DEFINING MARKETS FLEXIBLY ACROSS SPACE: AN APPLICATION TO PHYSICIAN SHORTAGES*

XUYANG XIA †

YUCI ZHOU ‡

November 16, 2025

[Click Here for the Latest Version]

Abstract

Rigid and broad market definitions can misrepresent consumption patterns and misguide resource allocation and policy decisions. We develop a new framework that flexibly defines overlapping markets at the consumer level by modeling consumption flows across space. Applied to the U.S. Health Professional Shortage Area (HPSA) designations, our approach estimates 11.6 million people lack adequate access to primary care physicians (PCPs) — far fewer than the 30 – 100 million identified by HPSAs. Simulations indicate that allocating 3,734 additional PCPs could eliminate shortages nationwide. These findings underscore the importance of flexible market delineation and offer actionable guidance for improving primary care access.

*We are particularly grateful to Jimmy Roberts, Janet Currie, Kate Ho, and Jakub Kastl for their continued guidance and generous support. We thank Patrick Bayer, Kate Bundorf, Emily Cuddy, Michael Dinerstein, Joshua Gottlieb, Adam Kapor, Riley League, Ryan McDevitt, David Ridley, Felix Tintelnot, Daniel Xu for invaluable advice, and Tirza Angerhofer, Arthur Carvalho, Yufei Li, Kate Musen, Nicolás Urdaneta, and seminar participants at ASHEcon, Duke University's IO and Labor Lunches, and Princeton University's Industrial Relations Section Graduate Student Workshop and Center for Health and Wellbeing Student Workshop for helpful discussions and comments. All errors are our own.

†Duke University. Email: xuyang.xia@duke.edu.

‡Princeton University. Email: yucizhou@princeton.edu.

1 Introduction

Many government programs target resources to areas defined by existing geographic boundaries, with limited attention to whether these divisions accurately reflect local market conditions. For example, Economic Development Districts allocate grants along multi-county lines, and housing voucher values are set at the metropolitan level. While targeting can, in principle, be more focused when markets are defined narrowly, standard measurement methods are not well equipped to account for interactions across boundaries — hence most programs rely on broad administrative divisions to approximate self-contained markets. This approach, however, often overlooks substantial within-area heterogeneity, resulting in overly broad and less effective targeting.

We develop a new framework that flexibly evaluates resource access at the consumer level by modeling consumption flows across arbitrarily small geographic units. We apply this framework to evaluate primary care physician (PCP) shortages in the United States using the ZIP Code Tabulation Areas (ZCTAs),[1] and show that existing government designations substantially overstate the extent of shortages and misidentify the locations most in need. Leveraging the network structure of patient travel, we propose practical heuristics for more effective physician allocation that enable larger, more targeted incentives to be deployed within existing budgets.

The shortage of PCPs has been a persistent public concern for decades: studies estimate that the shortfall of PCPs ranges between 20,000 and 50,000 nationwide (Petterson et al., 2015); according to the American Medical Association, as of 2023, approximately 83 million people in the U.S. live in areas with PCP shortages (American Medical Association, 2023; Landro, 2024). In response, federal and state governments have spent millions of dollars annually on programs aimed at addressing PCP shortages with mixed results (Shugarman and Farley, 2003; Khoury et al., 2025).

Despite the policy attention, the current PCP shortage measures exhibit critical limitations. The most widely used metric — federal designation of Health Professional Shortage Areas (HPSAs) — guides the distribution of federal and state resources by identifying shortage areas based on population-to-PCP ratios within fixed geographic boundaries. This approach ignores key factors such as traveling for care across boundaries, within-area heterogeneity in access, and patient travel costs (Guagliardo, 2004). As a result, it tends to overstate shortages near major care markets and understate them in regions with high inbound demand, which may lead to inefficient physician and funding allocations.

[1] The United States contains more than 33,000 ZIP Code Tabulation Areas, in contrast to approximately 3,000 counties.

Our framework addresses these limitations by tailoring the market to each small consumer location and allowing these markets to overlap arbitrarily. We use a small geographic division, the ZCTA, as the unit of analysis, and directly model and estimate visit flows across ZCTAs. We measure PCP access within patients' 30-minute travel radius, we call the "catchment area", which is consistent with HPSA's measurement principle, but our framework allows this area to be defined flexibly for each ZCTA based on its local travel network. Residents in each ZCTA access primary care resources within their catchment area, while physicians in the catchment area may also serve residents from other ZCTAs. Using administrative Medicare claims data, we estimate a gravity model of primary care visit flows between ZCTAs, following Dingel et al. (2023), which allows us to endogenize and predict the equilibrium total demand for PCPs in each catchment area. We also complement this framework with detailed ZCTA-level travel time data from Hu et al. (2020) to construct a more realistic travel network for patients. Our method naturally accounts for cross-boundary visits and allows for rich heterogeneity in local travel patterns, patient demand, and unobserved primary care characteristics.

Maintaining the HPSA criteria, which define markets as underserved when each PCP covers more than 3,500 people on average, we estimate that 11.58 million people lacked sufficient access to primary care in 2019 — well below the 83 million figure reported by industry groups and the 32.2 million residing in government-designated Geographic HPSAs. We also find that, conditional on the degree of shortage, HPSA designations tend to prioritize areas traditionally viewed as underserved, such as rural and socioeconomically disadvantaged communities, reflecting HPSA's tendency to overlook shortages in markets with high primary care inflows.

By incorporating the structure of the travel network, our framework also provides new insights into more efficient targeting strategies. We contrast two algorithms for assigning new PCPs using simulations: (1) add the PCP to the area that *has* the largest population in shortage, and (2) add the PCP to the area that *is connected to* the largest population in shortage through the travel network. Accounting for the network structure, the second algorithm eliminates national PCP shortages with 3,734 additional physicians — 8.5% fewer than under the first algorithm that only adds physicians to shortage areas. This estimate is significantly lower than the commonly cited shortage of over 20,000 PCPs (Petterson et al., 2015), which is based on the population-to-PCP ratios calculated in HPSAs. The Medicare HPSA Physician Bonus Program spent $150 million in 2019, on average around $3,000 per eligible PCP, which would be sufficient to instead fund annual subsidies of approximately $35,000 per

physician — for both existing and newly added physicians — in the ZCTAs targeted by our proposed plan.[2]

Overall, our findings suggest that current PCP shortage measures both overstate the extent of shortage and misidentify areas of need, and we highlight promising opportunities for improvement. The observed shortage reflects not only an insufficient supply of PCPs but also inaccuracies in shortage designations and inefficiencies in policy targeting. The extent of shortage in a given area depends on both local physician availability and the broader flow of patient visits — an equilibrium outcome shaped by demand, supply, and travel network structure. By directly modeling these flows, we can better diagnose the shortage landscape and evaluate potential policy responses. Moreover, our results show that resolving PCP shortages does not necessarily require placing physicians directly into shortage areas; instead, more strategic allocations that account for patient travel network can be more effective in resolving shortages.

More broadly, this paper highlights the importance of a well-founded measurement framework for assessing local conditions and designing targeted policies. A flexible delineation of markets yields more reliable assessments and informs more effective and efficient resource allocation.

The rest of this paper proceeds as follows. Section 2 reviews the related literature. Section 3 introduces the policy background of government-designated shortage areas. Section 4 describes the data we use. Section 5 explains our shortage measurement framework and relevant model specifications. Section 6 shows our estimates of primary care shortages and compares them with current government designations. Section 7 simulates and compares different plans for allocating PCPs to eliminate shortages. Sections 8 and 9 provide further discussions and conclude.

2 Related Literature

We contribute to multiple strands of literature: market definition, health care, and physician shortages. Defining and measuring consumer choices and access to resources is an essential task in economic research — ranging from geographic markets (Manning and Petrongolo, 2017; Houde, 2012) to online markets (Yu, 2024; Sullivan, 2025), and accounting for frictions such as behavioral and supply-side constraints (Abaluck and Adams-Prassl, 2021; Agarwal and Somaini, 2022). More realistic market definitions improve demand estimation, counterfactual prediction, and policy design. This paper ad-

[2] In 2017, the average (median) total individual income among 40- to 55-year-old PCPs is $282,300 ($235,300) (Gottlieb et al., 2025).

vances methods for defining and measuring access to resources across geographic space, enabling more precise estimation of local demand and more targeted, cost-effective place-based policies.

The definition of geographic markets is of particular importance given their widespread use in empirical research and policy design. Most existing approaches rely on the "little in from outside (LIFO)" and "little out from inside (LOFI)" principle, epitomized by the ElzingaHogarty shipment test widely used in antitrust analysis before the 2000s (Elzinga and Hogarty, 1973, 1978). Market delineations such as Commuting Zones (often treated as local labor markets; cf. Autor and Dorn, 2013), Hospital Referral Regions (The Center for the Evaluative Clinical Sciences, Dartmouth Medical School, 1996), and other healthcare markets (Schappe et al., 2022; Michelson et al., 2025) follow a similar logic, aiming to identify self-contained areas with fixed boundaries. However, these methods often yield overly broad markets, leading to mismeasured local conditions, making potentially harmful mergers appear benign (Capps et al., 2003; Dafny et al., 2019), and weakening policy effectiveness (Collinson and Ganong, 2018). In response, researchers and policymakers have developed more focused, context-specific alternatives. In antitrust, authorities now rely on the hypothetical-monopolist (SSNIP) test to identify the smallest set of products and locations over which a monopolist could profitably impose a small but significant, nontransitory price increase (FTC, 2023); labor economists have defined markets directly from worker mobility patterns (Nimczik, 2020); and, most closely related to our study, Manning and Petrongolo (2017) construct flexible, overlapping job-seeker markets in small geographic areas to better capture the local effect of demand shocks. Our framework aligns with these developments in aiming to flexibly define markets for small geographic units while capturing realistic consumption and substitution patterns. We contribute to this literature by developing a general framework that accommodates heterogeneous consumption behaviors across space and readily applies to other contexts where market boundaries are difficult to delineate.

The lack of reliable measurement tools for small geographic areas has posed substantial challenges for effective policy design and, at times, allowed market participants to game the system. For instance, Corinth et al. (2025) and Freedman et al. (2023) show that the Opportunity Zone program — which provides tax incentives for investments in low-income areas — has disproportionately subsidized projects in locations already experiencing strong economic growth among the eligible tracts. In this paper, we demonstrate that our measurement framework enables more precisely targeted interventions, substantially reducing fiscal burdens while allowing for stronger and more effective

incentives.

In the context of physician shortages, early studies questioned the need for policies aimed at redistributing physicians geographically (e.g., Newhouse et al., 1982a,b), arguing that the observed distribution may simply reflect market efficiency. However, the issue of healthcare access extends beyond efficiency considerations, as healthcare is widely recognized as a fundamental human right (World Health Organization, 2017). Measuring access to healthcare and addressing shortages is therefore a central policy concern, and this paper develops a reliable measurement framework to better inform and improve policy design.

Our paper thus contributes to a long-standing literature on primary care shortages, the designation of shortage areas, and physician allocation. HPSA designations incentivize physicians to practice in shortage areas, in return for benefits such as bonus payment and loan repayment options (see Section 3 for more details). There is mixed evidence on the effectiveness of such financial incentives in relocating physicians to underserved areas and improving access to care (Shugarman and Farley, 2003; Falcettoni, 2018; Mark, 2021; Sauve-Syed, 2025; Khoury et al., 2025). In particular, Khoury et al. (2025) shows that, after a county is fully or partially designated as an HPSA, the equivalent population-to-primary care physician ratio decreases from 2,571:1 to 2,493:1 in the medium run, almost all driven by early-career PCPs. On the other hand, Falcettoni (2018) shows that student loan forgiveness programs and other salary incentives, mostly tied to practicing in HPSAs[3], led to 1.2 percent more physicians choosing rural areas. Researchers have long been conjecturing that these limited effects are due to insufficient incentives, and Congress also once considered increasing the Medicare Bonus payment from 10% to 20% of the reimbursement rate (Shugarman and Farley, 2003). Our paper suggests that the poor targeting of HPSA designation could be limiting its effectiveness and sheds light on potential channels for improvement. If subsidies are better targeted at fewer areas, the federal government can allocate more funding to each area using the same budget.

Methodologically, our framework improves upon methods used in the transportation and geography literatures for measuring the spatial accessibility of healthcare resources (Guagliardo, 2004). Luo and Wang (2003) highlight the limitations of using administrative boundaries to define physician shortage areas, and adapt the "floating catchment method" from job accessibility studies (Peng, 1997; Wang, 2000) to measure physician access. The floating catchment method addresses the key

[3]It is part of the Affordable Care Act: https://www.hhs.gov/sites/default/files/v-healthcare-workforce.pdf.

limitation that fixed-boundary measures ignore across-boundary access. Meanwhile, tools such as kernel density functions (e.g., Guagliardo et al., 2004), simple gravity models (e.g., Joseph and Bantock, 1982), and hybrid approaches (e.g. Neng Wan and Sternberg, 2012) are used to capture patient travel patterns. These methods typically rely on oversimplified representations of visit flows; travel costs are often calibrated or estimated using simple statistical methods, while other dimensions of patient preferences are either ignored or reduced to a single factor, such as physician counts. However, studies have documented the significant influence of travel costs and local service quality on patients' care-seeking behavior (Buchmueller et al., 2006; Currie and Reagan, 2003; Avdic et al., 2024; Battaglia, 2025; Rossin-Slater, 2013). We improve upon existing methods by adding a detailed economic demand framework. We estimate a discrete choice model of patient care-seeking (à la Dingel et al., 2023) that incorporates both travel costs and destination-specific unobserved attractiveness. These local attractiveness terms are then projected onto a rich set of location characteristics and the number of physicians, endogenizing the relationship between patient choice and physician distribution in the travel network and allowing for more realistic counterfactual analyses.

3 Background

3.1 Healthcare Shortage Area Designations

Different types of healthcare shortage areas are designated by the federal Health Resources and Services Administration (HRSA). These designations identify areas, populations, or facilities with insufficient access to healthcare, and guide the allocation of federal and state resources for attracting and retaining healthcare providers — via bonus payments, student loan repayment, J-1 visa foreign residency requirement waiver, etc. — and for establishing community health centers and health maintenance organizations (HRSA Health Workforce, 2025b).

One of the widely used designations is the Geographic Primary Care Health Professional Shortage Area (HPSA), which identifies primary care shortages considering the entire population in the area.[4] Geographic Primary Care HPSAs are eligible for most government incentive programs. One such program is the Physician Bonuses in Health Professional Shortage Areas (HPSA Bonus), operated through Medicare, the public health insurance program that covers people aged 65 and older. The HPSA Bonus pays 10% quarterly bonuses to primary care providers who furnish Medicare services in

[4]Other types of HPSAs designate certain populations or healthcare facilities. See Appendix B for more details.

Geographic Primary Care HPSAs, on the basis of the amount billed to fee-for-service Medicare (CMS, 2025). Using Medicare claims data, we estimate that the Bonus program spent $150 million in 2019,[5] more than doubled from the $56 million (in 2019 dollars) in 1998 (Shugarman and Farley, 2003).

The objective of this paper is to provide policy guidance on measuring and alleviating geographic primary care shortage for the overall population. Therefore, we mainly benchmark our shortage measure to the Geographic Primary Care HPSA.[6] In the next section, we summarize and discuss the criteria used to define Geographic Primary Care HPSAs. We refer readers to Appendix B and Khoury et al. (2025) for more detailed discussions on other types of shortage designations, levels of HPSAs, and related government incentive programs.

3.2 Designation Criteria of Geographic Primary Care HPSA

In this section, we summarize the main ideas behind the designation criteria for Geographic Primary Care HPSA, and discuss its limitations and the motivation for developing an alternative framework. To be eligible for Geographic Primary Care HPSA, an area needs to meet three general conditions (HHS, 2025): i) the area is a *rational area* for primary care; ii) the area's population-to-physician ratio meets a certain requirement; iii) the area's *contiguous areas* cannot help alleviate its shortage situation.

Rational Service Area A Rational Service Area is the HPSA's definition for a primary care market. Generally, a rational area for primary care service can be a single county or a group of contiguous counties whose population centers are within 30 minutes of travel time, or a group of subcounties or census tracts where the combined area does not have contiguous resources within 30 minutes. In other words, people in a non-shortage area, abstracted to its population center, would have enough primary care physicians within 30 minutes.

Population to Physician Ratio After defining markets, the next step is to measure whether each market has enough primary care physicians. The main threshold HRSA uses is a population to full-time-equivalent primary care physician ratio of 3,500:1, with the exception of a 3,000:1 ratio for areas

[5]Estimated using the 20% sample of Medicare Part A and Part B Carrier file, by calculating the 10% of Line Provider Payment Amount of claims with Carrier Line Health Professional Shortage Area (HPSA)/Scarcity Indicator Code indicating HPSA designation. We then multiply the sum by 5 to estimate the total national amount. See Section 4 for details about the professional claims data.

[6]We acknowledge that, in practice, some shortage areas are designated as Population HPSAs instead: PCOs sometimes apply for a Population HPSA as it is easier to justify than a Geographic HPSA, according to our interview with a state PCO. Thus, whenever we compare our measure with Primary Care HPSAs, we also include a version comparing with both Geographic and Population Primary Care HPSAs in the Appendix.

with high needs for primary care. Using the 2019 HPSA designation data, we find that 64% of the designated area (and 60% of the designated population) used the 3,500:1 ratio as the target. Under our framework, places with high demand are directly reflected in the data, estimated with a high local attractiveness, and accounted for in our shortage measurement. Therefore, in our baseline measure, we use the standard 3,500:1 ratio for all areas, and include robustness checks using 3,000:1 in Table A.2.

Contiguous Area Considerations For each market, exceeding the cutoff (3,500 or 3,000) alone is not sufficient to designate a shortage area, since its population can still access primary care resources in adjacent areas — in other words, this market definition may misrepresent consumers' choice sets. To take this into account, HRSA specifies additional rules for the areas contiguous to the focus area.

However, these rules are ambiguous and may lead to multiple interpretations. First, there is no clear guideline on how to define the boundary of the contiguous area in which residents may access healthcare resources; these boundaries may be arbitrarily close or far from the area of consideration. Even if one assumes that the adjacent area is a "rational area" too, there may still be multiple ways to draw its boundaries.

Moreover, although it is inevitable to pick a ratio cutoff as the starting point (in this case, 3,500 or 3,000), the rule requires the contiguous area to have a population-to-physician ratio over 2,000:1 to indicate "that practitioners in the contiguous area cannot be expected to help alleviate the shortage situation in the area being considered for designation" (HHS, 2025), which appears ad hoc. It does not consider the relative market size of the area under consideration and the contiguous areas, distance and travel time, quality of services, or other important factors. In addition, the policy lacks credibility for evaluating whether providers in nearby areas can indeed "help alleviate the shortage situation." A more detailed discussion of these rules is in Appendix B.

As further discussed in Section 5, the HPSA designation inevitably leads to significant ambiguity especially when markets are defined based on areas with fixed boundaries: if the area is large, the heterogeneity across small units inside the area is ignored; if the area is small, it is likely that people inside the area also seek care in nearby areas, which, as seen in the HPSA case, complicates the criteria design. This challenge motivates the new framework we propose in this paper.

3.3 Current Measures on Primary Care Physician Shortage

After recasting all Geographic Primary Care HPSAs to ZIP Code Tabulation Areas (ZCTAs)[7], we estimate that Geographic Primary Care HPSAs cover 32.2 million people in 2019. To find the number of additional physicians needed to resolve the shortage, the widely cited measures (e.g., 14,000 from KFF[8]) compare the status quo to the target ratio (3,500:1 or 3,000:1), and then sum up the number of physicians needed to meet the target in all places.

This measure mechanically overestimates the physicians needed to resolve shortages, even if we assume all shortage areas are properly defined. Figure 1(A) shows the areas designated as Geographic Primary Care HPSAs in 2019. Many designated areas are adjacent to or near one another, so adding a physician to one place should relieve the shortage in nearby areas. However, the current measurement framework does not allow for such interactions, since each HPSA is considered in isolation and available statistics do not capture patient flow between areas. This limitation is another key motivation for our framework.

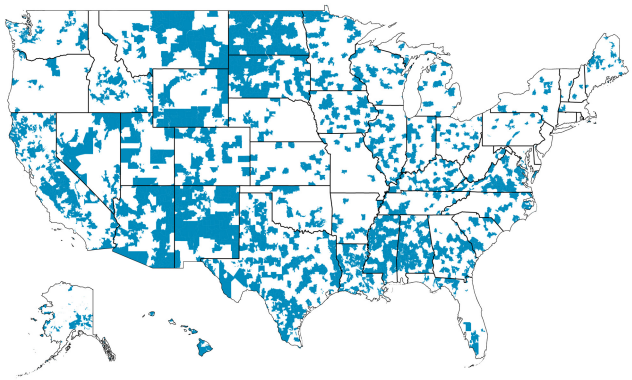
As a preview of our results, Figure 1(B) compares the Geographic Primary Care HPSAs with our shortage measure, which is described in later sections. The two measures disagree in many places: pink areas are places where we measure a shortage but are not designated as a Geographic Primary Care HPSA (5.6 million population), yellow areas are Geographic Primary Care HPSAs that we estimate not to have a shortage (26.3 million population), and blue areas are where the two measures agree on shortage designations (5.9 million population). Many blue areas are next to or inside yellow areas, suggesting that while some HPSAs do cover areas in shortage, the designation could be more precisely targeted. We also see notable variation across states in the pink areas, suggesting differences in how state Primary Care Offices survey and apply for HPSAs.

[7] We regard a ZCTA as designated if and only if more than 50% of its residential population lives in a designated area — see Section 4 for more details.

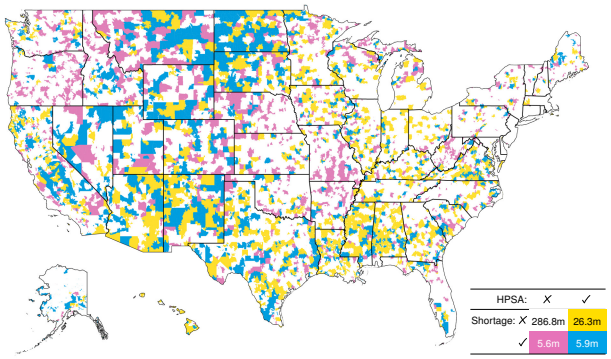
[8] https://www.kff.org/other/state-indicator/primary-care-health-professional-shortage-areas-hpsas. Accessed in May 2025 via Wayback Machine for record as of September 2019.

Figure 1: Distribution of Geographic Primary Care HPSAs and Shortages, 2019

(A) Geographic PC HPSA



(B) Our Shortage Designation vs. Geographic PC HPSA



Notes: Panel (A) shows, in blue, all ZCTAs that are designated as a Geographic Primary Care HPSA. For a partially covered ZCTA, it is regarded as designated if more than 50% of its residential population lives in an HPSA. Panel (B) shows the comparison between our shortage measure with Geographic Primary Care HPSAs at the ZCTA level. Our measurement framework is described in Section 5.

4 Data

Flows of Primary Care Visits To estimate patient preferences for primary care and to predict visit flows, we use a random 20% sample of Medicare fee-for-service beneficiaries between 2015 – 2019. We restrict attention to enrollee-years older than 65 and with a full year of Medicare Part A and Part B coverage, unless it is the enrollee's year of death. We observe patients' 5-digit ZIP codes of residence each year associated with their mailing addresses. The baseline sample has 7.41 million unique enrollees and 27.12 million unique enrollee-years.

From Medicare Part B professional claims, we observe the 5-digit ZIP code of the place of service.[9] Following Barnett et al. (2021), we define a primary care visit as a claim with associated procedure codes for initial visits, follow-up visits, and annual physical exams performed by a primary care physician. After crosswalking ZIP codes to ZCTAs for each calendar year, we calculate the number of patients from each ZCTA of residence who have a primary care visit in each ZCTA of service. If a patient visits multiple ZCTAs for service in a given year, we split the patient proportionally to her share of the number of visits.

[9]This variable is the 9-digit ZIP code associated with the place of service address. It is not available for professional claims before calendar year 2015.

Population We obtain the total population in each ZCTA from the American Community Survey (ACS). For each focus year, we use the corresponding 5-year ACS estimate with that year as the terminal survey year — for example, the 2015 – 2019 ACS estimate for populations in 2019. The Medicare sample is only used to estimate coefficients in the demand flow model; all of our shortage measures are based on the total population from ACS.

Primary Care Physician (PCP) Sample To calculate the number of primary care physicians (PCPs) at each location, since our Medicare FFS claim-level dataset only contains a 20% beneficiary sample, we supplement it with the Medicare Physician & Other Practitioners by Provider dataset from the Center for Medicare and Medicaid Services (CMS). This is a public dataset derived from 100% of Medicare enrollment and FFS claims data, and contains the use, payments, submitted charges, beneficiary demographics, and physician characteristics aggregated at the physician-year level for the universe of physicians who provided services to Medicare FFS beneficiaries.

Following the literature (e.g., Khoury et al., 2025) and PCP definition used in HPSA designations, we define PCPs as physicians with specialties of internal medicine, family practice, general practice, geriatric medicine, pediatric medicine, and obstetrics and gynecology,[10] from the Medicare Physician & Other Practitioners by Provider datasets. Our final analysis sample has over 275,000 PCPs.

Our primary care physician sample does not include those who do not provide care to Medicare patients (i.e., over age 65), especially pediatricians — in 2019, there were estimated to be nearly 30,000 pediatricians in the US.[11] Thus, we consider fewer physicians than HPSA does. Our lower shortage population estimate indicates more severe overestimates by methods using HPSA. Meanwhile, since the majority of family care physicians also see children (Jetty et al., 2021), our measure better reflects the primary care access of the population.

PCP Locations The Medicare Physician & Other Practitioners by Provider dataset only contains the ZIP code of each PCP's primary business address registered on the National Plan and Provider Enumeration System (NPPES). However, previous studies (such as DesRoches et al., 2015) have documented that the addresses in NPPES may not reflect physicians' actual places of service, due to the lag of updates, registration under group practices, or other logistical hurdles. Therefore, for physicians

[10] These specialty are the ones eligible for Medicare's HPSA Bonus Payment Program.

[11] U.S. Bureau of Labor Statistics. https://www.bls.gov/oes/2019/may/oes291221.htm. Last accessed on May 27, 2025.

who we observe in the Medicare FFS claims data, we instead use the ZIP of their places of service observed in the claims. 81% of PCP-years in the Medicare Physician & Other Practitioners by Provider dataset can be observed in our 20% Medicare FFS sample from 2015 to 2019. We allow PCPs to practice in multiple places in the same year, by splitting them into ZIPs proportional to the number of claims they furnished in each ZIP. For the remaining 19% PCP-years that we do not observe in the Medicare claims sample, we keep using the ZIP of their place of service addresses registered in NPPES.

ZCTA-level Travel Time and Distance We use the ZCTA-level travel time dataset shared by Hu et al. (2020), which contains travel between all ZCTA pairs within 150 miles, measured via ArcGIS Network Analyst. For ZCTA pairs that are more than 30 minutes apart, we use distance to approximate patients' travel costs. Here, we use ZCTA centroid coordinates provided by the National Bureau of Economic Research (NBER).

HPSA Designation We obtain the historical record of HPSA designations via HRSA. This file contains all the HPSAs that were once designated, with their dates of designation and withdrawal, HPSA types (e.g., Geographical, Population), geographic unit types (e.g., single county, sub-country, census track), and the identifier of the corresponding type. The HPSA designation needs to be reviewed every three to five years, and the dataset contains the most recent update on variables such as designated population, until 2024. Since HPSA designation can be based on different geographic unit types and some ZIPs are partially designated, we convert all designations to the ZIP level (and later crosswalk to ZCTA level) using crosswalks of ZIP-census tract, ZIP-county, and ZIP-sub county, from the United States Department of Housing and Urban Development. These crosswalks all contain the population shares for one-to-multiple ZIP matches, which we use to calculate the population share in a ZIP being designated as HPSA. In the main analysis, we regard a ZIP as designated if and only if more than 50% of its residential population is designated. We also supplement this data with the HPSA Auto-ZIP data from CMS, which includes the list of ZIPs that are entirely covered by Geographic Primary Care HPSAs in each year. CMS produces these files for the distribution of the Medicare 10% HPSA bonus to PCPs. We let the HPSA Auto-ZIP file overwrite the Geographic Primary Care HPSAs identified via the designation file above, since the Auto-ZIP file is not subject to errors in population shares from the crosswalks and therefore is more accurate.

For more detailed descriptions of the variable construction and data sources, see Appendix A.

5 Research Design

5.1 Measurement of Primary Care Shortage

In this section, we formally introduce our framework for measuring the shortage of primary care physicians. Based on the discussion in Section 3, we keep our measure's key principles as close as possible to those of Geographic Primary Care HPSAs, while aiming to develop a more systematic and conceptually consistent procedure.

Our approach is built on three key general principles used in the definition of Geographic HPSAs. For each place j:

1. We aim to measure j's shortage level in a predetermined region around j, A_j, e.g., the region within 30 minutes of travel time from j;

2. We measure the shortage by a population-to-physician ratio, r_j;

3. We set a predetermined threshold $\bar{\eta}_j$ (e.g., 3,500 or 3,000), for the population-to-physician ratio to define shortage — j is in shortage if and only if $r_j > \bar{\eta}_j$;

These principles define the market for each place's patients, measure supply and demand in this market, and designate shortages relative to a pre-specified threshold of demand-to-supply ratio.

However, these three principles alone are insufficient for a consistently defined measurement procedure. Specifically, they are ambiguous about which population and physician counts should be used for calculating r_j — those within location j, within region A_j, or others — and what to do when two areas' corresponding regions, A_j and $A_{j'}$, overlap.

As discussed in Section 3, to handle these issues, HPSA introduces additional, somewhat arbitrary requirements on j's "contiguous" areas — such as a population-to-physician ratio larger than 2,000:1 — to indicate when nearby areas cannot "help alleviate the shortage" in j (HHS, 2025). First, there is no clear guidance on how to define "contiguous areas", which leads to ambiguity in measuring the shortage. Moreover, although it is inevitable to choose a shortage threshold (e.g., 3,500:1 or 3,000:1 used in HPSA) as the starting point, the choice of 2,000:1 is relatively arbitrary. Our framework resolves these ambiguities by providing clear and consistent definitions of measurement principles and objectives, and by directly modeling the visit flow across areas.

Definition 1 (Catchment Area)**.** *For any geographic region of interest U (e.g., the entire U.S.), let $\mathcal{P}$ be a given partition of U (e.g., the set of all ZCTAs). For any geographic unit $j \in \mathcal{P}$, its catchment area, A_j, is a subset of $\mathcal{P}$ containing all units in $\mathcal{P}$ that are within 30 minutes of travel to j:*

$$A_j := \{k \mid k \in \mathcal{P},\ t(j,k) \leq 30\ min.\}$$

where $t(j,k)$ is the shortest travel time between area j and k.

The catchment areas are defined by a pre-specified travel time threshold for measuring PCP coverage, set by the policymaker. Given this threshold, the geographic space can be represented as a travel network. Formally, let all geographic units $j \in \mathcal{P}$ form a finite, undirected graph $G = (\mathcal{P}, E)$, where each j corresponds to a node. Two nodes are connected if and only if they are within 30 minutes of travel time from each other; that is, $\forall\ j, j' \in \mathcal{P},\ (j,j') \in E \iff j' \in A_j$. This network representation captures the interconnection among areas and provides useful insights for policy design, as illustrated in the simulation exercises in Section 7.

The term "catchment area" is adopted from the transportation geography literature (e.g., Peng, 1997; Wang, 2000; Luo, 2004; Neng Wan and Sternberg, 2012) that uses similar concepts to study the geographic coverage of resources. As discussed in Section 2, we improve on these methods with an economic model of visit flows that allows for heterogeneity in patient choice sets and preferences. This approach allows us to study the equilibrium effect of adding and reallocating physicians in counterfactual analyses.

When calculating the travel time $t(j,k)$, we abstract each element ("geographic unit") of $\mathcal{P}$ to a single "representative point" and measure the shortest travel time between these points. In empirical applications, each geographic unit in $\mathcal{P}$ should ideally be small and internally connected, such that a point can represent the behavior of the whole population in the unit. In this paper, we set our geographic units to be ZCTAs, the finest location measure of patients in the Medicare claims data.[12]

We complete our measurement framework with the following definitions, and then provide a simplified example to illustrate these concepts.

Definition 2 (Catchment Demand under Full Capacity)**.** *Let D_j be the population living in the geographic*

[12]The within-ZCTA travel time estimated by Hu et al. (2020) has a median of 7 minutes (IQR 3.6 – 11), although we do not use the within travel time in our estimation, due to its tendency to overestimate the between ZCTA travel times. See Appendix A for more details.

unit $j \in \mathcal{P}$. The "catchment demand under full capacity", $\widetilde{D}_{A_j}$, of any catchment area A_j, is the total number of people seeking primary care in A_j, assuming all people from k, D_k, seek primary care at the same time and do so only in A_k for all $k \in \mathcal{P}$.

Definition 3 (Primary Care Physician Coverage Rate)**.** *Let S_j be the number of primary care physicians in geographic unit $j \in \mathcal{P}$. For any $j \in \mathcal{P}$, the Primary Care Physician Coverage Rate, r_j, is*

$$r_j := \frac{\widetilde{D}_{A_j}}{\sum_{k \in A_j} S_k}.$$

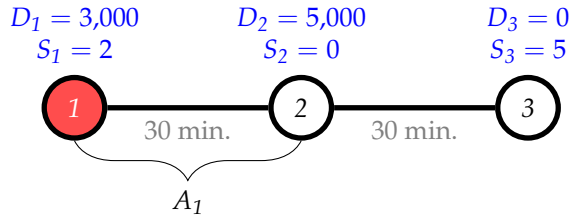
Definition 4 (Primary Care Physician Shortage)**.** *Given any shortage threshold, $\bar{\eta}$ (e.g., 3,500 or 3,000), a geographic unit $j \in \mathcal{P}$ is in primary care physician shortage if and only if*

$$r_j > \bar{\eta}.$$

Definition 2 to Definition 4 explicitly define the shortage measure as a counterfactual analysis: What would be the demand-to-physician ratio if everyone sought care at the same time, and only did so within their own A_j? These definitions formalize the implicit premise underlying the idea of using population-to-physician ratio in the traditional measure — namely, calculating the average physician workload (potential patients per physician) as if the whole population seeks care at the same time within the area.

The main idea of our shortage measure is simple and intuitive: to measure the shortage of people in area j, we account for both the number of physicians in all the places they can access, and, in these places, the total number of people who share the same physician resources. The next step is to calculate catchment demand under full capacity, $\widetilde{D}_{A_j}$, which requires predicting visit flows between many pairs of geographic units. Previous literature in transportation geography (e.g., Luo and Wang, 2003; Luo, 2004; Neng Wan and Sternberg, 2012) relies on simple decay functions with calibrated coefficients, or uses distance or physician counts in simple functional forms. In contrast, we estimate a demand model of primary care services with rich heterogeneity using national Medicare claims and travel time data, as described in Section 5.2. This model allows us to consider not only travel costs and physician counts, but also unobserved local primary care service attractiveness, real travel networks and constraints, and variation in primary care choice sets, which are crucial to measuring the inflow and outflow of nearby regions and delivering a reliable coverage measure.

Figure 2: An Illustration of Our Shortage Calculation



$$\widetilde{D}_{A_1} = D_{1\to1} + D_{1\to2} \quad (= D_1 = 3{,}000)$$
$$+ D_{2\to1} + D_{2\to2} \quad (= D_2 - D_{2\to3})$$
$$+ D_{3\to2} \quad (= 0)$$
$$r_1 = \frac{\widetilde{D}_{A_1}}{S_1 + S_2} \quad \left(= \frac{3{,}000 + 5{,}000 - D_{2\to3}}{2+0}\right)$$

Area 1 is in shortage (i.e., $r_1 > 3{,}500$) if and only if

$$D_{2\to3} < 1{,}000$$

We present an illustrative example of our measurement framework in Figure 2. Consider three ZCTAs, labeled *1*, *2*, and *3*, positioned on a straight line such that {*1*,*2*} and {*2*,*3*} are each 30 minutes apart. Let D_j denote the population and S_j is the number of primary care physicians in each ZCTA $j \in \{1,2,3\}$. We focus on evaluating whether *1* is in shortage. In this example, *1*'s catchment area is $A_1 = \{1,2\}$, and the calculation of *1*'s coverage rate r_1 is shown in the figure.

Since *3* is accessible to residents in *2*, the shortage status of *1* depends on the number of people from *2* seeking primary care in *3*, $D_{2\to3}$. For example, if $D_{2\to3} = 0$, then $r_1 = \frac{3000+5000}{2} > 3{,}500$ and *1* is in shortage; *1* is in shortage if and only if $D_{2\to3} < 1{,}000$. Intuitively, whether people in *1* live in a shortage area depends on how many people from *2* and *3* share physicians with them in *1* or *2*. In this simplified example, since *3* has no population, only the flow from *2* to *3* matters. If *3* had a positive population, we would also need to account for people from *3* seeking care in *2*, since these people would directly share physicians with people in *1*.

Instead, if one used areas with closed boundaries to define markets and shortages, the shortage designation would entirely depend on the arbitrarily-drawn boundaries that partition the space. If we regard *1* as its own area, its coverage rate is $\frac{3000}{2} < 3500$, so *1* is not in shortage; meanwhile, we fail to consider the fact that people in *2* may also seek care in *1*. If we combine *1* and *2* together as one area, then the population-to-physician ratio of the whole area is 4000, which is qualified as in shortage; however, we then fail to account for patients in *2* also seeking care from physicians in *3*. Our new measurement framework, therefore, resolves this ambiguity and provides a consistently defined measure for any interconnected units.

This example highlights the advantage of our framework and the importance of predicting visit

flows between ZCTAs in our framework. The next section describes our methodology for modeling these flows.

5.2 Model of Primary Care Demand

Following Dingel et al. (2023), we build a gravity model of primary care demand to recover the revealed-preference measure of patients' travel costs and the attractiveness for each location in each year. The model is estimated using the observed flow of primary care visits between ZCTAs for sampled Medicare fee-for-service beneficiaries. We then combine model estimates and population data from the national census to evaluate PCP shortages. Throughout this section, we use "patients" to denote all sample beneficiaries, including those who did not seek care.

Let d ("destination") index the location of providers, o ("origin") index the location of beneficiary k, and t index calendar year. Patient k's indirect utility from getting primary care in location d in year t is

$$u_{dkt} = \log \delta_{dt} + \log \rho_{o(k)d} + \varepsilon_{dkt}$$

where δ_{dt} is the attractiveness of location d at time t, $\rho_{o(k)d}$ is the travel cost between o and d, and ε_{dkt} is an i.i.d. Type-I Extreme Value shock. The outside option $d = 0$ is not getting care, which we normalize to $\delta_{0t} = 1$ for all years. Patients have homogeneous outside options and agree on the mean attractiveness of each location — that means we do not model patient heterogeneity in health and patient-provider matches.[13] The choice probability of patients living in o choosing care at d is then

$$s_{odt} = \frac{\exp(\log \delta_{dt} + \log \rho_{o(k)d})}{\sum_{d' \in \mathcal{C}_o \cup 0} \exp(\log \delta_{d't} + \log \rho_{od'})},$$

where $\mathcal{C}_o$ is the set of potential destinations of people in o.

Let N_{ot} be the total number of patients residing in location o at time t, and Q_{odt} be the number of patients from o who got care in location d at time t. The primary care flow between locations is the share of N_{ot} choosing location d at t:

$$Q_{odt} = \delta_{dt}\rho_{od} \frac{N_{ot}}{\sum_{d' \in \mathcal{C}_o \cup 0} \delta_{d'}\rho_{od'}}.$$

[13] If patients also care not only about accessing any physician, but about specific matches, then our measure would differ from the real shortages — yet this idea is not incorporated in the HPSA either.

Taking the log of both sides of this equation, we get

$$\log Q_{odt} = \log \delta_{dt} + \log \rho_{od} + \log \theta_{ot}. \tag{1}$$

$\log \theta_{ot} = \log \left(\frac{N_{ot}}{\sum_{d' \cup 0} \delta_{d'} \rho_{od'}} \right)$ is an origin-year fixed effect and $\log \delta_{dt}$ is a destination-year fixed effect. The origin-year fixed effect, therefore, captures the inclusive value of origin ZCTA's choice set, or its overall accessibility in primary care; the destination-year fixed effect estimates the overall primary care attractiveness of the destination ZCTA. We estimate (1) using Poisson Pseudo-Maximum Likelihood (PPML; Silva and Tenreyro, 2006), which accommodates zero visit flows. We include a balanced panel of origindestination pairs that have any positive visit flows during our sample period. PPML allows us to retain yearpair observations with zero visits within these pairs. In estimation, we parameterize the travel cost as

$$\log \rho_{od} = \begin{cases} \gamma_0 & \text{if } o = d; \\ \gamma_1 \log[\text{travel time}_{od}] & \text{if travel time}_{od} \leq 30 \text{ mins}; \\ \gamma_2 \log[\text{distance}_{od}] & \text{if travel time}_{od} > 30 \text{ mins}. \end{cases} \tag{2}$$

We represent travel cost between two ZCTAs using the estimated drive time from Hu et al. (2020) if the driving takes less than 30 minutes, and using the great-circle distances for farther ZCTA pairs. The travel time estimates enable us to incorporate meaningful heterogeneity in the local road network and congestion; in Appendix Figure A.1, we show that the maximum distance patients can travel within 30 minutes in our sample indeed varies substantially across patient ZCTAs.

5.3 Factors Behind Primary Care Attractiveness

In this section, we model the local average attractiveness of primary care service, δ_{jt}, as being proportional to the number of primary care physicians and local healthcare amenities. This allows us to endogenize visit flows in the counterfactual simulation, where we change the physician distribution. In addition, $\widehat{\log \delta_{jt}}$ estimated from the visit flows may contain temporary demand shocks, which could bias our baseline shortage measure; the panel data and additional variables on local health amenities can help us separate these shocks.

For place j at time t, we model

$$\widehat{\delta_{jt}} = \widetilde{\delta_{jt}}\xi_{jt}$$
$$\widetilde{\delta_{jt}} = A_{jt}S_{jt}^{\alpha},$$

where $\widetilde{\delta_{jt}}$ is the baseline primary care attractiveness of place j at time t, and ξ_{jt} an i.i.d. log-normal shock with mean 1; A_{jt} is the local primary care amenity level, and S_{jt} is the number of primary care physicians.

We separate ξ_{jt} and estimate α using the following regression

$$\widehat{\log \delta_{jt}} = \alpha \log S_{jt} + g(X_{jt}) + \log \xi_{jt}. \tag{3}$$

In the absence of good instrumental variables that exogenously shift S_{jt},[14] we use a flexible function $g(\cdot)$ of observables X_{jt}, which contains ZCTA fixed effects, year fixed effects, and observed time-varying ZCTA characteristics such as the number of hospitals, median income, and the age distribution, to approximate the local healthcare amenity A_{jt} that is invariant to the number of physicians. Omitted variables may still bias our estimate of α, the coefficient on PCP counts S_{jt}, in either direction; it is reassuring that the estimate remains robust after flexibly accounting for local characteristics in different specifications, as discussed in Section 6.

We then have

$$\widehat{\log A_{jt}} = \widehat{g}(X_{jt})$$
$$\log \widetilde{\delta_{jt}} = \widehat{\alpha} \log S_{jt} + \widehat{g}(X_{jt}). \tag{4}$$

In the results and counterfactual simulations, we use $\log \widetilde{\delta_{jt}}$, instead of $\widehat{\log \delta_{jt}}$, to predict the expected visit flows. Some ZCTAs do not have estimated $\widehat{\log \delta_{jt}}$ because we observe neither providers nor visits in the Medicare claims data for those ZCTAs. We expect these places to have low primary care amenity levels, and therefore fill in their $\widehat{\log A_{jt}}$ with the minimum value estimated in the ZCTA's state in each year.

[14] Theoretically, the Primary Care HPSA could potentially be a good instrumental variable. However, as we discussed in Section 3, the policy has limited effect, so it easily fails the test for relevance.

6 Results

This section discusses two sets of results. We first show estimates of model parameters. We then report our measure of primary care shortage, and compare it with HPSA's designations.

6.1 Model Estimates

Table 1 shows the coefficients on travel costs from the demand models. All else equal, the travel time (log minutes) elasticity of primary care demand is -0.840 for areas within 30 minutes of travel, and the travel distance (log miles) elasticity is -1.609 for areas over 30 minutes of travel; both are precisely estimated. This result is similar in magnitude to Dingel et al. (2023) — their estimate of distance elasticity for specialist care between Hospital Referral Regions is -1.689, using Medicare claims data from 2017. The coefficient on seeking care in the patient's own ZCTA is negative since it is normalized to the outside option of not seeking care, for which the travel time is set to zero.

In addition, we estimate the revealed-preference measure of average attractiveness for primary care for all ZCTAs with at least one primary care physician in the years 2015 to 2019. Panel (A) of Figure 3 shows the distribution of estimated log attractiveness for 2019, with considerable variation. As shown in Panel (B) of Figure 3, estimated primary care attractiveness is positively associated with the number of PCPs in a ZCTA, and in Panel (C) with the number of acute care hospitals, an important indicator of local healthcare amenities.

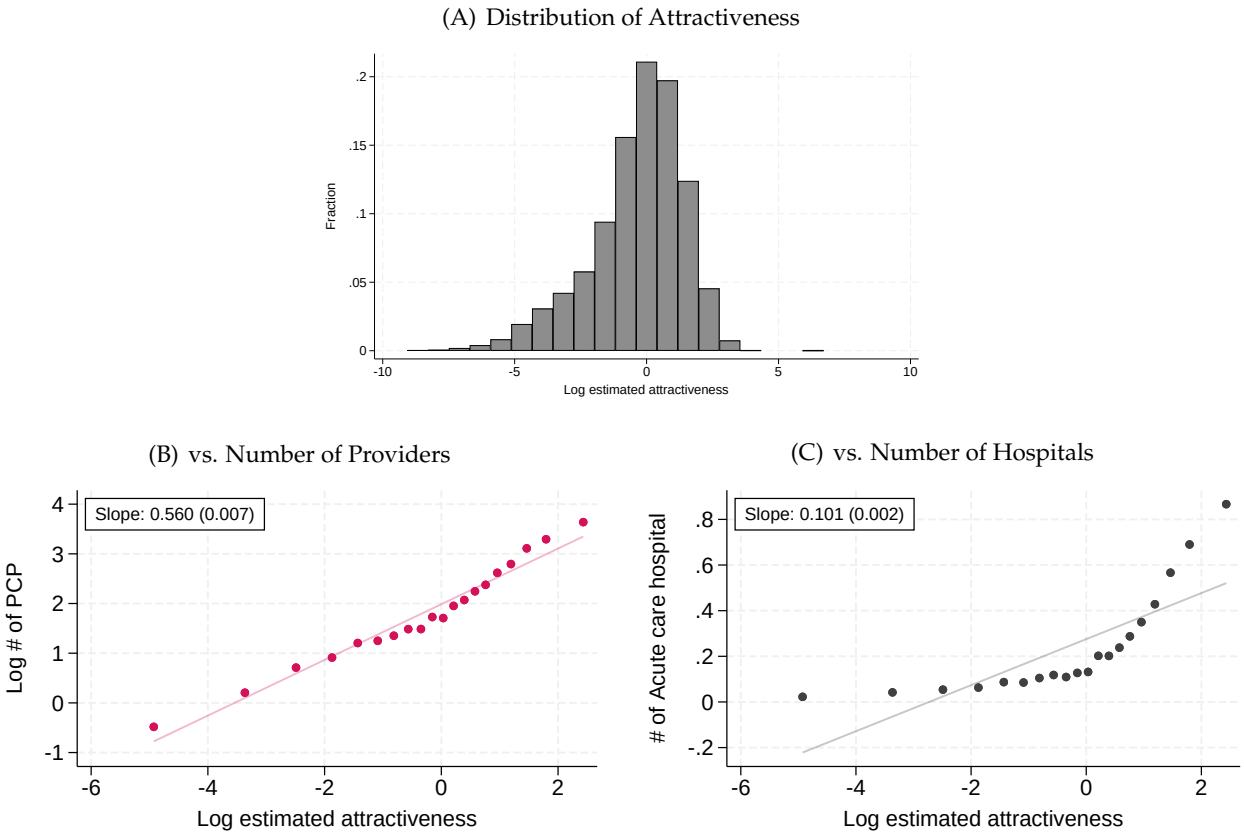
Table 2 shows the estimated coefficients in equation (4), a parameterized model projecting estimated attractiveness onto health amenity levels and the number of primary care physicians. Across specifications, the elasticity between PCP count and a location's primary care attractiveness is close to 0.7. Different specifications vary $g(X_{jt})$, the control variables included in this regression. In column (1), we only control for the number of acute care hospitals, children's hospitals, and children's acute care hospitals in a ZCTA. In column (2), we additionally include year fixed effects. Columns (3) and (4) use the post-double selection method to flexibly control various ZCTA characteristics (Belloni et al., 2014; Ahrens et al., 2018); a full list of ZCTA characteristics is in Appendix Table A.1.

Table 1: Parameter Estimates: Patient Travel Costs

	Dep. var: PC flow btwn. ZCTAs	
Own ZCTA	-0.944	(0.00887)
Log distance (if $>$ 30 min)	-1.609	(0.00238)
Log time (if $\leq$ 30 min)	-0.840	(0.00244)
# Patient ZCTAs	32,507	
# Provider ZCTAs	14,427	
Obs.	9,028,665	

Notes: Table shows estimates of the gravity equation (1) under the parameterization of travel costs (2). We use the Poisson pseudo-maximum likelihood estimator (Silva and Tenreyro, 2006), estimated using the software package by Correia et al. (2020).

Figure 3: Estimated Primary Care Attractiveness (2019)



Notes: Panel (A) plots the histogram of log estimated attractiveness parameters for primary care from (1). Panel (B) and (C) are binned scatterplots of log estimated attractiveness against log number of providers, and log number of acute care hospitals, respectively. The slope for each line of best fit and its standard error from a regression in the data prior to binning are shown in the figure. We restrict attention to 2019 in all three panels.

Table 2: Parameter Estimates: Attractiveness

	(1)	(2)	(3)	(4)
Log PCP count (α)	0.746	0.735	0.736	0.731
	(0.0159)	(0.0159)	(0.0160)	(0.0160)
Year FE		✓		✓
No. of ACH, children, CAH	✓	✓		
PDS (Belloni et al. 2014), local chars.			✓	✓
Obs.	62,465	62,465	62,465	62,465

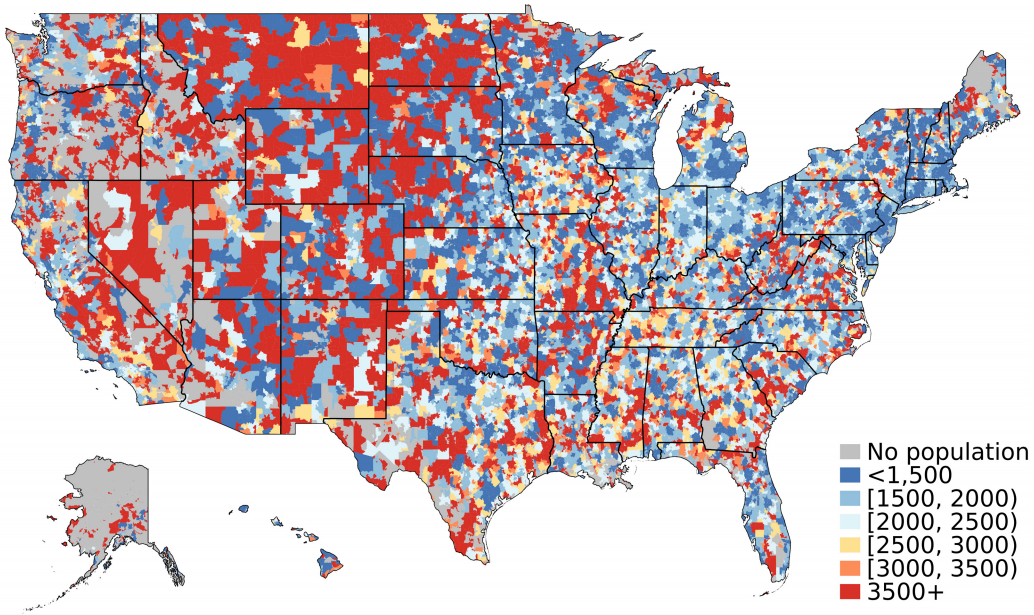
Notes: Table shows the estimated coefficient on log PCP count for equation (3). ACH: acute care hospitals. CAH: Children's acute care hospitals. PDS: post double-selection à la Belloni et al. (2014). Local characteristics variables used in the procedure are listed in Appendix Table A.1.

6.2 The Shortage Landscape

To calculate the coverage ratio of all ZCTAs, we predict the *catchment demand under full capacity*, as defined in Definition 2, for all ZCTAs with a positive population. We use the estimated baseline attractiveness from Equation (4) and the travel cost estimates to predict visit flow $D_{o\to d}$, the potential number of patients going from each origin ZCTA o to destination ZCTA d. Based on Definition 2, there are two distinctions between $D_{o\to d}$ and the $Q_{o\to d}$ used in the estimation: first, the market size is the population residing in ZCTA o, instead of the number of Medicare beneficiaries; second, we remove the option of not seeking care, and restrict the choice set for population in ZCTA o to A_o, the set of all ZCTAs within 30 minutes of travel from o. We then calculate the coverage rate, which equals the catchment demand under full capacity divided by the total number of primary care physicians in the catchment area A_o, as in Definition 3.

Figure 4 visualizes the primary care physician coverage rates for all ZCTAs in the U.S. Areas with coverage rates greater than 3,500 are in red; these are shortage areas defined following the standard of HPSA designation. These areas cover 11.58 million people in total, or 3.57% of the U.S. population, which is 64% fewer than the 32.2 million covered by the Geographic Primary Care HPSAs.

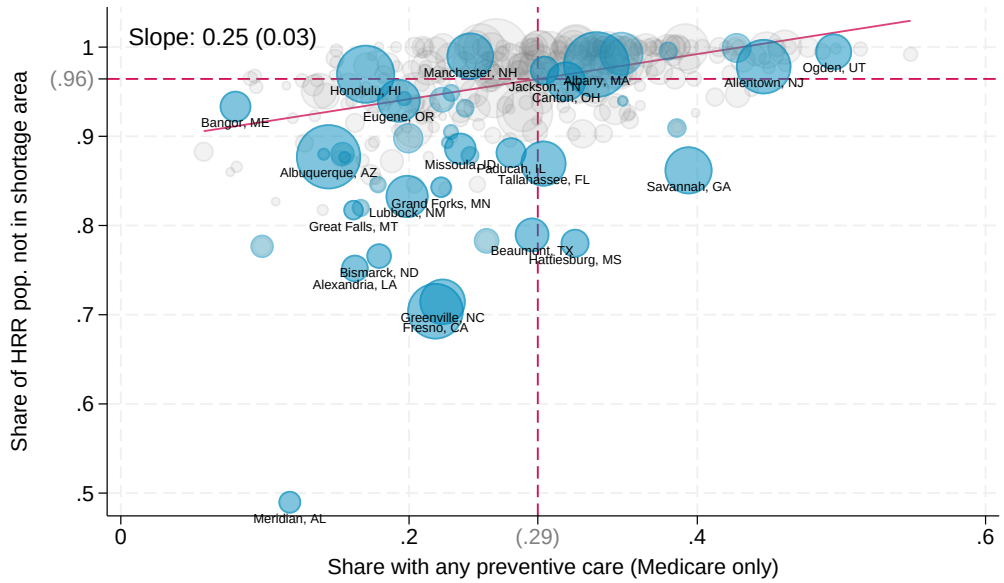
Figure 4: Geography of Primary Care Shortage Ratio, 2019



Notes: Total population in shortage areas (coverage rate 3,500+): 11.58 million (3.57% of national population). Figure plots the ZCTA-level primary care coverage rates, as defined in Section 5. ZCTA population data comes from the American Community Survey 2015 – 2019 5-year estimates.

Figure 5 investigates how our estimate of primary care shortage is associated with Medicare patients' preventive care use. To aid with interpretation, we conduct this analysis at the Hospital Referral Region (HRR) level, a common geographic unit used for analyzing tertiary care. For each HRR, we calculate the share of population that does not live in a shortage ZCTA, and compare it with the share of Medicare beneficiaries who used any preventive care. We find a positive association between the two (with the slope being 0.25, standard error 0.03). We also highlight, for each state, the HRR that has the largest share of population in primary care shortage, as defined by our framework. This uncovers the heterogeneity between states: for New England and Mid-Atlantic States, such as MA, NJ, ME, and NH, the HRRs with more shortage population are close to the national average of 96% living in non-shortage ZCTAs. The most severe shortage HRRs in mountain states, southern states, and geographically diverse states such as CA and TX, have over 20% of the population living in primary care shortage ZCTAs.

Figure 5: Estimated HRR-level Primary Care Coverage vs. Preventive Care Use



Notes: We calculate the share of population living in non-shortage ZCTAs (estimated by this paper) for each Hospital Referral Region (HRR) in 2019. We plot this variable against HRR's share of Medicare patients who had any preventive care visits. Highlighted in blue are HRRs that have the smallest y-variable of its state. For expositional clarity, not all highlighted HRRs are labelled with names.

As mentioned in Section 5, the geographic unit used in this framework should ideally be small. We use ZCTA since it is the smallest unit for which we can still obtain rich enough information for estimation, and we believe results generated at this level can provide useful lessons.

In our measure, we use the threshold of 3,500:1, the standard target for most HPSA designations. HPSA designations use the 3,000 threshold for areas with "high needs for primary care services or insufficient capacity of existing primary care providers". In our framework, such patterns are reflected in the visit flows data. Therefore, related local characteristics are embedded into estimates of the primary care demand model, and are accounted for in the shortage measurement and counterfactual analysis. Still, in Table A.2, we reevaluate the shortage using the 3,000:1 threshold as a robustness check, and find 18.28 million population living in shortage areas. Even with this more lenient cutoff, the estimate is still significantly lower than the 32.2 million population covered by Geographic HPSAs.

Table A.2 also presents other robustness checks that vary the travel time threshold and travel elasticities, demonstrating that HPSAs over-designate. As we decrease the travel time radius and

require patients to have access to closer PCPs, the estimated overall shortage naturally increases. However, even if we permit traveling to PCPs within only 25 minutes, the estimated total shortage population of 21.62 million is still lower than the current Geographic HPSA measure, which uses the 30-minute cutoff.

Moreover, the shortage measure exhibits limited changes even when we increase or decrease the travel cost for all ZCTAs by 50%, because the 30-minute travel time cutoff is already restrictive enough that, within this radius, the travel elasticity does not significantly affect where people seek care. These findings suggest that our measure provides a reliable estimate of the national PCP shortage, even if the travel cost elasticity estimated from the Medicare population may differ somewhat from that of the general population.

In the bottom panel of Table A.2, we reduce travel costs only for residents of shortage areas, simulating a simple travel subsidy policy, while keeping their travel network as-is. The changes to the shortage population are minimal, even when travel costs are reduced to just 10% of their current level. This suggests that shortage areas lack providers within the pre-set 30-minute travel network radius, which cannot be solved through encouraging travel within this radius. Rather, alleviating shortages requires more direct interventions: either increasing the supply of physicians near shortage areas — as we explore in the next section — or steering patients in non-shortage areas away from the physicians accessible to shortage areas.

6.3 Comparison to HPSA Designations

We benchmark our shortage measure with the current government shortage designation, the Geographic Primary Care Health Professional Shortage Areas. To do so, we take a snapshot of HPSA designations in 2019 and compare them with the ZCTA-level primary care coverage rates we calculated. We have previewed the differences in shortage classification in Figure 1, in which we see considerable disagreement between the two.

To investigate underlying factors that drive these disagreements, we regress an indicator for a ZCTA being in a Geographic HPSA on a vector of ZCTA characteristics, controlling for indicators for each decile of the ZCTA's coverage rate. In the regression, we include Z-scores of the variables, constructed by re-centering the variable around its mean, divided by its standard deviation. The coefficient on the Z-score can be interpreted as, for ZCTAs in the same coverage rate decile that we estimated, the association between a standard deviation increase of the variable of interest and the

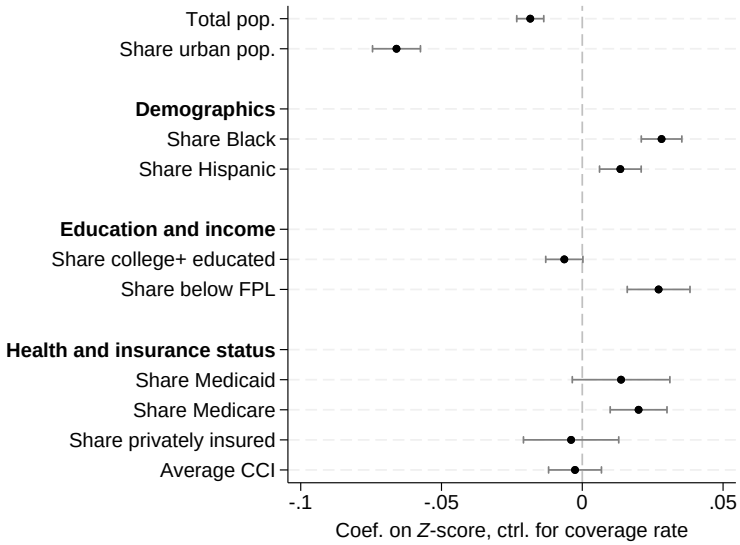
probability of it being designated as a Geographic Primary Care HPSA, holding values of all other variables fixed. The normalization to Z-scores makes coefficients comparable across variables.

Figure 6 plots these coefficients. Most notably, HPSA designations prioritize areas with low shares of urban population, high shares of racial and ethnic minority residents, and high shares of residents living in poverty or with Medicare, even conditional on our estimated coverage rate.

We provide two potential mechanisms. First, this can reflect the disadvantage of a fixed-boundary approach used by HPSA designations, where a population-to-physician ratio is calculated within the boundary of the area, and cross-boundary visit flows are ignored. The fixed-boundary approach tends to understate the shortage for areas with a large inflow of patients seeking primary care, which are likely to be urban areas with more providers. This can be a reason why urban areas and areas with fewer minority populations are less likely to be designated.

Second, these results may reflect some logistical complexities in the HPSA designation faced by state primary care offices (PCOs). As mentioned in Section 3, an HPSA designation is often initiated by the state PCO, which first needs to survey for potential shortage areas in the state, and then tries to justify the shortage using data available in the online application. Results from Figure 6 indicate that HPSAs tend to designate areas that are "stereotypically" in shortage, which reflects that state PCOs may sometimes need to form a prior on the shortage landscape before deciding which areas to survey, resulting in certain bias against areas that do not normally fit the prior.

Figure 6: HPSA Designations Priorities



Notes: We run an OLS regression of an indicator for a ZCTA being a geographic HPSA on a set of each ZCTA characteristic listed on the left of the figure, conditional on deciles of the coverage rate defined in section 5. In the regression, we include Z-scores of the variables, constructed by re-centering the variable around its mean, divided by its standard deviation. Figure plots the coefficient on each Z-score variable. CCI: Charlson Comorbidity Index, a score ranging from 0-37 that captures a patient's mortality risk based on age and chronic conditions. We take the ZCTA-level average of it among Medicare beneficiaries included in our sample.

Compared to HPSA designations, our measure also better predicts primary care access observed in the data. At the ZCTA level, it explains more variation (higher R^2) in the share of beneficiaries with at least one primary care visit within 30 minutes — conditional on having any primary care visits — than the HPSA designations (Figure A.2).

7 Simulation of Allocation Policies

The new framework and the model estimates allow us to simulate different policies for adding physicians and compare their effectiveness. Deriving the optimal plan that resolves all shortages with the fewest number of physicians is an *NP*-hard problem.[15] Therefore, in this section, we simulate adding physicians to places based on various heuristics, and compare their efficiencies in resolving shortages.

[15]Similar to the *p-median* problem in Auerbach and Dix (2019). Our problem is reduced to the classic *NP*-hard dominating set problem or set cover problem (Karp, 2009) when the shortage cutoff becomes infinity, i.e., an area is defined not in shortage as long as there is one doctor within 30 minutes.

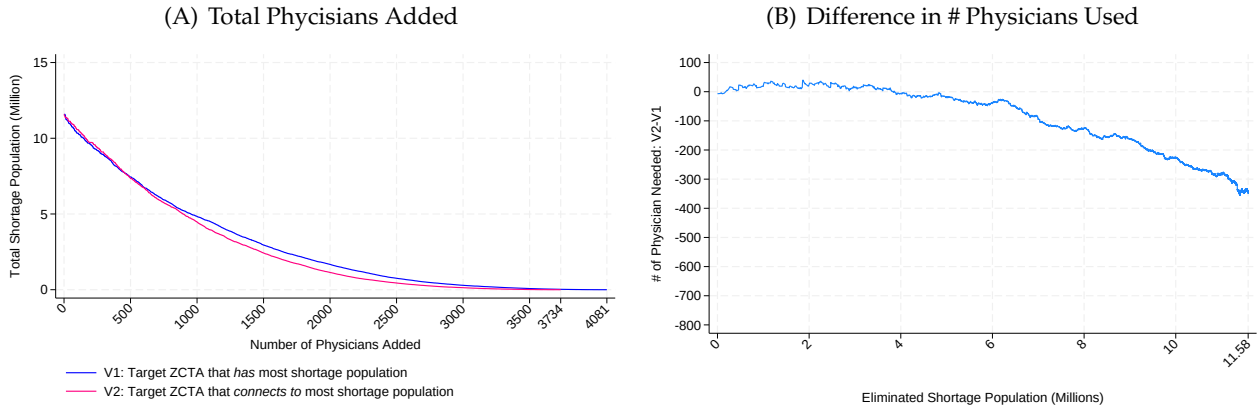
As mentioned in Section 5, we represent the geographic space as a travel network, where each ZCTA is a node and two nodes are connected if and only if they are within a 30-minute travel time. This network informs us how markets overlap with each other. We can then simulate different algorithms while holding the structure of this travel network fixed.

In each iteration, we add one primary care physician to a ZCTA. We test two different heuristics for choosing the target ZCTA:

- $V1$: Add the physician to the ZCTA that *has* the most population in shortage;
- $V2$: Add the physician to the ZCTA that *connects to* the most population in shortage;

Heuristic $V1$ assumes the policymaker does not have any knowledge of the travel network structure, and aims to maximize the static return on adding the additional physician only using the "node-by-node" information. Heuristic $V2$ leverages the simplest local network structure, incorporating information from adjacent nodes. This policymaker still maximizes the static return of the additional physician, but considers the population being covered in adjacent areas. Mechanically, $V1$ only adds physicians to areas that are labeled as a shortage area, which is similar to the current policies that only provide eligibility for government programs to designated HPSAs. $V2$, however, can potentially add physicians to non-shortage ZCTAs when some of its adjacent ZCTAs are in shortage.

Figure 7: Comparison of Heuristics for Adding Physicians

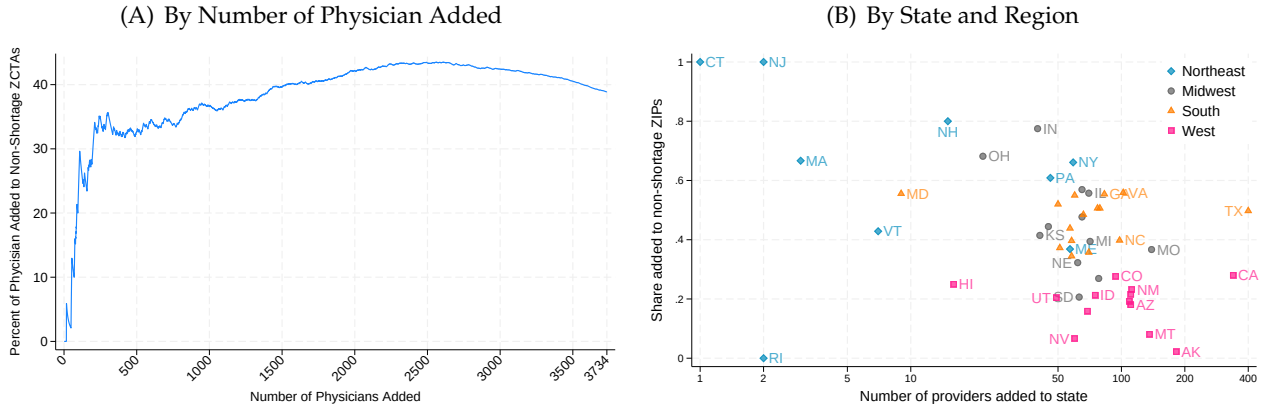


Notes: Panel (A) plots the shortage population remaining after cumulatively adding the number of PCPs shown on the x-axis. Panel (B) compares the effectiveness of the two heuristics. It shows the difference in the number of PCPs used, after eliminating the number of shortage population shown on the x-axis.

Figure 7 shows the main simulation result. Panel (A) shows that $V1$ requires 4,081 additional

physicians to eliminate the primary care physician shortage in the U.S., and $V2$ only requires 3,734. The heuristic that leverages the simplest network structure ($V2$) uses 8.5% fewer physicians. Panel (B) calculates the difference in additional physicians used at each level of the resolved shortage population. $V2$ uses >100 fewer physicians, compared to $V1$, when eliminating 8 million shortage population in the U.S., and >300 fewer physicians when eliminating 11 million shortage population.

Figure 8: Share Added to Non-Shortage Areas



Notes: Panel (A) shows the cumulative percentage of PCPs added to non-shortage ZCTAs after adding the number of physicians shown on the x-axis. The y-axis of Panel (B) is the share of physicians added to non-shortage ZCTAs for each state; the x-axis is the total number of providers added to each state. States are color-coded by their Census Regions.

Adding a physician to a node in the travel network affects *all* connected markets in equilibrium. Therefore, the number of physicians needed to resolve the shortage depends on the targeting policy, and some knowledge of the network structure can provide meaningful improvement on the policy's efficiency. This efficiency gain mainly stems from the fact that $V2$ adds physicians to areas not in shortage but connected to multiple shortage areas. Figure 8 shows the share of physicians added to non-shortage areas. Panel (A) shows that $V2$ adds more than 30% of physicians into non-shortage areas when it adds the 500th physician, and by the end, when shortages are fully eliminated, 40% of the 3,734 additional physicians are added to ZCTAs with no shortage.

Our finding also highlights the importance of accounting for heterogeneity in factors not related to healthcare, such as road networks, in designing shortage policies. In Panel (B) of Figure 8, for each state, we plot the share of physicians added to shortage areas when all shortages are eliminated. We see that states from the West, although requiring a similar level of additional physicians as states in the South and the Midwest, added fewer physicians to ZCTAs with no shortage. One potential

explanation is that shortage areas in the West are more isolated and have poorer road connections to nearby areas. In this case, the only way to resolve shortages in those areas is to add physicians directly to them. Conversely, in states that are densely populated, flat, or have better road networks, such as MD, IN, and GA, the simulation adds more physicians to non-shortage areas that connect to more shortage population. The current HPSA designations use a simple correspondence between different terrain types to appropriate travel distances to account for heterogeneity in travel times: 20 miles for normal conditions with access to primary roads, 15 miles for mountains or areas with only secondary roads, and 25 miles for flat areas or areas connected by interstate highways (HHS, 2025). Our finding shows the importance and advantage of incorporating richer travel patterns using a data-driven approach.

Meanwhile, the status quo shortage distribution (Figure 4) has many shortage and non-shortage areas close to each other, which provides an opportunity for efficiency gains by placing a doctor who serves multiple shortage areas. This pattern provides a new perspective on shortage policy design: *Resolving the physician shortage does not always require adding physicians to shortage areas*.

As shown in Figure 1 and Section 6, our shortage areas disagree with the Geographic Primary Care HPSAs in many places, with the latter covering more areas and populations. Still, many physicians added in our simulation are in areas that are not designated as Geographic Primary Care HPSAs, which further indicates the potential for improvement in current targeting policies. Figure A.5 shows more details at the state level.

Finally, we acknowledge that $V2$ is unlikely to be the optimal algorithm as this problem is *NP*-hard. $V2$ does not account for dynamic effects within each iteration, where adding one physician may influence demand beyond adjacent nodes; nor for effects across iterations, where adding one physician may alter the effectiveness of adding the next. Nonetheless, we view this policy as one of the most practical improvements available to policymakers due to its simplicity, and our simulation suggests it delivers promising gains. Using the Medicare claim data, we also estimate that the Medicare HPSA Physician Bonus Program spent $150 million in 2019, which could instead be used to subsidize around $35,000 per year for each physician working in the ZCTAs targeted in $V2$.

8 Discussions

Our study has several limitations. First, our demand model is estimated using the Medicare fee-for-service sample. Although this sample only consists of adults above 65 years old, it is the only dataset available with national coverage and rich geographic information on both patients and providers. Because Medicare fee-for-service plans do not restrict patients to a provider network, we are unable to measure non-geographic barriers of accessing care, such as being uninsured or having a private insurance plan with limited provider networks. Second, we do not model details in the providers' labor market, which may include, but are not limited to: (1) heterogeneity in providers and patient-provider matching at the individual level, which Mark (2021) studies extensively in the context of a payment reform for primary care providers in Northern Ontario, Canada; (2) physicians' labor supply, as well as substitutability and/or complementarity between primary care and specialty care providers, which is the central focus of Falcettoni (2018). Finally, our counterfactual simulations may ignore potential practical and logistical complexities involved in allocating physicians. Incorporating these factors and exploring their interactions with shortage-relief policies is an important direction for future research.

Nurse practitioners (NPs) can also provide basic primary care services, though the scope of their practice varies across states (see Currie et al., 2023). This paper benchmarks against Health Professional Shortage Areas (HPSAs), which consider only primary care physician (PCP) coverage. A comprehensive assessment of how NPs can compensate for limited primary care access is beyond the scope of this paper. Nonetheless, we find that areas with insufficient PCPs also tend to have few NPs. Among all ZCTAs where our simulations add PCPs, 87.4% of PCPs are added to ZCTAs that currently have fewer NPs than the number of additional PCPs required, and 64.8% of PCPs are added to ZCTAs with no NPs at all. These patterns suggest that existing NP capacity may be insufficient to substantially mitigate PCP shortages.

9 Conclusions

In this paper, we developed and formalized a new framework for defining sensible geographic markets, using flexible catchment area designs and a spatial model of demand flows. We apply this framework to measuring primary care physician shortage in the U.S. Our method improves the reliability and consistency of the current measures used for various government subsidy programs that

define shortages independently for areas with fixed boundaries. Combining our measurement framework and a model estimated on rich data, including Medicare fee-for-service claims, location characteristics in the American Community Survey, and a dataset on travel times between ZCTAs, we estimated that the U.S. had 11.58 million people living in areas with primary care shortage in 2019, which is 64% lower than the counterpart measure of shortage using the Geographic Primary Care Health Professional Shortage Areas (HPSAs). We found that, conditional on our measured shortage levels, the Geographic Primary Care HPSAs were more likely to designate rural, low-income areas and areas with high population of racial and ethnic minorities, reflecting potential behavioral bias in states' designation processes and the methodological bias in the HPSA measure, which understates the importance of patient travel and tends to overlook areas that receive more primary care inflows.

Simulation analysis revealed that an additional 3,734 primary care physicians can eliminate primary care shortages nationwide, and we provide a simple heuristic for efficiently adding these physicians. By leveraging the interconnectedness of patient choice sets, the heuristic uses 8.5% fewer physicians compared to a naive approach that only adds physicians to shortage areas, similar to the current targeting policies for HPSAs. We also document that the Medicare HPSA Physician Bonus Program spent $150 million in 2019 — an amount that could instead fund annual subsidies of around $35,000 per physician for all physicians in the ZCTAs targeted under our proposed plan. The simulations provided a useful insight: *resolving the physician shortage does not necessarily require adding physicians directly to shortage areas.* This could potentially lead to higher effectiveness of incentive programs, since, by physicians' revealed preference, they would be more willing to go to non-shortage places than shortage ones.

On top of a conceptually consistent measurement design, the development and application of our new framework are facilitated by the use of high-quality data. Using the Medicare claims data with national coverage, complemented with detailed geographic, patient, and physician characteristics from various data sources, as well as travel times data provided by Hu et al. (2020), we estimate a flexible demand model with rich heterogeneity in patient preferences and unobserved local service attractiveness. This combination of granular data and methodological innovation allows us to make significant improvements over previous approaches, including those in fields outside of economics (e.g., Luo and Wang, 2003; Luo, 2004; Neng Wan and Sternberg, 2012).

A central challenge in designing policies to relieve physician shortages is that both primary care

physicians and government subsidy budgets are scarce resources. The policymaker must balance the risks of missing potential shortage areas and spreading limited resources too thinly across too many places. If policies are not well-targeted, related incentive programs may become ineffective everywhere.

Our analysis suggests that the current shortage designation is overly broad and lacks efficient targeting, and we propose practical heuristics that demonstrate promising gains in allocation efficiency. The current outcry over the primary care physician shortages (e.g., Landro, 2024) reflects not just a shortage of physicians, but a shortage of reliable shortage measures and effective targeting policies. We believe this study highlights some promising directions for improvement — in data, methodology, and policy designs — which warrant further discussion and collaboration from academics and policymakers.

References

Abaluck, Jason and Abi Adams-Prassl (2021), "What do consumers consider before they choose? identification from asymmetric demand responses." *The Quarterly Journal of Economics*, 136, 1611–1663.

Agarwal, Nikhil and Paulo J Somaini (2022), "Demand analysis under latent choice constraints." Technical report, National Bureau of Economic Research.

Ahrens, Achim, Christian B. Hansen, and Mark E Schaffer (2018), "PDSLASSO: Stata module for post-selection and post-regularization OLS or IV estimation and inference." Statistical Software Components, Boston College Department of Economics, URL https://ideas.repec.org/c/boc/bocode/s458459.html.

American Medical Association (2023), "Ama president sounds alarm on national physician shortage." https://web.archive.org/web/20250127155132/https://www.ama-assn.org/press-center/press-releases/ama-president-sounds-alarm-national-physician-shortage. Archived at the Wayback Machine; accessed 2025-05-20.

Auerbach, Shane and Rebekah Dix (2019), "A neighborhood search heuristic for the p-median problem with continuous demand." *Working Paper*, URL https://rebekahanne.github.io/files/NeighborhoodSearchForContinuousPMedian.pdf.

Autor, David H and David Dorn (2013), "The growth of low-skill service jobs and the polarization of the us labor market." *American economic review*, 103, 1553–1597.

Avdic, Daniel, Petter Lundborg, and Johan Vikström (2024), "Does health care consolidation harm patients? evidence from maternity ward closures." *American Economic Journal: Economic Policy*, 16, 16089, URL https://www.aeaweb.org/articles?id=10.1257/pol.20200734.

Barnett, Michael L, Asaf Bitton, Jeff Souza, and Bruce E Landon (2021), "Trends in outpatient care for medicare beneficiaries and implications for primary care, 2000 to 2019." *Annals of internal medicine*, 174, 1658–1665.

Battaglia, Emily (2025), "The effect of hospital maternity ward closures on maternal and infant health." *American Journal of Health Economics*, 11, 201–246, URL https://doi.org/10.1086/727738.

Belloni, Alexandre, Victor Chernozhukov, and Christian Hansen (2014), "High-dimensional methods and inference on structural and treatment effects." *Journal of Economic Perspectives*, 28, 2950, URL https://www.aeaweb.org/articles?id=10.1257/jep.28.2.29.

Buchmueller, Thomas C., Mireille Jacobson, and Cheryl Wold (2006), "How far to the hospital?: The effect of hospital closures on access to care." *Journal of Health Economics*, 25, 740–761, URL https://www.sciencedirect.com/science/article/pii/S0167629605001116.

Capps, Cory, David Dranove, and Mark Satterthwaite (2003), "Competition and market power in option demand markets." *RAND Journal of Economics*, 737–763.

CMS (2025), "Physician bonuses in health professional shortage areas." https://www.cms.gov/medicare/payment/fee-for-service-providers/physician-bonuses-health-professional-shortage-areas-hpsas. Last accessed on Apr 30, 2025.

Collinson, Robert and Peter Ganong (2018), "How do changes in housing voucher design affect rent and neighborhood quality?" *American Economic Journal: Economic Policy*, 10, 62–89.

Corinth, Kevin, David Coyne, Naomi E Feldman, and Craig Johnson (2025), "The targeting of place-based policies: The new markets tax credit versus opportunity zones." Technical report, National Bureau of Economic Research.

Correia, Sergio, Paulo Guimar aes, and Tom Zylkin (2020), "Fast poisson estimation with high-dimensional fixed effects." *The Stata Journal*, 20, 95–115, URL https://doi.org/10.1177/1536867X20909691.

Currie, Janet, Anran Li, and Molly Schnell (2023), "The effects of competition on physician prescribing." Technical report, National Bureau of Economic Research.

Currie, Janet and Patricia B Reagan (2003), "Distance to hospital and children's use of preventive care: is being closer better, and for whom?" *Economic Inquiry*, 41, 378–391.

Dafny, Leemore, Kate Ho, and Robin S Lee (2019), "The price effects of cross-market mergers: theory and evidence from the hospital industry." *The RAND Journal of Economics*, 50, 286–325.

DesRoches, Catherine M, Kirsten A Barrett, Bonnie E Harvey, Rachel Kogan, James D Reschovsky, Bruce E Landon, Lawrence P Casalino, Stephen M Shortell, and Eugene C Rich (2015), "The results are only as good as the sample: assessing three national physician sampling frames." *Journal of general internal medicine*, 30, 595–601.

Dingel, Jonathan I, Joshua D Gottlieb, Maya Lozinski, and Pauline Mourot (2023), "Market size and trade in medical services." Technical report, National Bureau of Economic Research.

Elzinga, Kenneth G. and Thomas F. Hogarty (1973), "The problem of geographic market delineation in antimerger suits." *Antitrust Bulletin*, 18, 45–81.

Elzinga, Kenneth G. and Thomas F. Hogarty (1978), "The problem of geographic market delineation revisited: The case of coal." *Antitrust Bulletin*, 23, 1–18.

Falcettoni, Elena (2018), "The determinants of physicians location choice: Understanding the rural shortage." *Available at SSRN 3493178*.

Freedman, Matthew, Shantanu Khanna, and David Neumark (2023), "Jue insight: The impacts of opportunity zones on zone residents." *Journal of Urban Economics*, 133, 103407.

FTC (2023), "Merger guidelines." https://www.ftc.gov/system/files/ftc_gov/pdf/2023_merger_guidelines_final_12.18.2023.pdf. See discussion of market definition and use of SSNIP-style tools.

Gottlieb, Joshua D, Maria Polyakova, Kevin Rinz, Hugh Shiplett, and Victoria Udalova (2025), "The earnings and labor supply of u.s. physicians*." *The Quarterly Journal of Economics*, qjaf001, URL https://doi.org/10.1093/qje/qjaf001.

Guagliardo, Mark F (2004), "Spatial accessibility of primary care: concepts, methods and challenges." *International journal of health geographics*, 3, 1–13.

Guagliardo, Mark F, Cynthia R Ronzio, Ivan Cheung, Elizabeth Chacko, and Jill G Joseph (2004), "Physician accessibility: an urban case study of pediatric providers." *Health & place*, 10, 273–283.

HHS (2025), "Code of Federal Regulations, Title 42, Chapter I, Subchapter A, Part 5." https://www.ecfr.gov/current/title-42/chapter-I/subchapter-A/part-5. Accessed May 5, 2025.

Houde, Jean-François (2012), "Spatial differentiation and vertical mergers in retail markets for gasoline." *American Economic Review*, 102, 2147–2182.

HRSA Health Workforce (2025a), "Reviewing shortage designation applications." https://bhw.hrsa.gov/workforce-shortage-areas/shortage-designation/reviewing-applications. Last accessed on Apr 30, 2025.

HRSA Health Workforce (2025b), "What is shortage designation?" https://bhw.hrsa.gov/workforce-shortage-areas/shortage-designation. Last accessed on Apr 30, 2025.

Hu, Yujie, Changzhen Wang, Ruiyang Li, and Fahui Wang (2020), "Estimating a large drive time matrix between zip codes in the united states: A differential sampling approach." *Journal of transport geography*, 86, 102770.

Jetty, Anuradha, Max J. Romano, Yalda Jabbarpour, Stephen Petterson, and Andrew Bazemore (2021), "A cross-sectional study of factors associated with pediatric scope of care in family medicine." *The Journal of the American Board of Family Medicine*, 34, 196–207, URL https://www.jabfm.org/content/34/1/196.

Joseph, Alun E and Peter R Bantock (1982), "Measuring potential physical accessibility to general practitioners in rural areas: a method and case study." *Social science & medicine*, 16, 85–90.

Karp, Richard M (2009), "Reducibility among combinatorial problems." In *50 Years of Integer Programming 1958-2008: from the Early Years to the State-of-the-Art*, 219–241, Springer.

Khoury, Stephanie, Jonathan Michael Leganza, and Alex Masucci (2025), "Health professional shortage areas and physician location decisions." *American Journal of Health Economics*.

Landro, Laura (2024), "Seeing a doctor doesnt have to be so frustrating." *The Wall Street Journal*, URL https://www.wsj.com/health/wellness/doctor-visits-appointments-frustration-1c88eb9a. Accessed: 2025-05-20.

Luo, Wei (2004), "Using a gis-based floating catchment method to assess areas with shortage of physicians." *Health & place*, 10, 1–11.

Luo, Wei and Fahui Wang (2003), "Measures of spatial accessibility to health care in a gis environment: synthesis and a case study in the chicago region." *Environment and planning B: planning and design*, 30, 865–884.

Manning, Alan and Barbara Petrongolo (2017), "How local are labor markets? evidence from a spatial job search model." *American Economic Review*, 107, 2877–2907.

Mark, Nathaniel Denison (2021), *Three Essays on Access and Welfare in Health Care and Health Insurance Markets*. Columbia University.

Michelson, Kenneth A., Andrew D. Skol, John A. Graves, Katherine E. Remick, Patrick D. McMullen, Naveen Singamsetty, Danielle K. Cory, and Emily M. Bucholz (2025), "Development of an atlas for US pediatric acute care." *JAMA Pediatrics*, e254227, URL https://pubmed.ncbi.nlm.nih.gov/41143815/. Online ahead of print.

N.C. Department of Health and Human Services Office of Rural Health (2024), "North carolina health care provider placement services." https://www.ncdhhs.gov/nc-dhhs-orh-medical-placements-one-pager/open. Last accessed on Apr 30, 2025.

Neng Wan, Bin Zou and Troy Sternberg (2012), "A three-step floating catchment area method for analyzing spatial access to health services." *International Journal of Geographical Information Science*, 26, 1073–1089.

Newhouse, Joseph P., Albert P. Williams, Bruce W. Bennett, and William B. Schwartz (1982a), "Does the geographical distribution of physicians reflect market failure?" *The Bell Journal of Economics*, 13, 493–505, URL http://www.jstor.org/stable/3003469.

Newhouse, Joseph P., Albert P. Williams, Bruce W. Bennett, and William B. Schwartz (1982b), "Where have all the doctors gone?" *JAMA*, 247, 2392–2396, URL https://doi.org/10.1001/jama.1982.03320420042030.

Nimczik, Jan Sebastian (2020), "Job mobility networks and data-driven labor markets." *Unpublished Working Paper.[648]*.

Peng, Zhong-Ren (1997), "The jobs-housing balance and urban commuting." *Urban studies*, 34, 1215–1235.

Petterson, Stephen M., Winston R. Liaw, Carol Tran, and Andrew W. Bazemore (2015), "Estimating the residency expansion required to avoid projected primary care physician shortages by 2035." *The Annals of Family Medicine*, 13, 107–114, URL https://www.annfammed.org/content/13/2/107.

Rossin-Slater, Maya (2013), "Wic in your neighborhood: New evidence on the impacts of geographic access to clinics." *Journal of Public Economics*, 102, 51–69, URL https://www.sciencedirect.com/science/article/pii/S0047272713000741.

Sauve-Syed, Jessica (2025), "The impacts of shortage area designations on mortality." *Contemporary Economic Policy*.

Schappe, Tyler, Sarah Peskoe, Nrupen Bhavsar, L. Ebony Boulware, Jane Pendergast, and Lisa M. McElroy (2022), "Geospatial analysis of organ transplant referral regions." *JAMA Network Open*, 5, e2231863, URL https://jamanetwork.com/journals/jamanetworkopen/fullarticle/2796274.

Shugarman, Lisa R and Donna O Farley (2003), "Shortcomings in medicare bonus payments for physicians in underserved areas." *Health Affairs*, 22, 173–178.

Silva, J. M. C. Santos and Silvana Tenreyro (2006), "The log of gravity." *The Review of Economics and Statistics*, 88, 641–658, URL http://www.jstor.org/stable/40043025.

Sullivan, Michael (2025), "Fee optimality in a two-sided market."

The Center for the Evaluative Clinical Sciences, Dartmouth Medical School (1996), *The Dartmouth Atlas of Health Care*. American Hospital Publishing, Inc., Chicago, IL, URL https://data.dartmouthatlas.org/downloads/atlases/96Atlas.pdf. Principal Investigator and Series Editor: John E. Wennberg; Editor: Megan McAndrew Cooper.

Wang, Fahui (2000), "Modeling commuting patterns in chicago in a gis environment: A job accessibility perspective." *The Professional Geographer*, 52, 120–133.

World Health Organization (2017), "Fact sheet: Human rights and health." https://web.archive.org/web/20250510011642/https://www.who.int/news-room/fact-sheets/detail/human-rights-and-health. [Accessed 13 May 2025].

Yu, Chuan (2024), "The welfare effects of sponsored product advertising." *Available at SSRN 4817542*.

A Additional Data Descriptions

Crosswalk from ZIP to ZCTA The places of service for PCPs are identified via ZIP codes, yet most of our demographic, travel time, and distance data are at the ZIP Code Tabulation Area (ZCTA) level. Therefore, we crosswalk all ZIP codes in our data to ZCTA using the crosswalk from the Health Resources and Services Administration's HCP GeoCare Navigator. [16]

ZCTA-level Travel Time and Distance We use the ZCTA-level travel time dataset shared by Hu et al. (2020), which contains travel between all ZCTA pairs within 150 miles, measured via ArcGIS Network Analyst. [17] This dataset also calculates the within-ZCTA travel time for all ZCTAs in the US. We compare the travel time from this dataset with a small sample of ZCTAs using our own ArcGIS system, and found that including the within times of both origin and destination ZCTAs from Hu et al. (2020) often overestimate the total travel time. Therefore, we only use the between-ZCTA travel times from Hu et al. (2020) as a conservative measure. For ZCTA pairs that are more than 30 minutes apart, we use distance to approximate patients' travel costs. Here, we use ZCTA centroid coordinates provided by the National Bureau of Economic Research (NBER)[18] to calculate the great-circle distances between coordinates. We hand-queried the coordinate of 37 ZCTAs that are not matched with the centroid data file and have more than 4 physicians in any year from 2015 to 2019 using Google Maps.

B Details of Shortage Area Designation and HPSA

B.1 Healthcare Shortage Area Designations

Different types of healthcare shortage areas are designated by the federeal Health Resources and Services Administration (HRSA), an agency of the United States Department of Health and Human Services (HHS), to identify areas, populations, or facilities with insufficient access to healthcare. These designations guide the allocation of federal and state resources for attracting and retaining healthcare providers — via bonus payments, student loan repayment, visa exemption, J-1 visa foreign residency requirement waiver, etc. — and for establishing community health centers and health maintenance organizations (HRSA Health Workforce, 2025b).

For primary care, the main designation used is the Primary Care Health Professional Shortage Area (HPSA), which provides designated areas the eligibility for most federal programs that incentivize primary care physicians to practice in the area, and also helps guide many state-level incentive programs.[19]

The designation of HPSA is initiated by states. A state's Primary Care Office (PCO), usually under a state's Department of Health, surveys areas in its state for primary care shortages, and files an online

[16] https://geocarenavigator.hrsa.gov. Last accessed in July 2024.

[17] Hu et al. (2020) also used a multi-layer differential sampling approach to estimate the travel time of all ZCTA pairs that are more than 150 miles apart, which we do not use in this paper.

[18] https://www.nber.org/research/data/zip-code-distance-database. Accessed in October 2024.

[19] For example, North Carolina's Provider Recruitment and Placement program "provides recruitment efforts for medical, dental and psychiatric health professionals to North Carolina's rural and underserved areas and offers incentives to eligible providers." See N.C. Department of Health and Human Services Office of Rural Health (2024).

application via a portal provided by HRSA. In the application, the state PCO can justify the shortage using public datasets included in the online portal, such as physician information from the National Plan and Provider Enumeration System (NPPES), the online registry of all licensed providers in the U.S., geography data from the Environmental System Research Institute (ESRI), demographic data from the Census Bureau, and aggregate health measures from the Centers for Disease Control and Prevention (CDC) (HRSA Health Workforce, 2025a). State PCOs can also supplement these datasets with their own area or facility-specific data, although, based on our interview with one state PCO, collecting additional data is often logistically challenging and requires coordination between the PCO and local health providers. HRSA then reviews and approves PCO's applications based on certain criteria, which we discuss in the next section.

Primary care HPSA can be designated under three different categories: geographic, population, and facility. Geographic Primary Care HPSA measures the access of primary care services for *the entire population* in a geographic area; Population Primary Care HPSA designates *a certain group of population* in a geographic area, such as the low-income or homeless population; Facility Primary Care HPSA identifies specific facilities that "serve a population or geographic area with a shortage of providers", such as medical care facilities (rural hospitals) or correctional facilities (HRSA Health Workforce, 2025b).

Different types of Primary Care HPSAs are not always eligible for all federal programs related to HPSAs, but in general, Geographic Primary Care HPSAs are eligible for most programs. One such program is the CMS Physician Bonuses in Health Professional Shortage Areas (HPSA Bonus), which pays 10% quarterly bonuses to primary care providers who furnish Medicare services in Geographic Primary Care HPSAs, on the basis of the amount billed to fee-for-service Medicare (CMS, 2025). Using Medicare claims data, we estimate that the Bonus program spent $150 million in 2019, more than quadrupled from $36 million in 1998 (Shugarman and Farley, 2003)[20]

Our objective is to provide policy guidance on measuring and alleviating geographic primary care shortage. Therefore, we mainly benchmark our shortage measure to the Geographic Primary Care HPSA. Due to logistical reasons, rather than negligence by the state PCOs or the federal agency, the current Geographic Primary Care HPSA may not reflect all HPSA designations based on geographic shortage. One PCO officer informed us that the office sometimes applies for Population instead of Geographic HPSA because it is difficult to justify a Geographic HPSA designation: data at an aggregate geography may not reflect local shortages, yet collecting additional data is challenging. Thus, whenever we compare our measure with Primary Care HPSAs, we also include a version comparing with both Geographic and Population Primary Care HPSAs in the Appendix.

In the next section, we summarize and discuss the criteria used to define Geographic Primary Care HPSAs. We refer readers to Khoury et al. (2025) for more detailed discussions on other types of shortage designations, levels of HPSAs, and government incentive programs.

[20] All claims eligible for the Bonus Payment are flagged in the Medicare fee-for-service professional claims. We summed all such claims in the 20% Medicare Beneficiaries sample, and multiplied it by five. See Section 4 for details about the professional claims data.

B.2 Designation Criteria of Primary Care Health Professional Shortage Area (HPSA)

In this section, we summarize the main idea of the designation criteria for Geographic Primary Care HPSA, and discuss its drawbacks and the motivation for developing an alternative framework.

The U.S. Department of Health and Human Services (HHS) sets rules for designating health professional shortage areas (HHS, 2025). As mentioned above, Geographic Primary Care HPSA identifies geographic areas with a shortage of Primary Care physicians, considering the entire population in the area, which we use as the main benchmark in this paper.

To be eligible for Geographic Primary Care HPSA, an area needs to meet three general conditions, quoting HHS (2025):

1. *The area is a rational area for the delivery of primary medical care services.*
2. *One of the following conditions prevails within the area:*
    - *The area has population to full-time-equivalent primary care physician ratio of at least 3,500:1.*
    - *The area has a population to full-time-equivalent primary care physician ratio of less than 3,500:1 but greater than 3,000:1 and has usually high needs for primary care services or insufficient capacity of existing primary care providers.*
3. *Primary medical care manpower in contiguous areas are overutilized, excessively distant, or inaccessible to the population of the area under consideration.*

Then, the rule details how each item should be measured. We provide a simplified summary below and refer readers to HHS (2025) for more details.

Rational Service Area Generally, a rational area for primary care service can be a single county or a group of contiguous counties whose population centers are within 30 minutes of travel time, or, a group of subcounties or census tracts where the combined area does not have contiguous resources within 30 minutes. The rule also provides a guideline in determining the 30-minute travel time: 20 miles for normal conditions with access to primary roads, 15 miles for mountains or areas with only secondary roads, and 25 miles for flat areas or areas connected by interstate highways. The main takeaway here is that HHS uses *within 30 minutes* as a measure of access to primary care. In other words, people in a non-shortage area, abstracted to its population center, would have enough primary care physicians within 30 minutes.

Population to Physician Ratio After defining markets, the next step is to measure whether each has enough primary care physicians. The main threshold HRSA uses is a population to full-time-equivalent primary care physician ratio of 3,500:1, with the exception of a 3,000:1 ratio for areas with high needs for primary care. Using the HPSA designation data in 2019, we find that 64% of the designated area (and 60% of the designated population) used the 3,500:1 ratio as the target. Meanwhile, under our framework, places with high needs can be directly reflected in the data, estimated with a high local attractiveness, and accounted for in the shortage measurement. Therefore, we measure all shortages using the standard 3,500:1 ratio, and include robustness checks using 3,000:1 in Table A.2.

Contiguous Area Considerations Up to this point, we have defined a rational area with a specific boundary and calculated its population-to-physician ratio. However, exceeding the cutoff ratio (3,500 or 3,000) alone is not sufficient to establish a shortage, since its population can still have access to primary care resources in adjacent areas. To take this into account, HHS has to specify additional criteria for the areas contiguous to the area to be considered. The rule requires that the contiguous area must meet one of the following three conditions:

- The primary care resource in the contiguous area is more than 30 minutes of travel away;
- The population to full-time-equivalent primary care physician ratio in the contiguous area is over 2,000:1, which indicates "that practitioners in the contiguous area cannot be expected to help alleviate the shortage situation in the area being considered for designation." (HHS, 2025)
- The primary care resource in the contiguous area is inaccessible to people in the area under consideration, due to barriers such as socio-economic characteristics, income, Medicaid status, etc.

These rules are ambiguous and may lead to multiple interpretations. Firstly, there is no clear guideline on how to define the boundary of the contiguous area in which residents may access healthcare resources; these bounds may be arbitrarily close or far from the area of consideration. Should we assume that the adjacent area is a "rational area" for the delivery of primary care services, there may still be multiple ways to draw its boundaries. These areas may have different population-to-provider ratios, therefore leading to ambiguity and discretion in HPSA designations.

Also, although it is inevitable to pick a ratio cutoff as the starting point (in this case, 3,500 or 3,000), the choice of 2,000:1 for the contiguous area seems very arbitrary. It doesn't consider at all the relative market sizes of the area under consideration and the contiguous areas, the distance and travel time, the quality of services, etc., not to mention its credibility for evaluating whether providers in nearby areas can indeed "help alleviate the shortage situation."

As further discussed in Section 5, this convolution is conceptually inevitable when we define shortage based on areas with fixed boundaries: if the area is large, the heterogeneity across small units inside the area is ignored; if the area is small, it is very likely that people inside the area also seek care in nearby places, and then, as we see in the case of HPSA, the criteria design can get complex. This is the main challenge we aim to resolve in this paper by using a new framework.

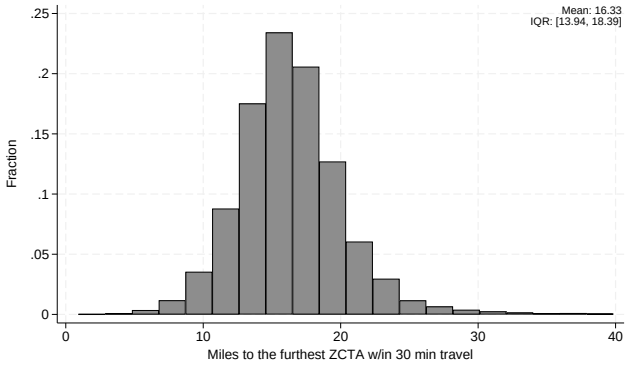
C Supplemental Exhibits

Appendix Table A.1: List of ZCTA Characteristics

Data Source	Variable
CMS Hospital Compare	Acute care hospitals
	Children's hospital
	Critical access hospital
American Community Survey (5-year estimates by ZCTA)	Total population
	Share under age 18
	Share age 18–64
	Share White
	Share Black
	Share Hispanic
	Share with high school degree
	Share with some college education
	Share with college degree and above
	Median income
	Share in poverty
	Share employed
	Share with any insurance
	Share privately insured
	Share on Medicaid
	Share on Medicare
Medicare Beneficiary Summary File	Share White (Medicare)
	Share Black (Medicare)
	Share Asian (Medicare)
	Share Other (Medicare)
	Share Hispanic (Medicare)
	Share North American Native (Medicare)
HPSA designation	Share living in geographic HPSA
	Share living in geographic and population HPSA

Notes: Table lists all variables fed into the post double-selection procedure in estimating primary care amenity levels in Equation (3), along with their following transformations: log, up to five degrees of polynomial, and deciles. Estimates are reported in Table 2.

Appendix Figure A.1: Heterogeneity in Crow-fly Distance Within 30 Minutes Travel



Notes: For each ZCTA, we calculate the maximum of crow-flying distance between all ZCTAs within its catchment area of 30 minutes travel. Figure plots the histogram.

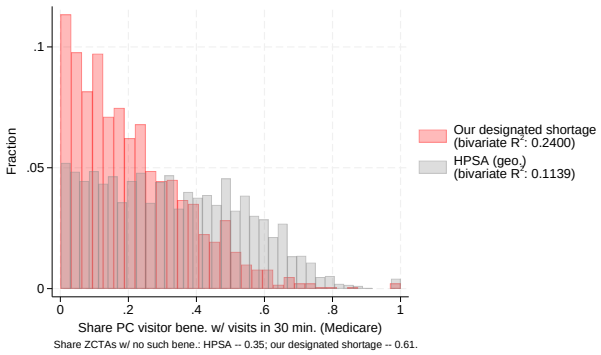
Appendix Table A.2: Robustness Check: Shortage Population under Different Model Parameter Values

	Coverage rate threshold	
	> 3,500:1	> 3,000:1
Baseline (30 min. travel network)	11.58m	18.28m
Change travel network		
25 min.	21.62m	29.83m
29 min.	13.45m	20.10m
40 min.	3.18m	6.87m
Change travel elasticities		
$\times 0.5$	11.46m	18.09m
$\times 1.5$	12.24m	19.16m
Change travel elasticities only for shortage areas		
$\times 0.5$	11.47m	18.18m
$\times 0.1$	11.39m	18.27m

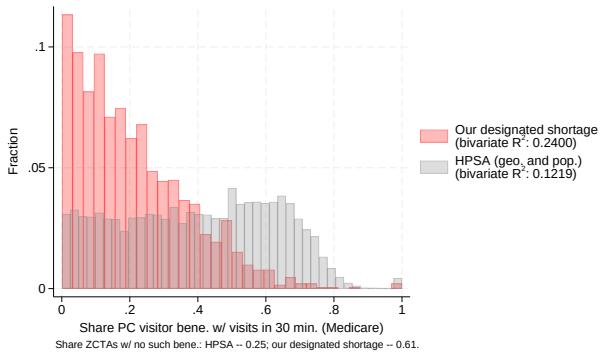
Notes: We calibrate parameter values and calculate alternative primary care coverage rates for ZCTAs in 2019. For each calibration specification, table shows the total number of population living in ZCTAs exceeding the coverage rate threshold listed in each column title. In panel "Changing travel network", we vary the threshold $\bar{\eta}$ for defining patients' choice sets. In panel "Changing travel elasticities", we vary the estimated travel cost coefficients reported in Table 1. In panel "Changing travel elasticities only for shortage areas", we vary the estimate travel cost coefficients reported in Table 1 *for ZCTAs designated as shortage in the base line specification.*

Appendix Figure A.2: Predictive Power of Shortage Measures on Realized Care

(A) Geographic HPSA



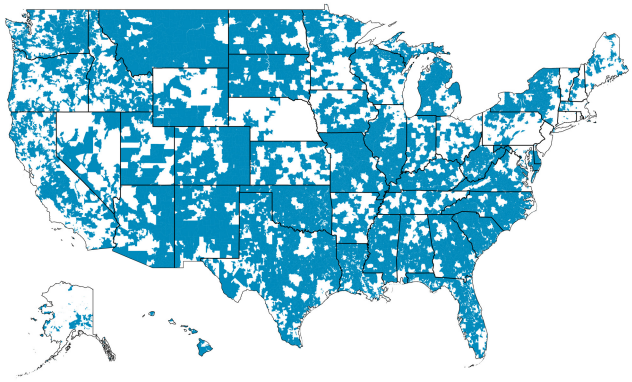
(B) Geographic + Population HPSA



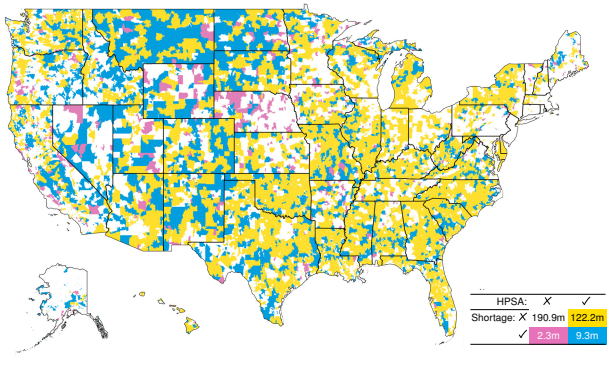
Notes: We plot the histogram of share of Medicare beneficiaries with any PCP visits who has a visit within 30 minutes from home for ZCTAs that are HPSAs and ZCTAs that have coverage rate greater than 3,500, respectively. At the right of the figure, we show the R^2 of a bivariate regression, with the any PCP visits within 30 minutes from home as the outcome variable, and an indicator for HPSAs and for ZCTAs that have coverage rate greater than 3,500, respectively, as the explanatory variable. Panel (A) uses geographic HPSAs and Panel (B) uses geographic or population HPSAs.

Appendix Figure A.3: Distribution of Providers and Geographic + Population HPSAs

(A) Health Professional Shortage Areas

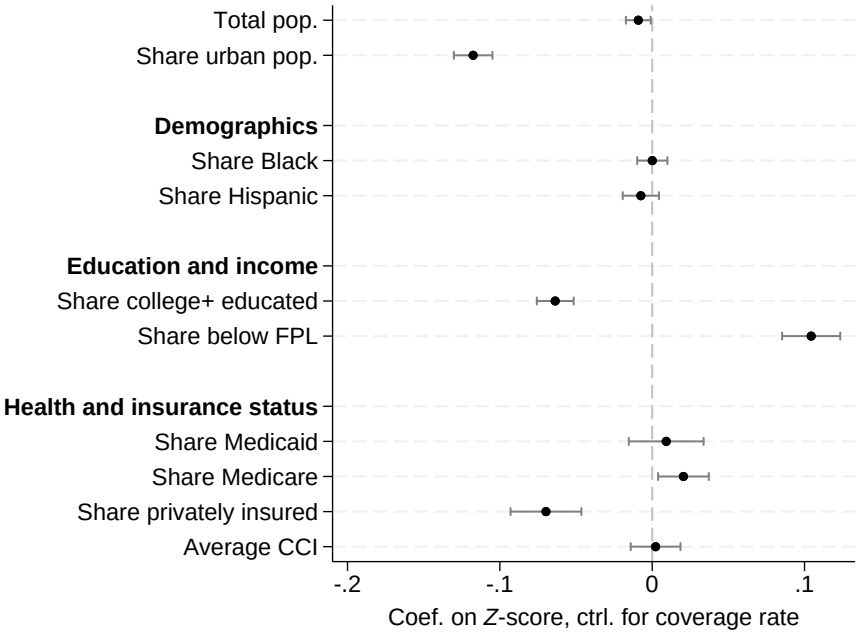


(B) Our Measure vs. Geographic/Population PC HPSA



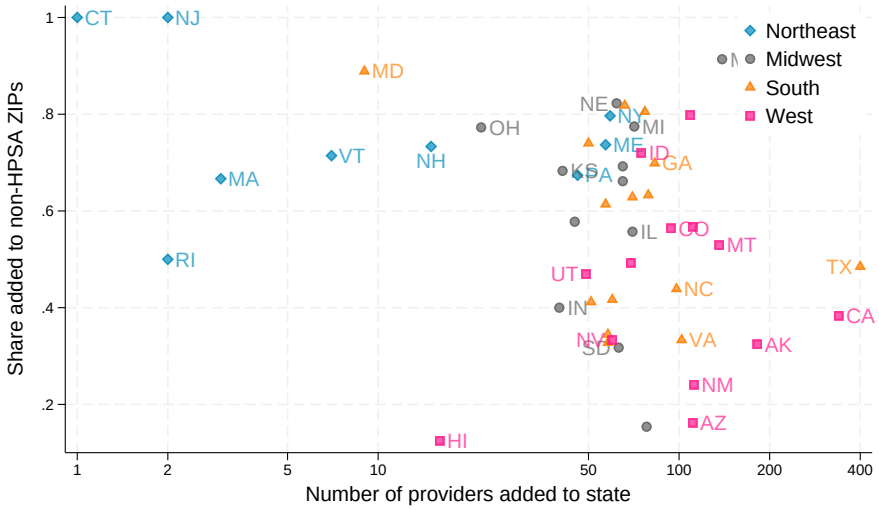
Notes: Panel (A) shows, in blue, all ZCTAs that are designated as a Geographic or Population Primary Care HPSA. Panel (B) shows the comparison between our shortage measure with Geographic or Population Primary Care HPSAs at the ZCTA level. Our measurement framework is described in Section 5.

Appendix Figure A.4: HPSA (Geographic + Population) Designations Priorities



Notes: We run an OLS regression of an indicator for a ZCTA being a geographic or population HPSA on a set of each ZCTA characteristic listed on the left of the figure, conditional on deciles of the coverage rate defined in section 5. In the regression, we include Z-scores of the variables, constructed by re-centering the variable around its mean, divided by its standard deviation. Figure plots the coefficient on each Z-score variable. CCI: Charlson Comorbidity Index, a score ranging from 0-37 that captures a patient's mortality risk based on age and chronic conditions. We take the ZCTA-level average of it among Medicare beneficiaries included in our sample.

Appendix Figure A.5: Percent Physicians Added to Non-Geographic HPSAs from Simulation $V2$



Notes: On the y-axis is the share of physicians added to ZCTAs that are not Geographic HPSAs in each state; on the x-axis is the total number of providers added to each state. States are color-coded by their Census Regions.