The Impact of Hospital-Plan Vertical Integration on Healthcare Utilization

Click here for the latest version

PRELIMINARY

Alberto Cappello[1] and Ira Abraham[2]

[1]Boston College
[2]Boston College

November 24, 2025

Abstract

The healthcare sector has become increasingly vertically integrated, with hospital systems integrating healthcare financing by launching their own health plans. This paper investigates the causal effect of hospital health-plan vertical integration on healthcare utilization and health outcomes using data on Medicare Advantage beneficiaries. We distinguish three causal objects: (i) the effect on patient outcomes of being enrolled in a provider-sponsored MA plan (PSHP), and (ii) the effect of receiving care at a vertically integrated (VI) hospital, along with (iii) the interaction of the two. For the plan channel, we adopt an instrumental variables approach that that combines (a) the quasi-random differences in PSHPs availability within zip codes that span more than one county. For the hospital channel, we exploit an ambulance-based instrument for hospital assignment. Using MA enrollment and claims, hospital characteristics, and ambulance claims, we estimate the impacts of plan-provider integration on different measures of healthcare utilization.

Keywords: Health Economics, Medicare Advantage, Vertical Integration, Provider-Sponsored Health Plan.

1 Introduction

Health care market consolidation in the United States has evolved over time. Historically, consolidation consisted primarily of hospitals mergers and insurers mergers, while more recently hospital and insurers are becoming increasingly vertically integrated. This latter form of integration is not only the result of mergers and acquisitions, but comes in the form of large health systems that integrate the insurance business with healthcare delivery. In recent years, there has been a rise in hospitals and health systems offering their own health plans. As of 2019, approximately 37 percent of Medicare beneficiaries are enrolled in a MA plan and 20 percent of MA plans are sponsored by a health system. See 1-4.

The literature on hospital-insurer vertical integration is primarily focused the trade-off between efficiency and market foreclosure. Instead, little is know about the effect of hospital sponsored health plans on healthcare utilization and health outcomes. Proponents argue that the vertical integration of providers and insurers can lead to administrative and clinical integration, possibly achieving high efficiency and quality (Orszag). However, opponents suspect that profit motives are the driving force behind vertical integration and believe that the vertical integration of providers and insurers may increase health care utilization and costs without necessarily improving quality (Geruso).

Medicare Advantage beneficiaries may be exposed to different level of vertical integration (VI) depending on the plan they are enrolled in and the provider from which they receive care. First, beneficiaries are not exposed to VI when they are on a non-VI plan and receive care from a non-VI provider. Second, some beneficiaries are on an integrated plan but admitted to a nonintegrated hospital. Third, some beneficiaries are on an non-VI plan and are admitted to an integrated hospital. Finally, some beneficiaries who experienced full integration when they are on a VI plan and receive care from the hospital that sponsors that plan.

The objective of this paper is to estimate the impact of insurer-provider vertical integration on healthcare utilization measures like hospital length of stay, diagnosis intensity, 30-day readmission and discharge location. In particular, we disentangle the differential impact due to the different channels of exposure to plan-provider integration:

1. a plan-side channel: the effect of enrolling in a provider-sponsored MA plan (PSHP),
2. a provider-side channel: the effect of receiving care at a hospital that is vertically integrated (VI) with a plan sponsor, and
3. a full-integration channel: a PSHP enrollee treated at their own plan's integrated hospital

We implement an empirical strategy to disentangle these channels. The novelty is to bring together two sources of quasi-experimental variation tailored to each margin. For the hospital channel, we adopt an ambulance-preference instrument that shifts emergency hospital assignment. For the plan channel, we exploit two sources of variation: first, we rely on the plausibly quasi-random differences in PSHPs availability across adjacent counties, and we use the shift in PSHPs attractiveness that is driven by the relative distance between the nearest non-VI hospital and the nearest VI hospital. For the full integration channel, we simply take the product of the hospital channel and the plan channel. The combined design delivers separate local average treatment effects (LATEs) for the plan-side and provider-side exposures and their interaction for emergency episodes. By separating plan and provider margins, our estimates inform whether observed differences arise from insurance design and network management (plan-side) or from practice patterns at integrated delivery sites (provider-side).

Relation to the literature. Our design builds on four literatures. First, a growing body of evidence compares integrated and non-integrated MA plans, documenting higher star ratings and premiums as well as lower non-claims (administrative) costs in integrated plans, with mixed equity implications (e.g., **?**). Second, encounter-level comparisons in integrated contexts show differences in coding intensity, ICU use, mortality, and plan retention (e.g., **?**). These associations motivate but do not identify causal effects. Third, the health economic theory of hospital insurer vertical integration emphasizes a trade-off between coordination benefits and upcoding incentives. In particular, (**?**) shows that in the MS market, VI plans' coding intensity is substantially higher than that of non-VI plan. Fourth, the emergency-care identification strategy that exploits ambulance-company hospital preferences provides a credible way to exogenously shift *hospital choice* for acute cases (**?**). We adapt the latter to instrument provider-side exposure (VI hospital) while separately instrumenting plan-side exposure (PSHP enrollment) using differential distance.

Contribution First, we provide a unified framework to separate plan-side and provider-side VI effects and their interaction. Second, we adapt ambulance-preference IV to the VI context, enabling credible estimation for emergency episodes. Third, we connect encounter-level effects to plan-level features documented in the MA literature (premiums, non-claims costs, quality), clarifying the mechanisms by which integration may affect value.

Methodologically, we show how to jointly identify two endogenous exposures and their synergy at the episode level, tying plan design to site-of-care assignment. Substantively, focusing on emergencies ensures that plan networks and beneficiary preferences are least able to confound hospital assignment, isolating provider-side behavior at the moment of acute care while still allowing plan-side design to operate through post-acute pathways.

2 Institutional Setting

2.1 Medicare Advantage and provider-sponsored plans

Medicare Advantage (MA) is the alternative to traditional fee-for-service Medicare and covers its beneficiaries through contracts with private plans. Integration between the provider and the health plan is becoming a prominent feature of the Medicare Advantage (MA) program. As MA plans are paid on a capitated basis and receive additional bonus payments based on quality performance, there may be an increased incentive to provide integrated care across care settings in order to maintain healthcare cost below the capitated amount and improve their enrollees' health to achieve the quality-based bonus.

Provider-sponsored health plans (PSHPs) are vertically integrated entities where a provider organization owns and operates its own insurance plan to provide healthcare services. This integration combines the payer-side (insurance) and provider-side (hospitals, doctors) functions within a single organization. This organization often takes the form of a health system, but there are instances of health plans that are sponsored by provider organizations that do not include hospitals.

Medicare Advantage (MA) plans have specific service areas composed of set of nearby counties. MA beneficiaries may enroll only in plans filed for their county of residence. As shown in Figure 3, there is a lot of variation in the number of PSHPs' enrollment share in nearby counties. In particular, we rely on the fact that there are several zip codes whose area span more than one county. This scenario is illustrated in Figure 1. MA beneficiaries who live on the upper side of the border (yellow area) have the option to enroll in a PSHP, whereas beneficiaries who live in the lower side of the border do not have that option. Thus, two otherwise similar neighborhoods that straddle a county line can face different choice sets in terms of PSHP availability. This plausibly exogenous variation in the PSHPs availability within zip codes that span more than one county is a essential for our identification strategy.

2.2 Emergency care and ambulance assignment

The locus of treatment for emergency hospitalizations is, to a large extent, determined by pre-hospital factors, including ambulance transport decisions and patient location. Moreover, there is wide evidence that large city areas are often served by multiple ambulance

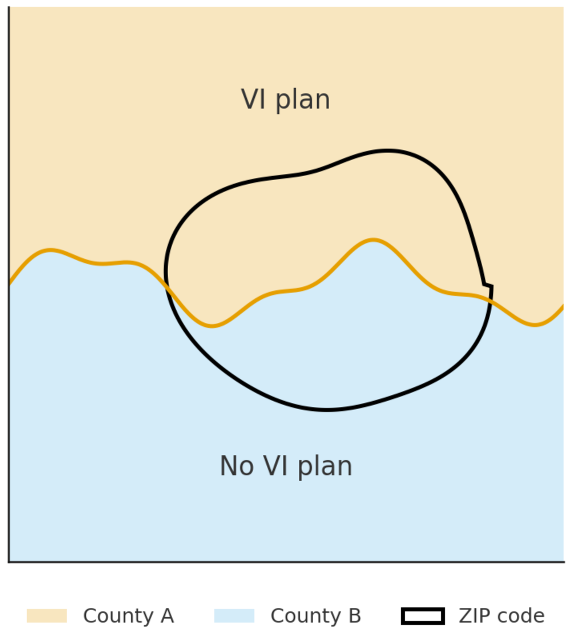

Figure 1: Within zip code variation in PSHP availability

companies, whereas smaller rural with a single ambulance company, neighboring companies provide service when the principal ambulance units are busy under so-called "mutual aid" agreements. Thus, in both cases, the assignment of the ambulance company to the patient is effectively random. (Doyle 2015, Chiang, et al., 2006 Ragone 2012).

In addition to plausibly exogenous assignment, ambulance companies are expected to have preferences for particular hospitals. In survey work described in Doyle et al. (2015), we found that paramedics have developed relationships with local emergency departments. For example, Skura (2001) studied ambulance assignment in the wake of a new system of competition between public and private ambulances in New York City. He found that patients living in the same ZIP code as public Health and Hospital Corporation hospitals were less than half as likely to be taken there when assigned a private, non-profit ambulance (29%) compared to when the dispatch system assigned them to an FDNY ambulance (64%). In most cases, the private ambulances were operated by non-profit hospitals and stationed near or even within those facilities, so they tended to take their patients to their affiliated hospitals. More broadly, with the exception of acute trauma care for which there are often defined local protocols for hospital assignment (Kahn et al. 2008) transport assignment for other emergencies is more likely to be driven by idiosyncratic preferences. Thus, in this context patient transport decisions are more likely to be made in ways that are not systematically tied to the underlying health of the patient.

Figure 2 illustrates the key source of quasi-random variation in hospital assignment that

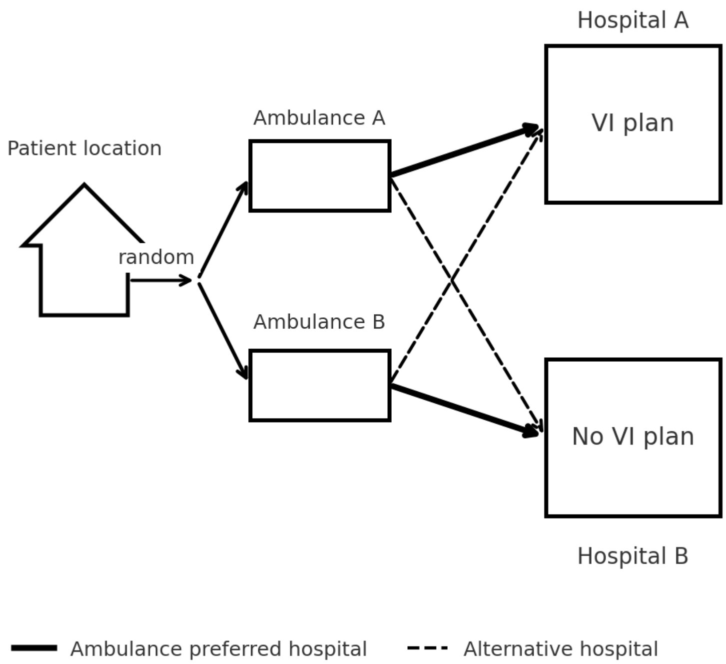

Figure 2: Hospital Assignment and Ambulance Preferences

underlies the empirical strategy. When a medical emergency occurs, the dispatch system typically assigns a patient to the nearest available ambulance, which is effectively random from the patient's perspective. Conditional on the patient's location, the ambulance dispatched may have systematically different "preferred" hospitals due to historical routing patterns, contractual relationships, or local familiarity.

As shown in Figure 2, two ambulances (A and B) serve the same geographic area but exhibit different propensities to transport patients to different hospitals. Ambulance A tends to deliver patients to Hospital A, which is vertically integrated (VI) with a Medicare Advantage plan, whereas Ambulance B tends to deliver patients to Hospital B, which is not integrated. Importantly, ambulances occasionally transport patients to the alternative hospital as well, generating natural variation in destination choice.

This setup creates a quasi-experimental shock in exposure to VI hospitals: patients who happen to be picked up by an ambulance with stronger ties to the VI hospital are more likely to be admitted there, even though the ambulance choice is conditionally random. This type of variation has been widely used in the hospital choice and emergency-care literature because it mimics an instrumental variable: ambulance identity predicts hospital assignment but is plausibly orthogonal to patient unobservables.

In this paper, I leverage this setting to isolate the causal effects of treatment within

VI systems. Conceptually, the identification compares patients who live in the same area and experience similar emergencies, but who—due to the effectively random dispatch of a given ambulance—are differentially likely to be routed to a vertically integrated versus non-integrated hospital system.

In other words, patients typically do not choose the ambulance, and dispatch/rotation protocols and ambulance company service areas generate quasi-random assignment to companies with heterogeneous hospital destination preferences. This yields plausibly exogenous variation in the admitting hospital conditional on residential location and origin, forming the basis for ambulance preference instruments (**?**).

2.3 Outcome variables

We pre-specify four episode-level outcomes that capture in-hospital intensity, near-term utilization, post-acute care pathways, and documentation intensity. Unless noted, measures pertain to the *index emergency admission* and its immediate aftermath and are constructed from Medicare Advantage encounters.

1. **Length of stay (LOS).** Hospital length of stay (LOS) is the duration a patient spends in a hospital from admission to discharge. It is a key metric used to measure a hospital's efficiency, resource utilization, and patient outcomes. Factors influencing LOS include a patient's condition, comorbidities, the quality of care provided, and staffing levels.

2. **30-day unplanned readmission.** Indicator for any unplanned acute inpatient readmission within 30 days of index discharge to any acute-care hospital. We exclude same-day transfers and identify planned readmissions using a CMS-style algorithm. Robustness includes a 90-day window and a composite of readmission-or-death within 30 days.

3. **Diagnosis (coding) intensity.** Our primary proxy counts *secondary* ICD-10-CM diagnoses on the index inpatient claim (excluding the principal diagnosis and external-cause codes). Robustness measures include: (a) the number of new Hierarchical Condition Categories (HCCs) activated during the admission relative to the prior 12 months and (b) the Elixhauser comorbidity count derived from secondary diagnoses.

4. **Post-acute care discharge.** Mutually nonexclusive indicators for discharge to: (i) skilled nursing facility (SNF), (ii) initiation of home health agency (HHA) services within two days of discharge, and (iii) inpatient rehabilitation facility (IRF) or long-term care hospital (LTCH). Primary coding uses discharge status and is corroborated

with subsequent post-acute encounters (e.g., HHA start-of-care claims). As secondary measures, we report 90-day SNF days and HHA visit counts.

3 Data and Descriptive Evidence

Provider-Sponsored Health Plans To identify integrated MA plans sponsored by health systems, we used the AHRQ Compendium of Health Systems which contains the list of all health pans offered by hospitals and health systems. We then linked all of these hospitals to their CMS provider identification numbers (CCN) to identify them in AHA database and add information on hospital characteristics. To identify health plans sponsored by other provider organizations, we searched through the full list of MA contracts to flag those whose sponsor is a provider organization.

Table 1 shows some descriptive statistics on the Enrollment Share of PSHP across counties. Each county is weighed by the number of eligible beneficiaries. Panel A shows the distribution among counties where MA penetration rate is below 50%. Among these counties, the PSHP share is almost zero in more than half the counties. The mean is significantly larger than the median which implies that there is large variation across counties. Panel B shows the distribution among counties where MA penetration rate in above 50%. We observe that PSHPs plans' enrollment share is much larger among HMO plans. For the latter, the mean and the median are similar and there is less variation compared to the other set of counties.

Figures 3 and 4 in Appendix A complement these summary statistics by mapping the county-level number of PSHP per thousands of beneficiaries enrolled in MA. Figure 3 (Figure 4) shows the number of HMO (PPO) plans sponsored by health system per thousands of beneficiaries enrolled in HMO (PPO) plans in each county. Vertically integrated plans are highly concentrated in specific regions—such as California, the Pacific Northwest, the Upper Midwest, and parts of the Southwest—while many neighboring counties exhibit little to no PSHP presence. In particular, the figure reveals sharp discontinuities in PSHP availability between adjacent counties. This stark cross–county heterogeneity, often occurring over very short geographic distances, is precisely the spatial variation that underlies our identification strategy, as beneficiaries living in the same ZIP code but assigned to different counties face meaningfully different exposure to PSHP availability.

Medicare Advantage Claims Data I use administrative data from the Centers for Medicare & Medicaid Services (CMS) for the years 2019 and 2022. The main data source for plan enrollment is the Master Beneficiary Summary File (MBSF), which contains individual-level

Table 1: Descriptive Statistics of County PSHP Availability

Plan type	**Median**	**Mean**	**SD**	**Within-ZIP SD**
Panel A: Counties with MA penetration $\leq 50\%$				
HMO	0.029	0.149	0.509	32.7%
PPO	0.004	0.049	0.202	30.6%
Total	0.010	0.098	0.390	31.4%
Panel B: Counties with MA penetration $> 50\%$				
HMO	0.027	0.060	0.145	17.5%
PPO	0.003	0.020	0.080	14.6%
Total	0.016	0.040	0.119	15.8%

Notes: Table reports the first quartile, median, mean, standard deviation, and third quartile of county-level VI plan enrollment shares, separately for HMO and PPO plans. Statistics are weighted by the number of eligible beneficiaries in each plan type at the county level. Panels split counties by MA penetration at county level: Panel A includes counties with penetration below 50%, and Panel B counties with penetration above 50%. The column "Within-ZIP share" is computed as $1 - R^2$ from a eligible-weighted regression of the county-level VI-plan measure on ZIP fixed effects (sample restricted as indicated). This is the fraction of the observed, county-eligible-weighted variation attributable to differences within the same ZIP across counties.

data on a 5% random sample of Medicare beneficiaries. The MBSF records enrollment information for each beneficiary, including whether they are enrolled in Traditional Medicare or a Medicare Advantage (MA) plan. The MBSF also includes detailed beneficiary characteristics such as age, gender, race, income measures, and county of residence. In 2019, total Medicare enrollment was approximately 61 million, so the 5% sample includes about 3 million individuals. Roughly 37% of beneficiaries were enrolled in MA plans. I exclude ESRD patients from the analysis and also special needs plans, pace plans, employer-sponsored plans or PFFS plans which have a different reimbursement scheme from CMS which is different from the standard plans. The main data source for the patients' health claims is the Inpatient Encounter Data, from which we can obtain the ambulance company and the hospital identifiers where patients were treated, the ICD-10 diagnosis codes, the procedure codes, the length of hospital stay, and the discharge location. In addition, I use the Inpatient, Outpatient, Skilled Nursing Facility (SNF), Home Health Agency (HHA), Durable Medical Equipment (DME), and Carrier Encounter files to calculate the average cost of care for Medicare beneficiaries.

Sample construction We construct the analysis sample in several steps. First, we restrict attention to non-discretionary emergency conditions —such as acute myocardial infarction, stroke, and hip fracture— since they represent high-acuity conditions for which the choice of hospital is effectively non-deferrable and largely outside the patient's control. These emergency conditions are clinically significant and account for a substantial share of inpatient hospital spending among Medicare beneficiaries. We also restrict the estimation sample to emergency hospitalizations arriving via ambulance transport. This ensures that hospital assignment is driven by the ambulance dispatch and routing process rather than by the patient or their family.

To exploit within–ZIP-code geographic variation in exposure to vertically integrated (VI) Medicare Advantage plans, we further restrict the sample to beneficiaries who reside in ZIP codes that span more than one county or are located immediately adjacent to a county boundary. Because MA plan availability is defined at the county level, but beneficiaries' residential ZIP codes often cross county borders, this restriction isolates ZIP codes where individuals are plausibly similar yet face different MA plan menus due to their county assignment. This geographic mismatch generates the key source of within- ZIP variation in VI plan exposure used in the empirical strategy.

Applying these restrictions yields a beneficiary-by–episode-of-care dataset. The unit of observation is an emergency episode for beneficiary i in calendar year t, defined as an inpatient admission originating in the emergency department (or an observation stay subsequently converted to inpatient) with no transfer-in from another acute-care facility. For each episode, we observe the beneficiary's MA plan enrollment at time t, the identity and characteristics of the admitting hospital, and the utilization measure during the inpatient saty.

Our 2019 sample contains approximately 16 thousands emergency episodes meeting the above criteria. In ongoing work, we will expand the dataset to include 2022 Medicare claims, which will significantly increase statistical power and allow us to assess whether patterns of VI hospital effects have changed in the post-COVID period.

4 Empirical Framework

The health plan chosen by the enrollee i is denoted as $p(i)$ and the hospital where she receives care is $h(i)$. The plan level vertical integration is given by $\mathrm{VI}_i^p \in \{0, 1\}$ with

$$\mathrm{VI}_i^p = \begin{cases} 0 & \text{if plan } p(i) \text{ is non-PSHP} \\ 1 & \text{if plan } p(i) \text{ is PSHP} \end{cases}$$

and the hospital level vertical integration is given by $\text{VI}_i^h \in \{0, 1\}$ with

$$\text{VI}_i^h = \begin{cases} 0 & \text{if hospital } h(i) \text{ is not a plan sponsor} \\ 1 & \text{if hospital } h(i) \text{ is a plan sponsor} \end{cases}$$

Notice that it is possible patient i is on a PSHP plan ($\text{VI}_i^p = 1$) and receives care receives care from an integrated hospital ($\text{VI}_i^h = 1$), but the plan and the hospital are not integrated with each other. Thus, in order to capture the fully integrated case we cannot not just the product of VI_i^p and VI_i^h. We need to define a third indicator $\text{VI}_i^f \in \{0, 1\}$

$$\text{VI}_i^f = \begin{cases} 0 & \text{if hospital } h(i) \text{ is not the sponsor of plan } p(i) \\ 1 & \text{if hospital } h(i) \text{ is the sponsor of plan } p(i) \end{cases}$$

We observe $(\text{VI}_i^p, \text{VI}_i^h, \text{VI}_i^f)$ and the health outcome y_{it}, and estimate

$$y_{it} = \beta_1 \text{VI}_{it}^p + \beta_2 \text{VI}_{it}^p + \beta_3 \text{VI}_{it}^f + X_{it}'\Gamma + \alpha_{s(i)} + \delta_{\text{d}(i)} + \tau_t + \varepsilon_{it}, \tag{1}$$

where X_{it} includes demographics, risk score, HCC comorbidities, plan characteristics. We include state fixed effects, $\alpha_{s(i)}$ principal diagnosis fixed effects $\delta_{\text{d}(i)}$ and time fixed effects τ_t.

The indicators for the level of plan-hospital integration are endogenous: patient choice to choose VI plan and receive care from a VI hospital may be correlated with unobserved plan and hospital features that also affect the outcome. Therefore, we construct two instruments, one for the VI-hospital choice channel and the other for VI-plan choice channel.

Instrument for VI-Hospital choice First, we adopt an ambulance-preference instrument Z_i that captures the destination hospital style associated with the ambulance company that transports patient i, measured as the leave-one-out average of a hospital-level attribute among other patients previously transported by that company (details in Section **??**). For each ambulance company a, we compute the share of emergency transports that terminate at an integrated hospital (a hospital that is sponsor of a health plan) among other patients carried by a in the same market, year, and emergency case. We exclude i from the computation of the share to avoid mechanical correlation between the instrument and the index patient's outcome. Formally, for patient i transported by $a(i)$:

$$\text{Amb Share}_{it} = \frac{1}{N_{a(i),t}} \sum_{j:a(j)=a(i)} \mathbf{1}\{\text{VI}_{jt}^h = 1\} \tag{2}$$

Conditional on fixed effects and controls, we assume that ambulance preference affects outcomes only through hospital assignment in the emergency (and the resulting exposure to a VI hospital), not through unobserved severity or ambulance treatment.

Instrument for VI-Plan Choice Medicare Advantage plans' service area is county specific. This implies that beneficiaries who live nearby but residing in different counties might be offered different MA health plans. Our identification strategy relies on the fact that beneficiaries residing in the same zip code might be presented with different choice sets in terms of PSHP. Therefore, we restrict our sample to those zip codes that either span multiple counties, or are located within few kilometers from another county border. We leverage the variation in PSHPs availability across these zip codes as a quasi-exogenous shifter of the probability that a beneficiary enrolls in a PSHP. The identification assumption is that variation in availability of PSHPs in these zip codes is not correlated with within zip code differences in health and demographic characteristics.

Let $c(i)$ and $z(i)$ denote the beneficiary i county and zip code respectively. To compare PSHP availability across counties, we compute the number of PSHP per thousand MA beneficiaries. We denote this by "Availability Index", $\text{Avail}_{i,t} \in [0, 1]$. The key source of exogenous variation is that beneficiaries who live in the same zip code $z(i)$, but reside in different counties $c(i)$, are exposed to different levels of $\text{Avail}_{i,t}$. [1]

We therefore use $\text{Avail}_{i,t}$ as instrument for PSHP_{it} choice. Thus, the first stage regression for plan enrollment is given by:

$$\text{VI}^p_{it} = \gamma \,\text{Avail Index}_{i,t} + X'_{it}\Gamma + \alpha_{s(i)} + \delta_{\text{d}(i)} + \tau_t + u_{1it}. \tag{3}$$

The exclusion restriction requires that, among the set of zip codes that we consider for estimation, $\text{Avail}_{i,t}$ affects our health outcomes measures only through PSHP_{it}.

We also need to instrument the fully integrated case VI^f_i. The latter cannot be written as $\text{VI}^p_i \times \text{VI}^h_i$ since PSHP enrollees can receive care from a VI hospital that is not the sponsor of the plan. However, we do not observe any instances of this in the data. Therefore, we use $\text{Amb Share}_{it} \times \text{Avail}_{i,t}$ as an instrument for VI^f_i.

[1] We exclude SNPs from both the instrument construction and the analytic sample to keep the menu relevant to non-dual beneficiaries and avoid plan types with distinct eligibility rules.

Baseline specification We estimate a two-stage least squares (2SLS) regression. The first stage is given by

$$\mathrm{VI}^{p}_{it} = \gamma_{10}\,\text{Avail Index}_{it} + X'_{it}\Gamma_0 + \alpha_{s(i)} + \delta_{\mathrm{d}(i)} + \tau_t + u_{0it}, \tag{4}$$

$$\mathrm{VI}^{h}_{it} = \gamma_{11}\text{Amb VI Share}_{it} + X'_{it}\Gamma_1 + \alpha_{s(i)} + \delta_{\mathrm{d}(i)} + u_{1it} \tag{5}$$

$$\mathrm{VI}^{f}_{it} = \gamma_{12}\big(\text{Amb VI Share}_{it} \times \text{Avail Index}_{it}\big) + X'_{it}\Gamma_2 + \alpha_{s(i)} + \delta_{\mathrm{d}(i)} + u_{2it} \tag{6}$$

and the second stage:

$$y_{it} = \beta_0 + \beta_1\widehat{\mathrm{VI}^{p}_{it}} + \beta_2\widehat{\mathrm{VI}^{h}_{it}} + \beta_3\widehat{\mathrm{VI}^{f}_{it}} + \Gamma X_{it} + \alpha_{s(i)} + \delta_{\mathrm{d}(i)} + \tau_t + \varepsilon_{it} \tag{7}$$

where X_i includes demographics and comorbidities; $\alpha_{s(i)}$ are state-level fixed effects, $\delta_{d(i)}$ are principal diagnosis fixed effects, and τ_t are time fixed effects. We cluster standard errors at the ambulance-company or hospital level.

5 Preliminary Results

The estimation results presented in this section should be considered preliminary. At this stage, we rely only on 2019 data, which limits the available identifying variation—particularly for estimating the effect of full integration, where plan and hospital are jointly vertically integrated. As we incorporate additional years of Medicare Advantage encounters (including 2022), we expect substantially greater identifying variation and statistical power. All results will be updated and refined over the coming weeks.

All outcome models control for beneficiary demographics (age, sex, race), dual eligibility, observable health characteristics (risk score and HCC comorbidities), and plan characteristics. We include principal-diagnosis fixed effects so comparisons are within clinical condition. Our geographic fixed effects are at the state level since our current sample size does not allow us the have more granular fixed effects. Standard errors are clustered at the ambulance-company or health system level. We interpret these results as consistent we the fact that VI plans implement tighter care coordination and stronger post-acute management within integrated systems. This echoes existing evidence that integrated provider–plan structures reduce fragmentation and create incentives to invest in follow-up management.

Table 2 shows the results for 30-day all-cause readmissions. The dependent variable is a binary indicator for whether the beneficiary is readmitted to any hospital within 30 days after her discharge date. Our results indicate that beneficiaries enrolled in a PHSP plan have approximately 33% lower likelihood of being readmitted. The impact of being treated by a

VI hospital is not significant. This pattern is consistent with integrated systems' incentives to manage inpatient utilization more aggressively, particularly since VI plans internalize hospital costs.

Table 2: 30-day all cause Readmission

	Odds ratio	Conf. Interval
VI plan	0.669***	[0.502, 0.893]
VI hospital	1.170	[0.935, 1.464]
Observations		16,082

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Notes: (i) **Sample restriction**: beneficiaries living in zip codes that span more than one county or are located near the county border. (ii) **Dependent variable**: indicator for 30-day all-cause readmission. (iii) principal diagnosis and state FE are included

Table 3 shows the results for the inpatient Length of Stay (LOS). The dependent variable in the logarithm of the day difference between the discharge date and the admission date. Our results indicate that the LOS among beneficiaries enrolled in a PHSP plan is approximately 0.6% shorter than that of beneficiaries enrolled in other plans. The impact of being treated by a VI hospital is not significant.

Table 3: Inpatient Length of Stay (log)

	Coefficient	Robust Std. Error
VI plan	-0.600**	(0.267)
VI hospital	0.089	(0.055)
Observations		16,082

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Notes: (i) **Sample restriction**: beneficiaries living in zip codes that span more than one county or are located near the county border. (ii) **Dependent variable**: log length of stay for inpatient admissions. (iii) principal diagnosis and state FE are included

Table 4 shows the results for the diagnosis intensity. The dependent variable in the logarithm the number of ICD-10 diagnosis codes. Our results indicate that the diagnosis intensity is lower among beneficiaries enrolled in a PHSP plan by approximately 0.55%, and 0.17% higher among beneficiaries treated by a VI hospital.

Table 5 shows the results for the discharge location. The dependent variable is an indicator of whether the beneficiary discharged to another hospital, to a post acute care (PAC) facility or home. The impact of enrolled in a PHSP plan are approximately 65% more likely to be discharge to a PAC facility. The impact of receving care from a VI hospital is not significant.

Table 4: Diagnosis Intensity (log)

	Coefficient	Robust Std. Error
VI plan	-0.551***	(0.206)
VI hospital	0.171***	(0.042)
Observations		16,082

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Notes: (i) **Sample restriction**: beneficiaries living in zip codes that span more than one county or are located near the county border. (ii) **Dependent variable**: log number of ICD-10 diagnosis. (iii) principal diagnosis and state FE are included

Table 5: Discharge Location (PAC)

	Odds ratio	Conf. Interval
VI plan	1.652*	[1.057, 2.582]
VI hospital	0.977	[0.794, 1.153]
Observations		16,082

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Notes: (i) **Sample restriction**: beneficiaries living in zip codes that span more than one county or are located near the county border. (ii) **Dependent variable**: indicator for PAC discharge. (iii) principal diagnosis and state FE are included

Taken together, these results indicate that VI plans tend to have lower healthcare utilization in terms of readmission, inpatient length of stay and diagnosis intensity. We think that these patterns are consistent with the fact that vertically integrated MA plans internalize both the cost of hospital care and the downstream financial implications of readmissions, inpatient stay and high-intensity procedures. A second explanation for this higher efficiency is the larger utilization of PAC services.

6 Conclusions

References

Appendix A

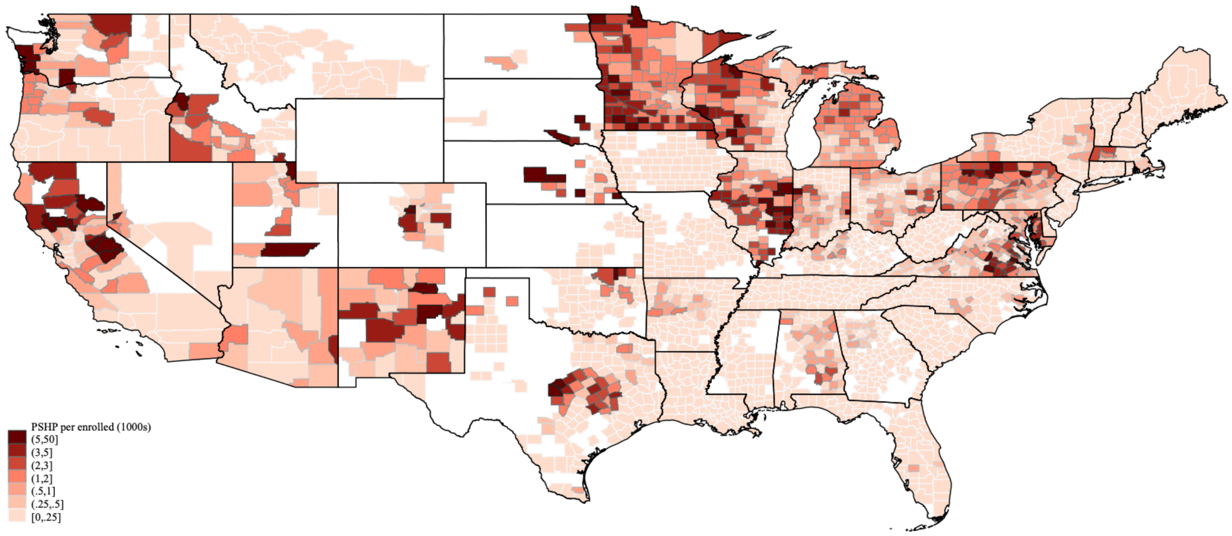

Figure 3: Number of HMO-PSHP per thousands of county HMO enrollees

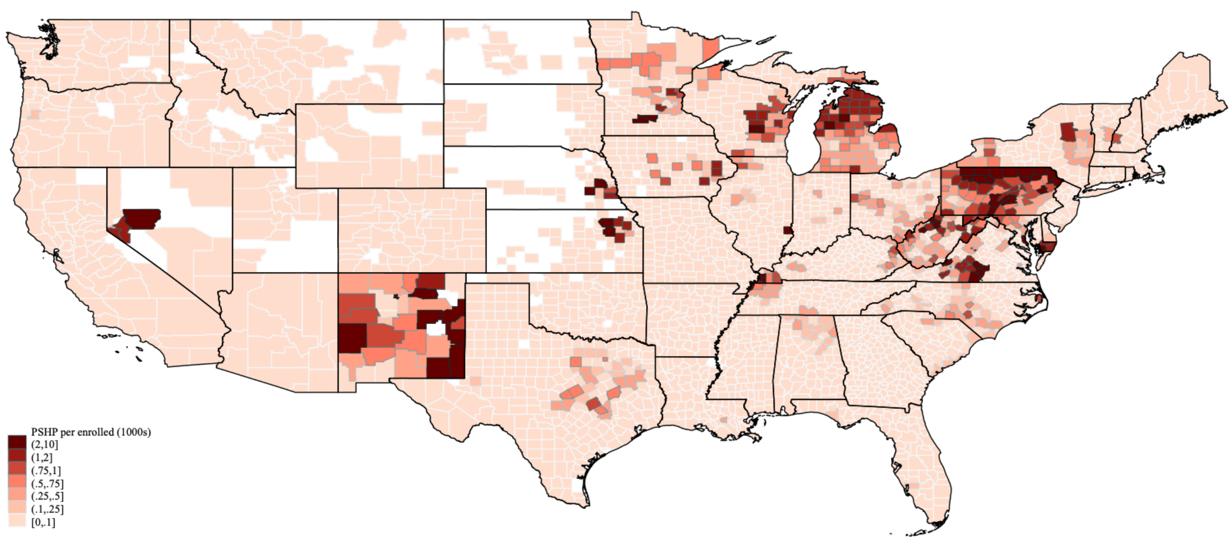

Figure 4: Number of PPO-PSHP per thousands of county PPO enrollees