED visits, and clinician encounters do not change in interpretation when we exclude small HSAs. On the other hand, excluding small HSAs changes the interpretation of the impact estimates for Medium HSAs for total cost of care. While there is a negative association for the Medium group with total cost of care in the full sample, in a sample excluding small HSAs the Medium group is not associated with the total cost of care, potentially due to confounding factors. Therefore, the full sample impact estimates on the total cost outcomes for the Medium group is subject to reliability concerns from small HSAs.

Excluding small HSAs does not change the interpretation of impact estimates for rural HSAs with High non-behavioral telehealth intensity for quality, overall access, and total cost outcomes.

Exhibit 176. Effect of Excluding Small HSAs on the Impact Estimates for Non-Behavioral Telehealth HSAs

Outcome	Full Sample (Medium vs. Low)	Excluding Small HSAs (Medium vs. Low)	Full Sample (High vs. Low)	Excluding Small HSAs (High vs. Low)
Urban HSAs				
ACS hospitalizations per 1,000 beneficiaries per semester	-0.01 (-0.06%)	0.05 (0.23%)	0.54 (2.45%)	0.57 (2.58%)
ACS ED visits per 1,000 beneficiaries per semester	0.62 (1.56%)	0.57 (1.44%)	1.72 (5.28%) **	1.63 (5.01%) **
Clinician encounters per beneficiary per semester	-0.13 (-1.18%) **	-0.14 (-1.35%) **	-0.03 (-0.24%)	-0.04 (-0.31%)
Total cost of care per beneficiary per semester (in dollars)	-68.24 (-1.13%)	-66.67 (-1.11%)	144.73 (2.16%) ***	140.31 (2.09%) ***
Rural HSAs				
ACS hospitalizations per 1,000 beneficiaries per semester	0.51 (2.17%)	0.48 (2.06%)	-0.01 (-0.06%)	0.18 (0.74%)
ACS ED visits per 1,000 beneficiaries per semester	-0.75 (-1.53%)	-0.73 (-1.48%)	-0.94 (-1.80%)	-1.38 (-2.65%)
Clinician encounters per beneficiary per semester	-0.03 (-0.35%)	-0.04 (-0.39%)	0.03 (0.35%)	-0.03 (-0.32%)
Total cost of care per beneficiary per semester (in dollars)	-87.34 (-1.46%) **	-57.00 (-0.96%)	59.18 (0.94%)	27.10 (0.44%)

Note. Estimates presented in the exhibit are coefficients from DID regressions along with the standard error and statistical significance. ACS = ambulatory care sensitive; ED = emergency department; HSAs = Hospital Service Areas. Impact estimates are provided separately for the High and Medium non-behavioral telehealth intensity groups, and are relative to the Low behavioral telehealth intensity group. The denominator for the percentages is the baseline average. *** and ** denote statistical significance at 1 and 5 percent levels, respectively.

3.3. Main Findings: Behavioral Telehealth

This section presents DID estimates of the association between HSAs with High versus Low behavioral telehealth intensity and HSAs with Medium versus Low behavioral telehealth intensity. In addition to presenting the impact estimates, we discuss whether the estimates for telehealth impact on outcomes are valid based on PTTs and robustness checks.

3.3.1. The Parallel Trends Assumption

We checked whether there is a statistically significant difference in outcomes between HSA groups with Medium and Low behavioral telehealth intensity and between HSA groups with High and Low behavioral telehealth intensity during the PTT period. We used multivariate

regression analysis to adjust for the differences in covariates that are included in the DID models (see **2.5. EMPIRICAL STRATEGY**). As previously discussed, to pass parallel trends, the p-value of the joint hypothesis test must not be significant at the 1 percent, 5 percent, or 10 percent levels (i.e., should have p-value > 0.10). Therefore, our approach for testing parallel trends counts even small baseline differences (with low statistical significance at the 10 percent significance level) between groups as indicative of PTT failure. Exhibit 18 summarizes parallel trends comparing HSAs with High and Medium to Low behavioral telehealth intensity. Note that for behavioral telehealth, we do not assess the impact of behavioral telehealth intensity in urban and rural HSAs on quality outcomes (ACS hospitalizations and ACS ED visits per beneficiary per semester), as these outcomes are not conceptually related to behavioral healthcare.

Urban HSAs: Among urban HSAs with Medium behavioral telehealth intensity, the main outcomes (clinician encounters and total cost of care) passed the PTT. The main outcomes among urban HSAs with High behavioral telehealth intensity also passed the PTT. A visual inspection of the trends (Exhibit 8, third quadrant; and Exhibit 9, third quadrant) for urban HSAs during the PTT period shows that the trends for clinician encounters and total cost of care are largely similar for HSAs with High, Medium, and Low behavioral telehealth intensity.

Across all 13 outcomes evaluated for the impact of behavioral telehealth utilization, outcomes for urban HSAs with Medium behavioral intensity passed the PTT more frequently (85 percent; 11 out of 13 outcomes) than did outcomes for urban HSAs with High behavioral intensity (46 percent; 6 out of 13 outcomes).

Rural HSAs: Among rural HSAs, we found that the main outcomes for Medium behavioral telehealth intensity passed the PTT. In contrast, for rural HSAs with High behavioral telehealth intensity, one of the main outcomes (total cost of care) passed the PTT and another main outcome (clinician encounters) failed the PTT. A visual inspection of the trends (Exhibit 8, fourth quadrant; and Exhibit 9, fourth quadrant) supports the PTT findings.

Across all 13 outcomes evaluated for the impact of behavioral telehealth utilization, outcomes for rural HSAs with Medium intensity passed the PTT more frequently (92 percent; 12 out of 13 outcomes) than did outcomes for rural HSAs with High intensity (54 percent; 7 out of 13 outcomes).

Exhibit 18. Parallel Trends Test for Quality, Access, and Cost, for Behavioral Telehealth Intensity

	Urban HSAs		Rural HSAs	
	Medium relative to Low	High relative to Low	Medium relative to Low	High relative to Low
Clinician encounters per beneficiary	PTT Pass	PTT Pass	PTT Pass	
Primary care physicians	PTT Pass	PTT Pass	PTT Pass	
Specialists (including Hospitalists)	PTT Pass		PTT Pass	
APRNs/Pas	PTT Pass		PTT Pass	PTT Pass
Other practitioners	PTT Pass			PTT Pass
Total cost of care per beneficiary	PTT Pass	PTT Pass	PTT Pass	PTT Pass
Inpatient	PTT Pass	PTT Pass	PTT Pass	PTT Pass
Outpatient	PTT Pass		PTT Pass	PTT Pass
Skilled nursing facility	PTT Pass	PTT Pass	PTT Pass	PTT Pass
Home health			PTT Pass	
Hospice	PTT Pass		PTT Pass	PTT Pass
Physician	PTT Pass	PTT Pass	PTT Pass	
Durable medical equipment			PTT Pass	

Note. APRNs = advanced practice registered nurses; PA = physician assistant. Green (or PTT Pass) denotes that the PTT passed, as the differences between HSAs with Medium or High behavioral telehealth intensity and HSAs with Low intensity were not statistically significant (p-value > 0.10). Blank denotes that the PTT failed, as the differences between HSAs with Medium or High behavioral telehealth intensity and HSAs with Low intensity were statistically significant (p-value ≤ 0.10). To statistically test for parallel trends, a joint hypothesis test was conducted to ensure differences in outcomes between High and Low groups or Medium and Low groups over three semesters (2018 semester 1, 2018 semester 2, and 2019 semester 1) were similar and not statistically significant from the second semester in 2019.

3.3.2. Impact Estimates

Exhibit 19 presents a summary of the impact estimates related to the impact of behavioral telehealth intensity on access and costs for urban HSAs separately from rural HSAs. All impact estimates presented are significant at the 5 percent level or 1 percent significance level. 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 19. Impact of Behavioral Telehealth Intensity on Access and Cost

	Urban HSAs		Rural HSAs	
	Medium relative to Low	High relative to Low	Medium relative to Low	High relative to Low
Clinician encounters per beneficiary	↘			
Primary care physicians	↘		↘	
Specialists (including Hospitalists)	↘			
APRNs/Pas		↘		
Other practitioners				
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-shaded cells with a blue arrow pointing up (or ↗) denote a statistically significant increase, PTT passed. White-shaded cells with a blue arrow pointing up (or ↗) denote a statistically significant increase, PTT failed. Orange-shaded cells with a red arrow pointing down (or ↘) denote a statistically significant decrease, PTT passed. White-shaded cells with a red arrow pointing down (or ↘) denote a statistically significant decrease, PTT failed. Grey-shaded cells denote no association, PTT passed. Blank (white) cells denote no association, PTT failed. We only consider DID results significant if the DID coefficient has a p-value ≤ 0.05. Impact estimates show the change between the baseline period (average of the second semesters of 2018 and 2019) and the treatment period (second semester of 2022) relative to Low behavioral telehealth intensity.

Preview of Findings for Access Outcomes in Urban HSAs: Urban HSAs with Medium behavioral telehealth intensity are associated with a decrease in clinician encounters. In contrast, urban HSAs with High behavioral telehealth intensity are not associated with clinician encounters.

Exhibit 20 presents these findings. For context, the overall trend between the baseline period and the second semester of 2022 across urban HSAs in the Low behavioral group is a decline, which is not statistically significant, of 0.13 clinician encounters per beneficiary per semester (1.29 percent of the baseline rate).

Urban Medium behavioral telehealth intensity: Urban HSAs with Medium behavioral telehealth intensity are associated with a decrease of 0.17 clinician encounters per beneficiary per semester, which is 1.60 percent of the baseline rate. For encounters with clinician types, the Medium group is associated with a 0.05 per beneficiary per semester decrease in encounters with primary care physicians (a 2.89 percent decrease relative to the baseline rate) and a 0.11 per beneficiary per semester decrease in encounters with specialists (a 1.76 percent decrease relative to the baseline). The Medium group is not associated with APRN/Pas and other practitioner encounters.

Urban High behavioral telehealth intensity: Urban HSAs with High behavioral telehealth intensity are not associated with overall clinician encounters and encounters with primary care physicians. Urbans HSAs with High behavioral telehealth intensity are associated, potentially due to confounding factors, with a decrease in encounters with APRNs/Pas. Finally, the High group is not associated, potentially due to confounding factors, with specialists and other practitioner encounters.

Preview of Findings for Access Outcomes in Rural HSAs: Rural HSAs with Medium behavioral telehealth intensity are not associated with overall clinician encounters. However, the rural Medium group is associated with fewer primary care physician encounters. Rural HSAs with High behavioral telehealth intensity are not associated, potentially due to confounding factors, with overall clinician encounters.

Exhibit 20 presents the findings for rural HSAs. For context, the overall trend across rural HSAs in the Low group is a decline of 0.26 (not statistically significant) clinician encounters per beneficiary per semester between the baseline period and the second semester of 2022, which is 2.80 percent of the baseline rate.

Rural Medium behavioral telehealth intensity: Rural HSAs with Medium behavioral telehealth intensity are not associated with total clinician encounters per beneficiary. When considering encounters by clinician type, the Medium group is associated with a decrease in primary care physician encounters by 0.04 encounters per beneficiary per semester (2.52 percent of the baseline rate). The Medium group was not associated with encounters with specialists

(including hospitalists) and APRNs/Pas. In addition, the Medium group was not associated, potentially due to confounding factors, with other practitioner encounters.

Rural High behavioral telehealth intensity: Rural HSAs with High behavioral telehealth intensity are not associated, potentially due to confounding factors, with total clinician encounters per beneficiary. For encounters by clinician type, the High group is not associated with APRN/PA and other practitioner encounters. The High group is not associated, potentially due to confounding factors, with primary care physician and specialist (including hospitalist) encounters.

Exhibit 20. Change in Clinician Encounters per Beneficiary per Semester for Behavioral Telehealth HSAs

	Low	Medium relative to Low	High relative to Low
Urban HSAs			
All clinicians	-0.13 (-1.29%)	-0.17 (-1.60%) **	-0.12 (-1.01%)
Primary care physicians	-0.07 (-3.54%)	-0.05 (-2.89%) **	-0.01 (-0.70%)
Specialists (including Hospitalists)	-0.13 (-2.23%)	-0.11 (-1.76%) **	-0.02 (-0.26%)
APRNs/Pas	0.22 (16.35%) **	-0.01 (-0.54%)	-0.10 (-9.04%) ***
Other practitioners	-0.13 (-9.99%) **	-0.01 (-0.74%)	0.02 (1.05%)
Rural HSAs			
All clinicians	-0.26 (-2.80%)	-0.04 (-0.39%)	-0.06 (-0.66%)
Primary care physicians	-0.16 (-9.85%) **	-0.04 (-2.52%) **	-0.04 (-2.58%)
Specialists (including Hospitalists)	-0.08 (-1.58%)	-0.01 (-0.21%)	-0.01 (-0.19%)
APRNs/Pas	-0.03 (-1.81%)	0.02 (1.54%)	0.02 (1.78%)
Other practitioners	-0.09 (-6.85%)	0.01 (0.37%)	-0.02 (-1.14%)

Note. Estimates presented in the exhibit are coefficients from DID regressions along with the standard error and statistical significance. APRNs = advanced practice registered nurses; Pas = physician assistants. Low, Medium, and High denote behavioral telehealth intensity. Estimates show the change between the baseline period (average of second semesters of 2018 and 2019) and the second semester of 2022. The denominator for the percentages is that group's average in the baseline period. *** and ** denote statistical significance at 1 and 5 percent levels, respectively.

Preview of Findings for Cost Outcomes in Urban HSAs: Urban HSAs with Medium behavioral telehealth intensity are associated with a decrease in total cost of care per beneficiary. Urban HSAs with High behavioral telehealth intensity are not associated with total cost of care per beneficiary.

Exhibit 21 presents these findings. For context, the overall trend across HSAs in the Low group is an increase in the average total cost of care per beneficiary per semester of $950.69 between the baseline period and the second semester of 2022 (15.53 percent of the baseline rate).

Urban Medium behavioral telehealth intensity: Urban HSAs with Medium behavioral telehealth intensity are associated with a decrease in the total cost of care per beneficiary per semester by $132.45 (2.18 percent of the baseline). For cost of care by claim type, the Medium group is associated with a decrease in skilled nursing facility costs by $37.14 (9.55 percent of the baseline average), hospice costs by $8.76 (6.70 percent of the baseline average), and physician costs by $39.99 (2.13 percent of the baseline average). The Medium group is also associated, potentially due to confounding factors, with a decrease in home health costs.

Urban High behavioral telehealth intensity: Urban HSAs with High behavioral telehealth intensity are not associated with total cost of care. For cost of care by claim type, the High group is not associated with inpatient, skilled nursing facility, and physician costs. The High group is associated, potentially due to confounding factors, with a decrease in home health costs. Finally, the High group is not associated, potentially due to confounding factors, with outpatient, hospice, and durable medical equipment costs.

Preview of Findings for Cost Outcomes in Rural HSAs: Rural HSAs with Medium behavioral telehealth intensity are associated with a decrease in the total cost of care per beneficiary per semester. Rural HSAs with High behavioral telehealth intensity are associated with a decrease in the total cost of care per beneficiary per semester.

Exhibit 21 presents the findings for rural HSAs. For context, the overall trend across rural HSAs with Low behavioral telehealth intensity is an increase in the average total cost of care per beneficiary per semester of $1,090.58 between the baseline period and the second semester of 2022 (17.49 percent of the baseline rate).

Rural Medium behavioral telehealth intensity: Rural HSAs with Medium behavioral telehealth intensity are associated with a decrease in the total cost of care per beneficiary per semester by $108.70 (1.82 percent of the baseline average). For cost of care by claim types, the Medium group is associated with a decrease in per beneficiary per semester costs for skilled nursing facility by $37.66 (7.80 percent of the baseline average), home health by $11.48 (6.72 percent of the baseline average), and hospice by $8.79 (9.21 percent of the baseline average). The Medium group is not associated with inpatient, outpatient, and physician costs.

Rural High behavioral telehealth intensity: Rural HSAs with High behavioral telehealth intensity are associated with a decrease in the total cost of care per beneficiary per semester by $118.30 (2.01 percent of the baseline average). For cost of care by claim types, the High group is associated with a decrease in per beneficiary per semester costs for hospice by $15.37 (17.99 percent of the baseline average). The High group is associated, potentially due to confounding factors, with a decrease in the home health costs. The High group is not associated with inpatient, outpatient, or skilled nursing facility costs. Finally, the High group is not associated, potentially due to confounding factors, with physician, and durable medical equipment costs.

Exhibit 21. Change in Total Cost of Care per Beneficiary per Semester (in Dollars) for Behavioral Telehealth HSAs

	Low	Medium relative to Low	High relative to Low
Urban HSAs			
All claim types	950.69 (15.53%) ***	-132.45 (-2.18%) **	0.48 (0.01%)
Inpatient	289.33 (16.86%) ***	-17.54 (-1.00%)	13.15 (0.67%)
Outpatient	153.39 (8.99%) **	-7.13 (-0.47%)	-5.43 (-0.36%)
Skilled nursing facility	140.99 (32.79%) ***	-37.14 (-9.55%) **	-11.15 (-2.41%)
Home health	60.18 (25.02%) ***	-19.86 (-8.50%) ***	-11.64 (-4.42%) **
Hospice	-15.37 (-11.06%)	-8.76 (-6.70%) **	-3.42 (-2.78%)
Physician	302.27 (17.65%) ***	-39.99 (-2.13%) **	19.67 (0.94%)
Durable medical equipment	19.95 (11.17%)	-2.03 (-1.16%)	-0.71 (-0.47%)
Rural HSAs			
All claim types	1,090.58 (17.49%) ***	-108.70 (-1.82%) **	-118.30 (-2.01%) **
Inpatient	31.54 (1.89%)	-6.45 (-0.39%)	14.54 (0.87%)
Outpatient	623.69 (30.60%) ***	-21.88 (-1.11%)	-46.84 (-2.26%)
Skilled nursing facility	201.93 (35.49%) ***	-37.66 (-7.80%) **	-30.64 (-7.01%)
Home health	59.20 (30.30%) ***	-11.48 (-6.72%) **	-16.82 (-10.72%) **
Hospice	52.28 (49.89%) ***	-8.79 (-9.21%) **	-15.37 (-17.99%) **

	Low	Medium relative to Low	High relative to Low
Physician	71.69 (4.92%)	-17.94 (-1.25%)	-22.64 (-1.74%)
Durable medical equipment	50.26 (25.21%) ***	-4.50 (-2.34%)	-0.54 (-0.31%)

Note. Estimates presented in the exhibit are coefficients from DID regressions along with the standard error and statistical significance. Low, Medium, and High denote behavioral telehealth intensity. Estimates show the change between the baseline period (average of second semesters of 2018 and 2019) and the second semester of 2022. The denominator for the percentages is that group's average in the baseline period. *** and ** denote statistical significance at 1 and 5 percent levels, respectively.

3.3.3. Robustness Checks

3.3.3.1 Controlling for Geographic Adjustment Factors

As a robustness check, we included a measure of the hospital wage index from the IPPS and three measures of GPCIs from the PFS as covariates in the DID regressions for all outcomes. As discussed in Section 3.2.3.1, the purpose of this check is to assess if estimated impacts are driven by the differences in the GAFs as opposed to differences in resource use between areas of varying telehealth intensities. In Exhibit 22, we present the DID results unadjusted for the GAFs and the DID results adjusted for GAFs for HSAs with Medium behavioral telehealth intensity and HSAs with High behavioral telehealth intensity compared to HSAs with Low behavioral telehealth intensity.

Urban HSAs: Adding the IPPS hospital wage index and the three GPCIs in the DID regressions does not change the interpretations of the impact estimates for urban HSAs with Medium behavioral telehealth intensity on clinician encounters and total cost of care. The Medium group is still associated with a decrease in clinician encounters and total cost care after adjusting for GAFs. Therefore, for the urban HSAs in the Medium group, impact estimates for clinician encounters and total cost of care are robust to any potential differences in GAFs. Adjusting for GAFs did not substantially change the interpretation of the impact estimates on clinician encounters and total cost of care for urban HSAs with High behavioral telehealth intensity.

Rural HSAs: Adding the IPPS hospital wage index and the three GPCIs in the DID regressions does not change the interpretation of impact estimates on clinician encounters and total cost of care for rural HSAs with Medium behavioral telehealth intensity. The Medium group continues to be associated with a decrease in total cost of care after adjusting for differences in GAFs. Similarly, rural HSAs with High behavioral intensity continue to be associated with a decrease in total cost care after adjusting for differences in GAFs. Therefore, impact estimates on the total

cost of care for the rural HSAs in the Medium and High groups are robust to any potential differences in GAFs.

Exhibit 22. Effect of Controlling for GAFs on the Impact Estimates for Behavioral Telehealth HSAs

Outcome	Unadjusted impact (Medium vs Low)	Adjusted Impact (Medium vs Low)	Unadjusted impact (High vs Low)	Adjusted Impact (High vs Low)
Urban HSAs				
Clinician encounters per beneficiary per semester	-0.17 (-1.60%) **	-0.11 (-1.06%) **	-0.12 (-1.01%)	-0.09 (-0.80%)
Total cost of care per beneficiary per semester (in dollars)	-132.45 (-2.18%) **	-106.84 (-1.76%) **	0.48 (0.01%)	10.46 (0.16%)
Rural HSAs				
Clinician encounters per beneficiary per semester	-0.04 (-0.39%)	-0.03 (-0.27%)	-0.06 (-0.66%)	-0.08 (-0.92%)
Total cost of care per beneficiary per semester (in dollars)	-108.70 (-1.82%) **	-122.83 (-2.05%) **	-118.30 (-2.01%) **	-131.87 (-2.24%) **

Note. Estimates presented in the exhibit are coefficients from DID regressions along with the standard error and statistical significance. GAFs = geographic adjustment factors. Impact estimates are for the High and Medium behavioral telehealth intensity group, and are relative to the Low behavioral telehealth intensity group. The denominator for the percentages is the baseline average. *** and ** denote statistical significance at 1 and 5 percent levels, respectively.

3.3.3.2. Controlling for In-Person Utilization

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. In Exhibit 23, we present the DID results unadjusted for in-person utilization and the DID results adjusted for in-person utilization for HSAs in Medium and High behavioral telehealth intensity groups compared to HSAs in the Low behavioral telehealth intensity group. As previously discussed in Section 3.2.3.2, this check is to assess whether the effect of telehealth usage may be confounded with the effect of general health care utilization.

Urban HSAs: Adjusting for in-person utilization does not change interpretation of the impact estimates on the total cost of care for urban HSAs with Medium and High behavioral telehealth intensity. In particular, the Medium group continues to be associated with a decrease in total cost of care after controlling for in-person utilization. Therefore, the impact estimate of urban HSAs in the Medium group on total cost of care is robust and not susceptible to omitted variable bias due to the exclusion of in-person utilization.

Rural HSAs: For rural HSAs, adjusting for in-person utilization in the DID regressions for the total cost of care does not change the interpretation of impact estimates on the total cost of care for the Medium and High groups. Both the Medium and High groups continued to be associated with a decrease in total cost of care after adjusting for in-person utilization. Impact estimates for rural HSAs with Medium and High behavioral telehealth intensity on the total cost of care are robust and not susceptible to omitted variable bias by excluding in-person utilization.

Exhibit 23. Effect of Controlling for In-Person Utilization on the Impact Estimates for Behavioral Telehealth HSAs

Outcome	Unadjusted Impact (Medium vs. Low)	Adjusted Impact (Medium vs. Low)	Unadjusted Impact (High vs. Low)	Adjusted Impact (High vs. Low)
Urban HSAs				
Total cost of care per beneficiary per semester (in dollars)	-132.45 (-2.18%) **	-103.17 (-1.70%) **	0.48 (0.01%)	50.91 (0.78%)
Rural HSAs				
Total cost of care per beneficiary per semester (in dollars)	-108.70 (-1.82%) **	-108.96 (-1.82%) **	-118.30 (-2.01%) **	-118.88 (-2.02%) **

Note. Estimates presented in the exhibit are coefficients from DID regressions along with the standard error and statistical significance. Impact estimates are for the High and Medium behavioral telehealth intensity groups, and are relative to the Low behavioral telehealth intensity group. The denominator for the percentages is the baseline average. *** and ** denote statistical significance at 1 and 5 percent levels, respectively.

3.3.3.2. 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 (i.e., HSAs with fewer than 500 beneficiaries) from the sample and repeated the analysis. In Exhibit 24, we present the DID results for the full sample and the DID results restricted to the sample excluding small HSAs for HSAs with Medium and High behavioral telehealth intensity compared to HSAs with Low behavioral telehealth intensity.

Urban HSAs: Excluding small HSAs does not change the interpretation of the impact estimates on clinician encounters and total cost of care for urban HSAs with Medium behavioral telehealth intensity. Therefore, the impact estimates of the urban HSAs in the Medium group on clinician encounters and total cost of care in the full sample is robust to reliability concerns from small HSAs.

Similarly, excluding small HSAs does not change the interpretation of the impact estimates on clinician encounters and total cost of care for urban HSAs with High behavioral telehealth intensity.

Rural HSAs: Excluding small HSAs does not change the interpretation of impact estimates on total cost of care for rural HSAs with Medium behavioral telehealth intensity. However, excluding small HSAs changes the interpretation of the impact estimates on clinician encounters for rural HSAs in the Medium group. While the Medium group is not associated with clinician encounters in the full sample, excluding small HSAs, the Medium group is associated with a 0.17 per beneficiary per semester decrease in clinician encounters.

Excluding small HSAs changes the impact estimates on clinician encounters for rural HSAs with High behavioral telehealth intensity. In the full sample, the rural High group is not associated, potentially due to confounding factors, with clinician encounters. Excluding small HSAs, the rural High group is associated, potentially due to confounding factors, with a decrease in clinician encounters. Impact estimates for the total cost of care outcome are similar across both specifications for the rural High group.

Exhibit 24. Effect of Excluding Small HSAs on the Impact Estimates for Behavioral Telehealth HSAs

Outcome	Full Sample (Medium vs Low)	Excluding Small HSAs (Medium vs Low)	Full Sample (High vs Low)	Excluding Small HSAs (High vs Low)
Urban HSAs				
Clinician encounters per beneficiary per semester	-0.17 (-1.60%) **	-0.17 (-1.56%) **	-0.12 (-1.01%)	-0.09 (-0.77%)
Total cost of care per beneficiary per semester (in dollars)	-132.45 (-2.18%) **	-127.74 (-2.11%) **	0.48 (0.01%)	6.18 (0.09%)
Rural HSAs				
Clinician encounters per beneficiary per semester	-0.04 (-0.39%)	-0.07 (-0.79%) **	-0.06 (-0.66%)	-0.10 (-1.13%) **
Total cost of care per beneficiary per semester (in dollars)	-108.70 (-1.82%) **	-90.59 (-1.53%) **	-118.30 (-2.01%) **	-135.06 (-2.32%) **

Note. Estimates presented in the exhibit are coefficients from DID regressions along with the standard error and statistical significance. HSAs = Hospital Service Areas. Impact estimates are for the High and Medium behavioral telehealth intensity group, and are relative to the Low behavioral telehealth intensity group. The denominator for the percentages is the baseline average. *** and ** denote statistical significance at 1 and 5 percent levels, respectively.

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 estimated the effect of telehealth usage separately for urban and rural HSAs. Within each HSA type, we examined the effect of two types of telehealth usage: non-behavioral telehealth and behavioral telehealth. In each case, we estimated the effect of telehealth usage by comparing the average change in an outcome for HSAs with Medium or High telehealth intensity with the average change in that outcome for HSAs with Low telehealth intensity between the baseline period (second half of 2018 and 2019) and the second half of 2022.

This study updates our previous analysis (AIR, 2023) which informed MedPAC's June 2023 report to Congress. In addition to looking at the effects of telehealth separately for urban and rural beneficiaries and differentiating between behavioral and non-behavioral telehealth use, this study uses the most recently available data from a time when there were fewer extreme surges of COVID-19 cases in the treatment period, to address the confounding effects of COVID-19 present in the previous study.[10] We discuss our findings in the following subsections.

4.1. Interpretation of the Results

To help frame the overall findings, it is helpful to note two points. First, we find better comparability between HSAs with Medium and Low telehealth intensity than between HSAs with High and Low telehealth intensity at baseline. This is the case when HSAs are grouped according to both non-behavioral and behavioral telehealth intensity. This pattern is evidenced by looking at the magnitude of differences in baseline characteristics across HSAs with Low, Medium, and High telehealth intensity (see Section 3.1 Descriptive Statistics) and by looking at the proportion of outcomes that pass the PTT when outcome trends are compared across HSAs in Medium and Low groups versus HSAs in High and Low groups (see Sections 3.2.1 and 3.3.1, which cover results of the PTTs for non-behavioral and behavioral telehealth). **Given that DID results can be interpreted reliably only if the parallel trends assumption is satisfied, in this study, drawing conclusions regarding the effects of telehealth are more reliable when Medium and Low groups are compared than when High and Low groups are compared.**

[10] The treatment period in AIR (2023) was the second half of 2021, and the treatment period in this study is the second half of 2022.

Second, we find a strong overlap in HSAs classified as Low, Medium, and High telehealth intensity when grouped according to utilization of non-behavioral telehealth and behavioral telehealth services, especially for urban HSAs. Among urban HSAs, 79 percent of the same HSAs are High intensity telehealth for both non-behavioral telehealth and behavioral telehealth. This figure is 50 percent for Medium intensity HSAs and 45 percent for Low intensity HSAs. The correlation coefficient between the two treatments is also high, 0.59. Among rural HSAs, 35 percent of the same HSAs are High intensity telehealth for both non-behavioral telehealth and behavioral telehealth. This figure is 41 percent for Medium intensity HSAs and 64 percent for Low intensity HSAs. The correlation coefficient between the two treatments is 0.32, suggesting moderate to low correlation. **Overall, this indicates that it is challenging to interpret the effects of non-behavioral telehealth and behavioral telehealth as distinct treatments, especially in urban areas. This means, for example, that outcomes in HSAs with High behavioral telehealth intensity may be reflective of high utilization of both behavioral telehealth and non-behavioral telehealth. Therefore, attributing the outcome effects to behavioral telehealth, even when the DID estimates pass the PTT, would be misleading, as they are likely also substantially affected by the use of non-behavioral telehealth.**

We next summarize our findings for each set of outcomes and discuss plausible interpretations. In Section 3 (Results) we discuss impact estimates for HSAs with Medium and High telehealth intensity. However, to draw overall conclusions about the findings of this study, we take a holistic approach that considers whether most of the outcomes pass the PTT, based on the Medium versus Low and High versus Low analyses. Based on this approach, we conclude that the impact estimates for HSAs with Medium versus Low telehealth intensity can reasonably be interpreted as reflecting reliable associations, because in most cases more than two-thirds of the outcomes passed the PTT across both treatments and for urban and rural HSAs.[11] However, we also conclude that the impact estimates for High versus Low telehealth intensity should not be interpreted as reflecting reliable associations given the high proportion of outcomes that did not pass the PTT (in most cases 50 percent or fewer outcomes passed the PTT).[12]

[11] Considering only the main outcomes in the Medium versus Low analysis, the PTT passed for 75 percent of the main outcomes for urban HSAs with non-behavioral telehealth intensity, 50 percent of the main outcomes for rural HSAs with non-behavioral telehealth intensity, 100 percent of the main outcomes for urban HSAs with behavioral telehealth intensity, and 100 percent of outcomes for rural HSAs with behavioral telehealth intensity. Considering all outcomes, including main outcomes and sub-outcomes in the Medium versus Low analysis, the PTT passed for 80 percent of outcomes for urban HSAs with non-behavioral telehealth intensity, 67 percent of outcomes for rural HSAs with non-behavioral telehealth intensity, 85 percent of outcomes for urban HSAs with behavioral telehealth intensity, and 92 percent of outcomes for rural HSAs with behavioral telehealth intensity.

[12] Considering only the main outcome, in the High versus Low analysis, the PTT passed for 25 percent of outcomes for urban HSAs with non-behavioral telehealth intensity, 25 percent of outcomes for rural HSAs with non-behavioral telehealth intensity, 100 percent of outcomes for urban HSAs with behavioral telehealth intensity, and 50 percent of outcomes for rural HSAs with behavioral telehealth intensity. Considering all outcomes, including main outcomes and sub-outcomes, in the High versus Low analysis, the PTT passed for 40 percent of outcomes for urban HSAs with non-behavioral telehealth intensity, 33 percent of outcomes for HSAs with rural non-behavioral telehealth intensity, 46 percent of outcomes for urban HSAs with behavioral telehealth intensity, and 54 percent of outcomes for rural HSAs with behavioral telehealth intensity.

Given this holistic approach, the findings discussed below are largely based on impact estimates from HSAs with Medium telehealth intensity and on outcomes that passed the PTT among HSAs with Medium telehealth intensity.

Telehealth and quality. We do not observe an association between telehealth intensity and quality outcomes (ACS hospitalizations and ACS ED visits). When comparing urban HSAs in the Medium and Low groups, there is no association between non-behavioral telehealth intensity and ACS hospitalizations and ACS ED visits.[13] For rural HSAs, there is no association between non-behavioral telehealth intensity and ACS ED visits. There is some evidence of an increase in ACS ED visits in the urban HSAs with High telehealth intensity, but this increase is potentially due to confounding given the lack of baseline parallel trends.

The overall lack of association between telehealth intensity and quality outcomes is consistent with some other findings in the literature. For instance, Wilcock et al. (2023) compared patients receiving mental health care at practices that almost exclusively switched to telemental health service with those receiving care at practices that largely used in-person visits and found no differences in quality metrics—such as acute hospital and ED encounters, all-cause mortality, and medication adherence—between the two groups.

It is also important to note that our previous study (AIR, 2023) found greater telehealth intensity to be associated with increased ACS hospitalizations (a similar finding was also observed in Li et al., 2022). The treatment period of our previous study was still heavily confounded by the effects of the COVID-19 pandemic. As we discussed in AIR (2023), differences across HSAs—in terms of 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—very likely had an impact on both the telehealth intensity in an area and outcomes, such as ACS hospitalizations, making it difficult to cleanly interpret the observed increases in ACS hospitalizations. The estimated relationship in this study, which shows a lack of association between non-behavioral telehealth intensity and ACS hospitalizations and ACS ED visits, may be more reliable as the treatment period of this study (second half of 2022) was less confounded by effects of the COVID-19 pandemic.

Telehealth and access. In some of our analysis, we find evidence that higher telehealth intensity is associated with fewer clinician encounters. For urban HSAs, for both non-behavioral and behavioral telehealth, HSAs with Medium telehealth intensity are associated with a decrease in clinician encounters per beneficiary when compared to HSAs in the Low group. The magnitude of the association is small relative to baseline clinician encounters, ranging from a

[13] Recall that we do not estimate quality outcomes for behavioral telehealth intensity. It is also important to note that one quality outcome (ACS hospitalizations) in rural HSAs did not pass the PTT.

decline of 1.18 percent (0.13 fewer clinician encounters per beneficiary) to a decline of 1.60 percent (0.17 fewer clinician encounters per beneficiary) for urban and rural HSAs, respectively.

For rural HSAs, for behavioral telehealth, HSAs with Medium telehealth intensity are associated with a decrease in primary care physician encounters per beneficiary when compared to HSAs in the Low group. One speculative interpretation of this finding is that HSAs with Medium behavioral intensity benefited from new telehealth options for behavioral health services, thus alleviating the demand on primary care for such services. Rural HSAs with Medium behavioral telehealth intensity are also associated with a decrease in overall clinician encounters per beneficiary once small HSAs are excluded—a decline of 0.79 percent of the baseline rate (0.07 fewer clinician encounters per beneficiary).

These impact estimates can also be expressed in terms of the difference in telehealth use between HSAs with Medium and Low telehealth intensity. Urban HSAs with Medium telehealth intensity have, on average, 0.2 non-behavioral telehealth visits per beneficiary and HSAs in the Low group have 0.12 non-behavioral telehealth visits per beneficiary. Our DID estimates show that 0.08 additional non-behavioral telehealth visits per beneficiary in Medium HSAs translates to 0.13 fewer clinician encounters per beneficiary. This suggests that 1 additional non-behavioral telehealth visit per beneficiary is associated with 1.63 fewer clinician encounters per beneficiary in urban HSAs with Medium telehealth intensity. Similarly, among urban Medium HSAs, 1 additional behavioral telehealth visit per beneficiary is associated with 5.67 fewer clinician encounters per beneficiary.[14] Another calculation suggests that 1 additional behavioral telehealth visit per beneficiary is associated with 2.33 fewer clinician encounters per beneficiary, once small HSAs are excluded, in rural HSAs with Medium telehealth intensity.[15]

The decrease in clinician encounters we observe may relate to a transitory effect of the COVID-19 pandemic. The second half of 2022, which is the treatment period of this study, was a period when the pandemic had largely subsided in the majority of the United States.[16] A relative decline in clinician encounters in HSAs with Medium telehealth intensity compared to those with Low intensity could reflect less pent-up demand for care in these HSAs, as these areas were able to see providers more regularly during the pandemic. In other words, the DID findings for 2022 may be reflecting the fact that beneficiaries in HSAs with Medium telehealth intensity did not need to catch up with postponed care as much as beneficiaries in HSAs with

[14] The 5.67 fewer clinician encounters figure is obtained by dividing the relevant point-estimate (0.17) by the difference in the average number of behavioral telehealth visits per beneficiary in Medium versus Low HSAs (0.03).
[15] The 2.33 fewer clinician encounters figure is obtained by dividing the relevant point-estimate (0.07) by the difference in the average number of behavioral telehealth visits per beneficiary in Medium versus Low HSAs (0.03).
[16] In May 2022, the Centers for Disease Control and Prevention (2023) documented that 71.52 percent of the U.S. population was in a location with low levels of COVID-19 community transmission; 20.73 percent had medium levels and 7.76 percent had high levels of COVID-19 community transmission.

Low telehealth intensity, because telehealth capacity helped these areas reduce the need to postpone or cancel care.

It is also worth noting that our previous study (AIR, 2023) found greater telehealth intensity to be associated with slightly higher clinician encounters per beneficiary. The difference in the post-periods across the two studies could help explain this difference, as the post-period of the previous study (second half of 2021) was confounded by the COVID-19 pandemic. In that study, we observed areas with higher telehealth utilization being associated with more utilization as measured not only by higher rates of clinician encounters but also by more ACS hospitalizations and higher total cost of care per beneficiary. In the previous study, when the two forms of utilization (telehealth and outcome measures) went in the same direction of an increase, we were unable to fully rule out that both variables were just reflecting a general increase in utilization, as they were going in the same direction. In this study, the fact that telehealth utilization increases go hand in hand with healthcare utilization decreases, suggests that the analysis may be capturing a real relationship as opposed to merely reflecting the fact that some areas utilize more health care than others. Nevertheless, given the heterogeneity in the findings, caution should be exercised in strongly interpreting any given study result as a causal estimate of the impact of telehealth.

Telehealth and costs. In some of our analyses, we find evidence that higher telehealth intensity is associated with a decrease in the total cost of care, though the findings related to telehealth and costs are mixed. Urban HSAs with Medium non-behavioral telehealth intensity are not associated with the total cost of care per beneficiary outcome. However, rural HSAs with Medium non-behavioral telehealth intensity are associated with a decrease in the total cost of care per beneficiary, but this decrease is not robust to the exclusion of small HSAs. While there is no association with total cost, rural HSAs with Medium non-behavioral telehealth intensity are associated with a decrease in outpatient costs. In urban and rural HSAs, HSAs with Medium behavioral telehealth intensity are associated with a decrease in total cost of care per beneficiary. The magnitude of the association ranges from a decrease of 2.18 percent of the baseline rate for total cost of care in urban HSAs (132.45 fewer dollars per beneficiary) to a decline of 1.82 percent of the baseline rate in rural HSAs (108.70 fewer dollars per beneficiary).

In urban HSAs the decrease in the total cost of care is also accompanied by a decrease in physician costs. However other cost components—such as skilled nursing facility and hospice costs—also show a decrease, and these cost components can be less confidently tied to the intensity of behavioral telehealth treatment. The relationship between telehealth intensity and costs, when looking at HSAs with High telehealth intensity, is generally not reliable. Even though several cost outcomes pass the PTT in the High group, the direction and significance of estimated effects varies substantially across the various subgroups examined, likely reflecting the overall lack of comparability between the High and Low groups.

Overall, the evidence is consistent with other research suggesting that the telehealth expansion has not led to runaway healthcare spending or utilization (Ellimoottil, 2023) and with a previous study suggesting that telehealth is serving as a substitute for specific in-person encounters and that availability of telehealth has not led to additional primary care visits (Dixit et al., 2022).

4.2. Options for Future Work

In this section, we discuss two options for extending and strengthening the current analysis. First, HSAs with High and Low telehealth intensity in this study are not comparable at baseline, which renders any DID analysis based on comparing these groups unreliable. To address this, MedPAC could consider propensity score weighting, as the selection of more comparable groups at baseline, in terms of observable characteristics, may be more likely to yield parallel baseline trends in study outcomes between these groups. The analysis in Section 3.1.2 suggests that some of the observable characteristics would be important to match on, including age, race, the percentage of dual eligibles, ADI, and population size. For example, for urban HSAs grouped by non-behavioral telehealth intensity, HSAs with High intensity have an average ADI of 38.26, and HSAs with Low intensity have an average ADI of 66.87.[17] As another example, for urban HSAs grouped by non-behavioral telehealth intensity, HSAs with High intensity have an average population size (in 10,000 people) of 140.14, and HSAs with Low intensity have an average population size (in 10,000 people) of 11.67. However, it is also useful to note that in the previous iteration of this study, propensity score weighting improved parallel trends only marginally.

Second, it is difficult to cleanly attribute changes observed in outcomes in the current analysis to either the non-behavioral telehealth or behavioral telehealth treatments, as there is a fairly strong overlap in HSAs classified as High, Medium, or Low telehealth intensity for each treatment type, especially in urban HSAs (see Section 4.1 for a detailed discussion). One way to get around this could be to control for non-behavioral telehealth intensity in an HSA in a regression estimating the effects of behavioral telehealth intensity, and vice-versa. Another option would be to look at outcomes that are more specific to each type of treatment—for example, looking at ED visits or hospitalizations with behavioral health-related diagnosis codes for behavioral telehealth regressions and looking at non-behavioral health-related ED visits or hospitalizations for non-behavioral telehealth regressions. If telehealth is not permanently expanded for non-behavioral health services after 2024, then the majority of telehealth use will likely be for behavioral health services, which would provide an opportunity to specifically assess the impacts of behavioral telehealth.

[17] The ADI data in this study is an index that varies from 1–100, with 1 being the lowest ADI (most affluent) and 100 being the highest ADI.

References

Ahmad, F. B., Cisewski, J. A., Xu, J., & Anderson, R. N. (2023). COVID-19 mortality update—United States, 2022. *Morbidity and Mortality Weekly Report, 72*(18), 493–496. http://dx.doi.org/10.15585/mmwr.mm7218a4

American Institutes for Research. (2023). *Using population-based outcome measures to assess the impact of telehealth expansion on Medicare beneficiaries' access to care and quality of care.* Medicare Payment Advisory Commission (MedPAC). https://www.medpac.gov/wp-content/uploads/2023/06/Jun23_Telehealth_MedPAC_CONTRACTOR_SEC.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

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. (2023, March 15). *CDC museum COVID-19 timeline.* https://www.cdc.gov/museum/timeline/covid19.html

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

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

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

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

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

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), e2031640. 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/

Ellimoottil, C. (2023, November 14). *Ensuring Medicare beneficiary access: A path to telehealth permanency* [Testimony before the U.S. Senate Finance Committee, Subcommittee on Health Care]. https://www.finance.senate.gov/imo/media/doc/1114_ellimoottil_testimony.pdf

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.* Medicare Payment Advisory Commission (MedPAC). https://www.medpac.gov/wp-content/uploads/import_data/scrape_files/docs/default-source/contractor-reports/august2019_riskadjusted_ah_av_measures_contractor_sec.pdf

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

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://doi.org/10.1111/jrh.12128

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

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. (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. (2022a, 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. (2022b, July). Section 2: Medicare beneficiary demographics. In *Health care spending and the Medicare Program: A data book* (pp. 15–22). https://www.medpac.gov/wp-content/uploads/2022/07/July2022_MedPAC_DataBook_Sec2_SEC.pdf

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

Medicare Payment Advisory Commission. (2023b, June). *Report to the Congress: Medicare and the health care delivery system.* https://www.medpac.gov/document/june-2023-report-to-the-congress-medicare-and-the-health-care-delivery-system/

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

Office of the 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

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

Truelove, S., Smith, C. P., Qin, M., Mullany, L. C., Borchering, R. K., Lessler, J., Shea, K., Howerton, E., Contamin, L., Levander, J., . . . Viboud, C. (2022). Projected resurgence of COVID-19 in the United States in July—December 2021 resulting from the increased transmissibility of the Delta variant and faltering vaccination. *eLife*, 11, e73584. https://doi.org/10.7554/elife.73584

U.S. Department of Health & Human Services. (2023, August). *Telehealth policy changes after the COVID-19 public health emergency.* https://telehealth.hhs.gov/providers/telehealth-policy/policy-changes-after-the-covid-19-public-health-emergency

U.S. 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

Wilcock, A. D., Huskamp, H. A., Busch, A. B., Normand, S. L. T., Uscher-Pines, L., Raja, P. V., ... & Mehrotra, A. (2023, October). Use of telemedicine and quality of care among Medicare enrollees with serious mental illness. *JAMA Health Forum, 4*(10), e233648–e233648. https://doi.org/10.1001/jamahealthforum.2023.3648

Appendix A. Detailed Data Sources

Exhibit A-1.7 Telehealth Codes

Category	HCPCS Codes	Notes
Medicare-approved telehealth services in the Physician Fee Schedule (PFS)	Codes available from https://www.cms.gov/Medicare/Medicare-General-Information/Telehealth/Telehealth-Codes that were updated 6/17/22; and downloaded on 8/9/2022) All HCPCS codes and Current Procedural Terminology codes for telehealth **behavioral** services can be found here: https://telehealth.hhs.gov/providers/best-practice-guides/telehealth-for-behavioral-health/billing-for-telebehavioral-health	Claims must also meet one or more of the following requirements: HCPCS modifier code = GQ or 95 or GT or G0 or FQ or 93. 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] FQ or 93 modifiers are included for audio-only services starting in 2022. 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 service codes specific to Innovation Center models	NGACO: G9481–G9489; G0438–G0439[c] BPCI Advanced: G9978–G9986[d] ACO REACH: G9868–G9870[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 in the PFS	Communication technology–based services: G2012, G2010, G2250, or G2251 E-visits: 99421–99423 or 98970-98972 Telephone assessment and management services by qualified nonphysicians: 98966–98968[f]	G2250 & G2251 are for clinicians who can't bill evaluation and management 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. During PFS rulemaking for CY 2021 CMS decided to use 98970, 98971, and 98972 in place of G2061, G2062, and G2063 since their descriptors were similar.[g]

Category	HCPCS Codes	Notes
FQHC and RHC telehealth and remote services	Remote services: G0071 Telehealth services: From July 1, 2020: HCPCS code of G2025 with 95 modifier (not required). Based on data exploration, G0466-G0468 with modifier of 95, FQ, or 93 also counted as telehealth. Behavioral Health related telehealth services for FQHCs (from 2022): HCPCS codes G0470, G0469, with modifiers 95, FQ, or 93; or service line 090X/091X revenue centers with modifiers 95, FQ, 93 Behavioral Health related telehealth services for RHCs (from 2022): Service line 090X/091X revenue centers with modifiers 95, FQ, 93	Prior to 2019, FQHCs and RHCs could not provide telehealth services. Starting January 1, 2019, HCPCS code G0071 was used for technology-based or remote evaluation services furnished by a FQHC or RHC practitioner. Between Jan 27 and June 30th 2020, as more flexibilities were provided for FQHCs and RHCs to provide telehealth services, the 95 modifier in conjunction with any HCPCS codes for billable telehealth services can identify telehealth services for FQHCs and RHCs.[h] Starting July 1, 2020, FQHCs and RCHs had to submit HCPCS code G2025 for any claims for telehealth services with the option of also including the 95 modifier (not required anymore).[i] Effective January 1, 2022, FQHCs and RHCs may provide mental health visits via telecommunications. Modifiers 95 or FQ or 93 required. Modifier CG is also required for correct RHC behavioral telehealth billing, but does not directly relate to telehealth so is omitted from our measure specification.[j] Paired claim lines (same beneficiary and claim ID and revenue center and date) were deduplicated.

Note. ACO REACH = ACO Realizing Equity, Access, and Community Health Model; BPCI=Bundled Payments for Care Improvement; CMS = Centers for Medicare & Medicaid Services; HCPCS = Healthcare Common Procedure Coding System; NGACO = Next Gen Accountable Care Organization; POS = Place of Service. "Nonphysicians" means anyone who can't bill evaluation and management services (i.e., in this usage, "nonphysicians" does NOT include PAs or NPs). HCPCS = Healthcare Common Procedure Coding System.

[a] Centers for Medicare & Medicaid Services. (2017, December 4). *Elimination of the GT modifier for telehealth services*. https://www.hhs.gov/guidance/sites/default/files/hhs-guidance-documents/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] Centers for Medicare & Medicaid Services. (2021, May). *Next generation ACO model telehealth expansion waiver*. https://innovation.cms.gov/files/x/nextgenaco-telehealthwaiver.pdf

[d] innoviHealth. (n.d.). *Coronavirus & telehealth cheatsheet.* Findacode. https://www.findacode.com/medical-code-sets/covid19-card.pdf

[e] Centers for Medicare & Medicaid Services. (2022, February 2022). *ACO Realizing Equity, Access, and Community Health (REACH) model: Request for applications.* https://www.cms.gov/priorities/innovation/media/document/aco-reach-rfa

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

[g] Centers for Medicare & Medicaid Services. (2020, December 31). *2021 annual update to the therapy code list* (CMS Manual System, Pub 100-04 Medicare Claims Processing, Transmittal 10542). https://www.cms.gov/files/document/r10542cp.pdf

[h] National Association of Rural Health Clinics. (2020, April). *CMS releases guidance on telehealth billing for RHCs.* https://www.narhc.org/News/28316/CMS-Releases-Guidance-on-Telehealth-Billing-for-RHCs

[i] Centers for Medicare & Medicaid Services. (2023, May 12). *New & expanded flexibilities for Rural Health Clinics & Federally Qualified Health Centers.* https://www.cms.gov/files/document/se20016-new-expanded-flexibilities-rhcs-fqhcs-during-covid-19-phe.pdf

[j] Centers for Medicare & Medicaid Services. (2023, May 23). *Mental health visits via telecommunications for Rural Health Clinics & Federally Qualified Health Centers.* https://www.cms.gov/files/document/se22001-mental-health-visits-telecommunications-rural-health-clinics-federally-qualified-health.pdf

Exhibit A-2. Encounters with FQHCs and RHCs

Category	HCPCS Codes
Encounters with Federally Qualified Health Centers, Rural Health Clinics and Critical Access Hospitals method II payments	FQHC claims from 100% Outpatient SAF files (2018-2022): Extracted variables Facility Type and Type of Service (FAC_TYPE and TYPESRVC) are both set to "7". RHC claims from 100% Outpatient SAF files (2018-2022): Extracted variables Facility Type and Type of Service (FAC and TYPESRVC) are set to "7" and "1", respectively. CAH method II payment claims from 100% Outpatient SAF files (2018-2022): Last 4 digits of extracted variable provider number (PROVIDER) in the range for CAHs ("1300 – 1399"). In addition, extracted variables Facility Type (FAC_TYPE) set to "8", and Type of Service (TYPESRVC) set to "5". Lastly, extracted variable revenue center (REV_CNTR) set to values indicating method II ("0960" "0961" "0962" "0963" "0964" "0965" "0966" "0967" "0968" "0969" "0970" "0971" "0972" "0973" "0974" "0975" "0976" "0977" "0978" "0979" "0980" "0981" "0982" "0983" "0984" "0985" "0986" "0987" "0988" "0989")

Note. CAH = Critical Access Hospitals; FQHCs = Federally Qualified Health Centers; RCHs = Rural Health Clinics.

Exhibit A-3. Data Sources for HSA Medicare Population Characteristics

Item	Description
A	**Attribution to HSAs:** Items #B to #L are calculated for each HSA, year (2018, 2019, and 2022), 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 under 65, 65–74, 75–84, and 85+:** Beneficiaries in #C are assigned to one of four groups (under 65, 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 #E. Sex is determined using the column SEX.
G	**Shares of FFS Medicare beneficiaries with White/Black/Hispanic/Asian/other/unknown race:** Analogous to #E. 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	**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.
K	**Average HCC risk scores squared for FFS Medicare beneficiaries:** Analogous to #J. HCC risk scores will be squared before taking the average.

Item	Description
L	**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. The ADI is not updated each year between 2018 and 2022. Therefore, we use ADI datasets that overlap with our period of analysis. The 2021 (used for HSA ADI assignment for beneficiary data from second semester of 2022), and 2020 (used for HSA ADI assignment for beneficiary data from second semester of 2018 and second semester of 2019) ADI datasets were obtained from https://www.neighborhoodatlas.medicine.wisc.edu/.

Exhibit A-4. Data Sources for HSA Market Characteristics

Item	Description
A	**Attribution to HSAs:** Item #B is calculated for each county, year (2018, 2019, 2020, 2021, 2022) and semester (January to June, July to December). Items #C-#F are calculated for each county, year (2018, 2019, 2022), and semester. 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 2022 are accounted for based on https://www.census.gov/programs-surveys/geography/technical-documentation/county-changes.html.
B	**Urban, rural micropolitan, rural adjacent, and rural non-adjacent counties:** Beneficiaries 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 counties are defined as UICs 1 and 2. Rural micropolitan counties are defined as UICs 3, 5, and 8. Rural adjacent counties are defined as UICs 4, 6, and 7. Rural nonadjacent counties 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.
C	**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 2022 are obtained from https://www2.census.gov/programs-surveys/popest/datasets/2020-2022/counties/totals/co-est2022-alldata.csv, the variable [POPESTIMATE2022]. Values for each semester of a year are set to the value of that year.
D	**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 2022 are obtained from https://raw.githubusercontent.com/nytimes/covid-19-data/master/us-counties-2022.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 #C and multiplied by 10,000. New cases and deaths for both semesters of 2018 and 2019 are set to zero.

Item	Description
E	**Average IPPS hospital wage index [robustness check variable]:** 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 2022 are obtained from https://www.cms.gov/medicare/payment/prospective-payment-systems/acute-inpatient-pps/fy-2022-ipps-final-rule-home-page. Values for each semester of a year are set to the value of that year .
F	**Average PFS geographic practice cost indexes (GPCIs) [robustness check variable]:** 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 2022 are obtained from RVU18D, RVU19C, and RVU22D. 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. Outcome Trends

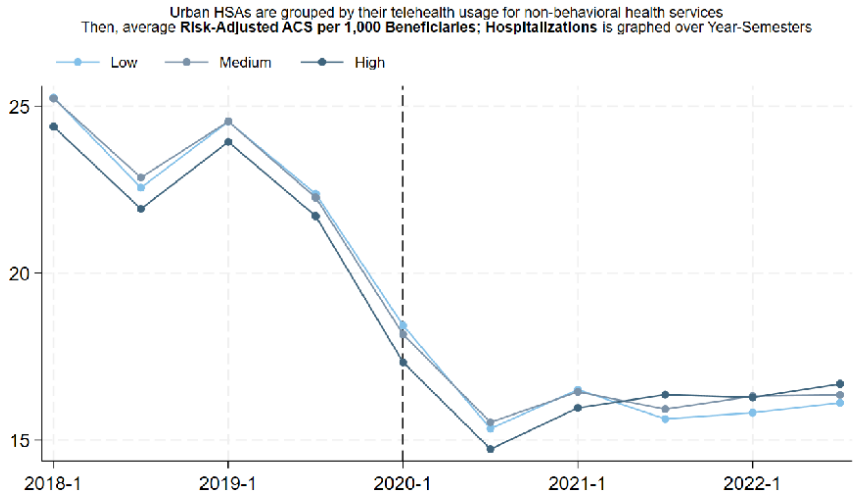
This appendix presents a time series of all outcomes between the first semester of 2018 and the second semester of 2022.

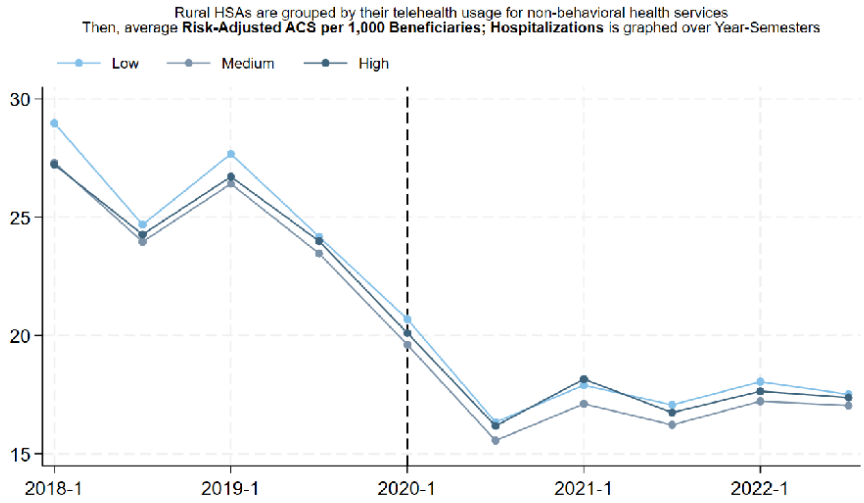
Outcome Trends by Year and Semester

Outcome Category: Quality

Outcome: Risk-Adjusted ACS per 1,000 Beneficiaries; Hospitalizations

Non-Behavioral Health

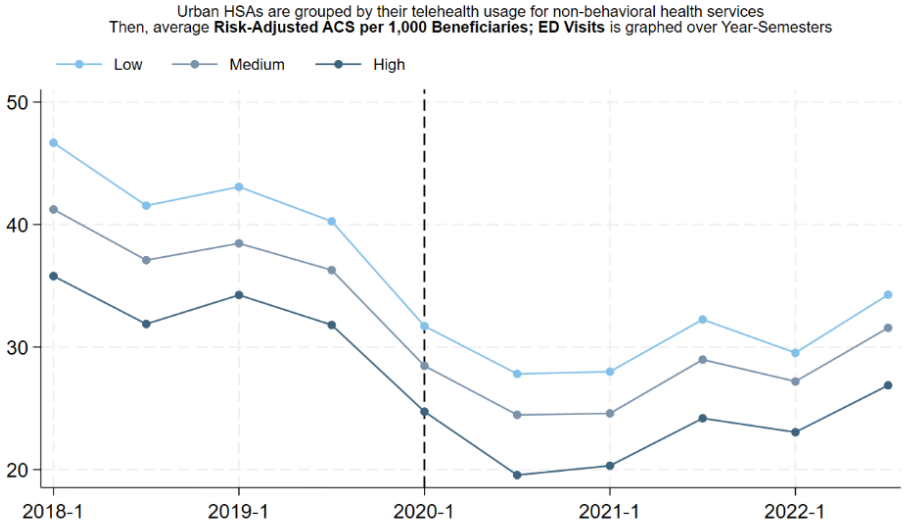


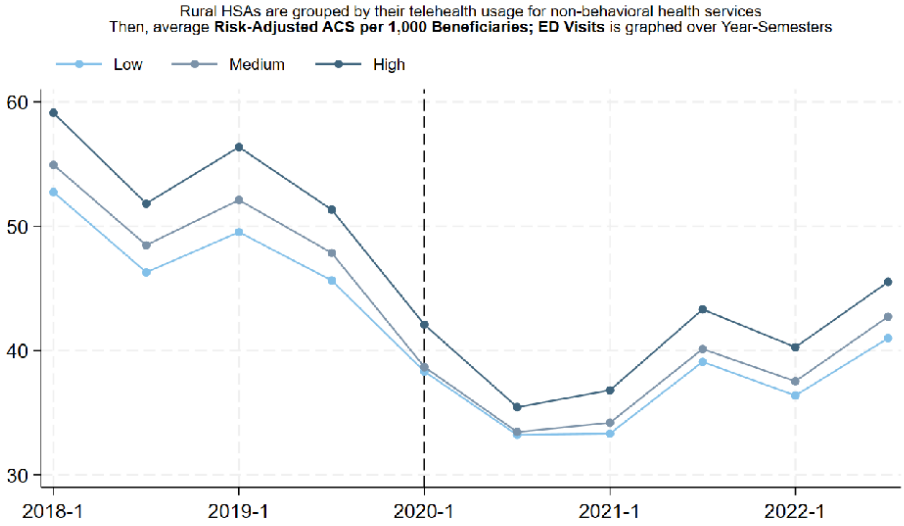


1

Outcome: Risk-Adjusted ACS per 1,000 Beneficiaries; ED Visits

Non-Behavioral Health



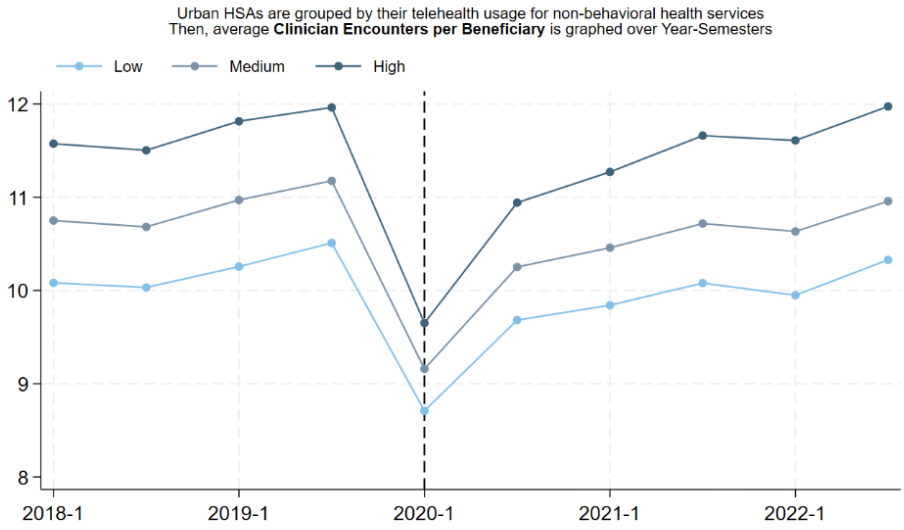


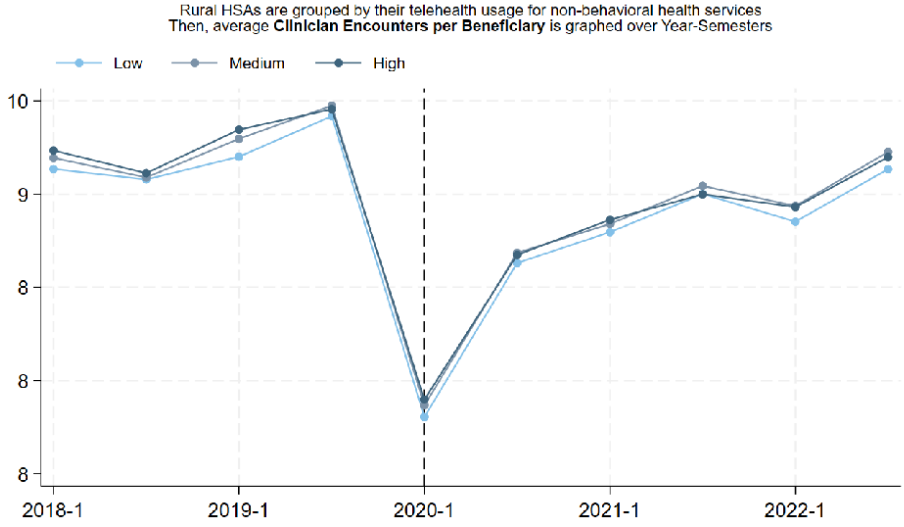
2

Outcome Category: Access

Outcome: Clinician Encounters per Beneficiary

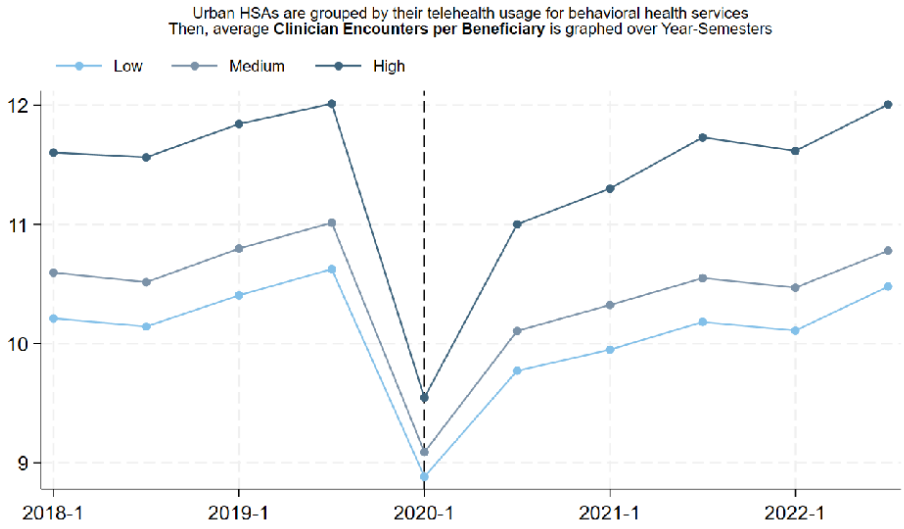
Non-Behavioral Health

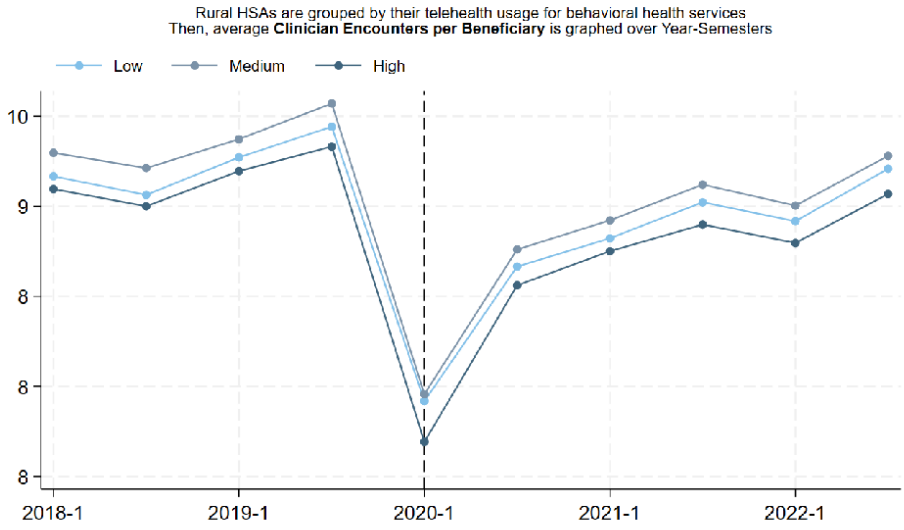




3

Behavioral Health

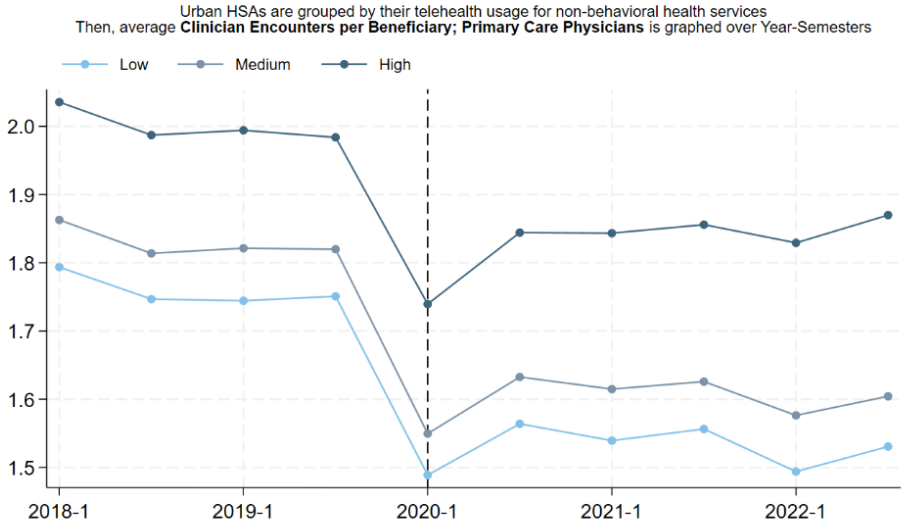


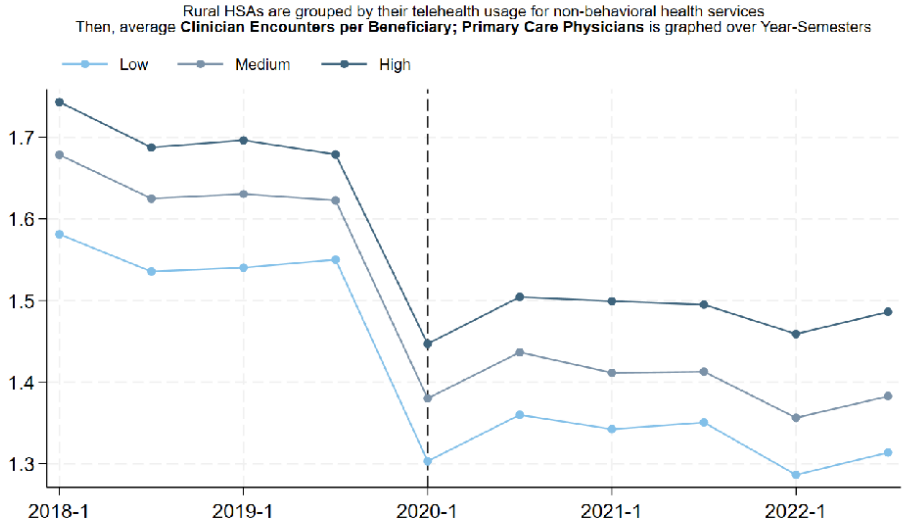


4

Outcome: Clinician Encounters per Beneficiary; Primary Care Physicians

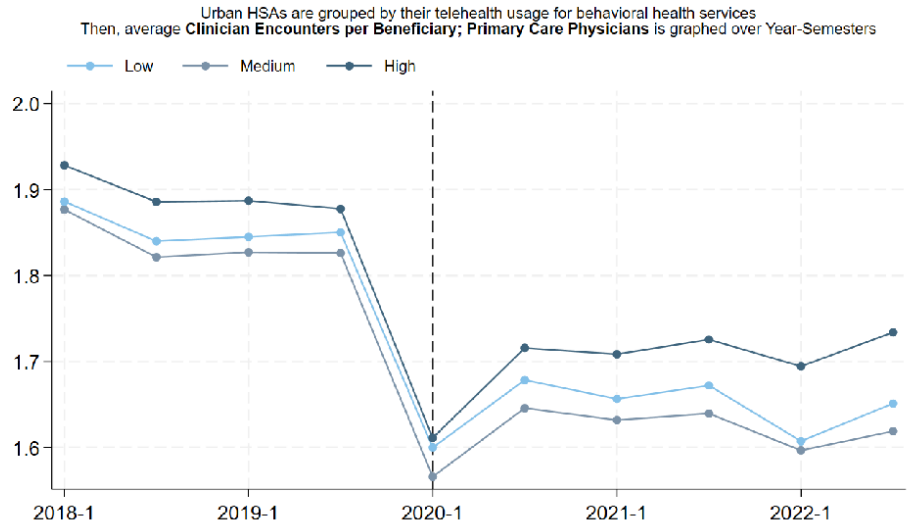
Non-Behavioral Health

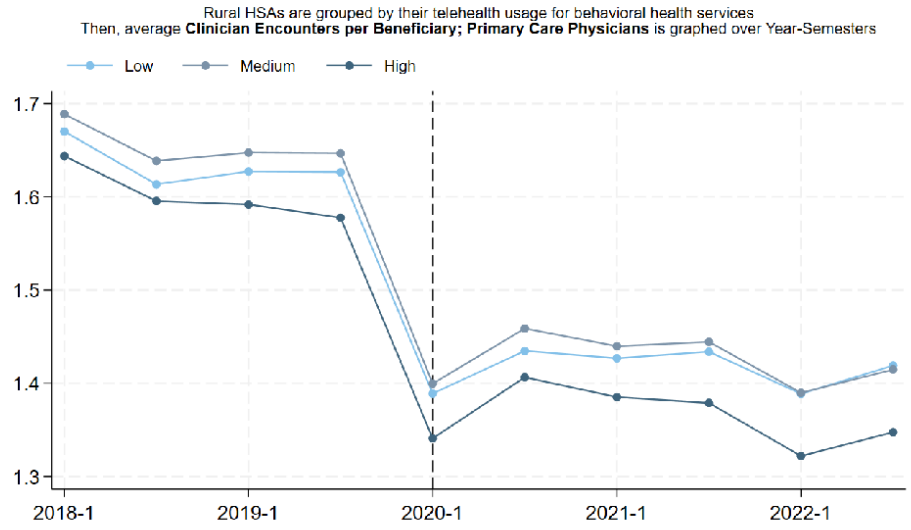




5

Behavioral Health

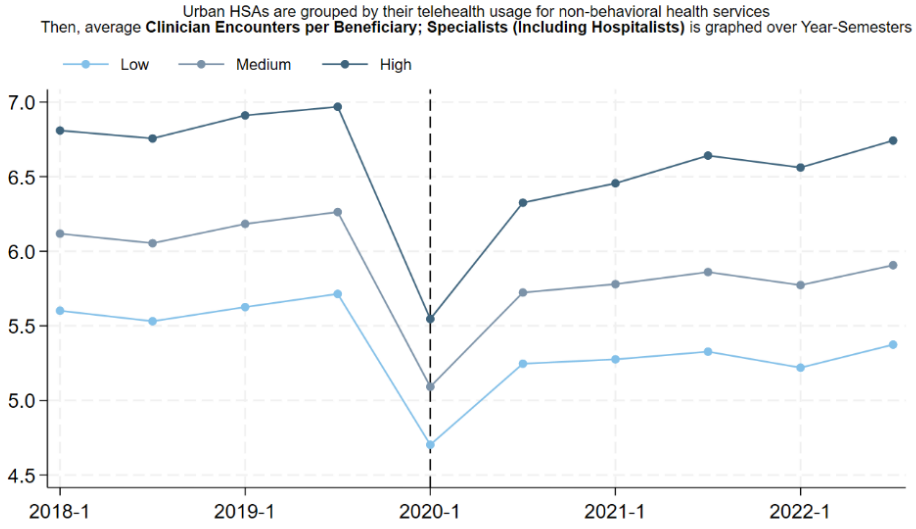


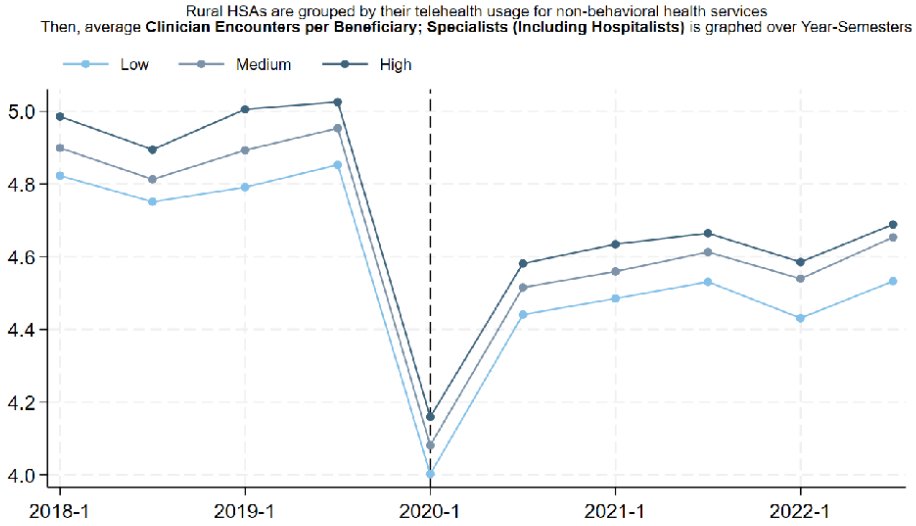


6

Outcome: Clinician Encounters per Beneficiary; Specialists (Including Hospitalists)

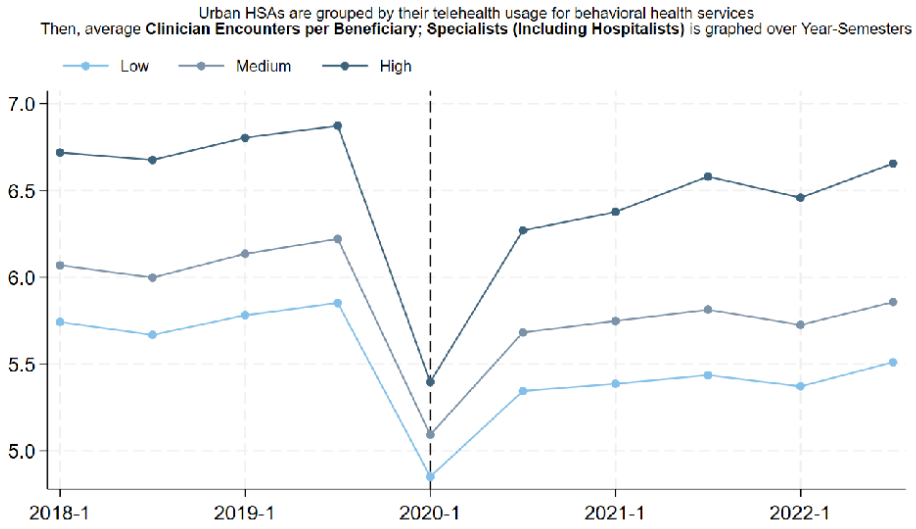
Non-Behavioral Health

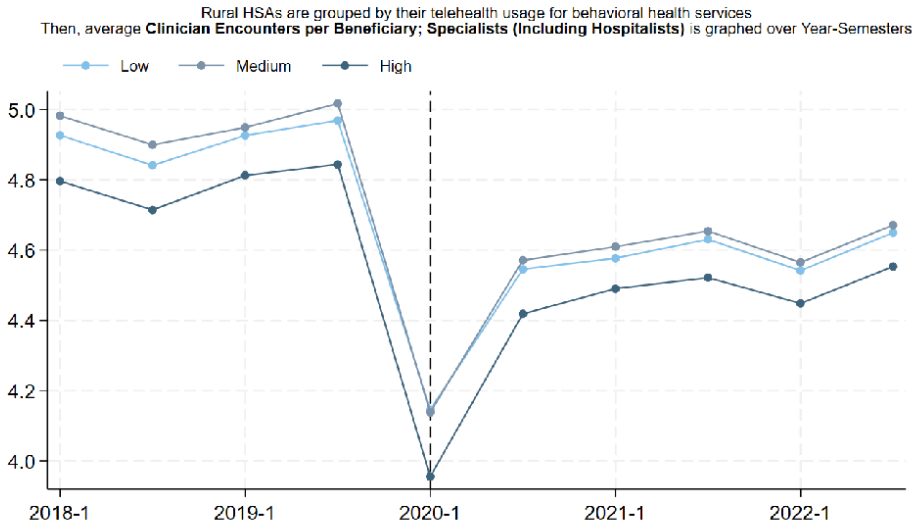




7

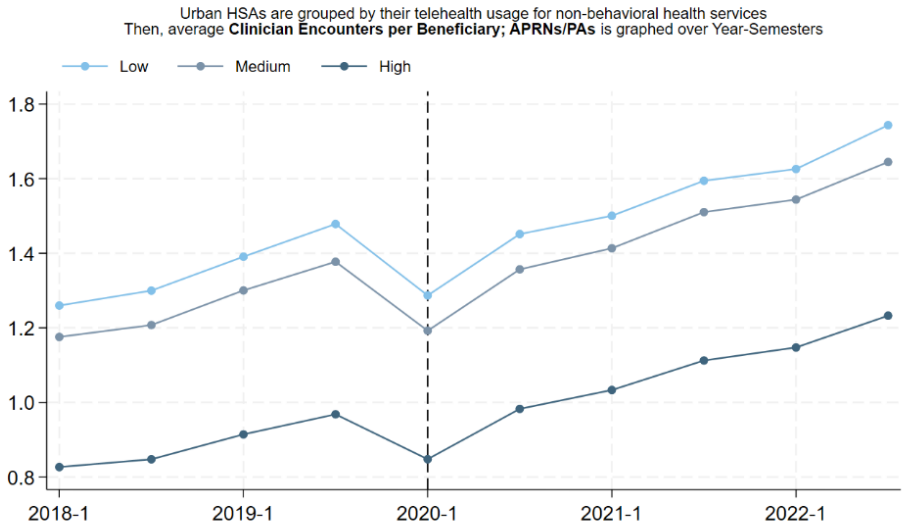
Behavioral Health

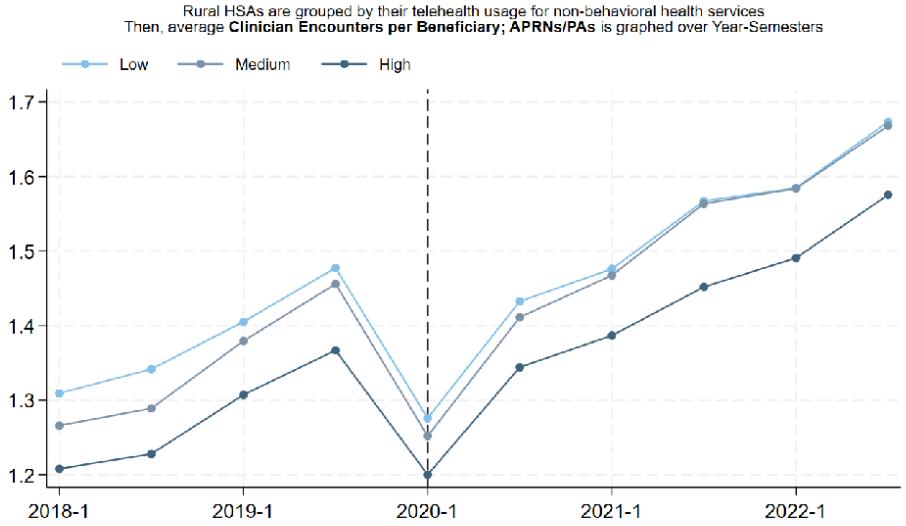




Outcome: Clinician Encounters per Beneficiary; APRNs/PAs

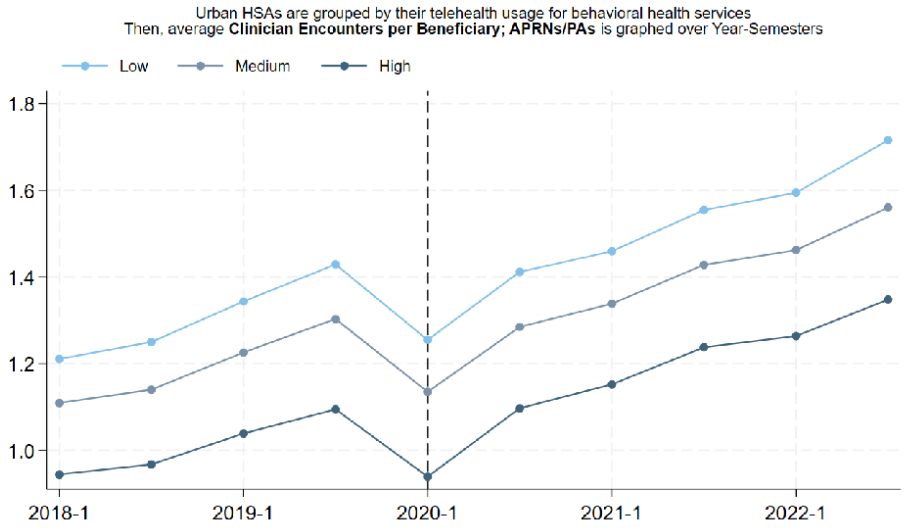
Non-Behavioral Health

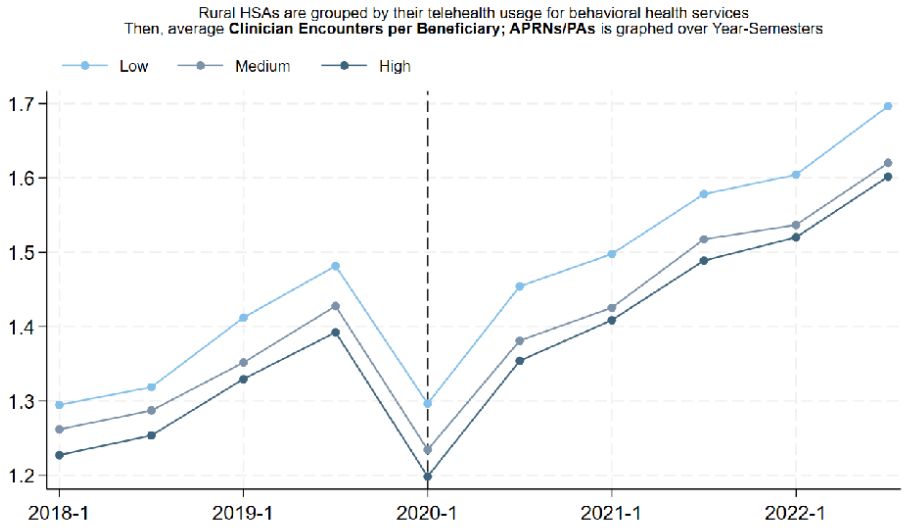




9

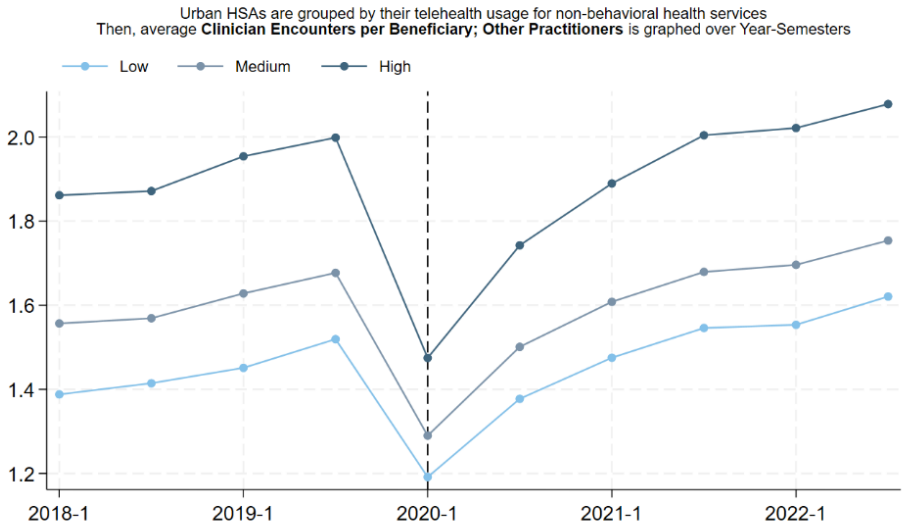
Behavioral Health

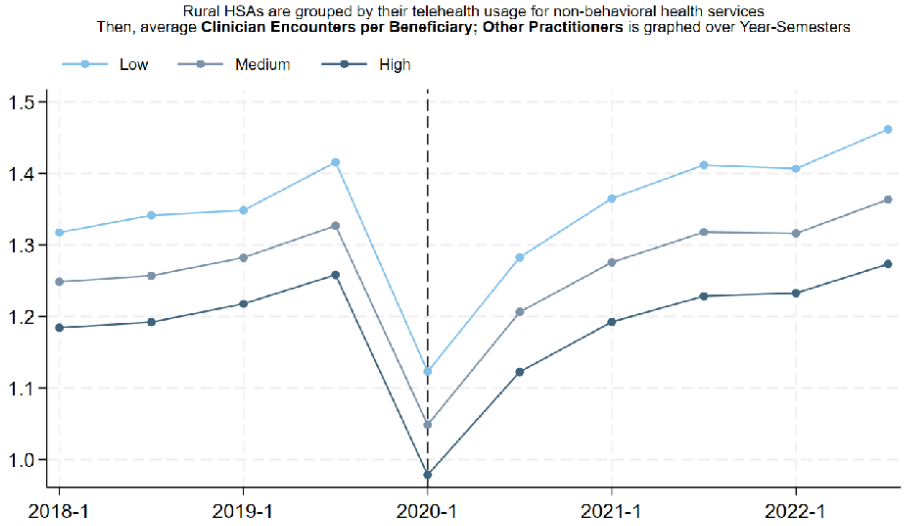




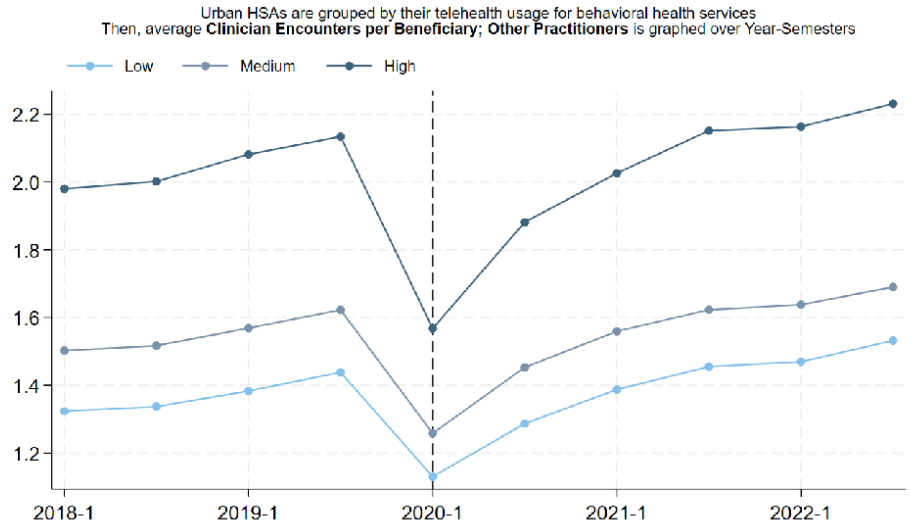
Outcome: Clinician Encounters per Beneficiary; Other Practitioners

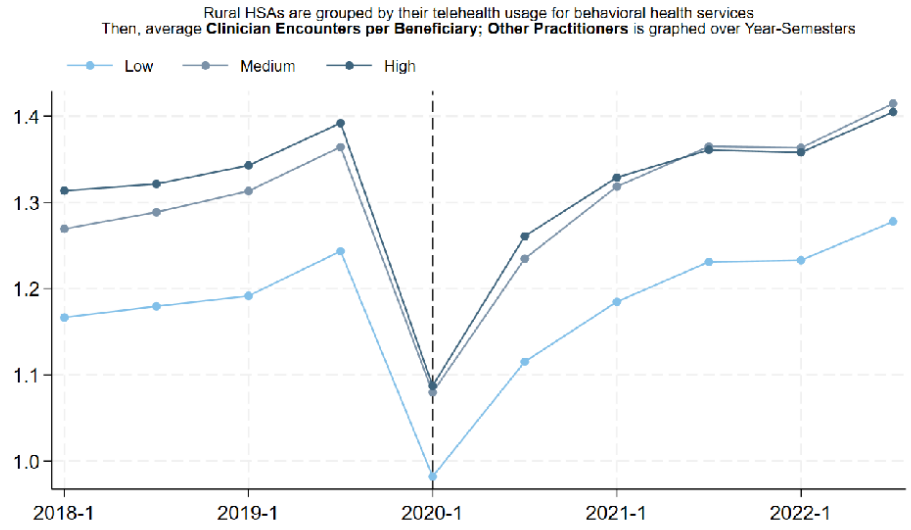
Non-Behavioral Health





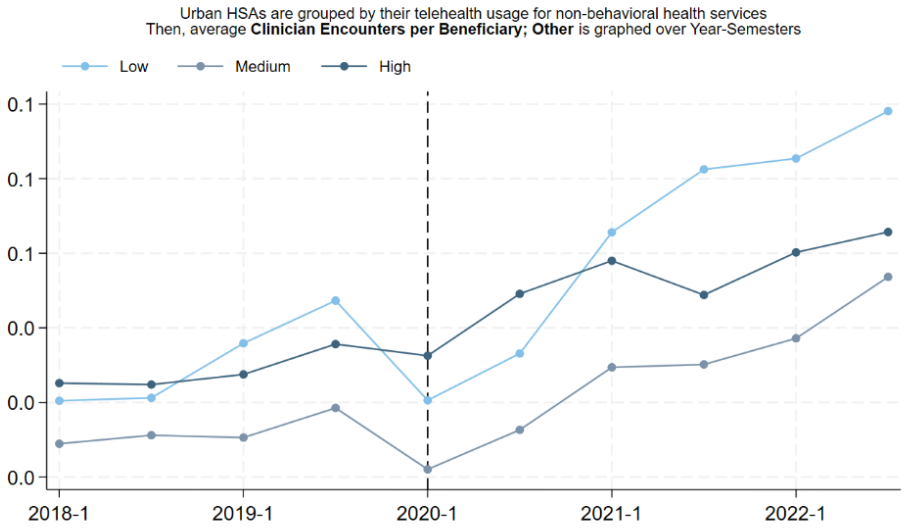
Behavioral Health

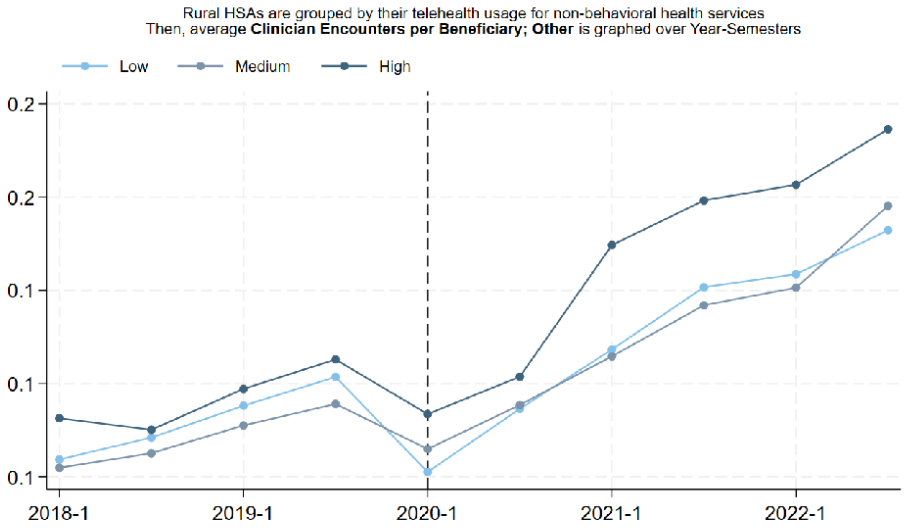




Outcome: Clinician Encounters per Beneficiary; Other

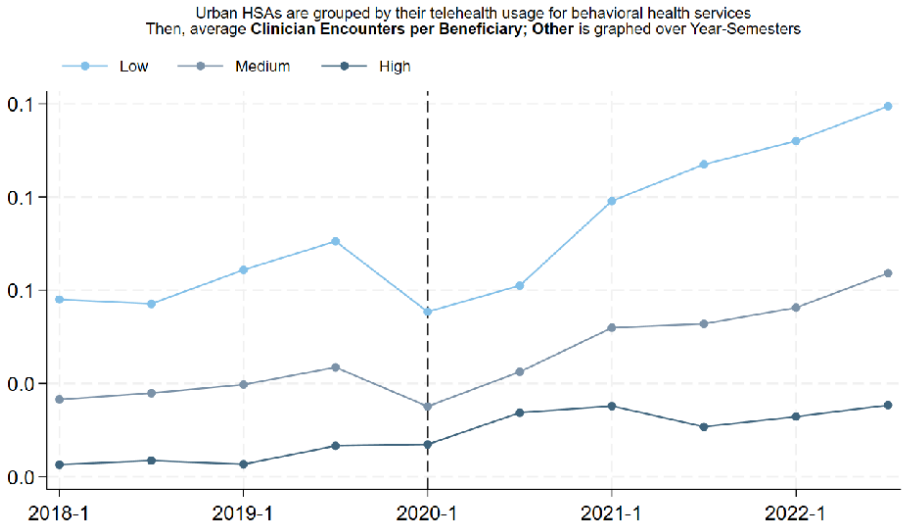
Non-Behavioral Health

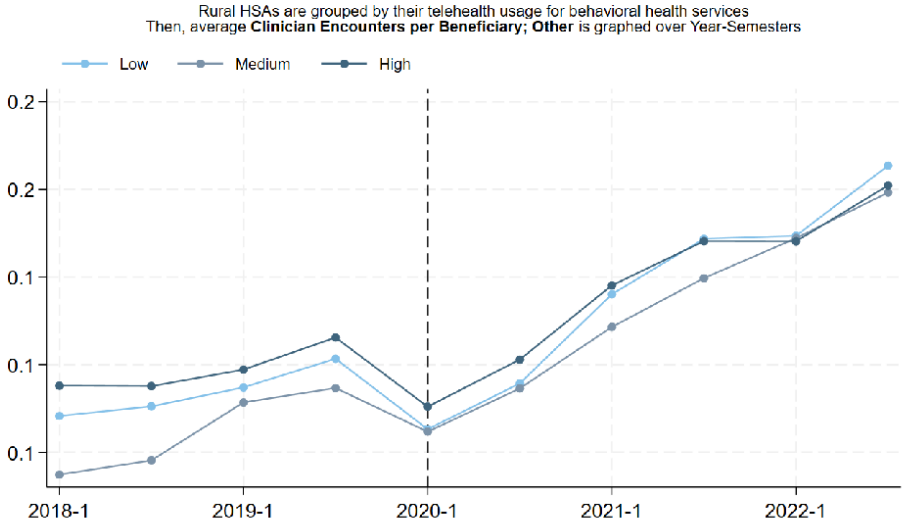




13

Behavioral Health



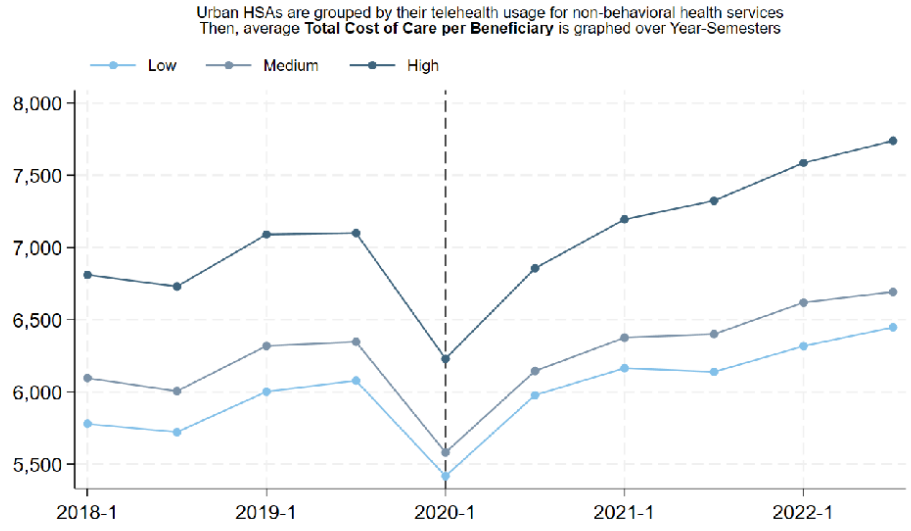


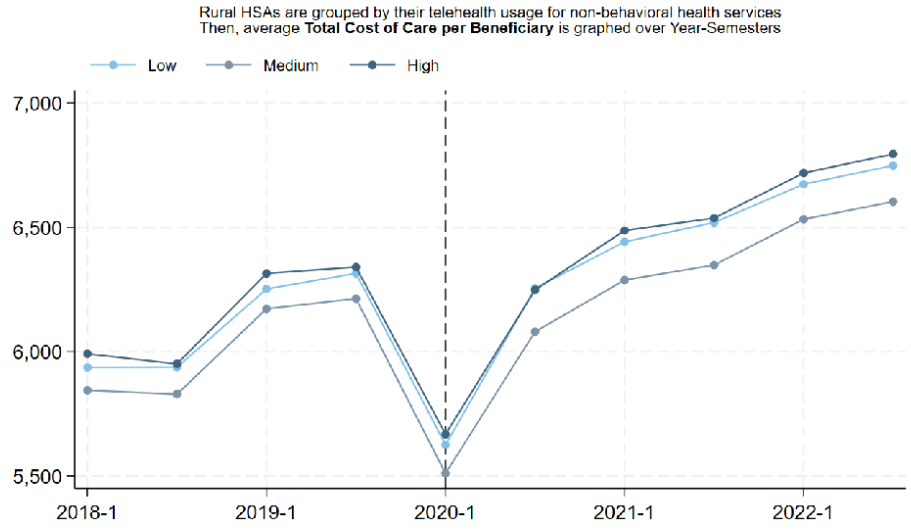
14

Outcome Category: TCoC/Beneficiary

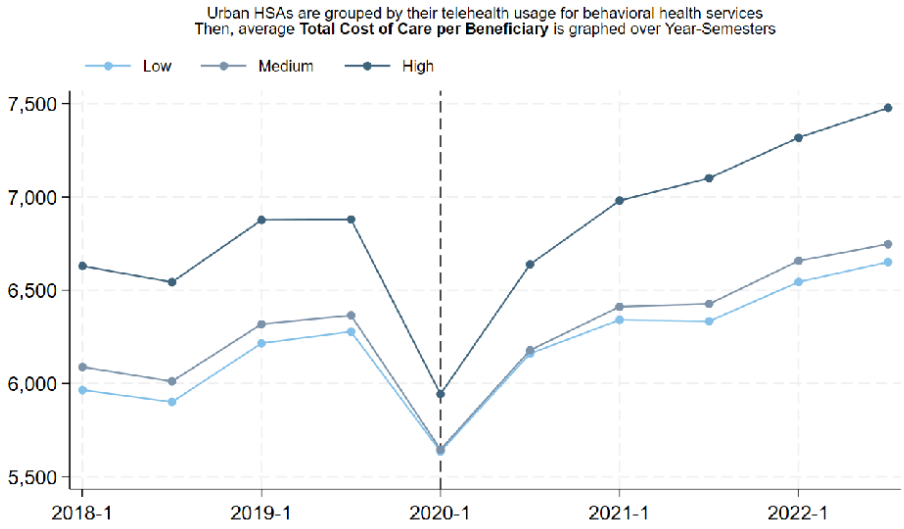
Outcome: Total Cost of Care per Beneficiary

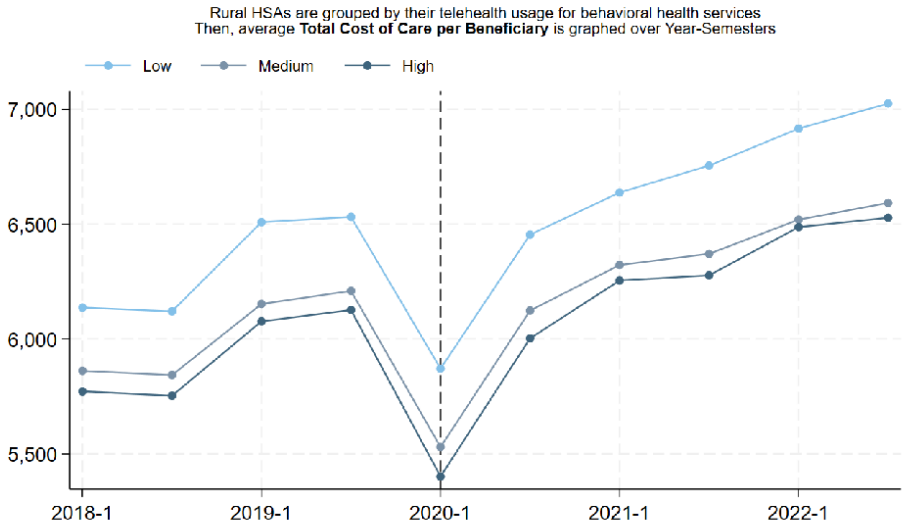
Non-Behavioral Health





Behavioral Health

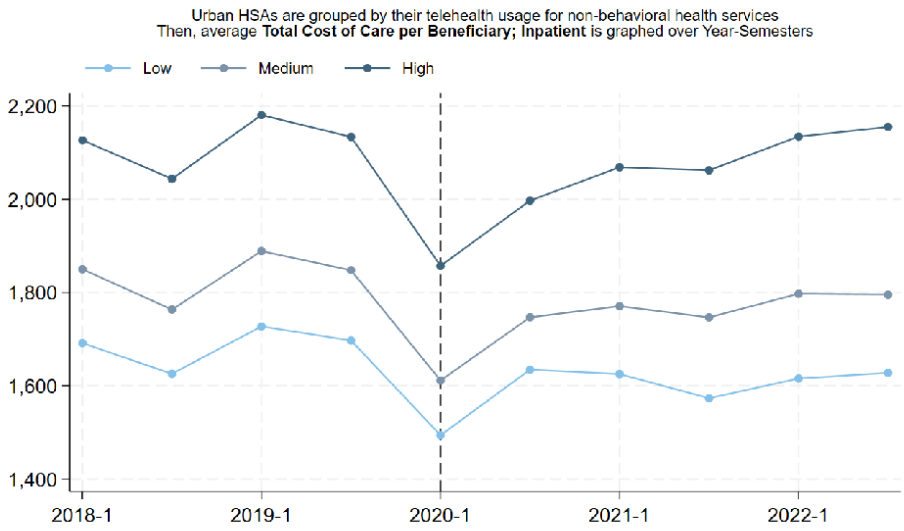


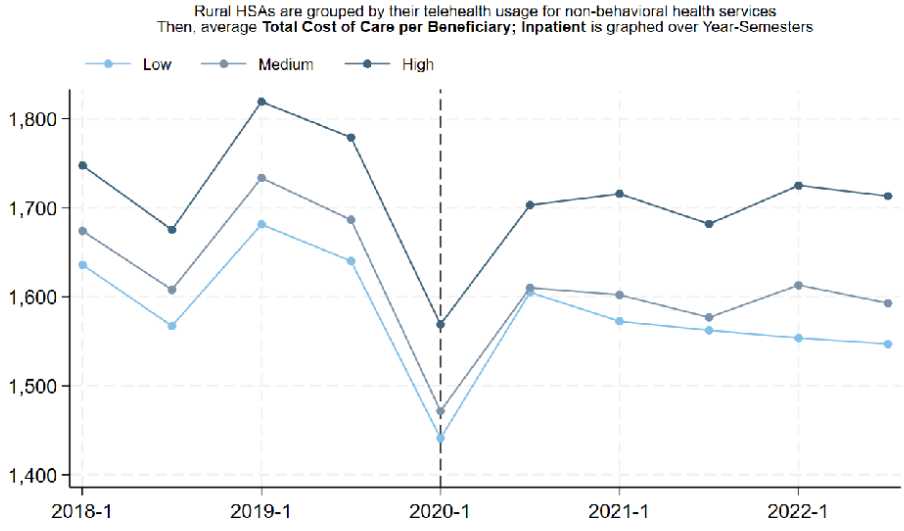


16

Outcome: Total Cost of Care per Beneficiary; Inpatient

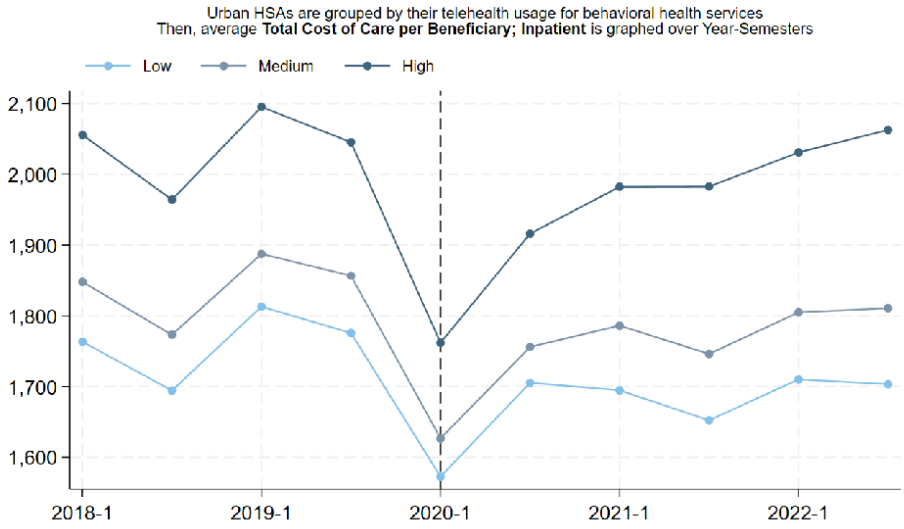
Non-Behavioral Health

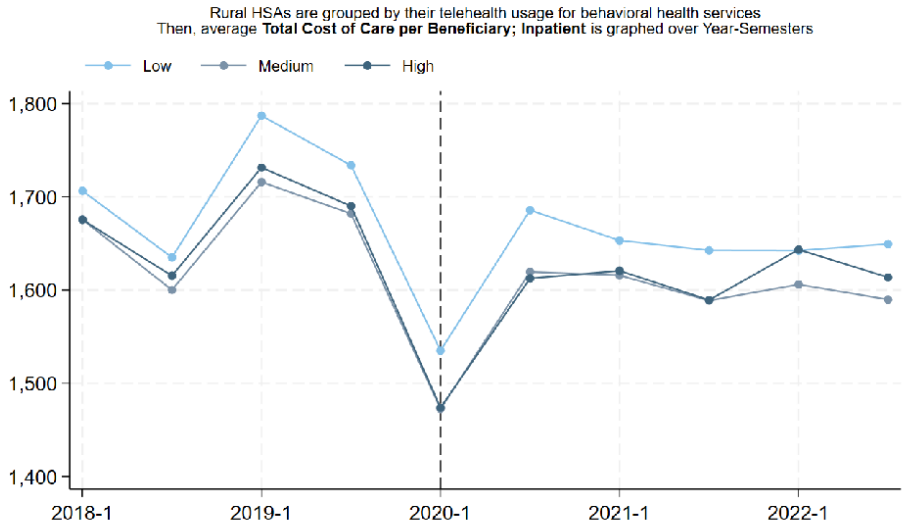




17

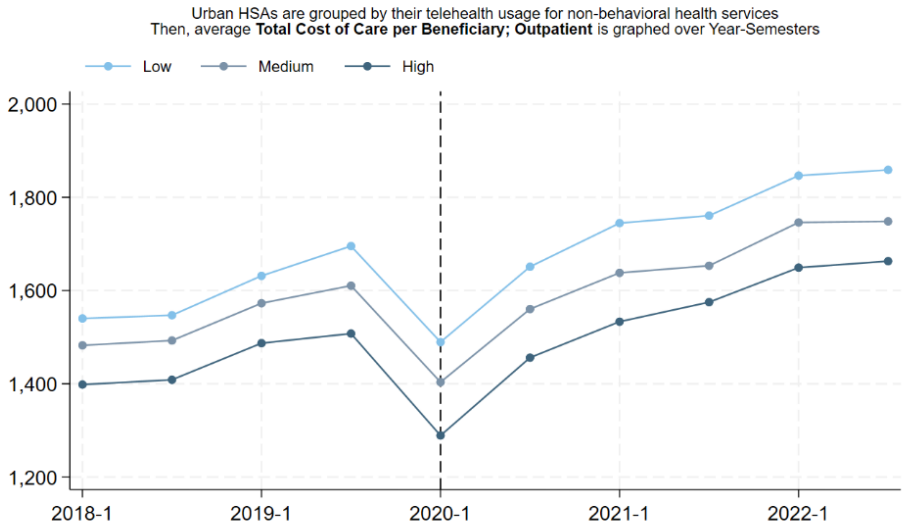
Behavioral Health

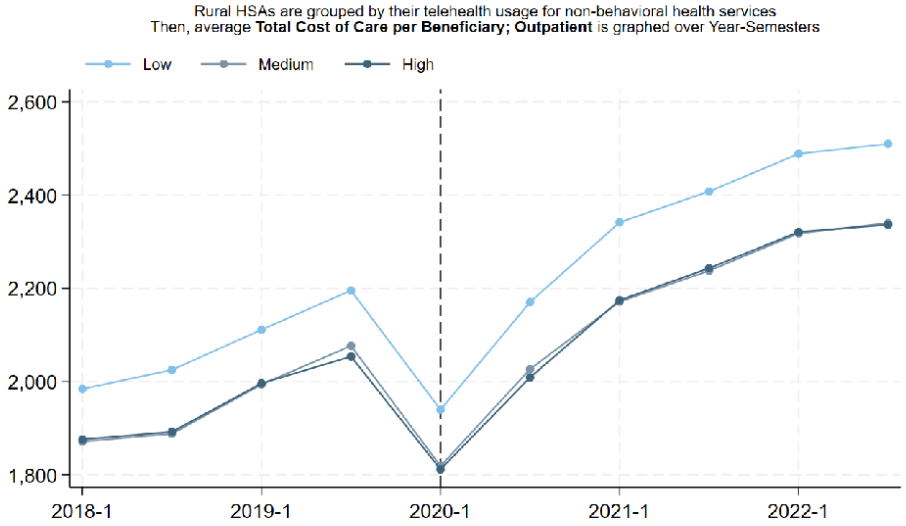




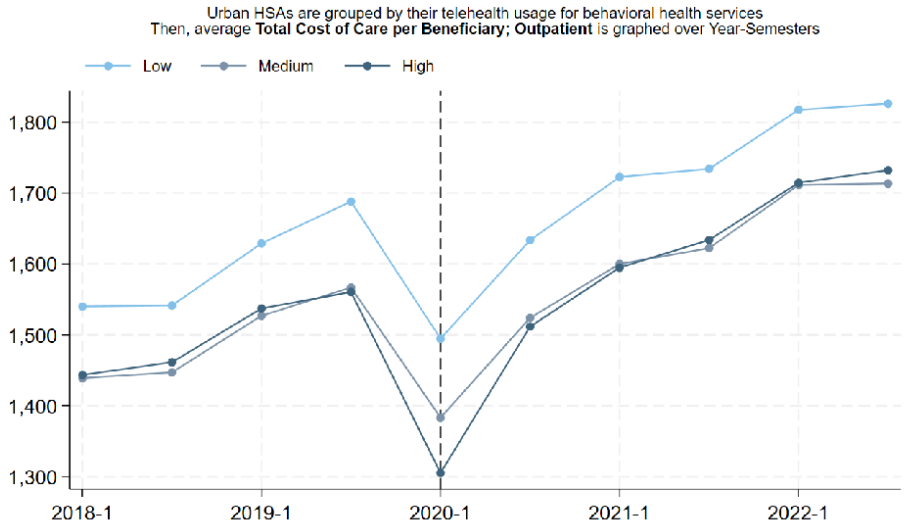
Outcome: Total Cost of Care per Beneficiary; Outpatient

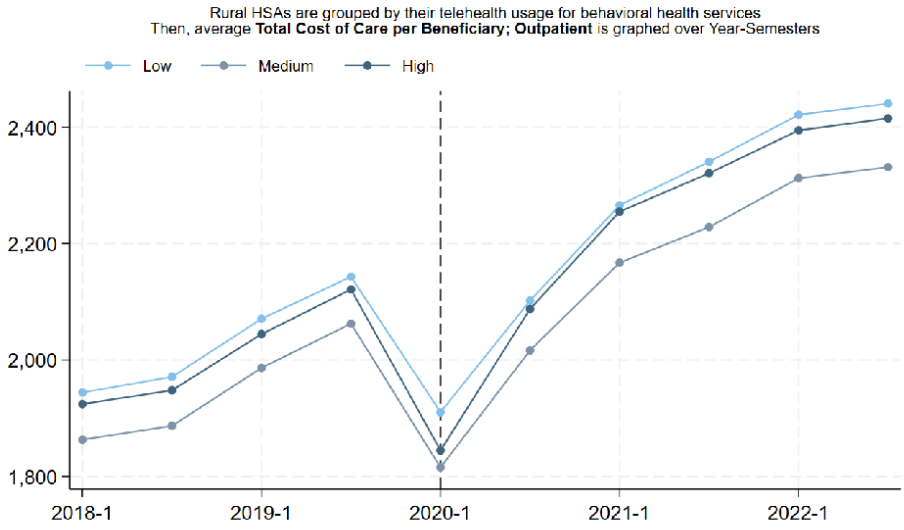
Non-Behavioral Health





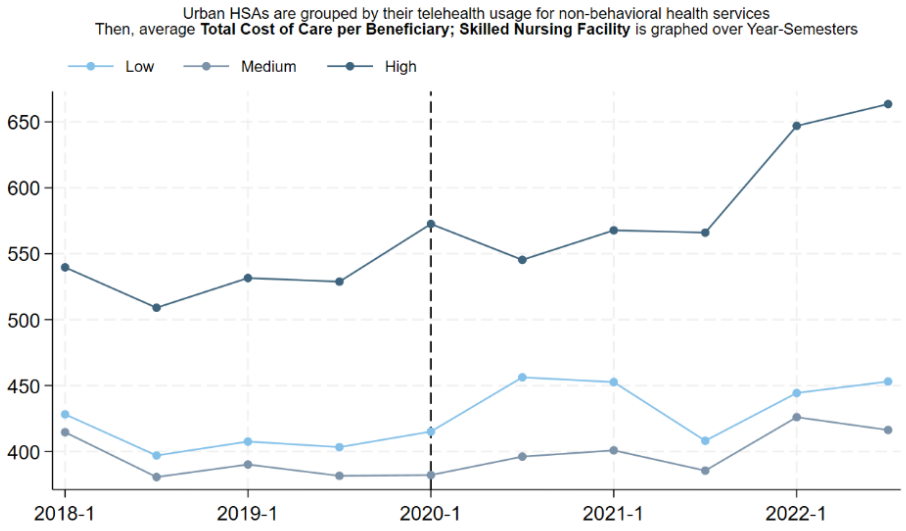
Behavioral Health

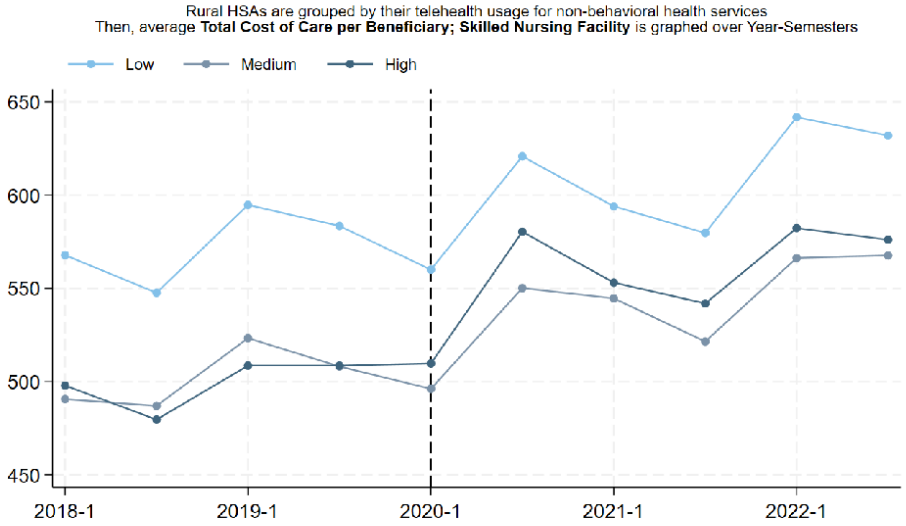




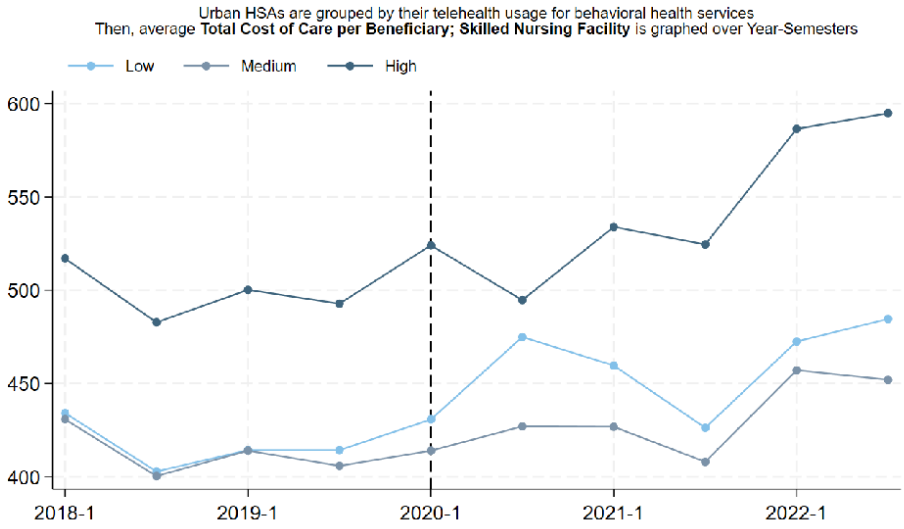
Outcome: Total Cost of Care per Beneficiary; Skilled Nursing Facility

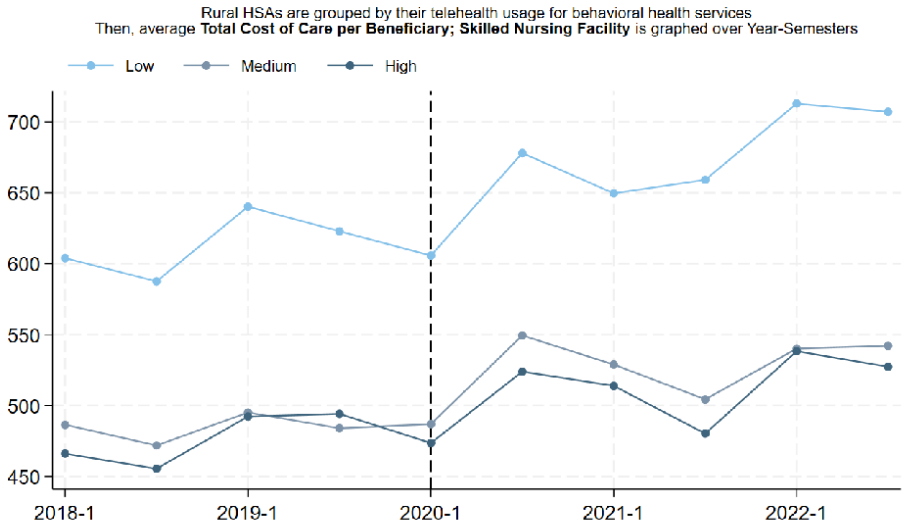
Non-Behavioral Health





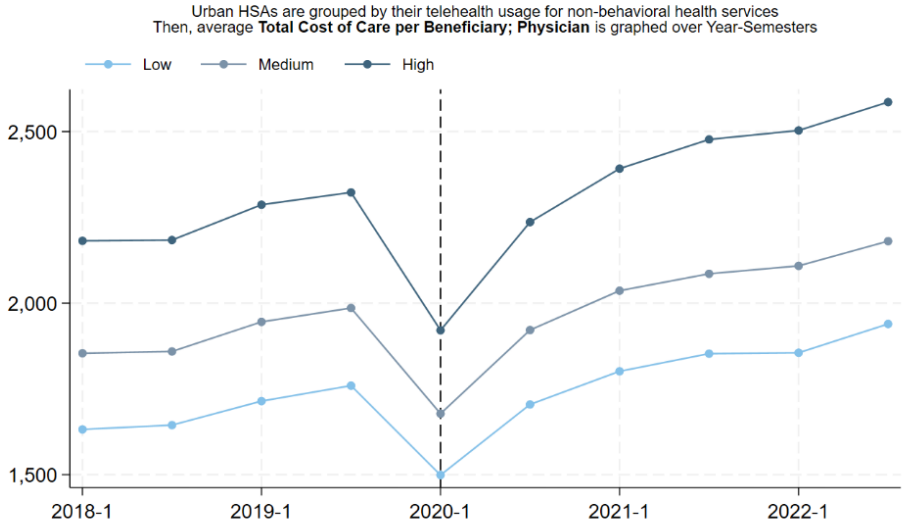
Behavioral Health

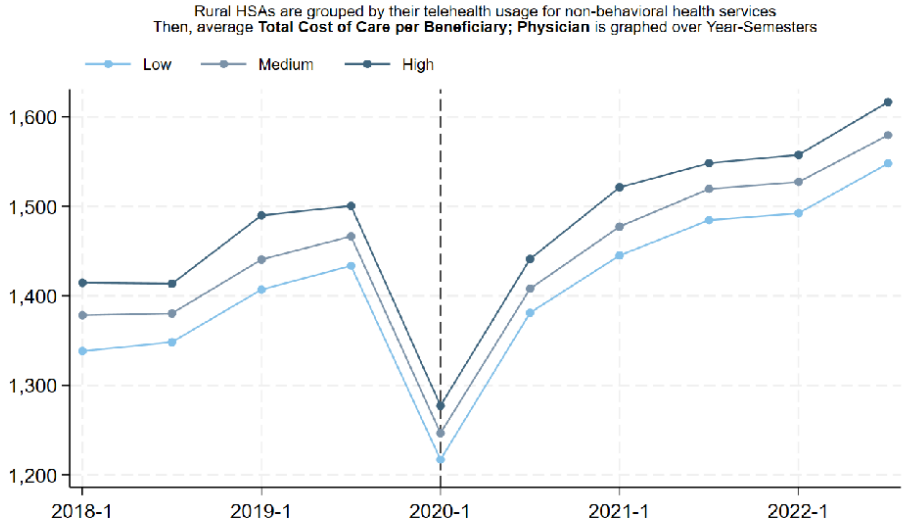




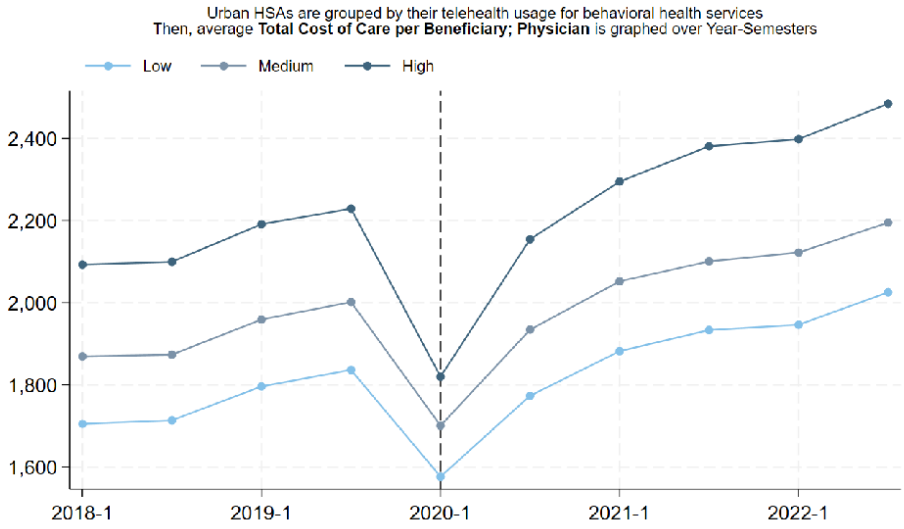
Outcome: Total Cost of Care per Beneficiary; Physician

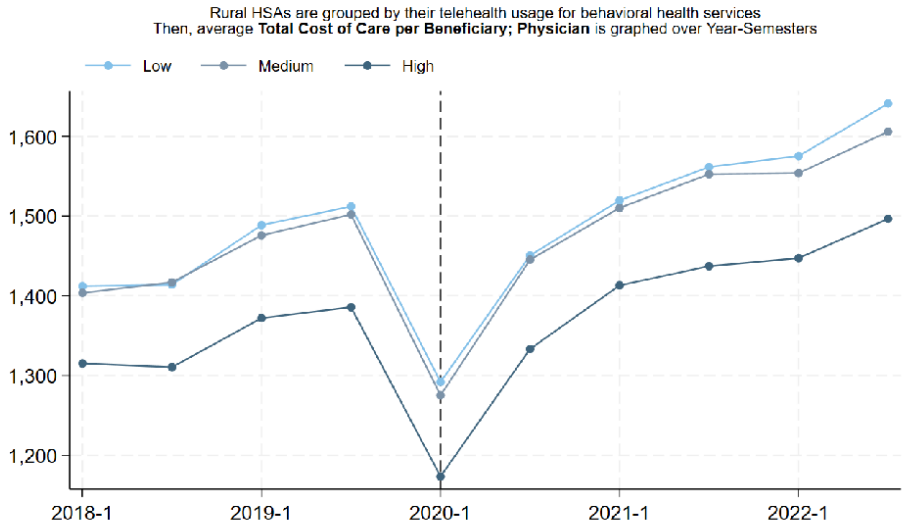
Non-Behavioral Health





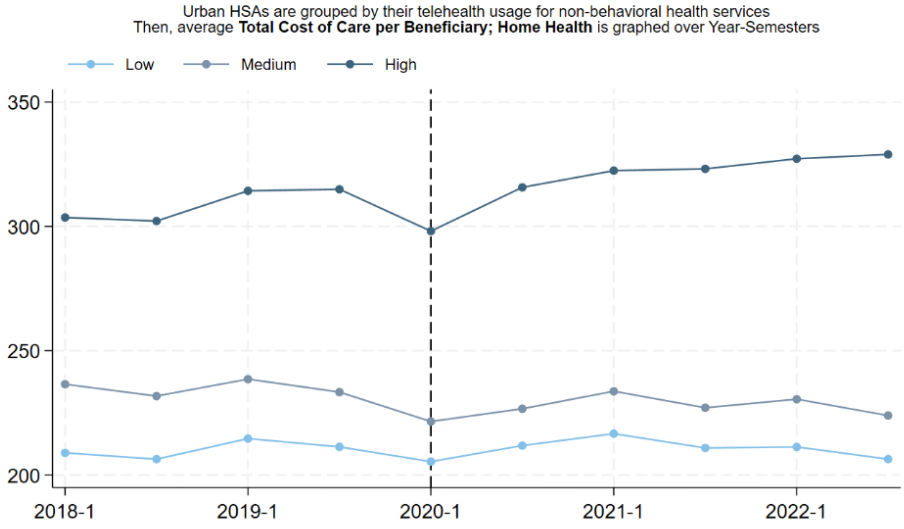
Behavioral Health

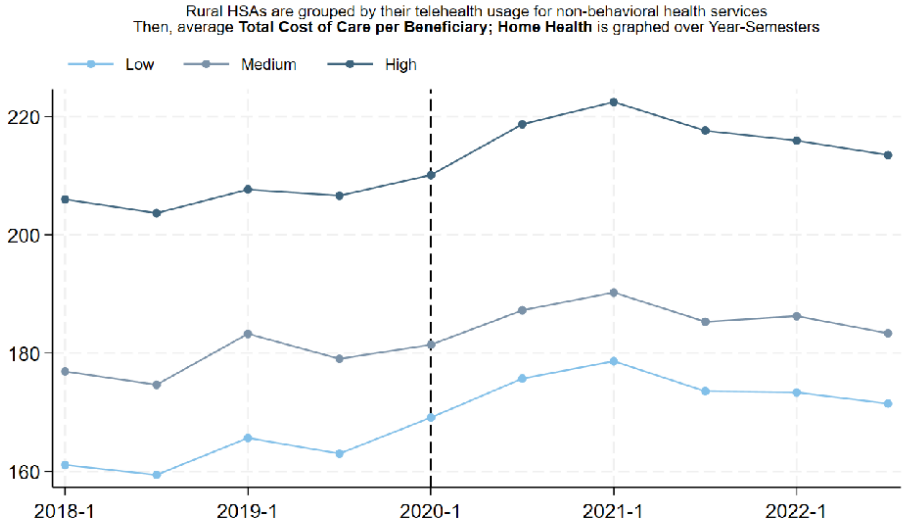




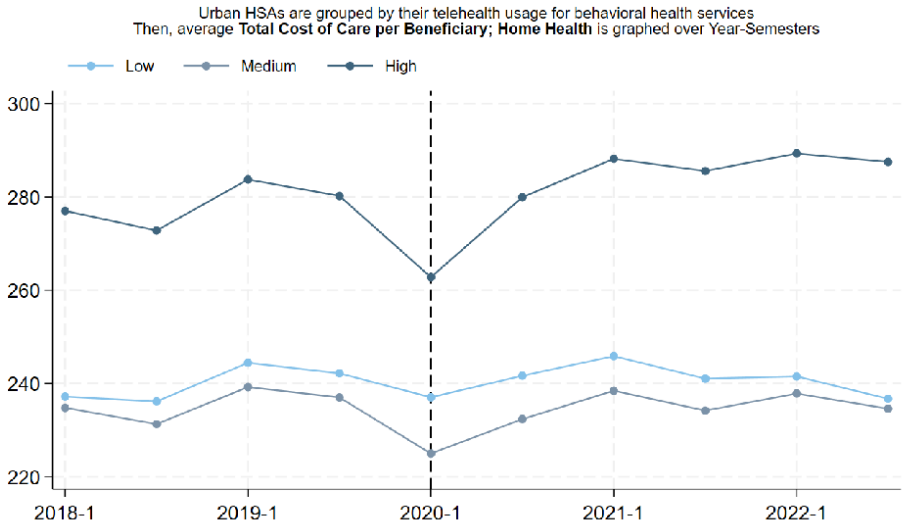
Outcome: Total Cost of Care per Beneficiary; Home Health

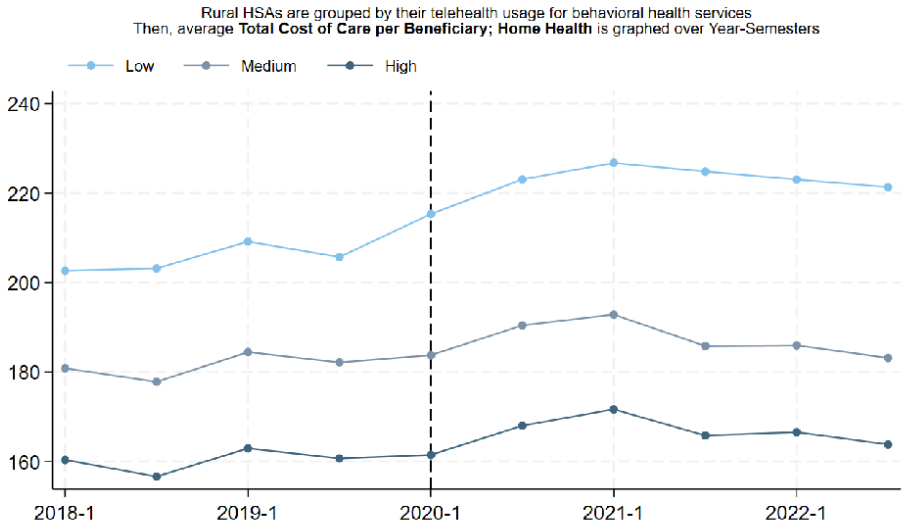
Non-Behavioral Health





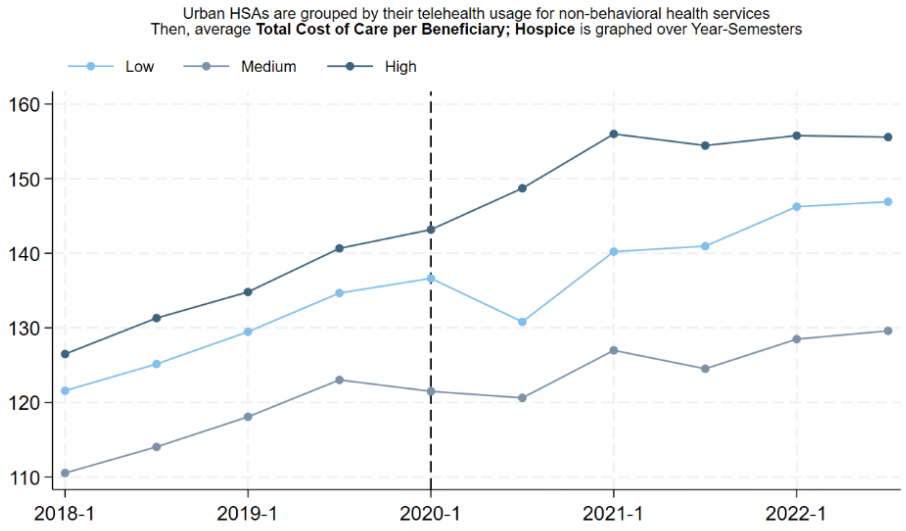
Behavioral Health

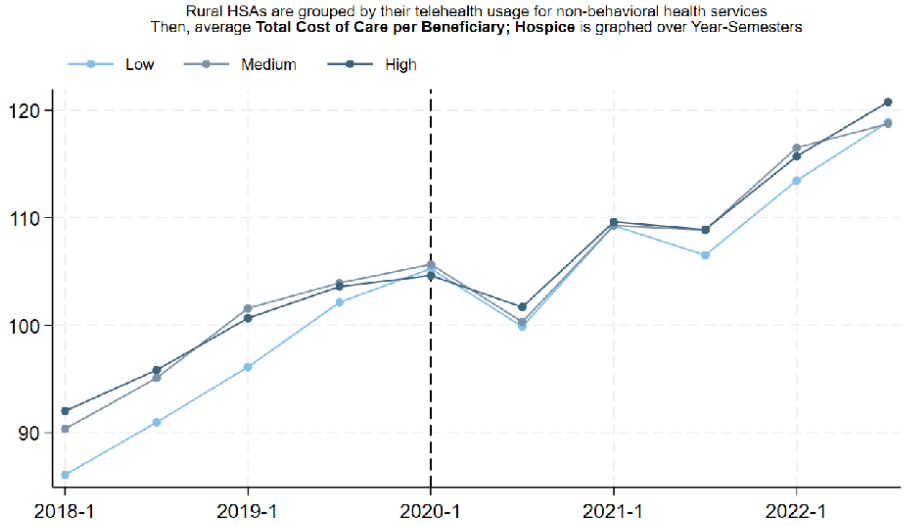




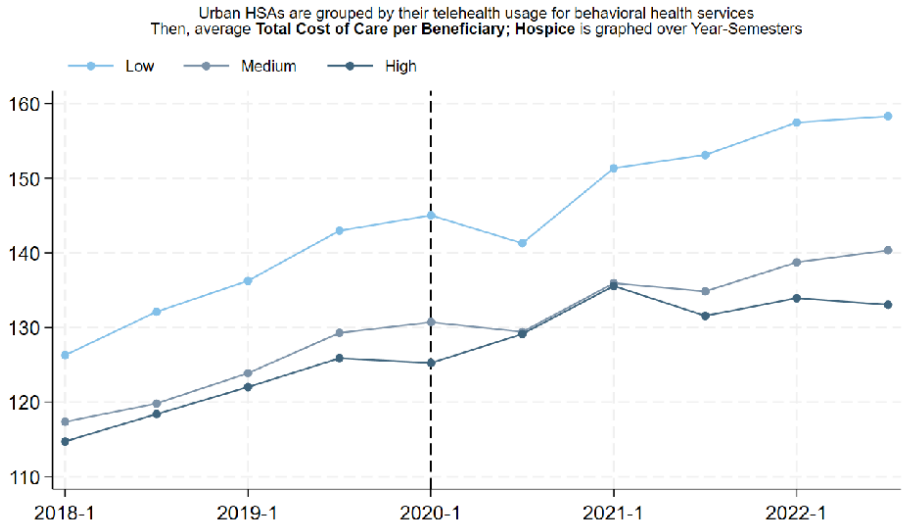
Outcome: Total Cost of Care per Beneficiary; Hospice

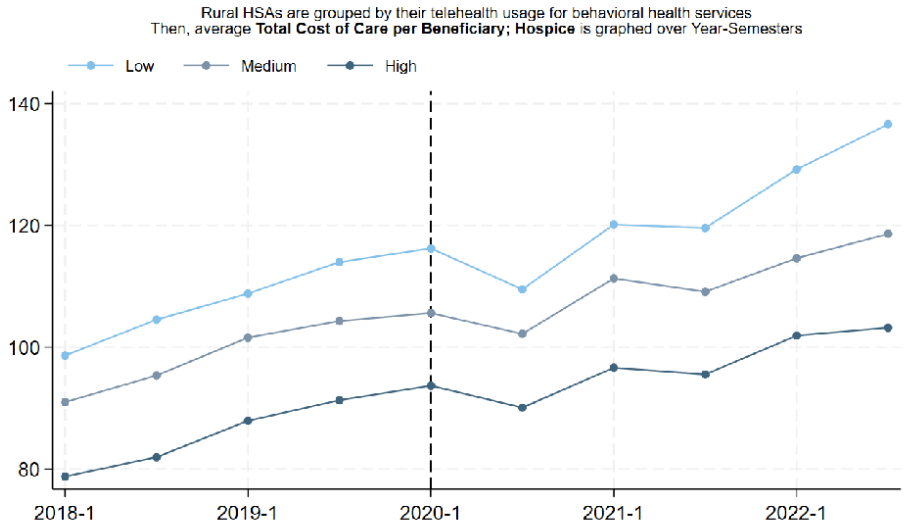
Non-Behavioral Health





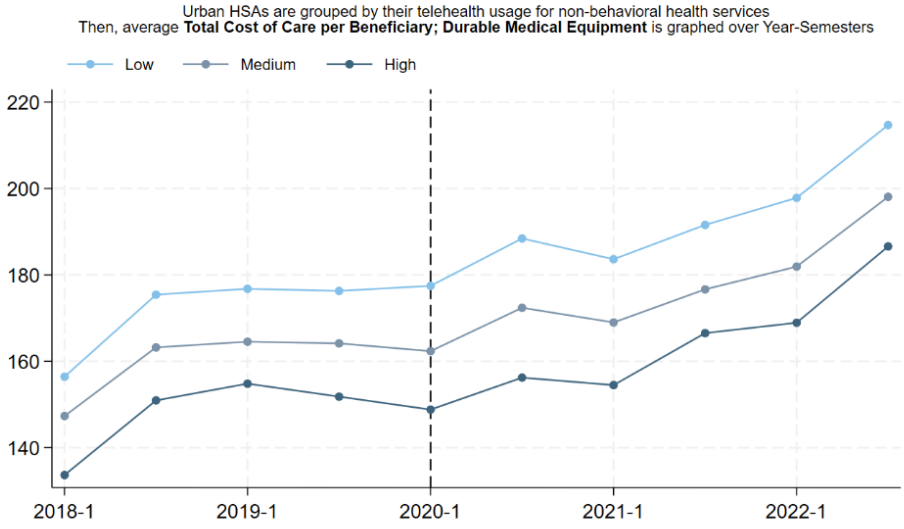
Behavioral Health

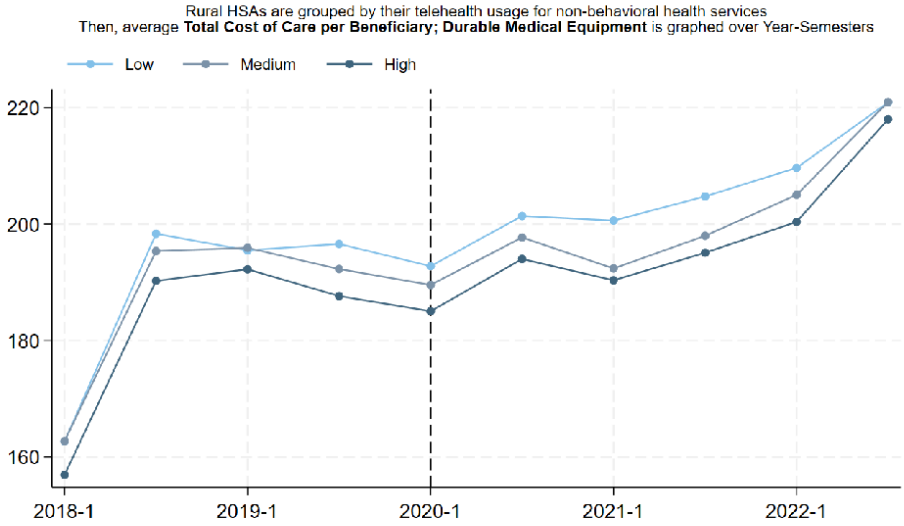




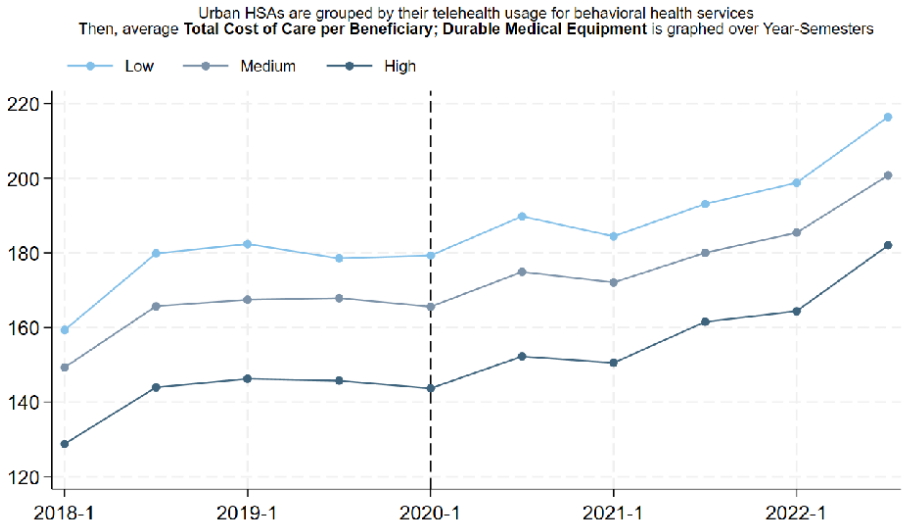
Outcome: Total Cost of Care per Beneficiary; Durable Medical Equipment

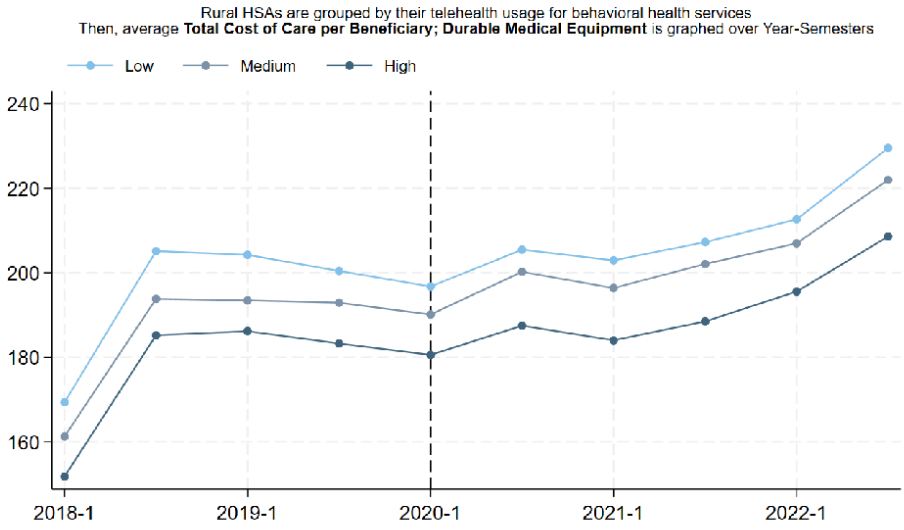
Non-Behavioral Health





Behavioral Health





30

About the American Institutes for Research®

Established in 1946, the American Institutes for Research® (AIR®) is a nonpartisan, not-for-profit institution 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**

23985_03/24