

Opportunity &  
Inclusive Growth  
INSTITUTE



FEDERAL RESERVE BANK OF MINNEAPOLIS



INSTITUTE WORKING PAPER  
No. 68

# Market Size and Trade in Medical Services

April 2023

---

**Jonathan I. Dingel**

*Chicago Booth, CEPR, and NBER*

**Pauline Mourot**

*Chicago Booth*

**Joshua D. Gottlieb**

*University of Chicago and NBER*

**Maya Lozinski**

*University of Chicago*

**DOI:** <https://doi.org/10.21034/iwp.68>

**Keywords:** Healthcare access; Market-size effects; Medicare claims data; Trade in services

**JEL classification:** F12, F14, I11, R12

The views expressed herein are those of the authors and not necessarily those of the Federal Reserve Bank of Minneapolis or the Federal Reserve System.

# Market Size and Trade in Medical Services\*

Jonathan I. Dingel  
Chicago Booth,  
CEPR, and NBER

Joshua D. Gottlieb  
University of Chicago  
and NBER

Maya Lozinski  
University of Chicago

Pauline Mourot  
Chicago Booth

March 2023

## Abstract

We measure the importance of increasing returns to scale and trade in medical services. Using Medicare claims data, we document that “imported” medical care—services produced by a medical provider in a different region—constitute about one-fifth of US healthcare consumption. Larger regions specialize in producing less common procedures, which are traded more. These patterns reflect economies of scale: larger regions produce higher-quality services because they serve more patients. Because of increasing returns and trade costs, policies to improve access to care face a proximity-concentration tradeoff. Production subsidies and travel subsidies can impose contrasting spillovers on neighboring regions.

*Keywords:* healthcare access, market-size effects, Medicare claims data, trade in services

*JEL codes:* F12, F14, I11, R12

---

\*We thank Cedric Elkouh, Feng Lin, Noah Sobel-Lewin, and especially Luke Motley and Vaidehi Parameswaran for excellent research assistance. We are grateful to Rodrigo Adão, Zarek Brot-Goldberg, Don Davis, Rebecca Diamond, Liran Einav, Ed Glaeser, Tali Han, Tom Holmes, Loukas Karabarbounis, Doug Miller, Sean Nicholson, Bruce Schakman, Martin Schneider, Bradley Setzler, Jon Skinner, Felix Tintelnot, discussants Jan David Bakker, Barthelemy Bonadio, Jessie Handbury, Naomi Hausman, Elena Patel, and many seminar participants for helpful feedback. We thank the Becker Friedman Institute at the University of Chicago for funding support and enabling us to access the Medicare claims data. We thank Antoine Levy for sharing CBSA-level bedrock-depth data. Dingel thanks the Cohen and Keenoy Faculty Research Fund at the University of Chicago Booth School of Business for support. Lozinski thanks the NIH (grants T32GM007281 and T32AG051146). Authors’ email addresses: [jdingle@chicagobooth.edu](mailto:jdingle@chicagobooth.edu), [jgottlieb@uchicago.edu](mailto:jgottlieb@uchicago.edu), [mayalozinski@uchicago.edu](mailto:mayalozinski@uchicago.edu), [pmourot@chicagobooth.edu](mailto:pmourot@chicagobooth.edu).

Rural Americans have worse health outcomes (Deryugina and Molitor, 2021; Finkelstein, Gentzkow, and Williams, 2021), but America’s doctors are disproportionately located in big cities (Rosenblatt and Hart, 2000). This contrast might suggest a spatial mismatch between consumers and producers of medical services, and arguments about whether physicians are geographically “maldistributed” go back decades (Newhouse et al., 1982*a*; Skinner et al., 2019). To evaluate this concern, we must consider two economic mechanisms: economies of scale and patients’ travel costs. We find that both are key to understanding spatial patterns of healthcare within the United States.

When medical services exhibit increasing returns to scale, there are benefits to geographically concentrating production. Indeed, medicine has long been suggested as an industry in which the division of labor is limited by the extent of the market (Arrow, 1963; Baumgardner, 1988). But if healthcare markets are geographically segmented, the only way to serve patients in smaller regions is to disperse production across space, foregoing the benefits of scale.<sup>1</sup> For time-sensitive emergency care, this assumption is plausible. But the vast majority of medical spending is not for such emergencies. For example, if patients with cancer can travel across regions in search of the ideal oncologist—one specialized in their particular type of cancer, one with a better reputation, or simply a better personal match—the economic geography of medical care may resemble other tradable industries. Society would face a proximity-concentration tradeoff: patients who import medical services produced elsewhere incur trade costs but benefit from higher quality generated by scale economies.

We quantify the roles of local increasing returns and trade costs in medical services. Using millions of Medicare claims, we find that “imported” medical procedures—defined as a patient’s consumption of a service produced by a medical provider in a different region—

---

<sup>1</sup>Many economists assume trade costs for medical services are prohibitively high. Hsieh and Rossi-Hansberg (2021): “Producing many cups of coffee, retail services, or health services in the same location is of no value, since it is impractical to bring them to their final consumers.” Jensen and Kletzer (2005): “Outside of education and healthcare occupations, the typical ‘white-collar’ occupation involves a potentially tradable activity.” Bartik and Erickcek (2007): “An industry can bring in new dollars by selling its goods or services to persons or businesses from outside the local economy (‘export-base production’). . . For health care institutions, demand for services tends to be more local.”

constitute about one-fifth of US healthcare consumption. Imports are a larger share of consumption for patients in smaller markets. “Exported” medical services are disproportionately produced in large markets. Larger regions specialize in producing less common procedures, and these procedures are traded more. These patterns are attributable to local increasing returns to scale: larger regions produce higher-quality services because they serve more patients. We estimate a model and use it to quantify how production or travel subsidies would affect patients’ access to care and the quality produced in each region. Spatially neutral policies affect regions differently depending on their size and trade patterns.

Section 1 develops a model of trade in medical services to guide our analysis. We adapt standard models of agglomeration and trade to a setting in which the government sets prices, so endogenous quality and travel patterns clear markets. If there are local increasing returns, larger markets produce higher-quality care and export it. When economies of scale are sufficiently strong relative to market size, the model predicts that larger markets will be net exporters of medical services. Market size matters more at smaller scales, so less common medical procedures respond more to differences in market size.

Section 2 describes our Medicare claims data. Medicare is the federal government’s insurance program for the elderly and disabled and the largest insurer in the United States. Medical service providers submit claims that report the treatment location, where the patient lives, and distinguish among thousands of distinct medical procedures.

Section 3 begins our empirical investigation by examining how production and consumption vary with market size. Production is geographically concentrated in larger markets, while consumption is much less so. This contrast implies that larger markets are net exporters of medical services to smaller markets. To test whether this pattern reflects a home-market effect—that is, larger demand causes larger regions to export medical services—we estimate a gravity model of bilateral gross trade flows (Costinot et al., 2019). Controlling for the geographic distribution of demand and travel distances, regions with larger residential populations export more medical care. Local increasing returns to scale are so strong that

greater demand induces a larger increase in exports than imports, making larger markets net exporters of medical care. We show that these scale effects cannot be attributed to larger markets having lower input costs or medical production raising population size.

Section 4 shows that trade and market size play a larger role in less common procedures. The imported share of consumption is 22% for above-median-frequency procedures and 35% for those below the median. Doctors performing rare procedures export their services more often and across a broader geographic scope, sometimes serving patients who reside thousands of kilometers away. For example, half of the patients having left ventricular assist devices (LVADs) inserted to restore their heart function come from outside the surgeon’s region, while only 15% of screening colonoscopies are imported. Consistent with the model, the home-market effect is substantially stronger for less common procedures: a larger residential population drives a greater increase in net exports for rarer services.

Section 5 shows that larger markets produce higher-quality services thanks to economies of scale. We recover revealed-preference estimates of regional service quality by estimating patients’ willingness to travel to each exporting region for medical services.<sup>2</sup> These estimates are positively related to external quality measures, such as hospital rankings published by *U.S. News and World Report*. Inferred quality rises considerably with the regional volume of production. We estimate the scale elasticity of production to be about 0.6: a region producing 10% more because of greater demand produces about 6% higher quality.

A variety of mechanisms could generate these local increasing returns to scale: finer specialization among physicians, sharing of lumpy capital equipment, knowledge diffusion, learning by doing, and greater availability of complementary inputs (Marshall, 1890). While we cannot test all these hypotheses, we find that physicians in larger markets are more specialized and more experienced in the procedures they perform. Trade enables patients from across regions to share in these benefits of scale: imports are more likely to be provided by a specialist—and the appropriate specialist—than locally produced services. Specializa-

---

<sup>2</sup>Regional quality estimates and other results may be downloaded at <http://jdingel.com>.

tion and learning by doing likely contribute to the local increasing returns that produce higher-quality medical care in larger markets.

We use our estimates of scale economies and trade costs to quantitatively explore the proximity-concentration tradeoff. Section 6 shows that policies affect regions differently depending on their size and trade patterns. A nationwide increase in reimbursements raises local output quality more in smaller regions, but these regions experience smaller increases in patients' market access because fewer of their patients consume local services. We then examine the implications of increasing access to care in one region by either increasing reimbursements or reducing travel costs. Increasing reimbursements has a higher return in more populous regions: the nationwide improvement in patient market access is about 15% higher per dollar of spending when raising reimbursements in the largest regions instead of the smallest regions. Increasing reimbursements in one region reduces output quality in neighboring regions, while improving patients' market access to the extent they import from the treated region. Reducing travel costs for one region increases its import demand, which improves both output quality and market access in neighboring regions. The rich pattern of consequences when subsidizing patients in low-output regions highlights the importance of trade and agglomeration for the incidence of these policies on patients and producers.

The higher-quality care available in larger markets may not benefit all patients equally. Patients of lower socioeconomic status are less likely to travel for better medical care. Gravity regressions show that patients from the lowest neighborhood-income decile exhibit a distance elasticity of -2.1, while those in the highest decile have a distance elasticity of -1.7. This finding is not driven by differences in the composition of care needed: these patients are more sensitive to distance even when we examine travel patterns within specific billing codes. Thus, the gains generated by local increasing returns do not benefit all patients equally.

This paper builds on research in urban, trade, and health economics. Urban economists have documented skill-biased agglomeration in production as knowledge workers have become more numerous and concentrated in skilled cities (Berry and Glaeser, 2005; Moretti, 2011;

Diamond, 2016; Davis and Dingel, 2020; Eckert, Ganapati, and Walsh, 2020). Connecting this to the production and trade of services has been more difficult. Most studies of the geography of services analyze restaurants and retailers (Davis et al., 2019; Agarwal, Jensen, and Monte, 2020; Allen et al., 2021; Miyauchi, Nakajima, and Redding, 2021; Burstein, Lein, and Vogel, 2022). We show that—even in a service-based economy—the sizes of both local and potential export markets influence production and quality. This suggests that healthcare can serve as an export base for large markets (Bartik and Erickcek, 2007).

The trade literature has examined market-size effects in manufacturing but investigated services much less. Davis and Weinstein (2003), Hanson and Xiang (2004), and Bartelme et al. (2019) link manufactures’ market size to export patterns, in line with the home-market effect of Krugman (1980) and Helpman and Krugman (1985). Dingel (2017) shows that market-size effects drive quality specialization across US cities. Market-size effects for pharmaceuticals have been estimated using demographic variation over time (Acemoglu and Linn, 2004) and across countries (Costinot et al., 2019). Services are much less studied, in part because of the paucity of reliable trade data (Lipsey, 2009; Muñoz, 2022). We advance this literature using the detailed procedure and location information in medical claims data.

The importance of medical care for health, life expectancy, and welfare generates substantial public-policy interest. Rural locations have worse health outcomes but fewer doctors per capita. An important series of papers by Newhouse et al. (1982*a,b,c*), Newhouse (1990), and Rosenthal, Zaslavsky, and Newhouse (2005) considered this issue and argued against targeting a uniform geographic distribution of physicians. Building on these studies, we measure interregional trade in medical services, estimate the impact of geography on patient access, and connect this trade to economies of scale. Importantly, we use modern trade theory to guide our modeling, estimation strategy, and counterfactual policy analysis.

# 1 Theoretical framework

This section develops a competitive model of trade in medical services tailored to our empirical analysis of US healthcare. Patients select quality-differentiated services and face trade costs. Local increasing returns cause the quality-adjusted cost of producing a service to decline with scale. The distinction between lower costs and higher quality is important in our empirical context. The US government plays a unique role in healthcare, purchasing a large share of all output and imposing substantial regulations. We focus on Medicare, the large federal program that purchases healthcare for the elderly and disabled at regulated prices. In this context, prices do not play their traditional role in clearing markets. Instead, quality of care and patients' distance from care bring this market towards equilibrium.

Beyond healthcare, this model speaks to agglomeration effects in other markets subject to price controls. We show that such circumstances can be captured by a modest modification to conventional trade models. Our model continues to deliver a gravity equation for trade flows and to predict home-market effects. This framework delivers testable predictions about spatial variation in services' quality and trade patterns when prices are fixed.

## 1.1 Demand

We use a logit model of individuals choosing providers for a given service. Providers and patients are in regions indexed by  $i$  or  $j$ , with  $\mathcal{I}$  denoting the set of regions. Let  $N_j$  denote the number of patients residing in region  $j$  who make a choice.<sup>3</sup> All providers in a region are identical. Utility has a provider-region-specific component, a region-pair component, and an idiosyncratic component: patient  $k$  in region  $j$  choosing a provider in region  $i$  obtains utility

$$U_{ik} = \ln \delta_i + \ln \rho_{ij(k)} + \epsilon_{ik}.$$

---

<sup>3</sup>Appendix A.1 extends the model to have multiple patient types.



The provider-region-specific component  $\delta_i$  would usually include a product’s characteristics and price. Since Medicare pays reimbursement rates that it sets administratively,<sup>4</sup> the  $\delta_i$  relevant for the patient is the quality of the providers in region  $i$ . The region-pair component  $\rho_{ij}$  represents bilateral inverse trade costs (proximity). The idiosyncratic component  $\epsilon_{ik}$  is independently and identically drawn from a standard Gumbel distribution, so the probability that patient  $k$  selects a provider in region  $i$  is

$$\Pr(U_{ik} > U_{i'k} \ \forall i' \neq i) = \frac{\exp(\ln \delta_i + \ln \rho_{ij(k)})}{\sum_{i' \in 0 \cup \mathcal{I}} \exp(\ln \delta_{i'} + \ln \rho_{i'j(k)})}.$$

There is an outside option denoted by  $i = 0$ , which represents individuals choosing to forgo care, and we normalize its common component to zero,  $\ln \delta_0 = \ln \rho_{0j(k)} = 0 \ \forall k$ .<sup>5</sup>

This choice probability implies a gravity equation for the quantity of trade between any two regions when we aggregate patients’ decisions. Let  $Q_{ij}$  denote the quantity of procedures supplied by providers in  $i$  to patients residing in  $j$ , and let  $Q_{0j}$  denote the number of patients in  $j$  selecting the outside option. Because each patient selects at most one provider,  $N_j = \sum_{i \in \mathcal{I} \cup \{0\}} Q_{ij}$ . The demand by patients in  $j$  for procedures performed in  $i$  is

$$Q_{ij} = \delta_i \frac{N_j}{\Phi_j} \rho_{ij}, \tag{1}$$

---

<sup>4</sup>Patients pay a share of these reimbursements through copayments and deductibles. But note that these cost-sharing rules are constant nationally, and most Medicare patients have a supplemental insurance (Medigap or Medicaid) which covers most or all of this cost-sharing.

<sup>5</sup>This formulation of demand is familiar from the hospital competition literature, which has studied competition among hospitals on price and quality. The literature tends to assume competition occurs within a specified geographic radius (*e.g.*, Kessler and McClellan, 2000; Cooper et al., 2018) or within a metropolitan area or similar geographic unit (*e.g.*, Ho, 2009; Gowrisankaran, Nevo, and Town, 2015; Clemens and Gottlieb, 2017; Lewis and Pflum, 2017; Ho and Lee, 2019; Dafny, Ho, and Lee, 2019; Garthwaite, Ody, and Starc, 2022). Data in this literature are often limited to certain states (*e.g.*, Town and Vistnes, 2001; Gaynor and Vogt, 2003; Capps, Dranove, and Satterthwaite, 2003; Lewis and Pflum, 2015; Ericson and Starc, 2015; Ho and Lee, 2017). Patients who are treated outside their home region may be dropped from the data or treated as choosing the outside option (as in Gaynor and Vogt, 2003). These definitions may be appropriate for modeling competition within specified markets (though they have been questioned by Gaynor, Kleiner, and Vogt, 2013; Dranove and Ody, 2016) and are natural if one assumes healthcare demand is local—as has been standard (see footnote 1). We assume all regions are in each patient’s choice set, so there are no “control” markets and modeling strategic interactions would be very computationally costly.

where  $\Phi_j \equiv \sum_{i' \in 0 \cup \mathcal{I}} \delta_{i'} \rho_{i'j}$  is the expected value of the choice set for patients in region  $j$ . We call this  $\Phi_j$  “patient market access.” Equation (1) is a gravity equation with an origin  $i$  component, a destination  $j$  component, and an  $ij$  pair component. Total demand for procedures produced in  $i$  is

$$Q_i = \delta_i \sum_j \frac{N_j}{\Phi_j} \rho_{ij}. \quad (2)$$

## 1.2 Production

We assume competitive production of services with free entry and local increasing returns that are external to the firm. That is, each price-taking provider chooses its output quality and quantity given total regional production, an exogenous factor price, and an exogenous productivity shifter. A provider in region  $i$  that employs  $L$  units of the composite input to produce service of quality  $\delta$  produces the following output quantity:

$$A_i \frac{H(Q_i)}{K(\delta)} L.$$

Improving quality is costly so  $K(\delta)$  is increasing. Local increasing returns to scale are a weakly increasing, concave function  $H(Q_i)$  of total regional production,  $Q_i$ , which competitive firms take as given (Chipman, 1970). The regional productivity shifter  $A_i$  captures any other influences, such as historical investments. Provider size  $L$  is indeterminate (and unimportant) given the linear production function, external economies of scale, and price-taking behavior. The composite input is supplied to region  $i$  at factor price  $w_i$ .<sup>6</sup> Thus, the unit cost of producing quality  $\delta$  in region  $i$  is

$$C(Q_i, \delta_i; w_i, A_i) \equiv \frac{w_i K(\delta_i)}{A_i H(Q_i)}.$$

---

<sup>6</sup>If the regional factor supply were upward-sloping rather than perfectly elastic, we would estimate increasing returns net of the cost of hiring additional inputs. That is, if the factor supply elasticity were  $\beta$ , our estimate of the scale elasticity  $\alpha$  from equation (4) below would instead be an estimate of the effective scale elasticity  $\tilde{\alpha} \equiv \alpha - \frac{\beta}{1+\beta}$ .

In our institutional setting, output prices are not an equilibrium object determined solely by the intersection of supply and demand. Instead Medicare sets “reimbursement rates” largely independent of quality, quantity, or region,<sup>7</sup> which we denote  $\bar{R}$ . Each provider that produces output of the highest quality produced in region  $i$  earns revenue  $\bar{R}$  per unit.

Provider optimization and free entry make the unit cost equal to the reimbursement rate in each region. Given the factor price  $w_i$  and productivity shifter  $A_i$ , the free-entry condition

$$C(Q_i, \delta_i; w_i, A_i) = \bar{R} \quad (3)$$

defines a regional isocost curve: the set of quantity-quality combinations for which the average cost of production equals the reimbursement rate. This isocost curve is the set of potential equilibrium production outcomes in region  $i$ . Local increasing returns make the isocost curve upward-sloping in  $(Q, \delta)$  space. With free entry and fixed prices, the benefits of scale are realized as higher-quality services in higher-output regions.

While our assumptions thus far suffice for qualitative results, we later specify functional forms for additional predictions and empirical quantification; specifically,  $K(\delta_i) = \delta_i$  and  $H(Q_i) = Q_i^\alpha$ , with a scale elasticity of  $\alpha \in (0, 1)$ . In this case, the free-entry condition (3) is

$$\bar{R} = \frac{w_i \delta_i}{A_i Q_i^\alpha}. \quad (4)$$

### 1.3 Equilibrium

Equilibrium equates supply and demand in each region,  $Q_i = \sum_j Q_{ij}$ . Given exogenous parameters  $\bar{R}$ ,  $\{w_i, A_i, N_i\}_{i \in \mathcal{I}}$ , and  $\{\rho_{ij}\}_{(i,j) \in (\mathcal{I}, \mathcal{I})}$ , an equilibrium is a set of quantities and qualities  $\{Q_i, \delta_i\}_{i \in \mathcal{I}}$  that simultaneously satisfy equations (2) and (3).

---

<sup>7</sup>While Medicare does have some quality incentive programs, the money at stake is a small share of Medicare’s overall spending (Gupta, 2021). Medicare has some spatial variation in physician reimbursements, but it is not very large and has diminished over time (Clemens and Gottlieb, 2014).

## 1.4 Scale effects in autarky

We first consider equilibrium in autarky: patients can choose whether to receive care, but they cannot travel between regions ( $\rho_{ij} = 0$  for  $i \notin \{0, j\}$ ). In this case, all demand is local and equation (2) simplifies to

$$Q_{jj} = \frac{\delta_j \rho_{jj}}{1 + \delta_j \rho_{jj}} N_j. \quad (5)$$

The autarkic equilibrium is at the intersection of the demand curve given by equation (5) and the free-entry isocost curve given by equation (3).<sup>8</sup> An increase in population size,  $\Delta N_j > 0$ , affects equilibrium outcomes by shifting the demand curve.

Figure 1 illustrates how greater demand affects quality in autarky. Panel 1(a) shows the role of increasing returns to scale. The vertical axis shows quality  $\delta_i$  and the horizontal axis shows quantity  $Q_i$  (on logarithmic scales). Higher quality attracts more patients, so demand is upward-sloping.<sup>9</sup> We draw two cases of the free-entry isocost curve defined by equation (4): the horizontal line depicts constant returns ( $\alpha = 0$ ) and the upward-sloping line depicts increasing returns ( $\alpha > 0$ ). With constant returns, a rightward shift in demand ( $\Delta N_j > 0$ ) causes a proportional increase in quantity produced and no change in output quality. With increasing returns, the demand shift elicits higher quality because producers move up the isocost curve and thus implies a more-than-proportional increase in quantity produced because the share of patients receiving care rises.

Panel 1(b) shows that an increase in demand raises quality more as the demand curve is increasingly elastic. The panel depicts two demand curves: the one on the left is more elastic, as we would expect for a less-common procedure.<sup>10</sup> Shifting each demand curve to the right raises the equilibrium quality of each procedure because of increasing returns to scale. This market-size effect is larger for the less common procedure with more elastic demand because

---

<sup>8</sup>For the equilibrium to be Marshallian stable, the demand curve must be steeper than the isocost curve at the intersection. There is a stable equilibrium because equation (5) means  $Q_{jj} \rightarrow N_j$  as  $\delta_j \rightarrow \infty$ .

<sup>9</sup>For visual clarity, we draw a log-linear demand curve. The logit demand function (5) is in fact log-convex, which is consistent with all the comparative statics illustrated in Figure 1.

<sup>10</sup>The demand function (5) is log-convex, so demand is indeed more elastic at lower quality. This is a fixed-price counterpart of Marshall's second law that demand is more elastic at higher prices.

the demand shift is amplified by a larger increase in quantity demanded.<sup>11</sup>

## 1.5 Market-size effects on trade flows

We now consider trade. With multiple regions and finite trade costs ( $\rho_{ij} > 0$ ), some patients will engage in trade—*i.e.*, select a provider located in another region. This trade stems from two sources. First, in the logit demand system with finite trade costs, patients have idiosyncratic preferences that yield a strictly positive probability of choosing every region. Second, when quality varies, regions producing higher-quality services attract more patients.

Fixing the qualities produced in other regions, an increase in one region’s demand affects its trade flows through three mechanisms. First, greater demand for services directly raises a region’s demand for imports through the  $N_j$  term in equation (1). A larger population translates proportionally to a greater demand for imports. Second, with increasing returns, an increase in  $N_i$  elicits an increase in quality  $\delta_i$ , which raises region  $i$ ’s *gross* exports to each region. Costinot et al. (2019) call this the “weak home-market effect.” Third, if increasing returns are sufficiently strong, the increase in quality  $\delta_i$  improves region  $i$ ’s patient market access  $\Phi_i$  so much that  $\ln \delta_i$  rises more than  $\ln \left( \frac{N_i}{\Phi_i} \right)$  does. That is, the increase in region  $i$ ’s gross exports exceeds any increase in its gross imports. This is the “strong” home-market effect: an increase in local demand raises a region’s *net* exports.

Figures 1(c) and 1(d) introduce trade and illustrate the distinction between weak and strong home-market effects.<sup>12</sup> Panel 1(c) depicts the quality and quantity produced in one region under two scale elasticities. Comparing points  $B$  and  $C$ , we see that a given increase in demand elicits a larger quality improvement when increasing returns are stronger. Panel 1(d) depicts equilibrium exports and imports as a function of the region’s demand shifter  $N_j$ . The import curves are upward-sloping because an increase in domestic demand raises demand for

---

<sup>11</sup>Alternatively, one could obtain this prediction by assuming that demand is log-linear and the isocost curve is log-concave. A rightward shift in demand would cause a larger (log) difference in quality for the low-volume procedure on the steeper part of the isocost curve.

<sup>12</sup>These diagrams are fixed-price analogues of Figures II and III in Costinot et al. (2019). See their discussion of the assumption that one region is large enough to affect its own quality but too small to affect the quality produced in other regions. This assumption is only made for this figure.

imports. The export curves are upward-sloping because of increasing returns: an increase in domestic demand causes an increase in quality, which causes an increase in gross exports. This is the weak home-market effect. When the scale elasticity  $\alpha$  is larger—the free-entry isocost curve in Figure 1(c) is steeper—greater demand elicits a larger increase in output quality, which steepens the export curve and flattens the import curve in Figure 1(d). When the export curve is steeper than the import curve, there is a strong home-market effect: the increase in demand raises exports more than imports.

We predict larger effects of market size for less common procedures. When two procedures have the same production function and trade costs, demand is more elastic at the rare procedure’s equilibrium quantity. As Figure 1(b) shows, an increase in demand raises quality more when the demand curve is more elastic, leading to a stronger home-market effect for the rarer procedure.

If rare procedures also have greater economies of scale (higher  $\alpha$ )—for example, because they require specialized equipment—that would amplify this contrast. This result motivates a difference-in-differences research design: we compare the market-size effects of common and rare procedures.

These results continue to hold when an increase in demand in one region affects equilibrium outcomes in all other regions. To demonstrate this, we consider the isoelastic special case with scale elasticity  $\alpha \in (0, 1)$  and examine the home-market effect in the neighborhood of a symmetric equilibrium. Suppose all regions are the same size,  $N_i = \bar{N} \forall i$ , and trade costs are symmetric:  $\rho_{ii} = 1$  and  $\rho_{ij} = \rho \in (0, 1) \forall i \notin \{0, j\}$ . There is a symmetric equilibrium, which has quality  $\bar{\delta}$  and patient market access  $\bar{\Phi}$  in each region. As detailed in Appendix A.2, we totally differentiate the system of equations in terms of  $\{d\delta_i, dN_i\}_{i=1}^I$  and evaluate this system with  $dN_1 > 0$  and  $dN_j = 0 \forall j \neq 1$  at the symmetric equilibrium.

With increasing returns of any magnitude, there is a weak home-market effect; with sufficiently strong increasing returns, there is a strong home-market effect. When  $\alpha > 0$ , an increase in the population size of region 1 elicits an increase in the quality of service

produced in region 1 relative to the other regions:

$$d \ln \delta_1 - d \ln \delta_{j \neq 1} = \left[ \frac{1 - \alpha}{\alpha} \frac{(\bar{\Phi} - 1)}{(1 - \rho)\bar{\delta}} + \frac{(1 - \rho)\bar{\delta}}{\bar{\Phi}} \right]^{-1} d \ln N_1 > 0.$$

This higher quality causes region 1 to export more to every other region:  $\frac{d \ln Q_{1j}}{d \ln N_1} > 0$ . The effect on the region's net exports is

$$d \ln Q_{1,j \neq 1} - d \ln Q_{j \neq 1,1} = \left[ \frac{1 - \frac{1 - \alpha}{\alpha} \frac{1 + (\mathcal{I} - 1)\rho}{1 - \rho}}{\frac{1 - \alpha}{\alpha} \frac{1 + (\mathcal{I} - 1)\rho}{(1 - \rho)} + \frac{(1 - \rho)\bar{\delta}}{1 + (1 + (\mathcal{I} - 1)\rho)\bar{\delta}}} \right] d \ln N_1. \quad (6)$$

Net exports increase if and only if

$$\frac{\alpha}{1 - \alpha} > \frac{1 + (\mathcal{I} - 1)\rho}{1 - \rho}.$$

When this inequality holds, the larger population size of region 1 makes it a *net* exporter of the medical procedure; *i.e.*, the procedure exhibits a strong home-market effect around the symmetric equilibrium. This occurs if increasing returns are sufficiently strong ( $\alpha$  is large enough) and trade costs are sufficiently large ( $\rho$  is small enough). Otherwise, there is a weak home-market effect but not a strong one. Given a strong home-market effect, the effect in equation (6) is diminishing in the number of potential patients  $\bar{N}$ , so we predict a stronger home-market effect for less common procedures.

While the existence of increasing returns seems likely—at least for some types of medical care—there is no guarantee they are sufficiently large to generate a strong home-market effect. When larger markets are net exporters, they produce care that smaller regions need. This trade can also support the larger markets' economies: rather than exporting manufactured goods, as in decades past, larger cities can reinvent themselves (Glaeser, 2005) and export medical services. Absent a strong effect, healthcare would be a net import, not an economic base, for larger regions.

## 2 Data description

Our primary dataset is 2017 claims data from Medicare, the US federal government’s insurance program for the elderly and disabled. Medicare is the largest health insurer in the United States. It does not directly employ physicians or run its own hospitals. Instead, it pays bills submitted by independent physicians, physician groups, hospitals, and other medical service providers. These bills—called “claims” in industry terminology—report the specific services provided using 5-digit codes from the Healthcare Common Procedure Coding System (HCPCS). There are over 12,000 distinct HCPCS codes, which identify individual procedures at a granular level.<sup>13</sup> Federal regulation determines the payment for each claim, rather than physicians’ or hospitals’ pricing decisions. In alternative analyses we use groupings of patient *diagnoses* to account for potential substitution between treatments.<sup>14</sup>

The claims data report the geographic location of both the physician providing the care and the patient receiving it, allowing us to construct a trade matrix for medical services. We study all medical care provided by physicians outside an emergency room, whether in an office or hospital facility.<sup>15</sup> Because Medicare rarely reimbursed telehealth in 2017, this trade involves traveling to receive a service delivered in-person.<sup>16</sup> We aggregate the ZIP-code-level information up to 306 hospital referral regions (HRRs), which are geographic units defined by the Dartmouth Atlas Project to represent regional health care markets for tertiary medical care based on 1992–93 data. We construct HRR-to-HRR trade flows by interpreting the patient’s residential HRR as the importing region and the service location’s HRR as the

---

<sup>13</sup>For instance, there are distinct codes for providing flu vaccines based on patient age, whether the vaccine protects against three or four strains of flu, and whether administration is intramuscular or intranasal. There are distinct codes for chest X-rays based on whether the images are of ribs, the breastbone, or the full chest, both sides or one side of the body, and the number of images taken (1, 2, 3, or 4+).

<sup>14</sup>We use the Clinical Classifications Software Refined (CCSR) diagnosis categories produced by the Agency for Healthcare Research and Quality’s Healthcare Cost and Utilization Project. CCSR aggregates over 70,000 ICD-10-CM diagnosis codes into “clinical categories,” of which 482 have at least 20 patients each in our data. We split these categories at the median frequency to separate common from rare diagnoses.

<sup>15</sup>Our results are robust to adding the value of hospital facility fees on top of physicians’ professional fees.

<sup>16</sup>In 2012, Medicare spent only \$5 million—less than 0.001% of its expenditures—on telehealth services (Neufeld and Doarn, 2015), lagging other insurers (Dorsey and Topol, 2016).



exporting region.<sup>17</sup> The Dartmouth Atlas Project defines HRRs by aggregating residential areas based on where patients were referred for major cardiovascular surgical procedures and for neurosurgery and requires each HRR to have at least one city where both major cardiovascular surgical procedures and neurosurgery were performed. Thus, the construction of these geographic units should tend to minimize trade between different HRRs.<sup>18</sup>

Physicians, hospitals, pharmacies, and other healthcare providers submit different types of claims. We use a random 20% sample of all physician claims paid by Traditional (fee-for-service) Medicare in 2017, selected randomly by patient.<sup>19,20</sup> One year of data from this sample contains 229 million services, representing \$19 billion in spending. The Medicare claims are not perfectly representative of all US healthcare, since Medicare beneficiaries are elderly or disabled. But the geographic distribution of Medicare beneficiaries is quite similar to the overall population, and Medicare alone finances one-fifth of medical spending. So it is likely to capture the key features of overall healthcare production and consumption.

Since we only see a sample of Medicare data—and hence an even smaller share of overall medical care—we might completely miss physicians or procedures so rare that a 20% sample includes none of them in a particular location. We use two other sources to address this concern. First, we use a less-detailed but more comprehensive extract of Medicare data (based on all Traditional Medicare patients) to replicate some of our analyses and obtain extremely similar findings.<sup>21</sup> Second, we use physician registry data to study the geographic

---

<sup>17</sup>The Medicare claims are US patients receiving care at US service facilities. These data do not report any international transactions. Throughout this paper, “imports” and “exports” refer to domestic transactions between regions of the United States.

<sup>18</sup>We have also used alternative geographies, including core-based statistical areas (CBSAs) and metropolitan statistical areas, a subset of CBSAs that excludes the smaller micropolitan areas. Because these yield consistent findings, we do not report all such estimates.

<sup>19</sup>We also use data from 2011 to 2016 to investigate trade patterns over time in Appendix Figure D.1.

<sup>20</sup>One-third of Medicare patients opt out of the traditional version of Medicare, where care is paid directly by the government, in favor of a private insurance scheme (“Medicare Advantage”). In these private schemes, the government pays the insurer a fixed amount per patient and the insurers are responsible for the patient’s care. Because Medicare does not pay claim-level bills in these private insurance schemes, the availability and quality of data for the privately insured patients is lower. We exclude these patients from our analysis.

<sup>21</sup>Appendix B.1 explains why we must use the 20% sample and uses the 100% data to confirm some of our measures. It also shows that the relative frequencies of services purchased by private insurance are similar to those in Medicare.

patterns of production by specialty. These data provide the ZIP code and specialty of all physicians registered to practice in the United States. Physician specialty is conceptually distinct from medical service—and there is not a one-to-many mapping of specialties to services, since many services can be provided by physicians of different specialties—but we expect many of the same economic forces to apply at the level of physician specialties.

### 3 Is there a home market effect in medical services?

This section estimates how scale economies and trade costs shape the geography of aggregate healthcare production and consumption. Section 3.1 documents size-related spatial variation in both production and consumption. Section 3.2 shows that bilateral trade declines with distance. Section 3.3 describes our empirical strategy, which identifies the consequences of market size using gravity equations to model bilateral trade flows of medical services. Section 3.4 reports the empirical estimates, which demonstrate a strong home-market effect.

#### 3.1 Spatial variation in production and consumption

Figure 2 shows maps of healthcare production and consumption across regions. The consumption map shows the substantial geographic variation that has been well-documented by the Dartmouth Atlas and related literature on geographic variation in healthcare (Fisher et al., 2003*a,b*; Finkelstein, Gentzkow, and Williams, 2016). The production map shows even more pronounced variation: more production in large urban agglomerations and less in rural areas. There is substantial variation in production even between neighboring regions, while spatial variation in consumption is smoother.

The subsequent panels show patterns of trade, which constitutes the difference between production and consumption. Nationally, 22.4% of production is exported to a patient in another region.<sup>22</sup> Panel 2(c) indicates whether each HRR is a net importer or net exporter.

---

<sup>22</sup>This value is nearly identical whether measured across HRRs or across CBSAs. Appendix Figure D.1 shows that the exported share rose steadily from 18.6% in 2011 to its 2017 level of 22.4%. For manufactured

Net-exporting regions tend to be major urban agglomerations, plus places such as Rochester, Minn. and Hanover, N.H. that specialize in healthcare. Panel 2(d) shows gross exports as a share of local production for each HRR. Three-quarters of services produced in the Rochester metropolitan area, home to the top-rated Mayo Clinic, are provided to patients from other regions, who travel an average of 545 km to Rochester. As a major healthcare exporter with a population of merely 220,000, Rochester is an outlier: larger regions are responsible for a disproportionate share of medical services production.

Figure 3 plots the average production and consumption per capita across HRRs of different sizes. Both rise monotonically with population. Production rises about twice as steeply, with a population elasticity of 0.13 versus 0.06 for consumption. The difference between production and consumption is net trade: larger markets are net exporters and smaller markets are net importers. Gross trade flows exceed net trade flows, with imports comprising about one-third of consumption in the smallest regions. Exports per capita are approximately flat, which means total exports are increasing with local population. Imports per capita decline with an elasticity of  $-0.25$  with respect to population.

### 3.2 Bilateral trade and bilateral distance

Despite the clear patterns in Figure 3, geographic variation in trade is far from entirely explained by market size. The four regions with the lowest export shares are Anchorage, Honolulu, and Yakima and Spokane, Wash., likely reflecting their isolated geographic locations. The highest export shares are in Rochester, Minn., Ridgewood, N.J. (just outside of New York City), and Hinsdale, Ill. (just west of Chicago). Other than Rochester—home to the Mayo Clinic—these exporting regions are all on the edge of major metropolitan areas and serve patients from those metros’ hinterlands. To ensure our analysis captures these geographic patterns, we next examine bilateral trade flows.

Figure 4 depicts how trade varies with the distance between the patient and place of

---

goods, the export share across CBSAs is about 68%.

service. Figure 4(a) shows the distribution of distances patients travel for care, distinguishing between care provided in the patient’s home region and other regions.<sup>23</sup> Within HRRs, there is a narrow distribution of distances that peaks around 10 km. When visiting providers in a different HRR, patients travel a great variety of distances. There is a local plateau between approximately 30–100 km, suggesting a fair amount of travel to nearby HRRs, perhaps indicating regional medical centers. There is another substantial peak at thousands of kilometers, demonstrating substantial long-distance travel for care.<sup>24</sup> Patients’ willingness to travel these distance underpin our revealed-preference estimates of regional service quality.

Figure 4(b) shows that trade declines with distance. The blue curve depicts trade volume against distance (for pairs of HRRs with positive trade flows) after removing fixed effects for each exporter and each importer.<sup>25</sup> This intensive-margin relationship is roughly log-linear. The red curve shows the extensive margin: the share of pairs with positive trade as a function of distance. This is 100% for nearby pairs and under 60% for the most distant pairs. These patterns motivate the inclusion of distance covariates in our gravity-based analysis.

Patients may vary in their ability or willingness to travel, especially by socioeconomic status. We quantify it here, to the extent feasible in our data, for use in counterfactual scenarios and interpreting welfare implications. Figure 4(c) depicts distance elasticities estimated separately by neighborhood income decile.<sup>26</sup> We find a strong, nearly monotonic relationship between socioeconomic status and the distance elasticity: patients from the highest neighborhood-income decile exhibit a distance elasticity 25% smaller than those in the lowest decile.<sup>27</sup> This means patients from higher-income neighborhoods are more amenable

---

<sup>23</sup>For travel within an HRR, we use the distance between the centroids of the patient’s residential ZIP code and the ZIP code of the service location. We obtain the centroid coordinates from the Census Bureau’s corresponding ZIP code tabulation areas (ZCTAs). For travel across HRRs, we use ZCTA-to-ZCTA distances when they are within 160 km, and (for computational ease) use HRR-to-HRR distances beyond 160 km.

<sup>24</sup>The average patient travels 500 km to Chicago and 605 km to New York City, compared with less than 135 km to Urbana-Champaign, Ill. or Charlottesville, Va. An older literature cited in Dranove and Satterthwaite (2000) finds that patients who travel farther to hospitals tend to incur higher treatment costs.

<sup>25</sup>This application of the Frisch-Waugh-Lovell theorem is only feasible for positive trade volumes.

<sup>26</sup>Our data do not contain patients’ wealth or income, so we use their residential ZIP code. We split ZIP codes into deciles by median household income and estimate equation (12) separately by decile.

<sup>27</sup>These estimates are consistent with the interaction that Silver and Zhang (2022) estimate between income and distance to care. These differences in distance elasticities are not driven by differences in the

to travel for medical care. Thus, the benefits of agglomeration—higher-quality rare care produced in major centers—may not be shared evenly. This is especially notable given the empirical setting: Medicare insures the near-universe of elderly and disabled Americans.

### 3.3 Gravity-based empirical strategy

We base our empirical examination of trade flows on a gravity equation that summarizes the geography of demand. We obtain this equation from the model by assuming the region-pair component in equation (1) satisfies  $\ln \rho_{ij} = \gamma X_{ij} + v_{ij}$ , where  $X_{ij}$  is a vector of observed trade-cost shifters and  $v_{ij}$  is an orthogonal unobserved component. Taking expectations and then logs yields gross bilateral trade flows:

$$\ln \mathbb{E}(\overline{RQ}_{ij}) = \ln \delta_i + \ln \left( \frac{N_j}{\Phi_j} \right) + \gamma X_{ij}. \quad (7)$$

The left side of (7) is the value of procedures exported from region  $i$  to patients residing in  $j$ . We specify the first two right-side regressors as either observable demand shifters or fixed effects in different specifications described below. We generally parameterize observed trade-cost shifters as containing log distance and a same-region dummy, so that  $\gamma X_{ij} = \gamma_1 \ln \text{distance}_{ij} + \gamma_0 \mathbf{1}(i = j)$ . Alternative specifications include  $(\ln \text{distance}_{ij})^2$  or replace these continuous distance covariates with indicators for distance deciles.

When using the total value of bilateral exports as the dependent variable in (7), we aggregate quantities across thousands of distinct medical procedures using the average national Medicare reimbursement rate for each procedure. This produces an expenditure measure independent of any spatial variation in reimbursement rates.<sup>28</sup> We also estimate procedure-level versions of (7) for selected procedures, such as LVAD insertion and screening colonoscopy. The dependent variable in this case is the procedure count and no aggregation

---

composition of procedures. When we estimate elasticities separately for rare and common services—or even for individual procedures (see Appendix Table D.1)—the income gradient of distance elasticities persists.

<sup>28</sup>Mechanically, we multiply the quantity of each procedure by the national average price for that procedure and denote the sum across all procedures by  $\overline{RQ}_{ij}$ .

is required. Since observed bilateral trade is zero for many pairs of regions, especially when looking at trade in individual procedures, we estimate (7) using Poisson pseudo-maximum-likelihood (PPML; Santos Silva and Tenreyro, 2006).

We test for a home-market effect in medical services using population as an observed demand shifter. Following Costinot et al. (2019), we differentiate the system of equations (2) and (3) around the symmetric equilibrium. This delivers the local relationship between trade and population, independent of market access  $\Phi_j$ . The estimating equation is

$$\ln \mathbb{E} [\overline{RQ}_{ij}] = \lambda_X \ln \text{population}_i + \lambda_M \ln \text{population}_j + \gamma X_{ij}. \quad (8)$$

Relative to (7), equation (8) replaces  $\ln \delta_i$  and  $\ln \left( \frac{N_j}{\Phi_j} \right)$  by log population in the producing and consuming regions, respectively. A positive coefficient  $\lambda_X > 0$  implies a weak home-market effect as defined in Costinot et al. (2019): *gross* exports increase with market size. If  $\lambda_X > \lambda_M > 0$ , the home-market effect is strong: *net* exports increase with market size.

One potential concern with estimating (8) directly is reverse causality. Suppose that success in exporting medical services serves as an employment base that raises current population size, as epitomized by “anchor institutions.” For example, William Worrall Mayo settling in Rochester, Minn. in the 1860s, and subsequent investment in medical care and reputation, helps explain Rochester’s current population (Clapesattle, 1969).

We use two instrumental variables to address this concern. First, we use historical population. Medicine was a far smaller industry in 1940, and it is implausible that it could have driven local population in the way it might today. Since population is persistent over time, population in 1940 predicts contemporary population, and we are interested in capturing any effects of historical population that operate through current population. We therefore instrument for both the exporting region’s and importing region’s contemporaneous log populations with the respective log populations in 1940.

Our second instrument goes farther back than 1940 and uses local geology to predict

population. Rosenthal and Strange (2008) and Levy and Moscona (2020) show that shallower subterranean bedrock makes construction easier, leading to higher population density. Bedrock depth also predicts population size, so we use this as a second instrument for local demand, again for both the importing and exporting regions.<sup>29</sup>

### 3.4 A strong home-market effect in medical services

Table 1 reports the results of estimating (8). The first column shows significant, positive coefficients on both patient and provider market population. The coefficient on provider-market population is two-thirds greater than that on patient-market population. This demonstrates what Costinot et al. (2019) term a *strong home-market effect*. Not only does a larger population increase gross exports, but it does so more than it increases gross imports by local patients. The distance elasticity of medical services trade between hospital referral regions is -1.7. This is substantially larger than the distance elasticity of -0.95 estimated for trade in manufactures between CBSAs (Dingel, 2017).<sup>30</sup> This suggests that trade in personal services incurs greater distance-related costs, relative to the degree of product differentiation across regions, than trade in manufactured goods. The most obvious difference is that patients themselves must travel to the provider.

The next two columns of Table 1 demonstrate that more flexible distance-covariate specifications do not alter the result. Column 2 introduces the square of log distance as an additional covariate. Column 3 replaces the parametric distance controls with dummies for deciles of distance. The result is stable across the columns: gross and net exports both increase with market size. The magnitudes are stable in columns 2 and 3, and the magnitude of gross (though not net) exports increases when excluding zeros.

---

<sup>29</sup>This instrument is currently only available for CBSAs, but not for HRRs. We demonstrate that our main results are robust to defining markets based on CBSAs and to using both instruments at this level. Levy and Moscona (2020) show that the instrument has ample first-stage power for predicting population density; the same is true for our endogenous variables (population levels).

<sup>30</sup>We find a distance elasticity of medical services trade between CBSAs of -2.3. The analogous elasticity of health care and social assistance services trade between Canadian provinces is -1.42 (Anderson, Milot, and Yotov, 2014). The distance elasticity of international trade is typically near -0.9 (Disdier and Head, 2008).

The last column of Table 1 uses the historical population instrument to address concerns about reverse causality. We obtain similar home-market-effect estimates to our baseline results. Appendix Table D.2 reports similar results estimated using CBSAs rather than HRRs as our geographic unit. It also shows the CBSA-based results are robust to instrumenting with either historical population or bedrock depth. Appendix Table D.3 reports similar results when adding facility payments on top of physician fees.

The primary competing explanation for these results is other factors that reduce the cost of production  $w_i$  in larger markets. If doctors prefer to live in big cities (Lee, 2010), as college graduates generally do (Diamond, 2016), they could accept lower nominal wages and thus reduce healthcare production costs in such cities.

We investigate whether this mechanism is sufficiently large quantitatively to drive a net cost reduction in larger markets. We use data from Gottlieb et al. (2020) to measure the population elasticities of doctors' earnings and the American Community Survey (Ruggles et al., 2022) to examine other healthcare workers' earnings and real estate costs.<sup>31</sup> We confirm that doctors are cheaper in larger markets (Gottlieb et al., 2020), but other costs rise with population size. Appendix Figure D.2 shows that the population elasticity of doctors' earnings is -0.01, but that for non-physicians is 0.045. To compute the population elasticity of labor costs, we use ACS data to estimate that non-physician labor's share of healthcare production is three times as much as physician labor's share. The population elasticity of labor costs is thus positive. The higher cost of real estate in larger markets reinforces these higher labor costs. This spatial variation in costs undercuts the idea that amenities make production cheaper in larger markets.

A number of related phenomena do not threaten our results. If doctors accept lower wages because they prefer the sort of work available in healthcare agglomerations, this is not a confound. Rather, it is a mechanism increasing profitability in healthcare agglomerations: greater scale lowers the cost of an input. Similarly, teaching hospitals are not a

---

<sup>31</sup>Appendix B.2 discusses subtleties of the income data.



confounder. Teaching hospitals tend to be large, suggesting an agglomeration benefit of combining training with treatment at scale. Indeed, medical training exposes trainees to a large volume of patients so that they learn clinical skills by practicing them. The most salient example is Cornell University: after an abortive attempt to have medical training in both Ithaca and New York City, the Cornell Trustees quickly closed down the Ithaca location and centered the medical school in New York—where the patients and doctors were more abundant—in the early 20th century (Flexner, 1910; Gotto and Moon, 2016). As this history illustrates, the potential local demand for care can drive the location of medical training.<sup>32</sup> If academic hospitals attract doctors, and their location is driven by market size, they are part of the agglomeration mechanism, not a confounder.

One final concern is measurement error in Medicare’s records of patients’ residences. To address this, Appendix B.4 first demonstrates our results’ robustness to excluding states with large seasonal populations. Second, we examine how far dialysis patients appear to travel. We find that residential measurement error is limited and does not drive our results.

## 4 Comparing rare and common services

Because our model predicts larger home-market effects for rarer procedures, comparing market-size effects by service frequency is a finer test of our theory. Section 4.1 examines how spatial variation in the production and consumption of each procedure relates to market size. Section 4.2 generalizes our gravity-based regression analysis to estimate home-market effects separately for rare and common procedures.

---

<sup>32</sup>In general education, in contrast, university placement induces economic growth (Moretti, 2004).

## 4.1 Spatial variation in production and consumption by frequency

We estimate the population elasticity of production and consumption per Medicare beneficiary for each procedure.<sup>33</sup> We find that production rises with market size more than consumption, especially for less common procedures.

### 4.1.1 Method

We first estimate the population elasticity of production per Medicare beneficiary for each procedure. Let  $Q_{pi}$  denote the count of procedure  $p$  produced in region  $i$  and its national volume be  $Q_p = \sum_i Q_{pi}$ . Let  $M_i$  denote the number of Medicare beneficiaries residing in  $i$ . For each procedure  $p$ , we estimate the following relationship across regions:

$$\ln \mathbb{E} \left[ \frac{Q_{pi}}{M_i} \right] = \zeta_p + \beta_p \ln \text{population}_i. \quad (9)$$

The estimated population elasticity of production per beneficiary,  $\hat{\beta}_p$ , describes how production varies with market size, and we estimate it using Poisson pseudo-maximum-likelihood.<sup>34</sup> If the quantity produced were simply proportional to population,  $\beta_p$  would be zero.

Our model suggests that scale effects play a larger role for rarer procedures. It predicts less common services will have higher population elasticities of production. We therefore estimate a linear regression relating  $\hat{\beta}_p$  to the total national volume of service  $p$ ,  $\ln Q_p$ .

To summarize size-linked variation in consumption patterns, we separately estimate the population elasticity of *consumption* per beneficiary for each procedure. That is, we estimate a Poisson model in which the outcome variable is the count of procedure  $p$  consumed by

---

<sup>33</sup>Davis and Dingel (2020) relate population elasticities to other measures of geographic concentration, such as location quotients, and estimate population elasticities of employment for various skills and sectors.

<sup>34</sup>In a robustness check, we have also estimated a zero-inflated Poisson model, to account for the possibility that fixed costs are especially important for the decision of whether to provide the first instance of a service in a region. These results (not reported here) are quite similar.

patients residing in region  $i$ ,  $G_{pi}$ , per Medicare beneficiary residing there:

$$\ln \mathbb{E} \left[ \frac{G_{pi}}{M_i} \right] = \zeta_p^C + \beta_p^C \ln \text{population}_i. \quad (10)$$

If  $\beta_p^C \neq \beta_p$ , there is size-predicted net trade in procedure  $p$ . Our model predicts that procedure frequency influences the pattern of trade, a prediction we test in Section 4.2.

#### 4.1.2 Results

Production per beneficiary rises with market size, especially for less common procedures. Figure 5(a) relates the population elasticity of production per beneficiary  $\hat{\beta}_p$  for each procedure to its national volume  $\ln Q_p$ . Across all volumes, procedure output per beneficiary increases with market size. Less common procedures have higher elasticities, consistent with economies of scale that decline with quantity.

This finding raises questions about patients' access to care. What happens to patients who live in smaller markets but need rare services? To investigate this question, we estimate equation (10), the population elasticity of *consumption* per beneficiary of each procedure.

The population elasticity of consumption per beneficiary is smaller for the vast majority of procedures and less steeply related to a procedure's national frequency. Figure 5(a) also plots the population elasticity of consumption per beneficiary  $\hat{\beta}_p^C$  for each procedure against its national volume  $\ln Q_p$ . While the relationship is negative, the slope for consumption is only one third that for production. Appendix Table D.7 reports the production, consumption and trade patterns for two exemplar procedures: screening colonoscopy and LVAD implantation. Colonoscopies are common and geographically dispersed, while LVAD procedures are rare, geographically concentrated, and traded over longer distances.

We have thus far modeled patients as demanding (and providers as producing) specific service codes. An alternative view is that patients have a particular medical condition that requires treatment, but the patients may not know what particular care they need; they

simply know they require care. As physicians might use different treatments across regions for the same condition, our estimates thus far could reflect substitution among procedures. We address this by conducting a similar analysis at the level of clinical condition.

Figure 5(b) shows production and consumption elasticities by diagnosis, rather than by procedure. The key patterns remain similar: production elasticities are higher than consumption and decline more rapidly with national patient volume. Both consumption and production elasticities have less steep relationships with national volume than for procedures. This could reflect measurement error within each category: the 482 diagnosis categories we use are far coarser than the 8,253 procedures in Figure 5(a). Alternatively, it could indicate true substitution among procedures within a condition that varies with location.

The contrasting population elasticities of production and consumption summarized in Figure 5 imply trade in medical services between markets of different sizes. Just as theories of trade with scale effects would predict, larger markets export rare procedures and smaller markets import them. For almost all procedures, production increases more than proportionately with market size. Consumption also increases more than proportionately with market size, but much less so than production. The differences between these elasticities mean net exports vary with market size. The implied net trade between markets of different sizes is particularly large for procedures that have small national volumes.

## 4.2 Market-size effects are stronger for less common procedures

Procedure-level variation in bilateral trade provides a finer test of how market-size effects depend on a procedure's frequency. Appendix Figure D.3(a) shows a wide distribution of imports as a share of consumption by procedure.<sup>35</sup> We divide procedures into two equal-sized groups, common and rare, based on the quantity produced nationally and show each group's distribution of import shares across regions in Panel D.3(b). The difference is dra-

---

<sup>35</sup>This kernel density plot exhibits a spike at just above 20%, indicating that trade is, quite common in most procedures. There is a long tail reaching all the way to 1 and also many procedures with few or even zero imports.

matic: rare procedures (those with national frequency below the median) have much higher import shares, while the common procedures are overwhelmingly lower.<sup>36</sup> To formally test for differences in home-market effects, we again employ gravity models.

### 4.2.1 Empirical strategy

To test the model’s difference-in-differences prediction for trade volumes, we estimate market-size effects separately for common and rare services. We compute trade flows between each HRR pair  $\bar{R}Q_{ijc}$  separately for these two categories of care,  $c \in \{\text{common}, \text{rare}\}$ . We thus have two observations for each  $ij$  pair, allowing us to estimate:

$$\begin{aligned} \ln \mathbb{E} [\bar{R}Q_{ijc}] &= \lambda_X \ln \text{population}_i + \lambda_M \ln \text{population}_j + \gamma X_{ij} \\ &+ (\mu_X \ln \text{population}_i + \mu_M \ln \text{population}_j + \psi X_{ij}) \cdot \mathbf{1}(c = \text{rare}). \end{aligned} \quad (11)$$

An alternate specification introduces  $ij$ -pair fixed effects, which absorb all the covariates not interacted with  $\mathbf{1}(c = \text{rare})$ . The theory from Section 1.5 predicts stronger market-size effects for rare procedures,  $\mu_X > 0$ .

### 4.2.2 Results

Table 2 reports estimates for a gravity regression in which each pair of location has two observations: one for rare services and one for common. Column 1 repeats our baseline regression from Table 1 but with this new structure and obtains identical results. Column 2 limits the sample to pairs of location that have positive trade in at least one of the two procedure groups, which is the estimation sample used in the remainder of the table. In columns 3 and following, we interact both provider-market and patient-market population with an indicator for rare services. We find significant and robust evidence that the home-

---

<sup>36</sup>Nationally, the imported share of consumption is 35% for below-median-frequency procedures and 22% for those above the median. Within both groups of procedures, there is substantial variation in import shares across hospital referral regions.

market effect is stronger for rare services. The coefficient on provider-market population increases by about 50% relative to common services. The coefficient on patient-market population shrinks by nearly half. Column 4 introduces location-pair fixed effects. Columns 5 and 6 are analogues of the previous two, but add a quadratic distance control. These results are statistically indistinguishable from the previous columns.

Table 3 shows that these results are robust to instrumenting for market size with either historical population or depth to bedrock. Columns 1 and 2 show estimates for common and rare services, respectively, when instrumenting for population in each region by its 1940 population. Columns 3 and 4 repeat the exercise using CBSAs rather than HRRs, and columns 5 and 6 switch to the bedrock-depth instrument. The results are consistent regardless of geographic unit or instrument. The estimates' stability suggests that neither the aggregate result nor the variation with procedure frequency is driven by anchor institutions or similar omitted variables.

The finding that less common procedures exhibit stronger home-market effects is robust to different ways of defining rare and common care. Table 4 demonstrates that our result holds when we look across diagnoses rather than procedures, and Appendix Table D.8 shows the same when including facility spending. As with the production and consumption elasticities in Figure 5(b), the magnitude of the difference between rare and common care shrinks. This could reflect substitution across care within a diagnosis or a less precise classification of diagnoses than of procedures. But the qualitative pattern holds and remains significant, consistent with the model's difference-in-difference prediction.

These findings reflect each procedure's national frequency, not how often an individual patient receives the same procedure. We call the latter concept the procedure's "engagement". If patients are less willing to travel for high-engagement services and these services are more common, higher engagement could drive the stronger home-market effect we observe for rare procedures. In fact, the national frequency of a service has a very low correlation

with various measures of engagement for that service, so it does not confound this result.<sup>37</sup> While the distance elasticity is more negative for high-engagement procedures, Appendix Table D.9 shows that separating high- from low-engagement procedures does not meaningfully alter the estimated differential impacts of population size for rare procedures.

Figure 6 returns to categorizing services by frequency, reporting estimates of (8) separately for each national frequency decile. The blue circles show estimated provider-market population elasticities, which decline monotonically from the least common to most common procedures. The red squares show patient-market population elasticities, which increase across the frequency distribution. The difference between the respective coefficients demonstrates a strong home-market effect for all deciles. This effect is stronger the less common the procedure. Appendix Table D.10 shows the same pattern among illustrative procedures.<sup>38</sup>

The potential concern about omitted cost shifters from Section 3.4 has an analogue here: Do the doctors who provide rare services benefit more from urban amenities than those providing common ones, lowering the cost of producing rare services in larger markets? This has facial plausibility if rare services are produced by elite specialists, who are higher-earning and more willing to pay for urban amenities through lower compensation.

Examining the population elasticities of physician earnings for each specialty alleviates this concern. If urban amenities drive specialists' locations, earnings elasticities should be negative, especially for rare specialties. But Appendix Figure D.4 shows that the income elasticities are close to zero on average and uncorrelated with the specialty's national abundance. However urban amenities affect physicians' choices, they do not exhibit the compensating differentials necessary to explain the relationship between market size and specialization.

---

<sup>37</sup>For example, the correlation between the share of patients who had more than one claim for the procedure in a given year and the procedure's national frequency is 0.14.

<sup>38</sup>We show two common procedures—screening colonoscopy and cataract surgery—along with four rare ones: two treatments for brain cancer, implantation of a left ventricular assist device (LVAD), and total colectomy. All six procedures exhibit strong home-market effects, but the differences between  $\hat{\lambda}_X$  and  $\hat{\lambda}_M$  are smaller for the common procedures than the rare ones.

## 5 Estimating the scale elasticity of quality

To estimate the scale elasticity of regional medical services production, we first estimate each region’s quality in Section 5.1. Section 5.2 describes our empirical strategy for estimating the scale elasticity, which Section 5.3 reports to be around 0.6 for aggregate medical services. Section 5.4 documents one mechanism linking these increasing returns and interregional trade: larger markets support a finer division of labor, and traded services are performed by more specialized and more experienced physicians.

### 5.1 Quality estimates

We use a two-step procedure, which begins by estimating a fixed-effects version of the gravity equation. In equation (7), the exporting region  $i$  component of the bilateral trade flow is its perceived service quality. We can thus estimate  $\ln \delta_i$  as the origin fixed effect in this gravity equation. Similarly,  $\ln \left( \frac{N_j}{\Phi_j} \right)$  can be estimated as a destination fixed effect, denoted  $\ln \theta_j$ . This implies the following estimating equation:

$$\ln \mathbb{E} (\bar{R}Q_{ij}) = \underbrace{\ln \delta_i}_{\text{exporter FE}} + \underbrace{\ln \theta_j}_{\text{importer FE}} + \gamma X_{ij}. \quad (12)$$

We interpret the exporter fixed effects  $\widehat{\ln \delta_i}$  as a revealed-preference measure of quality, an interpretation we validate using hospital rankings and measures of physician specialization. The importer fixed effects  $\widehat{\ln \theta_j}$ , plus an assumption about potential market size, enable us to compute  $\Phi_j = N_j / \hat{\theta}_j$ , a measure of patient market access for those who reside in location  $j$ . We also estimate (12) separately by service frequency, yielding  $\widehat{\ln \delta_i}^{\text{rare}}$  and  $\widehat{\ln \delta_i}^{\text{common}}$ .

To test whether  $\widehat{\ln \delta_i}$  reflects quality, the first three panels of Figure 7 compare the estimated exporter fixed effects to external measures of regional hospital quality. We count the number of times each region’s hospitals appear on *U.S. News Best Hospitals*.<sup>39</sup> We

---

<sup>39</sup>Appendix B.2 explains how we use these rankings.



also obtain Hospital Safety Grades from the Leapfrog Group and average them by HRR. The significant positive slopes in both Figures 7(a) and 7(b) show that patients prefer to obtain care from HRRs with better *U.S. News* rankings. There is also a positive relationship with Hospital Safety Grades, shown in Figure 7(c).<sup>40</sup> The positive relationships with both measures suggest that our estimates capture a meaningful measure of hospital quality.

The *U.S. News* rankings are intended to capture the “Best Hospitals,” a concept associated with providing highly specialized care. So it is natural that there is a stronger relationship between the *U.S. News* rankings and exporter fixed effects for rare services; the slope in Figure 7(b) is twice as large as that for common services in Figure 7(a).<sup>41</sup>

## 5.2 Empirical approach

We use the estimated exporter fixed effects  $\widehat{\ln \delta}_i$  to examine the determinants of regional service quality, in particular the scale elasticity,  $\alpha$ . In the free-entry condition (4), service quality in region  $i$  is an isoelastic function of the quantity produced, conditional on revenue, cost, and productivity shifters. Taking the log of (4) and rearranging terms yields an estimating equation for the quality-quantity relationship across locations:

$$\ln \delta_i = \alpha \ln Q_i + \ln \bar{R} - \ln w_i + \ln A_i. \quad (13)$$

Replacing  $\ln \delta_i$  with its estimate  $\widehat{\ln \delta}_i$  from (12) yields an estimating equation for  $\hat{\alpha}$ .<sup>42</sup>

One potential concern with estimating equation (13) by ordinary least squares is reverse causality. Shifts of the isocost curve would cause movements along the upward-sloping demand curve, biasing the estimated scale elasticity upwards. We address this with three

---

<sup>40</sup>The distance elasticity does not meaningfully vary with procedure frequency. This suggests that patients’ preference for a particular region loads onto the region fixed effects, consistent with our interpretation.

<sup>41</sup>In contrast, safety grades are not differentially relevant for rare services: Appendix Figures D.6(a) and D.6(b) show virtually identical slopes.

<sup>42</sup>Appendix B.5 quantifies the potential bias resulting from our observing only the quantity produced for Traditional Medicare beneficiaries, rather than the total quantity produced for all patients. It shows that the bias is small: the estimates in Table 5 should be deflated by about 5%.

instruments, starting with current population. Population is relevant for healthcare output and is valid if not correlated with healthcare quality other than by driving local demand. The “anchor institutions” concern discussed in Section 3.3 could violate this exclusion restriction, so we also use the historical population and bedrock-depth instruments.

Despite our instruments, other channels related to population size could generate the same relationship as the market-size effect we estimate. Most significantly, physicians might prefer to live in cities (Lee, 2010), regardless of patient demand. This could drive up quality in large markets, but through a different mechanism than the one we emphasize.

Before we address this problem, first note what is *not* a problem: physicians preferring to work in larger regions for job-related reasons. A larger population of patients allows physicians to specialize, conduct research, and train medical students. As discussed in Section 3.4, these forces operate through the scale of healthcare production in the region. Academic medical centers are often an important part of a region’s medical industry. If their scale attracts workers, this is an agglomeration benefit  $\alpha$  ought to capture.

The challenge to our interpretation arises if physicians prefer larger markets for non-professional reasons, and this labor supply shift increases quality. If urban amenities attract physicians—and higher-quality physicians in particular—this would represent variation in  $w_i$  or  $A_i$  that is correlated with population size and hence local output in equation (13). The analysis of local costs in Section 3.4 and Appendix Figure D.2 mitigates this concern.

### 5.3 Scale improves quality

Estimated service quality  $\widehat{\ln \delta_i}$  rises substantially with the regional volume of production  $\ln Q_i$ . Figure 7(d) depicts this relationship and Table 5 reports regression estimates. The estimated scale elasticity is around 0.6 and stable under various estimation approaches. The first row uses OLS, while subsequent rows instrument for output using contemporaneous or historical population. The first and third columns omit the diagonal  $Q_{ii}$  observations when estimating the gravity equation (12), to avoid any bias from having a region’s own

domestic consumption influence both the quality measures and output. The third and fourth columns control for spatial variation in reimbursements. Across twelve estimates, the lowest elasticity is 0.55 and the highest is 0.88. Instrumenting for output tends to reduce the estimated scale elasticity. Excluding the diagonal of the trade matrix when estimating quality tends to raise it. The results for CBSAs, reported in Appendix Table D.11, are also stable across specifications and when using the alternative bedrock instrument. While the existence of home-market effects implied local increasing returns, these estimates quantify their magnitude.<sup>43</sup> These estimates are central to our counterfactual calculations in Section 6.

The second panel of Table 5 estimates the scale elasticity for rare services. Section 1.5 shows that market-size effects are larger for rarer procedures even if all procedures have the same scale elasticity. These differences are amplified if rarer procedures have a larger scale elasticity than more common procedures. The scale elasticity is indeed substantially larger for rare services, with estimates centered around 0.9.

## 5.4 Scale facilitates the division of labor

One source of increasing returns—though certainly not the only one—could be division of labor among physicians. In particular, the specialized labor required to produce rare services could drive the patterns we found in Section 4 across treatments and diagnoses. Specialized services may require physicians with specific training, whom low demand in smaller HRRs may not support (Dranove, Shanley, and Simon, 1992).

### 5.4.1 Specialization as a source of local increasing returns

To study this mechanism, we estimate the population elasticity of physicians per capita for each specialization and relate it to the number of physicians in the specialization. Let  $Y_{si}$

---

<sup>43</sup>These estimates lie in the middle of other estimated agglomeration elasticities, Kline and Moretti (2013) estimate an elasticity of 0.4–0.47 from the Tennessee Valley Authority’s investments. In manufacturing, Greenstone, Hornbeck, and Moretti (2010) report an analogous elasticity above 1 (a 12% increase in total factor productivity caused by adding a plant representing 8.6% of the county’s prior output).

be the number of doctors of specialty  $s$  in location  $i$ .<sup>44</sup> We estimate a Poisson model,

$$\ln \mathbb{E} \left[ \frac{Y_{si}}{\text{population}_i} \right] = \zeta_s^S + \beta_s^S \ln \text{population}_i, \quad (14)$$

for each specialty  $s$  by maximum likelihood.

Figure 8(a) shows a clear negative relationship between a specialty’s per capita population elasticity  $\hat{\beta}_s^S$  and the national number of physicians in that specialization.<sup>45</sup> A natural explanation for rare procedures and rare specializations both being geographically concentrated in larger regions is that the size of the market limits the division of labor. To the extent that producing rare procedures requires specialized physicians, a larger volume of patients makes production economically viable.

Consistent with this idea, Appendix Figure D.5 shows the number of distinct procedures produced as a function of market size by procedure type. We group procedures into seven categories, count the number of procedures produced in each region, and project these onto regional population.<sup>46</sup> Larger regions produce a greater variety of procedures in all seven categories. If physicians specialize in particular procedures, this makes sense: larger markets have more specialties of physicians and thus a greater ability to provide rare procedures.

This evidence on specialization does not preclude other agglomeration mechanisms from also playing a role. Lumpy capital, knowledge diffusion (Baicker and Chandra, 2010), and thicker input markets could also be important productivity benefits of scale. We focus on specialization and physician experience because of their close link to the procedure-level

---

<sup>44</sup>Data come from the National Plan and Provider Enumeration System (NPPES) data, which cover all physicians, not just those serving Medicare patients. These data only report the number of doctors/specialists and their location, but contain no further information about procedures performed. We restrict attention to the 223 specializations within Allopathic & Osteopathic Physicians. We restrict attention to national provider identifiers of the “individual” entity type (as opposed to “organization”). We consider each physician’s primary specialty, as indicated in the NPPES file. Results (unreported) are similar when we allow for multiple specialties per physician, a common occurrence in the NPPES data.

<sup>45</sup>This pattern is not attributable to spatial sorting driven by rare specialties commanding higher earnings. In fact, a specialty’s number of physicians and mean earnings are uncorrelated. Appendix Table D.13 shows that controlling for a specialty’s earnings has no effect on the negative relationship between population elasticity and number of physicians across specialties.

<sup>46</sup>Appendix Table D.14 reports regression estimates for these relationships.

agglomeration we analyze and we can observe them in claims data.

#### 5.4.2 Imports are specialist-intensive

We next ask whether the distribution of specialties helps explain trade. Figure 8(b) shows the share of imports and of domestic consumption that are provided by specialists as a function of regional population.<sup>47</sup> Imports are significantly more specialist-intensive than domestic production. This difference is especially pronounced in the smallest regions, and it remains true throughout the population distribution.

Does trade match patients with the appropriate specialist? Among all specialty care, we determine the two most common specialties to provide each unique service and label these the “standard” specialties for that care. We then determine whether each instance of the treatment was provided by a standard or non-standard specialty.

Figure 8(c) shows the share of imports and of domestic consumption provided by the standard specialties. Imports are more likely to come from the standard specialist than domestic care, and the distinction is especially pronounced in the smallest regions. The difference is substantial: Domestic care in the smallest regions is 50% more likely to be provided by a non-standard specialist than in the largest regions (7.5% vs. 5%). When importing medical services, this share falls to 5%—indistinguishable from the largest regions’ domestic consumption.

We conduct a similar analysis based on provider experience. Using the public Medicare provider data (based on all Traditional Medicare patients), we count the number of times the physician billed for the specific service in the previous year. We divide this experience measure by the procedure’s national mean and average it across all procedures provided to patients in an HRR. Figure 8(d) shows that, at all population sizes, care imported from other regions is produced by more experienced providers than locally produced care.<sup>48</sup> Patients in

---

<sup>47</sup>We define “specialist” to mean all physicians except those whose primary specialty is internal medicine, general practice, or family practice.

<sup>48</sup>This comparison restricts attention to procedures that are performed in all hospital referral regions (143 procedures). Thus, regional variation does not reflect the fact that larger markets produce a greater number

larger regions see more experienced providers for both imported and locally produced care.

Specialists are disproportionately located in larger markets, as are physicians with more experience in any given procedure. Since imported care is predominantly specialty care, and provides patients access to this higher experience, we conclude that visiting the appropriate specialist based on training or experience is part of the value proposition for trade in medical care. This provides a second validation of our interpretation that trade reflects quality variation. Patients travel to regions with highly-ranked hospitals, which larger markets tend to have—along with the ability to provide rare services. This market-size effect strongly predicts gross and net exports. Together, this suggests that economies of scale play an important role in increasing the quality of care, and trade between regions enables patients from many regions to share the benefits of this agglomeration.

## 6 Tradeoffs and counterfactual scenarios

Given the estimated strength of local increasing returns, geographically concentrating health-care production has substantial benefits. Larger regions support specialists, house experienced physicians, and produce more specialized procedures. But this geographic concentration implies that patients in smaller regions may suffer from limited access to care. We use observed trade flows and our estimates of the scale elasticity  $\alpha$  and region-specific qualities  $\delta_i$  to quantify how various counterfactual policy scenarios would change each region’s patient market access. Our results underline the importance of distinguishing between the quality of locally produced services and the quality of services to which local residents have access.

We compute counterfactual equilibrium outcomes relative to the baseline equilibrium. For the baseline equilibrium, define export shares  $x_{ij} \equiv \frac{Q_{ij}}{\sum_{j'} Q_{ij'}}$  and import shares  $m_{ij} \equiv \frac{Q_{ij}}{N_j}$ . For every variable or parameter  $y$ , denote the ratio of its counterfactual value  $y'$  to its baseline value  $y$  by  $\hat{y} \equiv \frac{y'}{y}$ . Appendix C.1 shows how we solve for the relative counterfactual endogenous qualities ( $\hat{\delta}$ ) using baseline equilibrium shares  $(x_{ij}, m_{ij})$ , the scale elasticity ( $\alpha$ ), and of distinct codes (Appendix Figure D.5).

relative counterfactual exogenous parameters  $(\hat{A}, \hat{R}, \hat{w}, \hat{\rho}, \hat{N})$ . In particular, counterfactual qualities are given by a system of  $\mathcal{I}$  equations with unknowns  $\{\hat{\delta}_i\}_{i=1}^{\mathcal{I}}$ :

$$\hat{\delta}_i = \left( \hat{R}_i \hat{A}_i / \hat{w}_i \right)^{\frac{1}{1-\alpha}} \left( \sum_{j \in \mathcal{I}} \frac{x_{ij} \hat{\rho}_{ij} \hat{N}_j}{m_{0j} + \sum_{i' \in \mathcal{I}} m_{i'j} \hat{\delta}_{i'} \hat{\rho}_{i'j}} \right)^{\frac{\alpha}{1-\alpha}}.$$

The first term of this expression,  $\left( \hat{R}_i \hat{A}_i / \hat{w}_i \right)^{\frac{1}{1-\alpha}}$ , shows that the scale elasticity  $\alpha$  governs the effect of exogenous supplier shifters, including reimbursements  $\hat{R}_i$ , on quality produced in a region. Reimbursement rates shift the scale of production, and stronger scale economies (higher  $\alpha$ ) amplify these shifts. The second term shows how changes in other regions influence local outcomes through trade, combined with scale. Thus, our counterfactual scenarios rely on both our estimates of the scale elasticity  $\alpha$  and observed trade patterns.<sup>49</sup>

We first consider the impact of a nationwide change in reimbursements. Increasing reimbursements uniformly by 10% has heterogeneous effects. Figure 9(a) depicts the change in output quality in each region. Remote, rural areas tend to experience the largest increases in output quality  $\delta_i$ . Large regions such as Boston, New York, Atlanta, and Florida have the smallest increases, because they produce more care at baseline.

Figure 9(b) shows the impact on patient market access is nearly opposite: regions with the largest increase in output quality have the smallest improvements in market access. Their residents already had high import shares, so the least reliance on local production. The increase in local quality thus has limited impact on their overall market access. For patients who switch to consuming local care, the gains are modest as local production is still lower-quality than the care they otherwise import. In contrast, patients in Houston, Dallas, or Florida had limited reason to travel. The increase in  $\delta_i$  due to higher local reimbursements, even if modest, improves their access relatively more.

---

<sup>49</sup>In order to compute import shares, we assume that the number of potential patients is proportional to the number of enrolled Traditional Medicare beneficiaries. See Appendix C.2 for details. The qualitative and spatial patterns of counterfactual outcomes do not depend on what share of potential patients we assume choose the outside option. Appendices C.3 and C.4 generalize this method of computing counterfactual outcomes to the model with multiple types of patients introduced in Appendix A.1.

Figure 9(c) summarizes these contrasting changes in output quality and patient access. Regions with the lowest initial patient market access  $\Phi_i$  have the biggest increase in local production quantity and quality,  $\hat{Q}_i$  and  $\hat{\delta}_i$ , but the smallest increase in patient market access  $\hat{\Phi}_i$ . Appendix Figure D.7 conducts this exercise separately for rare and common services. The patterns are qualitatively similar, but the impacts on quality are much larger for rare services because of their larger scale elasticity ( $\alpha = 0.9$  rather than  $\alpha = 0.6$ ), more concentrated baseline production, and higher baseline trade shares.

These results help reconcile two notable aspects of US healthcare policy. First, a range of recent studies find medical outcomes that match our predictions: patients who travel farther for care in larger markets tend to have better outcomes (Petek, 2022; Fischer, Royer, and White, 2022; Battaglia, 2022). Second, there is nevertheless a major political and policy effort to subsidize production in rural areas.<sup>50</sup> Our contrasting results for output and access rationalize this pattern: *producers* in rural areas are especially dependent on high reimbursements. This naturally leads to political pressure to subsidize production in such places. But *patients* do not necessarily benefit. They would often benefit from traveling to larger markets for better care, suggesting that the emphasis on local production may not be efficient—even from the perspective of rural patients.

We next consider the impact of this nationwide reimbursement increase on different income groups, indexed by  $\kappa$ . We compute changes in market access for each region and income group, and rescale them into percentage changes,  $100(\hat{\Phi}_{j\kappa} - 1)$ . At the region-by-income-tercile level, Table 6 regresses these changes on income-group dummies (columns 1–3) and HRR fixed effects (columns 2–3). Column 1 shows that the market access gain for the highest income tercile is nearly 20% larger than for the lowest tercile. The lowest tercile experiences an 8.8% increase in patient market access (the constant in the regression). The highest tercile gains this plus an additional 1.6 percentage points. The difference is explained

---

<sup>50</sup>These policies include Critical Access Hospitals, Health Professional Shortage Areas, rural-biased adjustments to Medicare’s Geographic Practice Cost Index for physician work, hospital geographic reclassification for Medicare reimbursements, increasing residency slots in rural areas, and more federal and state programs. The effectiveness of these policies is not always clear (Khoury, Leganza, and Masucci, 2022; Falcettoni, 2021).



by differences in the groups' outside option shares,  $m_{0j\kappa}$ , as column 3 shows. Patients in the highest tercile are more likely to seek care, so benefit more from quality improvements.

Policies often target specific regions so we now examine how targeted production subsidies affect output quality and patient access. Figure 10 contrasts the consequences of raising reimbursements by 30% in Boston and in Paducah, Ky. Figure 10(a) depicts the impact on quality of care in each region relative to its baseline value in the Boston scenario. Free entry means that higher reimbursements translate to higher-quality care produced in Boston. Quality declines in the rest of New England as patients substitute away and scale economies translate lower volumes into lower quality (an “agglomeration shadow”, as in Fujita and Krugman, 1995). These effects diminish with distance to Boston.

Regions that experience larger declines in output quality due to Boston's expansion simultaneously experience larger improvements in patient market access. Figure 10(b) depicts the change in the value of patients' market access,  $\hat{\Phi}_i$ . Patients in Boston benefit the most from the higher reimbursement of their local production. Outside Boston, regional changes in patient market access are nearly opposite the changes in local output quality. The nearest regions import sufficient volumes that the benefits of improved quality in Boston exceed the declines in the quality of local production, causing their patient market access to improve. Regions closer to Boston experience larger declines in the quality of local production precisely because their residents' choice sets improve more, spurring more substitution. In more distant regions, the welfare impacts are neutral to ever-so-slightly negative.

We again see disproportionate gains for patients who live in higher-income neighborhoods, shown in columns 4–6 of Table 6. The value of market access increases by 0.098% for first-tercile patients nationwide; this is orders of magnitude lower than in columns 1–3 because only one region's reimbursement is increasing. Gains are 70% larger for third-tercile patients. Once again this can be explained by baseline trade shares, as column 3 shows.

The consequences of higher reimbursement rates in Paducah, Ky. exhibit very different spatial patterns than in Boston. Figures 10(c) and 10(d) depict the regional changes in

output quality and patient market access, respectively, caused by a 30% reimbursement increase in Paducah. Unlike Boston, Paducah is a net importer: its consumption of medical services exceeds local production by more than one-third. Higher reimbursements that improve output quality in Paducah cause Paducahans to reduce their imports from neighboring regions. This reduces the quantity produced in neighboring regions, lowering their output quality, similar to the regional spillovers in the Boston scenario. But Figure 10(d) shows that those regions where output quality declines more are the regions where patient market access declines more, contrary to the pattern of outcomes in the Boston scenario.

The contrasting outcomes reflect trade flows in the baseline equilibrium: Boston is a net exporter of medical services and Paducah is a net importer. Paducah imports one-third of its consumption, and Boston imports only six percent. Higher reimbursements in Boston cause output quality declines in nearby regions—largely because residents of those regions import more when Boston’s quality improves. In contrast, higher reimbursements in Paducah reduce neighboring regions’ output quality largely because Paducah residents demand fewer exports from these regions when Paducah’s quality improves. Nearby regions import little from Paducah, so they benefit little from its improved quality. Appendix Figure D.8 shows that the lessons from Boston and Paducah generalize: the pattern of spillovers from increasing reimbursements in one region is driven by that region’s net trade in medical care. To summarize, the spillover consequences of subsidizing production in one region depend on the pattern of trade; changes in regional output quality need not align with changes in regional patient market access.

The distributional consequences of region-specific subsidies depend on which region is subsidized. We compute the nationwide gains in market access from subsidizing production in each region, one at a time. Figure 11 shows this gain, scaled by the increase in total spending, as a function of region size. The aggregate gain in market access per dollar spent is higher in larger markets: further concentration of production has larger benefits. The graph also shows the gains per dollar separately by income tercile. Unlike the Boston scenario,

in which benefits accrue more to higher-income ZIP codes, subsidizing production in less populous regions benefits lower-income ZIP codes more. These contrasts reflect geographic divides in incomes: lower-income patients are more likely to live in and near smaller regions.

Rather than subsidizing local production, policies might improve patient market access in a particular region by facilitating trade. We examine the consequences of a policy that reduces travel costs for Paducahans obtaining care elsewhere (specifically,  $\hat{\rho}_{i,\text{Paducah}} = 1.3$  when  $i \neq \text{Paducah}$ ).<sup>51</sup> Figure 12 shows that, unlike an increase in Paducah reimbursements, this policy has positive spillovers on neighboring regions. These regions increase their exports to Paducah, and thus their own scale and quality. This improves their residents' market access.<sup>52</sup> Because lower-income patients are more sensitive to distance and are less likely to import care from other regions, a larger travel subsidy is necessary to achieve the same percentage improvement in their patient market access. To increase each income tercile's patient market access in Paducah by 7%, one would need to reduce trade costs by 40% for the first income tercile and by 37% for the third income tercile.

So this policy benefits both Paducah and its neighbors—though we do not estimate the costs of this travel subsidy. But facilitating travel reduces the quantity—and thus the quality—produced in Paducah. Analysts looking at the impact of travel subsidies on the quantity or quality of care provided in Paducah itself would reach very different conclusions than those looking at the impact on patient market access.

These counterfactual scenarios are subject to significant caveats, and we have not attempted to identify the optimal policy. Even so, this simple model rationalizes important aspects of the economics and politics of US healthcare policy. The counterfactual scenarios highlight our main findings: Healthcare production has substantial local increasing returns, and patient travel plays a meaningful role in enabling access to higher-quality care. Given these economic mechanisms, regional spillovers are larger when economies of scale

---

<sup>51</sup>The impact of this change on Paducah residents' market access  $\hat{\Phi}_{\text{Paducah}}$  is similar to an 8% increase in reimbursements in Paducah.

<sup>52</sup>Recall that our model assumes elastic supply. The short-run impact on exporters may be more complex if there are short-term diseconomies of scale due to crowding or queuing.

are stronger, depend on the pattern of trade flows, and differ depending on whether policies subsidize production or travel. This shows the importance of distinguishing between regional output quality and regional patient access when evaluating healthcare policies.

## 7 Conclusion

Smaller markets have fewer specialized physicians, produce less medical care per capita, and have worse health outcomes than larger markets. Thanks to trade in medical services, less production does not translate one for one into less consumption of medical services. Instead, trade affords patients who live in smaller markets access to higher-quality care. This higher quality comes in part from consuming services that would otherwise be unavailable, visiting appropriate specialists, and accessing experienced physicians.

This trade amplifies the scale advantages of large markets and hence the quality of care they produce. This means the healthcare industry can serve as an export base for large cities. Substantial scale economies also imply that policies to reallocate care across regions may impact the quality of care available. We simulate policies that aim to improve care access in “under-served” markets. The rich and varied patterns of welfare consequences when subsidizing production or travel highlight the importance of trade and agglomeration for the incidence of these policies on patients and producers.

## References

- AAPC. 2021. “What Is CPT?” Available online at <https://www.aapc.com/resources/medical-coding/cpt.aspx>.
- Acemoglu, Daron, and Joshua Linn. 2004. “Market Size in Innovation: Theory and Evidence from the Pharmaceutical Industry.” *The Quarterly Journal of Economics*, 119(3): 1049–1090.
- Agarwal, Sumit, J. Bradford Jensen, and Ferdinando Monte. 2020. “Consumer Mobility and the Local Structure of Consumption Industries.” NBER Working Paper 23616.
- Allen, Treb, Simon Fuchs, Sharat Ganapati, Alberto Graziano, Rocio Madera, and Judit Montoriol-Garriga. 2021. “Urban Welfare: Tourism in Barcelona.”
- Anderson, James E., Catherine A. Milot, and Yoto V. Yotov. 2014. “How much does geography deflect services trade? Canadian answers.” *International Economic Review*, 55(3): 791–818.
- Arrow, Kenneth J. 1963. “Uncertainty and the Welfare Economics of Medical Care.” *American Economic Review*, 53(5): 941–973.
- Baicker, Katherine, and Amitabh Chandra. 2010. “Understanding Agglomerations in Health Care.” *Agglomeration Economics*, 211–236. University of Chicago Press.
- Bartelme, Dominick G., Arnaud Costinot, Dave Donaldson, and Andrés Rodríguez-Clare. 2019. “The Textbook Case for Industrial Policy: Theory Meets Data.” National Bureau of Economic Research NBER Working Paper 26193.
- Bartik, Timothy, and George Erickcek. 2007. “Higher Education, the Health Care Industry, and Metropolitan Regional Economic Development: What Can ‘Eds & Meds’ Do for the Economic Fortunes of a Metro Area’s Residents?” W.E. Upjohn Institute for Employment Research Upjohn Working Papers 08-140.
- Battaglia, Emily. 2022. “The Effect of Hospital Closures on Maternal and Infant Health.” University of Delaware, mimeo. Available online at [https://emilybattaglia.github.io/Battaglia/Battaglia\\_JMP.pdf](https://emilybattaglia.github.io/Battaglia/Battaglia_JMP.pdf) (accessed June 29, 2022).
- Baumgardner, James R. 1988. “Physicians’ Services and the Division of Labor across Local Markets.” *Journal of Political Economy*, 96(5): 948–982.
- Berenson, Robert A., Jonathan H. Sunshine, David Helms, and Emily Lawton. 2015. “Why Medicare Advantage plans pay hospitals traditional Medicare prices.” *Health Affairs*, 34(8): 1289–1295.
- Berry, Christopher R., and Edward L. Glaeser. 2005. “The divergence of human capital levels across cities.” *Papers in Regional Science*, 84(3): 407–444.
- Burstein, Ariel, Sarah Lein, and Jonathan Vogel. 2022. “Cross-border shopping: evidence and welfare implications for Switzerland.”

- Capps, Cory, David Dranove, and Mark Satterthwaite. 2003. "Competition and market power in option demand markets." *The RAND Journal of Economics*, 737–763.
- Centers for Medicare and Medicaid Services. 2022. "National Health Expenditure Data."
- Chipman, John S. 1970. "External Economies of Scale and Competitive Equilibrium." *The Quarterly Journal of Economics*, 84(3): 347–385.
- Clapesattle, Helen. 1969. *The Doctors Mayo*. Mayo Foundation for Medical Education.
- Clemens, Jeffrey, and Joshua D. Gottlieb. 2014. "Do Physicians' Financial Incentives Affect Treatment Patterns and Patient Health?" *American Economic Review*, 104(4): 1320–1349.
- Clemens, Jeffrey, and Joshua D. Gottlieb. 2017. "In the Shadow of a Giant: Medicare's Influence on Private Payment Systems." *Journal of Political Economy*, 125(1): 1–39.
- Cooper, Zack, Stuart V. Craig, Martin Gaynor, and John Van Reenen. 2018. "The Price Ain't Right? Hospital Prices and Health Spending on the Privately Insured." *The Quarterly Journal of Economics*, 134(1): 51–107.
- Costinot, Arnaud, Dave Donaldson, Margaret Kyle, and Heidi Williams. 2019. "The More We Die, The More We Sell? A Simple Test of the Home-Market Effect." *The Quarterly Journal of Economics*, 134(2): 843–894.
- 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(2): 286–325.
- Davis, Donald R., and David E. Weinstein. 2003. "Market access, economic geography and comparative advantage: an empirical test." *Journal of International Economics*, 59(1): 1–23.
- Davis, Donald R., and Jonathan I. Dingel. 2020. "The comparative advantage of cities." *Journal of International Economics*, 123(C).
- Davis, Donald R., Jonathan I. Dingel, Joan Monras, and Eduardo Morales. 2019. "How Segregated Is Urban Consumption?" *Journal of Political Economy*, 127(4): 1684–1738.
- Deryugina, Tatyana, and David Molitor. 2021. "The Causal Effects of Place on Health and Longevity." *Journal of Economic Perspectives*, 35(4): 147–70.
- Diamond, Rebecca. 2016. "The determinants and welfare implications of US workers' diverging location choices by skill: 1980-2000." *American Economic Review*, 106(3): 479–524.
- Dingel, Jonathan I. 2017. "The Determinants of Quality Specialization." *Review of Economic Studies*, 84(4): 1551–1582.
- Dingel, Jonathan I., and Felix Tintelnot. 2021. "Spatial Economics for Granular Settings." National Bureau of Economic Research Working Paper 27287.

- Disdier, Anne-Célia, and Keith Head. 2008. “The Puzzling Persistence of the Distance Effect on Bilateral Trade.” *The Review of Economics and Statistics*, 90(1): 37–48.
- Dorsey, E. Ray, and Eric J. Topol. 2016. “State of Telehealth.” *New England Journal of Medicine*, 375(2): 154–161. PMID: 27410924.
- Dranove, David, and Christopher Ody. 2016. “Evolving measures of provider market power.” *American Journal of Health Economics*, 2(2): 145–160.
- Dranove, David, and Mark A. Satterthwaite. 2000. “The Industrial Organization of Health Care Markets.” *Handbook of Health Economics*, 1: 1093–1139.
- Dranove, David, Mark Shanley, and Carol Simon. 1992. “Is Hospital Competition Wasteful?” *The RAND Journal of Economics*, 23(2): 247–262.
- Eckert, Fabian, Sharat Ganapati, and Conor Walsh. 2020. “Skilled Scalable Services: The New Urban Bias in Economic Growth.” CESifo Working Paper Series 8705.
- Ericson, Keith Marzilli, and Amanda Starc. 2015. “Measuring consumer valuation of limited provider networks.” *American Economic Review*, 105(5): 115–19.
- Falchetti, Elena. 2021. “The Determinants of Physicians’ Location Choice: Understanding the Rural Shortage.” Mimeo, Federal Reserve. Available online at [https://papers.ssrn.com/sol3/papers.cfm?abstract\\_id=3493178](https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3493178).
- Finkelstein, Amy, Matthew Gentzkow, and Heidi Williams. 2016. “Sources of geographic variation in health care: Evidence from patient migration.” *Quarterly Journal of Economics*, 131(4): 1681–1726.
- Finkelstein, Amy, Matthew Gentzkow, and Heidi Williams. 2021. “Place-Based Drivers of Mortality: Evidence from Migration.” *American Economic Review*, 111(8): 2697–2735.
- Fischer, Stefanie J., Heather Royer, and Corey D. White. 2022. “Health Care Centralization: The Health Impacts of Obstetric Unit Closures in the US.” NBER Working Paper 30141.
- Fisher, Elliott S., David E. Wennberg, Threse A. Stukel, Daniel J. Gottlieb, F. Lee Lucas, and Etoile L. Pinder. 2003a. “The implications of regional variations in Medicare spending. Part 1: the content, quality, and accessibility of care.” *Annals of Internal Medicine*, 138(4): 273–287.
- Fisher, Elliott S., David E. Wennberg, Threse A. Stukel, Daniel J. Gottlieb, F. Lee Lucas, and Etoile L. Pinder. 2003b. “The implications of regional variations in Medicare spending. Part 2: health outcomes and satisfaction with care.” *Annals of Internal Medicine*, 138(4): 288–298.
- Flexner, Abraham. 1910. “Medical Education in the United States and Canada: A Report to the Carnegie Foundation for the Advancement of Teaching.” Carnegie Foundation Bulletin Number Four.

- Fujita, Masahisa, and Paul Krugman. 1995. “When is the economy monocentric?: von Thünen and Chamberlin unified.” *Regional Science and Urban Economics*, 25(4): 505–528.
- Garthwaite, Craig, Christopher Ody, and Amanda Starc. 2022. “Endogenous quality investments in the us hospital market.” *Journal of Health Economics*, 102636.
- Gaynor, Martin, and William B Vogt. 2003. “Competition among hospitals.” *The RAND Journal of Economics*, 34(3): 764–785.
- Gaynor, Martin S., Samuel A. Kleiner, and William B. Vogt. 2013. “A structural approach to market definition with an application to the hospital industry.” *The Journal of Industrial Economics*, 61(2): 243–289.
- Glaeser, Edward L. 2005. “Reinventing Boston: 1630–2003.” *Journal of Economic Geography*, 5(2): 119–153.
- Gottlieb, Daniel J, Weiping Zhou, Yunjie Song, Kathryn Gilman Andrews, Jonathan S Skinner, and Jason M Sutherland. 2010. “Prices don’t drive regional Medicare spending variations.” *Health Affairs*, 29(3): 537–543.
- Gottlieb, Joshua D., Maria Polyakova, Kevin Rinz, Hugh Shiple, and Victoria Udalova. 2020. “Who Values Human Capitalists’ Human Capital? Healthcare Spending and Physician Earnings.” University of Chicago, mimeo.
- Gotto, Antonio M., and Jennifer Moon. 2016. *Weill Cornell Medicine: A History of Cornell’s Medical School*. Ithaca: Cornell University Press.
- Gowrisankaran, Gautam, Aviv Nevo, and Robert Town. 2015. “Mergers when prices are negotiated: Evidence from the hospital industry.” *American Economic Review*, 105(1): 172–203.
- Greenstone, Michael, Richard Hornbeck, and Enrico Moretti. 2010. “Identifying agglomeration spillovers: Evidence from winners and losers of large plant openings.” *Journal of Political Economy*, 118(3): 536–598.
- Gupta, Atul. 2021. “Impacts of Performance Pay for Hospitals: The Readmissions Reduction Program.” *American Economic Review*, 111(4): 1241–83.
- Gupta, Atul, Amol Navathe, and Aaron Schwartz. 2022. “How Does Medicare Advantage Affect Health Care Use? Evidence from Beneficiary Transitions.” University of Pennsylvania, mimeo.
- Hanson, Gordon H., and Chong Xiang. 2004. “The Home-Market Effect and Bilateral Trade Patterns.” *American Economic Review*, 94(4): 1108–1129.
- Helpman, Elhanan, and Paul R. Krugman. 1985. *Market Structure and Foreign Trade*. MIT Press.

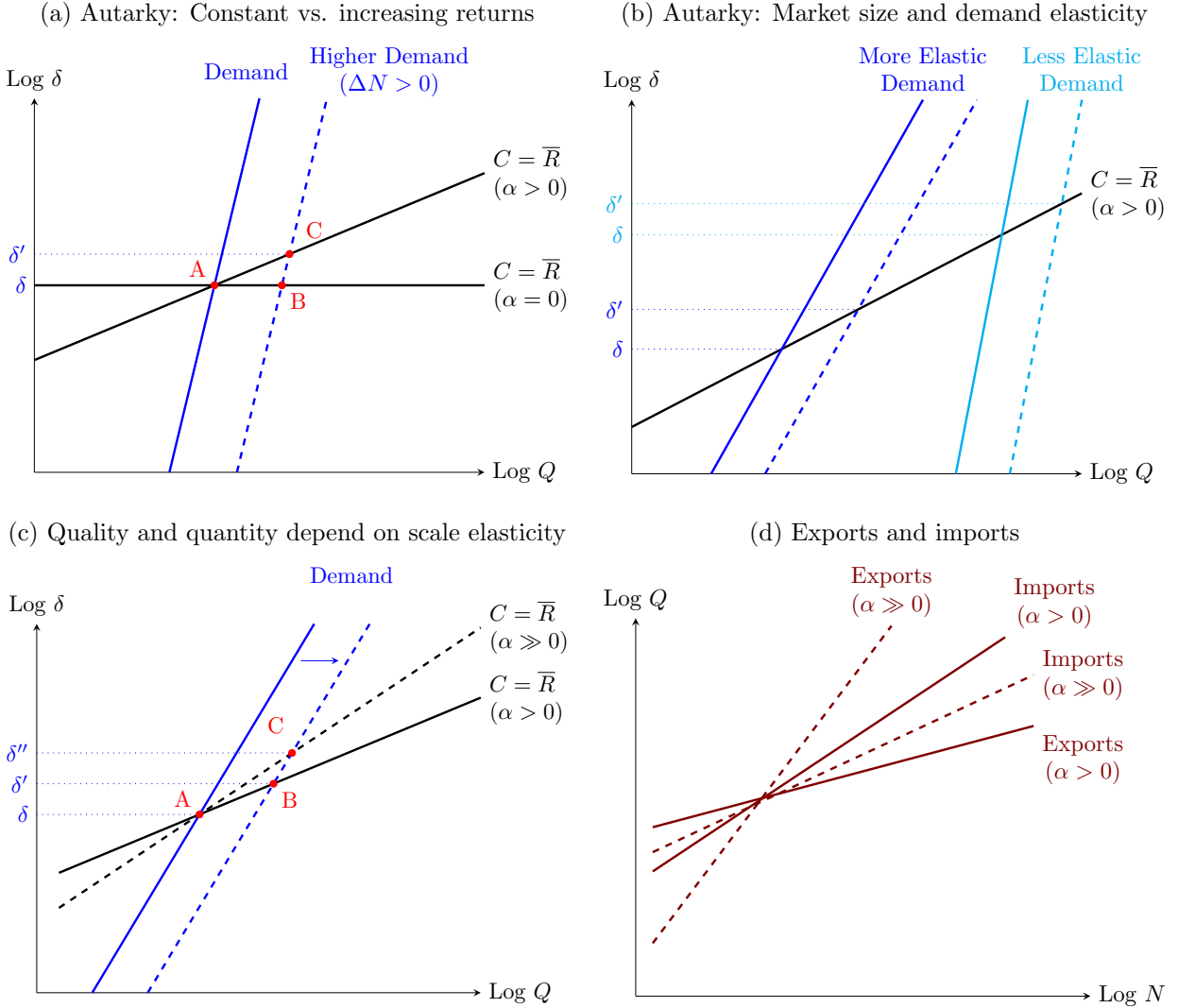


- Ho, Kate, and Robin S. Lee. 2017. “Insurer competition in health care markets.” *Econometrica*, 85(2): 379–417.
- Ho, Kate, and Robin S. Lee. 2019. “Equilibrium provider networks: Bargaining and exclusion in health care markets.” *American Economic Review*, 109(2): 473–522.
- Ho, Katherine. 2009. “Insurer-Provider Networks in the Medical Care Market.” *American Economic Review*, 99(1): 393–430.
- Hsieh, Chang-Tai, and Esteban Rossi-Hansberg. 2021. “The Industrial Revolution in Services.” Center for Economic Studies, U.S. Census Bureau Working Papers 21-34.
- Jensen, J. Bradford, and Lori G. Kletzer. 2005. “Tradable Services: Understanding the Scope and Impact of Services Offshoring.” *Brookings Trade Forum*, 75–116.
- Kessler, Daniel P., and Mark B. McClellan. 2000. “Is hospital competition socially wasteful?” *The Quarterly Journal of Economics*, 115(2): 577–615.
- Khoury, Stephanie, Jonathan M. Leganza, and Alex Masucci. 2022. “Health Professional Shortage Areas and Physician Location Decisions.” UCSD, mimeo. Available online at [https://papers.ssrn.com/sol3/papers.cfm?abstract\\_id=3701160](https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3701160).
- Kline, Patrick, and Enrico Moretti. 2013. “Local Economic Development, Agglomeration Economies, and the Big Push: 100 Years of Evidence from the Tennessee Valley Authority.” *The Quarterly Journal of Economics*, 129(1): 275–331.
- Krugman, Paul. 1980. “Scale Economies, Product Differentiation, and the Pattern of Trade.” *American Economic Review*, 70(5): 950–59.
- Lee, Sanghoon. 2010. “Ability sorting and consumer city.” *Journal of Urban Economics*, 68(1): 20–33.
- Levy, Antoine, and Jacob Moscona. 2020. “Specializing in Density: Spatial Sorting and the Pattern of Trade.” Mimeo, MIT. Available online at <https://economics.mit.edu/files/16986>.
- Lewis, Matthew S., and Kevin E. Pflum. 2015. “Diagnosing hospital system bargaining power in managed care networks.” *American Economic Journal: Economic Policy*, 7(1): 243–74.
- Lewis, Matthew S., and Kevin E. Pflum. 2017. “Hospital systems and bargaining power: Evidence from out-of-market acquisitions.” *The RAND Journal of Economics*, 48(3): 579–610.
- Lipsey, Robert E. 2009. “Measuring International Trade in Services.” In *International Trade in Services and Intangibles in the Era of Globalization.*, ed. Marshall B. Reinsdorf and Matthew J. Slaughter, 27–74. University of Chicago Press.
- Lopez, Eric, and Gretchen Jacobson. 2020. “How Much More Than Medicare Do Private Insurers Pay? A Review of the Literature.”

- Marshall, Alfred. 1890. *Principles of Economics*. London: Macmillan and Co.
- Miyauchi, Yuhei, Kentaro Nakajima, and Stephen J. Redding. 2021. “The Economics of Spatial Mobility: Theory and Evidence Using Smartphone Data.” National Bureau of Economic Research Working Paper 28497.
- Moretti, Enrico. 2004. “Estimating the social return to higher education: evidence from longitudinal and repeated cross-sectional data.” *Journal of Econometrics*, 121(1-2): 175–212.
- Moretti, Enrico. 2011. “Local labor markets.” In *Handbook of Labor Economics*. Vol. 4, 1237–1313. Elsevier.
- Muñoz, Mathilde. 2022. “Trading Non-Tradables: The Implications of Europe’s Job Posting Policy.”
- Newfeld, Jonathan D., and Charles R. Doarn. 2015. “Telemedicine Spending by Medicare: A Snapshot from 2012.” *Telemed J E Health*, 21(8): 686–693.
- Newhouse, Joseph P. 1990. “Geographic access to physician services.” *Annual review of public health*, 11(1): 207–230.
- 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(2): 493–505.
- Newhouse, Joseph P., Albert P. Williams, Bruce W. Bennett, and William B. Schwartz. 1982b. “The geographic distribution of physicians: Is the conventional wisdom correct?” Santa Monica: RAND Corp. Publ. No. R-2734.
- Newhouse, Joseph P., Albert P. Williams, Bruce W. Bennett, and William B. Schwartz. 1982c. “Where have all the doctors gone?” *JAMA*, 247(17): 2392–2396.
- Petek, Nathan. 2022. “The marginal benefit of hospitals: Evidence from the effect of entry and exit on utilization and mortality rates.” *Journal of Health Economics*, 86: 102688.
- Rosenblatt, R. A., and L. G. Hart. 2000. “Physicians and rural America.” *The Western Journal of Medicine*, 173(5): 348–351.
- Rosenthal, Meredith B., Alan Zaslavsky, and Joseph P. Newhouse. 2005. “The Geographic Distribution of Physicians Revisited.” *Health Services Research*, 40(6p1): 1931–1952.
- Rosenthal, Stuart S., and William C. Strange. 2008. “The attenuation of human capital spillovers.” *Journal of Urban Economics*, 64(2): 373–389.
- Ruggles, Steven, Sarah Flood, Ronald Goeken, Megan Schouweiler, and Matthew Sobek. 2022. “IPUMS USA.” Minneapolis, MN: IPUMS Version 12.0. <https://doi.org/10.18128/D010.V12.0>.

- Santos Silva, J. M. C., and Silvana Tenreyro. 2006. “The Log of Gravity.” *The Review of Economics and Statistics*, 88(4): 641–658.
- Silver, David, and Jonathan Zhang. 2022. “Impacts of Basic Income on Health and Economic Well-Being: Evidence from the VA’s Disability Compensation Program.” National Bureau of Economic Research Working Paper 29877.
- Skinner, Lucy, Douglas O. Staiger, David I. Auerbach, and Peter I. Buerhaus. 2019. “Implications of an aging rural physician workforce.” *New England Journal of Medicine*, 381(4): 299–301.
- Town, Robert, and Gregory Vistnes. 2001. “Hospital competition in HMO networks.” *Journal of Health Economics*, 20(5): 733–753.
- Zuckerman, Stephen, Laura Skopec, and Joshua Aarons. 2021. “Medicaid Physician Fees Remained Substantially Below Fees Paid By Medicare In 2019.” *Health Affairs*, 40(2): 343–348.

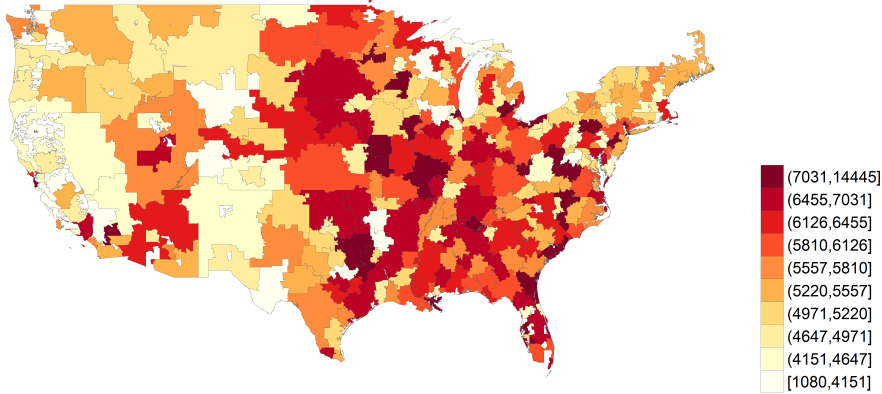
Figure 1: Illustrative model diagrams



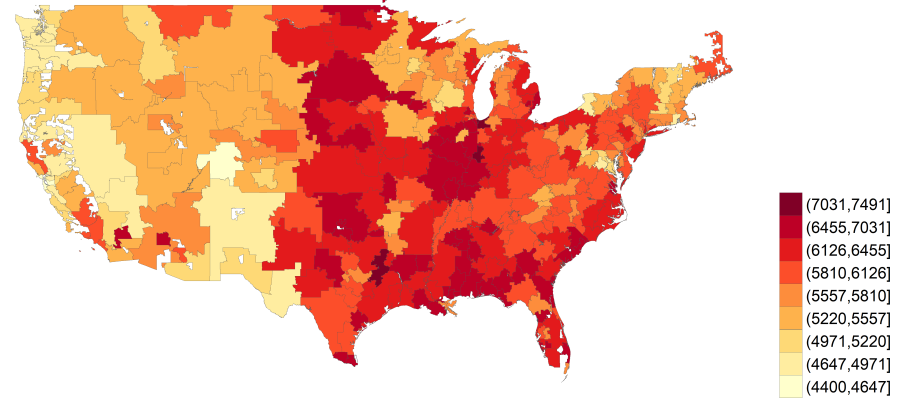
*Notes:* This figure depicts how increasing demand in one region affects its equilibrium outcomes. In panels (a)–(c), quantity produced  $Q$  is on the horizontal axis and service quality  $\delta$  is on the vertical axis. The black lines depict the free-entry isocost curve,  $C = \bar{R}$ , given by equation (3). The blue and cyan lines depict demand for the region’s service, which we depict as log-linear for visual clarity. (The logit demand function is actually log-convex, which is consistent with all the depicted comparative statistics.) Equilibrium is the intersection of the demand and isocost curves. An increase in demand is the rightward shift from the solid to the dashed demand curve. This shift increases equilibrium quality from  $\delta$  to  $\delta'$ . Panel (a) shows that higher demand elicits higher quality if there are increasing returns to scale. Panel (b) shows that this quality improvement is larger when demand is more elastic. Panels (c) and (d) introduce trade and compare the extent of quality improvement under two different magnitudes of increasing returns ( $\alpha > 0$  and  $\alpha \gg 0$ ). These magnitudes govern the patterns of interregional trade, shown in panel (d) as a function of the number of potential patients  $N$ . Imports from other regions rise with  $N$ . With increasing returns to scale ( $\alpha > 0$ ), exports to other regions also rise with  $N$  (a weak home-market effect). When the scale elasticity  $\alpha$  is larger ( $\alpha \gg 0$ ), the import curve is flatter and the export curve is steeper. With sufficiently strong increasing returns, an increase in domestic demand causes a greater increase in exports than imports (a strong home-market effect).

Figure 2: Production, consumption, and trade across regions

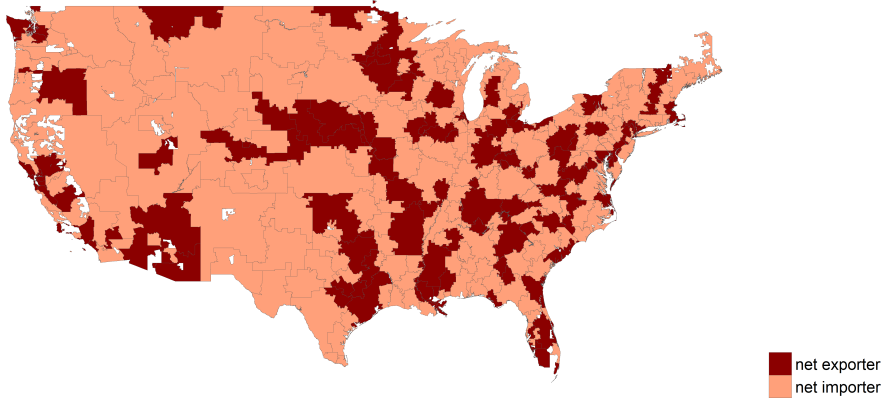
(a) Production per capita



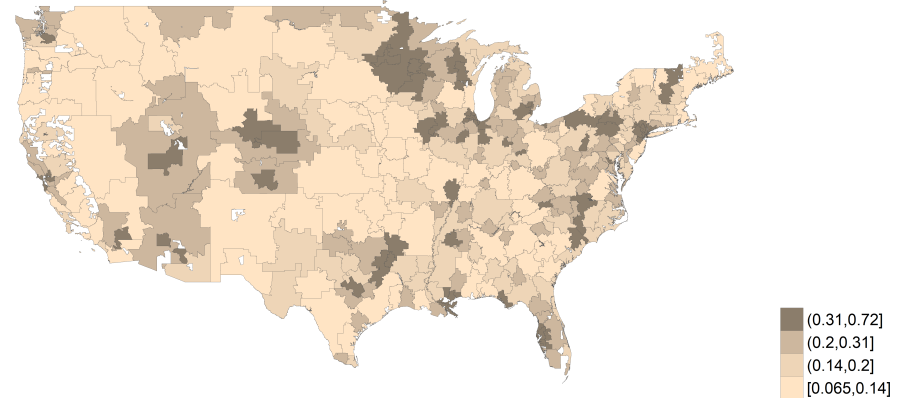
(b) Consumption per capita



(c) Net exporter vs net importer

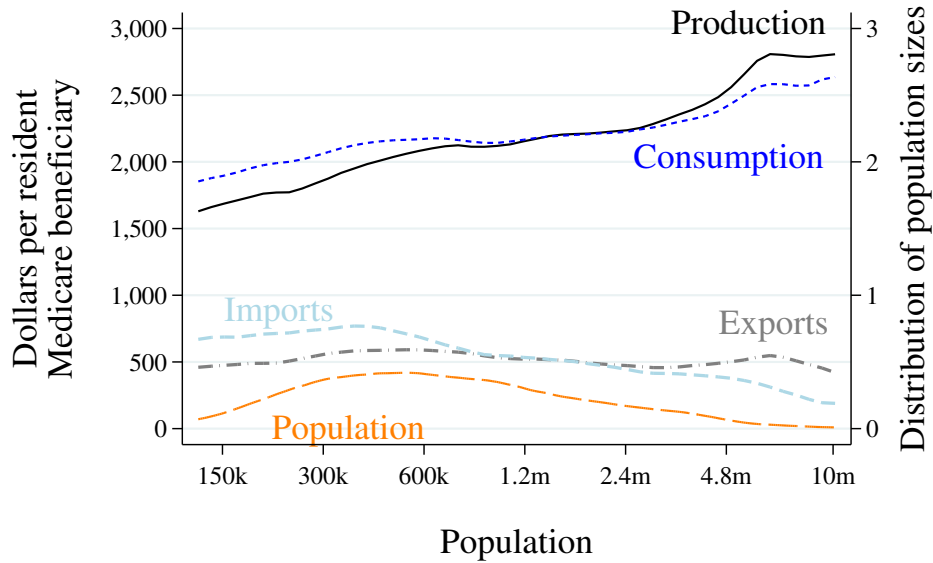


(d) Gross exports relative to production



Notes: Panel (a) shows production per capita, including professional and facility fees. The hospital referral region (HRR) of production is the location where the service is provided. Panel (b) shows consumption per capita, including professional and facility fees. The HRR of consumption is based on the patient's residential address. Colors depict deciles of production per capita in both Panels (a) and (b). Panel (c) contrasts HRRs that are net exporters with HRRs that are net importers for professional services. Panel (d) shows gross exports as a share of total production by HRR for professional services. Data come from the Medicare 20% carrier Research Identifiable Files. All calculations exclude emergency-room care and skilled nursing facilities. Expenditures are computed by assigning each procedure its national average price. HRR definitions are from the Dartmouth Atlas Project.

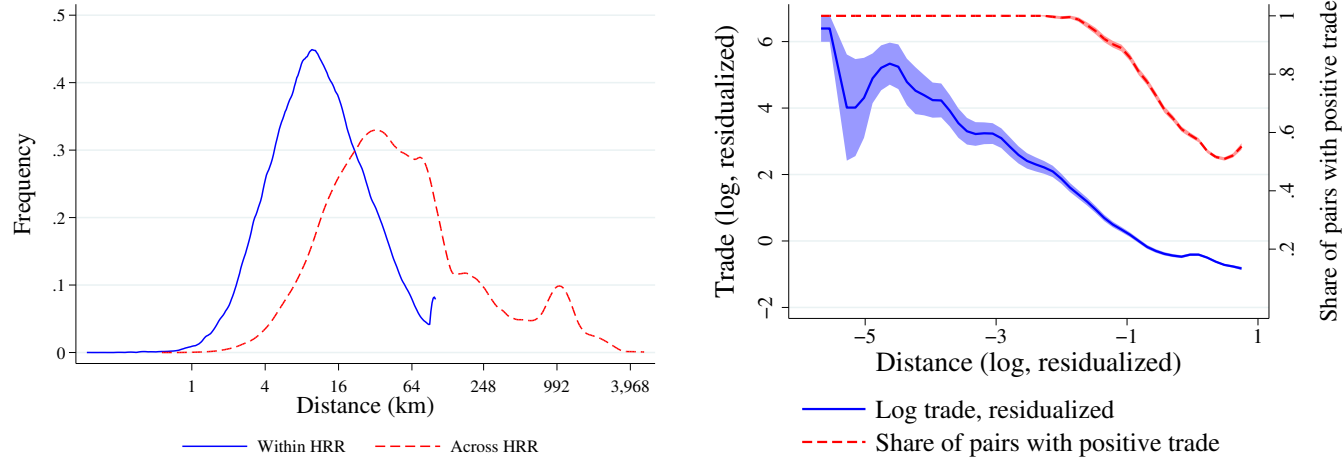
Figure 3: Production and consumption of medical care across regions and services



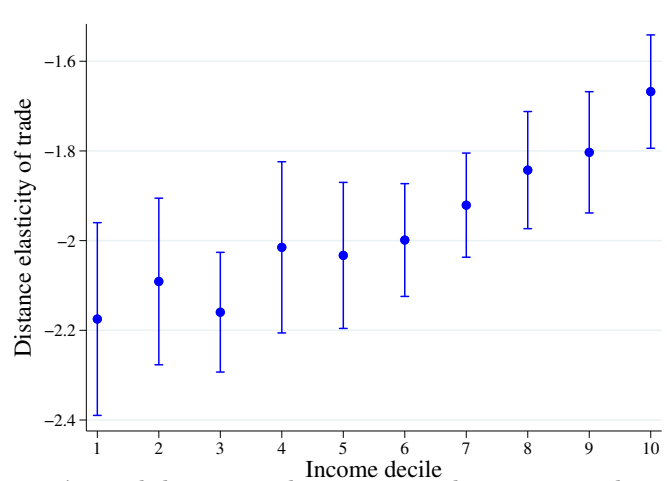
Population elasticity (log–log regression slope) of transactions per resident Medicare beneficiary:  
 Production: 0.13 (0.02), Consumption: 0.06 (0.01)  
 Exports: –0.00 (0.05), Imports: –0.25 (0.03)

*Notes:* This figure shows production, consumption, and trade per capita of Medicare services across hospital referral regions (HRRs) of different sizes, all smoothed via local averages. We use the Medicare 20% carrier Research Identifiable Files to compute the dollar value of physician services, excluding emergency-room care and assigning each procedure its national average price. The black series shows production of medical care per Medicare beneficiary residing in the HRR of production. The blue series shows consumption of medical care per Medicare beneficiary residing in the HRR of consumption. The dashed dark-gray series shows interregional “exports” of medical care and the dashed light-blue series shows interregional “imports” of medical care, again per Medicare beneficiary. The orange series depicts the distribution of HRR population sizes. HRR definitions are from the Dartmouth Atlas Project.

Figure 4: Patients travel between regions and trade declines with distance, more so for lower-income patients  
 (a) Distribution of travel distances within and across HRRs      (b) Trade volume and extensive margin by distance

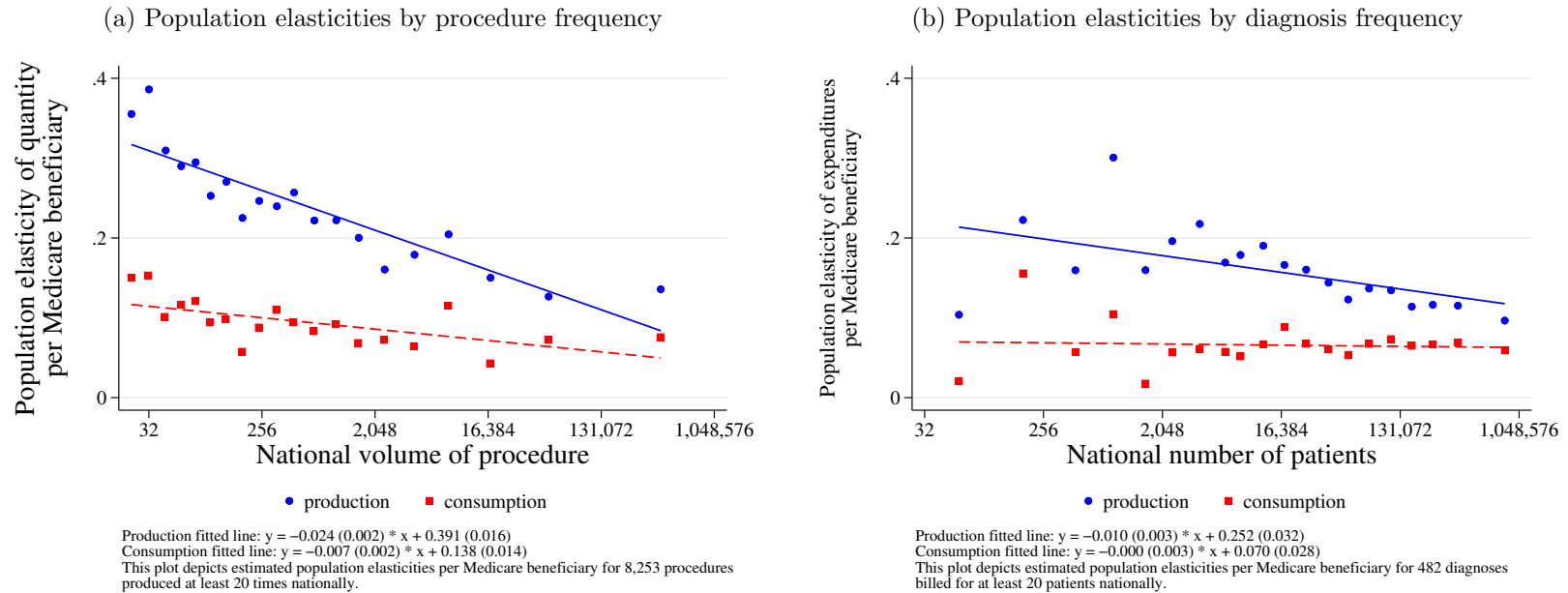


(c) Higher-income patients are less sensitive to distance



*Notes:* Panel (a) shows the distribution of patients’ travel distances when patients obtain care within their home HRR (blue distribution) and when they travel across HRRs (red distribution). Travel distances measure the distance between home and treatment locations. For travel within a hospital referral region, the distance measure reflects the distance between the centroid of the patient’s residential ZIP code and the ZIP code of the service location. We use ZCTA-to-ZCTA distances downloaded from the National Bureau of Economic Research; those exceeding 160 kilometers are winsorized at 160 kilometers. For travel across HRRs, we use ZCTA-to-ZCTA distances when they are within 160 kilometers and (for computational ease) use HRR-to-HRR distances beyond 160 kilometers. In Panel (b), the blue series depicts the volume of trade against distance, after conditioning out the fixed effects in equation (12), for positive-trade pairs of locations. The red series shows the share of HRR pairs with positive trade as a function of the distance between them, after conditioning out the importer fixed effects and exporter fixed effects, as in equation (12). Panel (c) depicts the coefficient on log distance obtained by estimating equation (12) separately for each decile of the national ZIP-level median-household-income distribution. The 95% confidence intervals are computed using standard errors two-way clustered by both patient HRR and provider HRR. Patients from higher-income ZIP codes are less sensitive to distance.

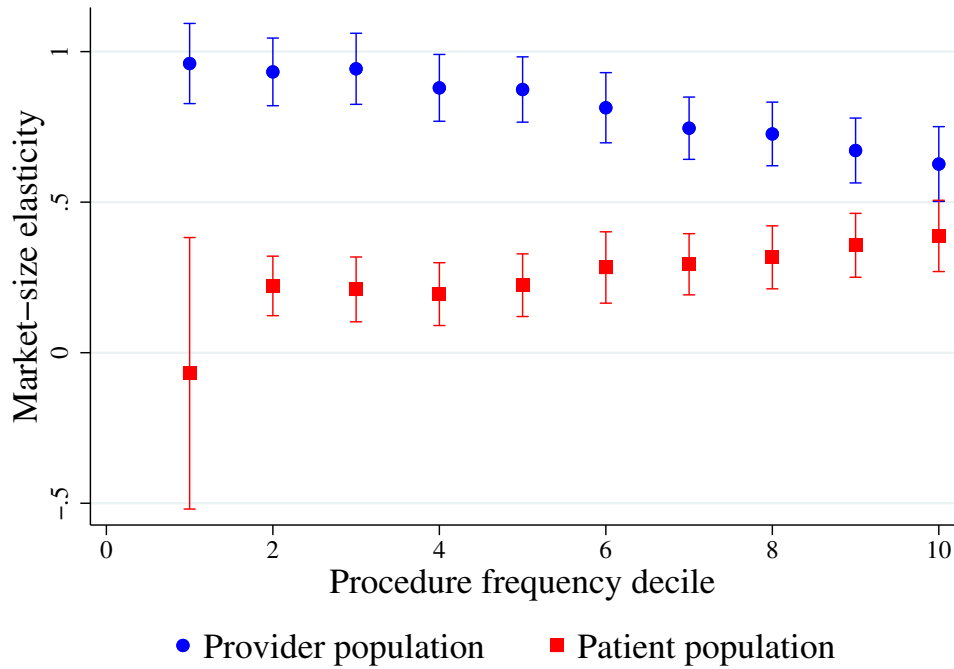
Figure 5: Population elasticities of production and consumption



*Notes:* The vertical axis of both panels plots the population elasticities of quantity of medical care produced and consumed per local Medicare beneficiary. The elasticities are computed using the Poisson models in equations (9) and (10) based on production location and patients' residential location, respectively. Panel (a) estimates these elasticities for each of the procedures provided at least 20 times nationally in the Medicare data. The horizontal axis shows the total national volume of physician services for the procedure. Panel (b) estimates the elasticities for care provided to treat each of the Clinical Classifications Software Refined (CCSR) diagnoses billed for at least 20 patients nationally in the Medicare data. Expenditures are computed from the Medicare 20% carrier Research Identifiable Files using the dollar value of physician services, excluding emergency-room care and assigning each procedure its national average price. The horizontal axis shows the total number of patients nationally with the diagnosis. In both panels, the blue dots are a binned scatterplot of the estimated population elasticity of production per beneficiary as a function of the national volume. The red dots are the same for consumption (residential location)-based estimates. There is a significant negative relationship for production, indicating that production elasticities are highest for rare services and rare diseases. The relationship for consumption is much more modest. The difference between these estimates must be driven by trade between locations.



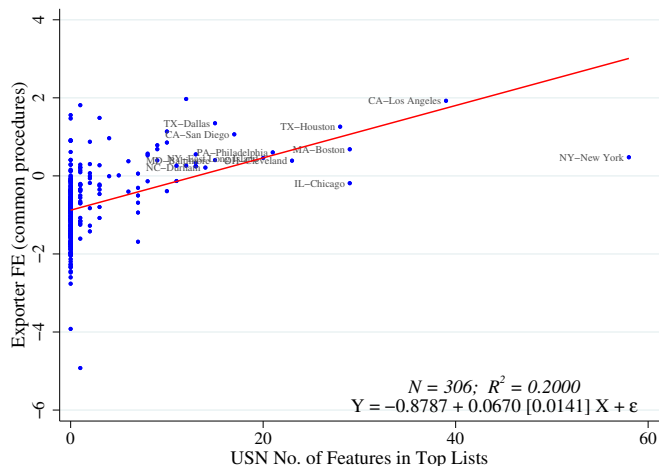
Figure 6: The home-market effect is stronger for rarer procedures



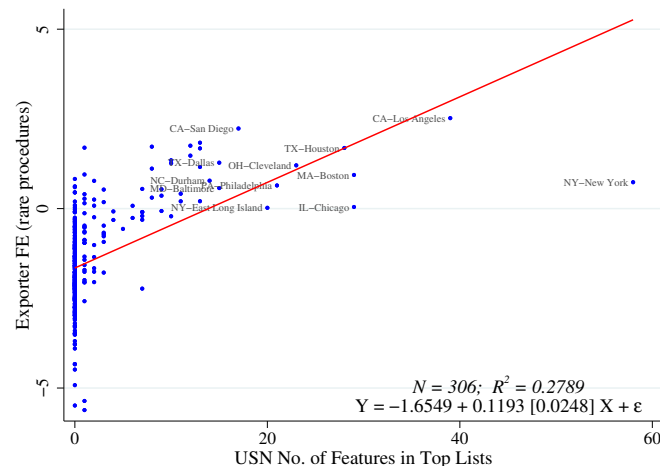
*Notes:* This figure groups non-emergency physician-provided services in the Medicare claims data into deciles based on the national frequency of each procedure. For each decile, we estimate equation (8), testing for a home market effect, and plot the estimated coefficients on provider and patient market log population with their 95% confidence intervals. The coefficients on provider-market size always exceed the respective coefficients on patient-market size, indicating a strong home-market effect. The coefficients on provider-market size monotonically decrease across the deciles. The coefficients on patient-market size monotonically increase across the deciles. Together, these two patterns show that the home-market effect is stronger the less common the procedure is, in line with the theoretical difference-in-difference prediction.

Figure 7: Estimated quality is positively correlated with total output and external quality metrics

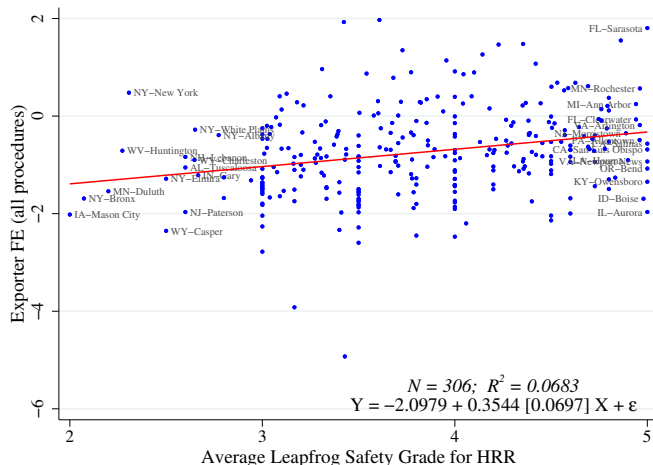
(a) *U.S. News* vs. quality estimated for common services



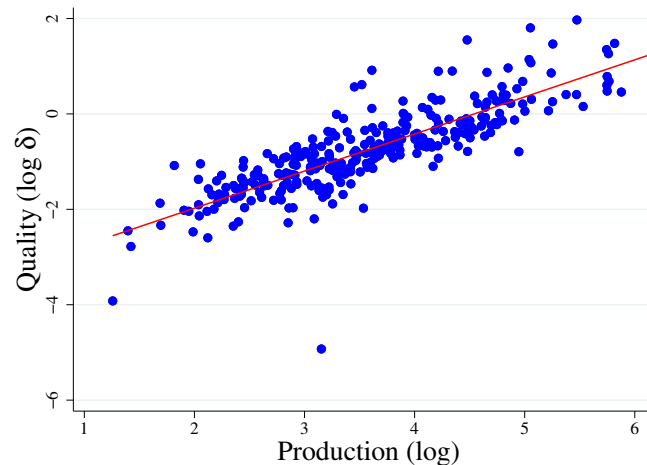
(b) *U.S. News* vs. quality estimated for rare services



(c) Leapfrog Safety Grade vs. estimated quality



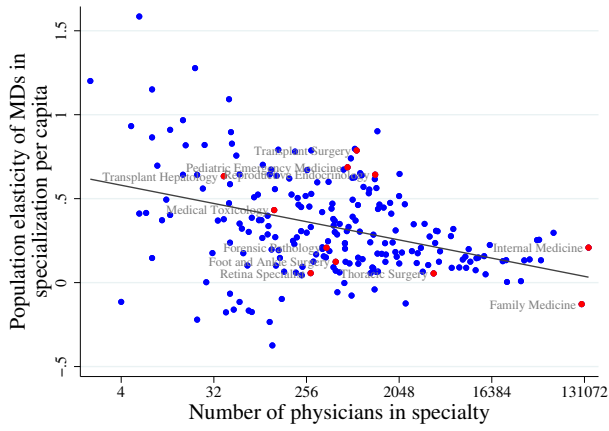
(d) Quality is higher in regions producing more output



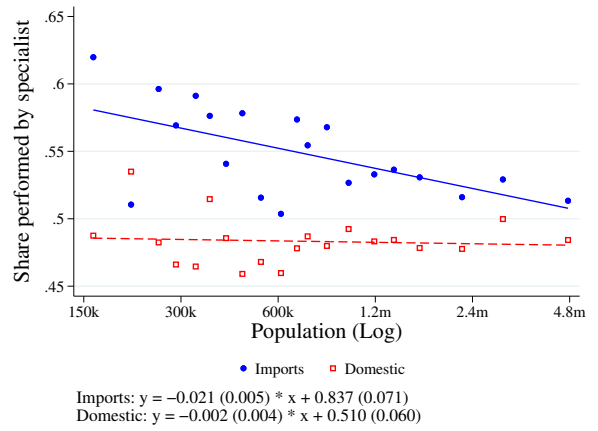
Notes: The first three panels show the relationship between the exporter fixed effects (our revealed-preference measure of quality) and external quality measures. The vertical axis shows the exporter fixed effects for each HRR estimated using trade in common services in Panel (a), using trade in rare services in Panel (b), and for all services in Panels (d) and (c). The horizontal axis in Panels (a) and (b) is a count of the number of times each region's hospitals appear on the *U.S. News* list of best hospitals. *U.S. News* produces an overall ranking as well as rankings for 12 particular specialties. We count the number of times each HRR's hospitals appear on any of these 13 lists. Both panels show a positive relationship, indicating that patients travel farther to obtain care from regions highly ranked by *U.S. News*. The relationship is stronger for rare services, as the slope is nearly double that for common services. The horizontal axis in Panel (c) is the average safety grade for hospitals in an HRR (mapping A=5, B=4, etc.), for grades determined by the Leapfrog Group. These are positively correlated with exporter fixed effects. Panel (d) shows the relationship between production and the exporter fixed effects from equation (12), across HRRs. HRR production is measured as Medicare output produced (in millions US dollars) for non-emergency physician services in the 20% carrier file.

Figure 8: Imports are specialist-intensive, especially in smaller regions

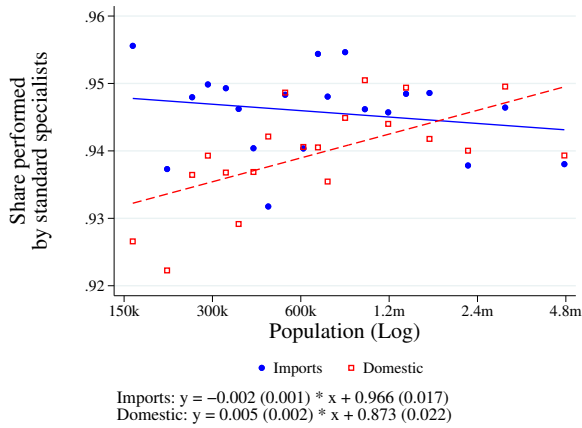
(a) Population elasticities of physician specializations



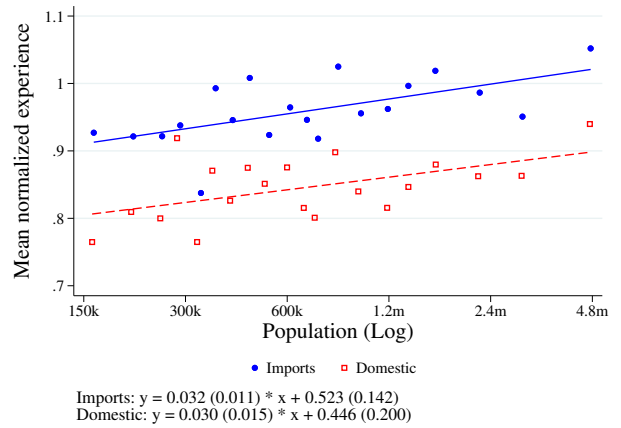
(b) Specialty care imports



(c) “Standard” specialty care

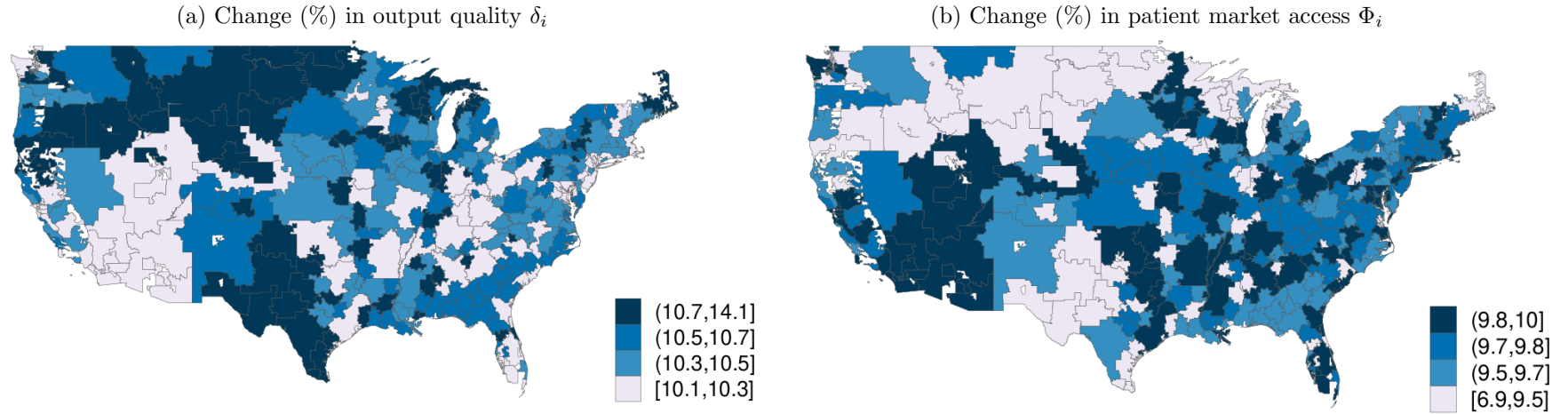


(d) Provider experience



*Notes:* The vertical axis of Panel (a) depicts the population elasticities of quantity of physicians in an HRR. The population elasticities are computed for each specialty using the Poisson model in equation (14). The horizontal axis shows the nationwide number of physicians in each specialty. The negative relationship indicates that rare specialties are disproportionately concentrated in high-population regions. Panel (b) shows the share of procedures that are performed by a specialist, for imports and locally produced procedures, by market size. We define generalists as internal-medicine, general-practice, and family-practice physicians and define specialists as all other physicians. Imports are more likely to be performed by a specialist, and smaller markets’ imports especially so. Panel (c) examines procedures that are typically performed by specialists, and classifies the “standard” specialists as the top two specialties performing the procedure nationally. It shows the shares of procedures performed by the “standard” specialties in imported specialty care and locally produced specialty care as a function of local population size. Imports are more likely to be performed by “standard” specialties, especially for smaller regions. Panel (d) shows the mean relative experience of providers for care produced locally and imported by population size of the patient’s region. This panel describes only procedures that are performed in all hospital referral regions (143 procedures). In public-use Medicare data, we define a provider’s experience for a given procedure as the number of times they performed the procedure for Traditional Medicare patients in the prior calendar year. Before aggregating to the regional level, we rescale experience in each procedure so that its mean is one. On average, patients in larger markets obtain treatment from more experienced providers. At all population levels, imported care is produced by more experienced providers than domestic care.

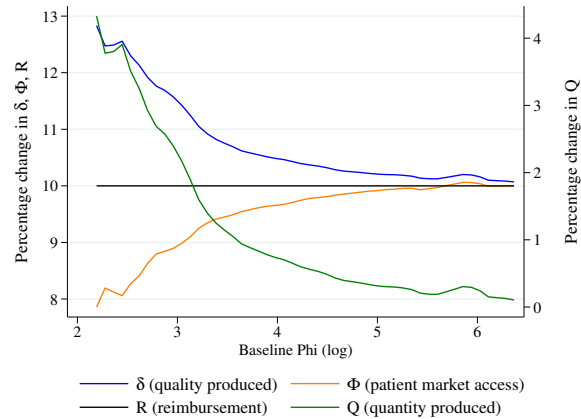
Figure 9: Counterfactual outcomes when reimbursements increase 10% everywhere



Cutoffs: Percentiles 25, 50, 75.

Cutoffs: Percentiles 25, 50, 75.

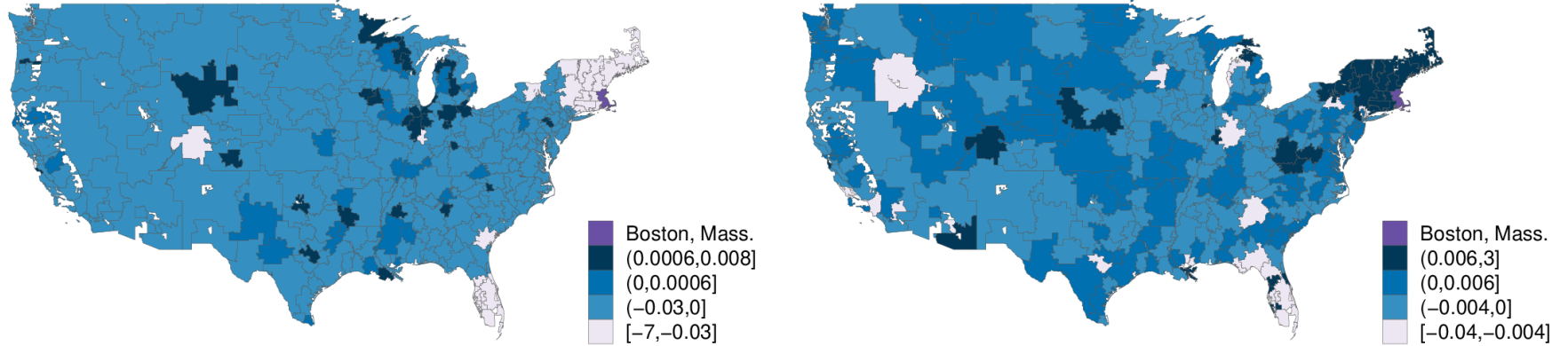
(c) Outcomes as a function of baseline patient market access  $\Phi_i$



Notes: Panels (a) and (b) show the impacts of increasing reimbursements by 10% everywhere ( $\hat{R}_i = 1.1 \forall i$ ) based on our estimated model. Panel (a) depicts the percentage change in quality of care  $\delta_i$  provided in each region. Panel (b) depicts the percentage change in the value of market access  $\Phi_i$  for patients who live in a region. Panel (c) shows local linear regressions of the percentage changes in  $\delta_i$ ,  $\Phi_i$ , and  $Q_i$  against the region's initial patient market access,  $\Phi_i$ . There is a negative relationship between the percentage changes in  $\delta$  and  $\Phi$  across regions. Patients who live in the regions with the largest quality increases in  $\delta$  tend to have the lowest gains in patients' market access,  $\Phi$ . The exercise is described in detail in Section 6.

Figure 10: Counterfactual outcomes for higher reimbursements in one region

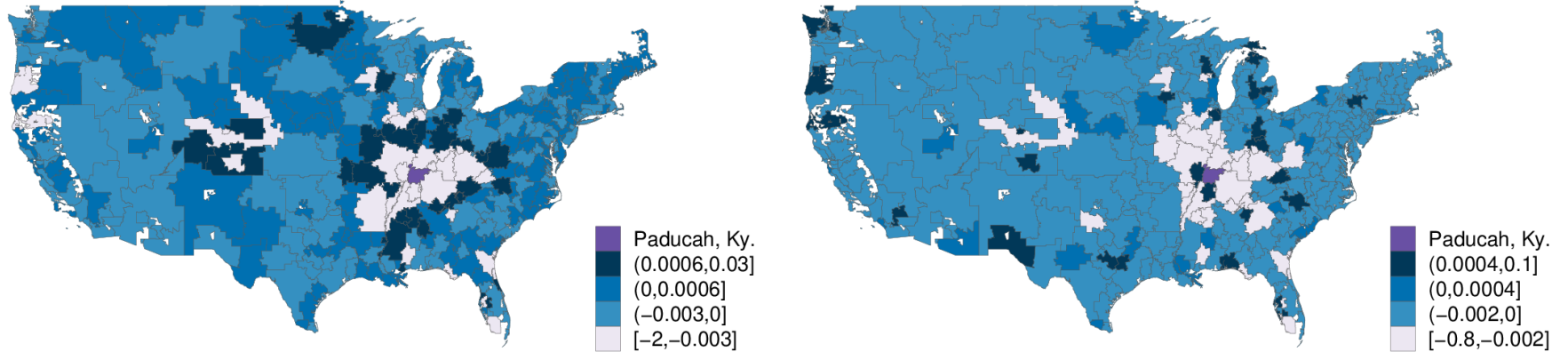
(a) Change (%) in output quality  $\delta_i$ : higher reimbursement in Boston, Mass. (b) Change (%) in market access  $\Phi_i$ : higher reimbursement in Boston, Mass.



Boston, Mass. = 34.6%

Boston, Mass. = 31.6%

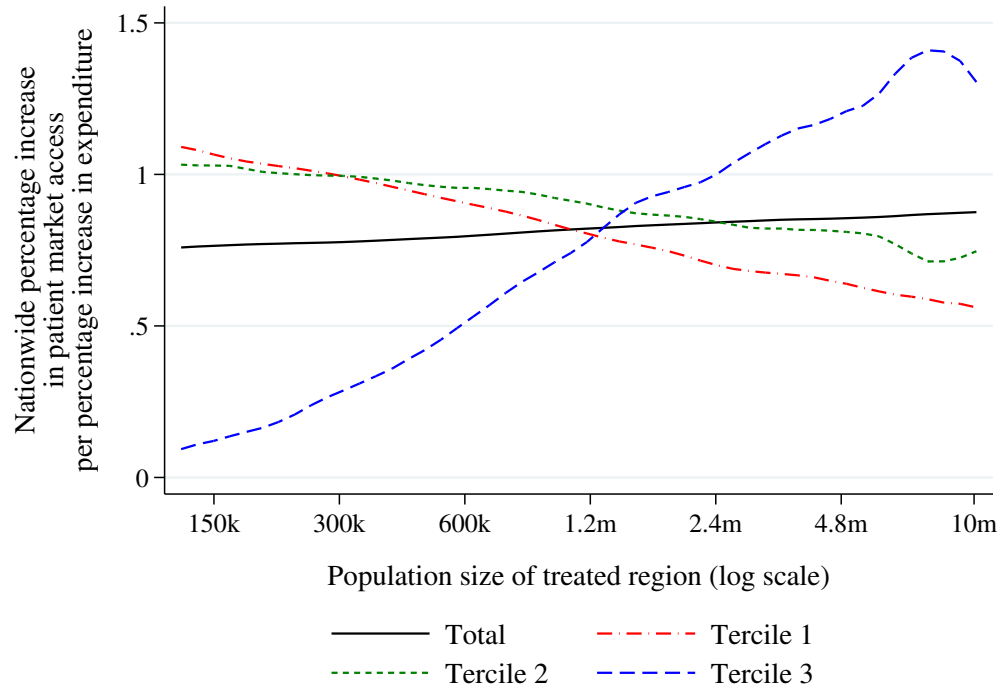
(c) Change (%) in output quality  $\delta_i$ : higher reimbursement in Paducah, Ky. (d) Change (%) in market access  $\Phi_i$ : higher reimbursement in Paducah, Ky.



Paducah, Ky. = 45.4%

Paducah, Ky. = 24.4%

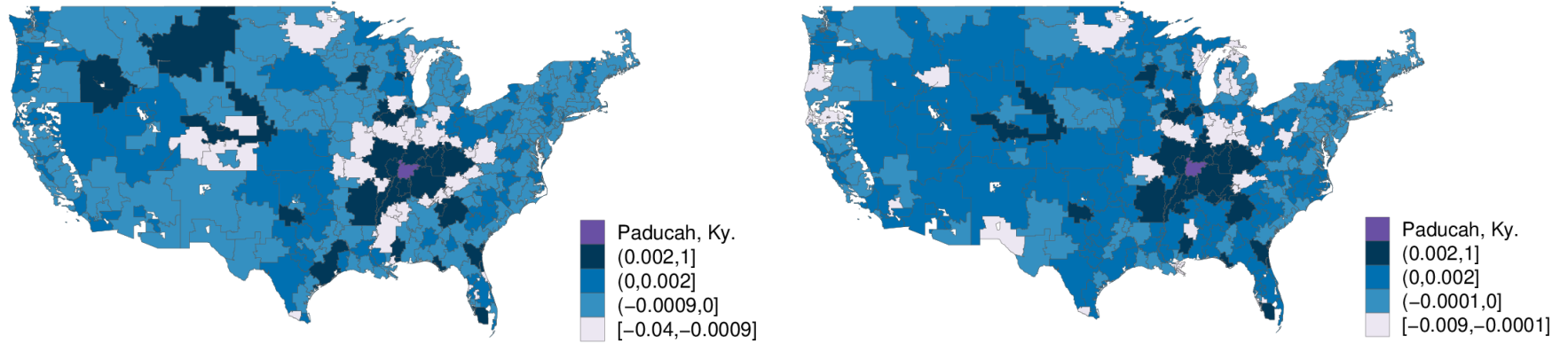
*Notes:* Panels (a) and (b) show the impacts of increasing reimbursements by 30% in the Boston, Mass. HRR ( $\hat{R}_i = 1.3$ ) based on our estimated model. Panel (a) illustrates the percentage change in quality of care  $\delta_i$  provided in each region. Panel (b) illustrates the percentage change in the value of market access  $\Phi_i$  for patients who live in an region. Panels (c) and (d) are analogous, but for a 30% increase in reimbursements in Paducah, Ky., a net importer. In all panels, the predicted change for the region whose reimbursement changes (“treated region”) is listed on the map itself. In both cases, the quality produced in neighboring regions declines (Panels (a) and (c)). Patients in regions near Boston benefit from increased access to the treated region (Panel (b)), so there is a negative relationship between the percentage changes in  $\delta$  and  $\Phi$  across regions. In contrast, patients in regions near Paducah suffer a decrease in access (Panel (d)). The contrasting outcomes stem from Boston being a net exporter and Paducah being a net importer in the baseline equilibrium. The exercise is described in detail in Section 6.

Figure 11: Changes in access  $\hat{\Phi}_{j\kappa}$  by income when increasing reimbursements

*Notes:* This figure summarizes the counterfactual outcomes of 30% higher reimbursements in one HRR as a function that HRR's population size. The nationwide return is the percentage increase in patient market access  $\sum_{\kappa} \sum_j N_{j\kappa} \hat{\Phi}_{j\kappa}$  per percentage increase in nationwide expenditures  $\sum_i Q_i R_i$ . The tercile-specific return is the increase in tercile-specific patient market access  $\sum_j N_{j\kappa} \hat{\Phi}_{j\kappa}$ . Increasing reimbursements in more populous HRRs has the highest return when measured as impact on aggregate market access. Subsidies in less populous regions favor lower-income patients, primarily because there are more low-income patients living in and close to smaller regions.

Figure 12: Counterfactual outcomes when changing travel costs for Paducah, Ky. residents

(a) Change (%) in quality  $\delta_i$ : reducing Paducah residents' travel costs by 30% (b) Change (%) in access  $\Phi_i$ : reducing Paducah residents' travel costs by 30%



Paducah, Ky. = -7.1%

Paducah, Ky. = 5.8%

Notes: Both panels show the impact of a 30% fall in travel costs for Paducah residents ( $\hat{\rho}_{ij} = 1.3 \forall i \neq \text{Paducah}$ ). Panel (a) illustrates the percentage change in quality of care  $\delta_i$  provided in each region. Panel (b) illustrates the percentage change in the value of market access  $\Phi_i$  for patients who live in a region. The note shows the change for Paducah itself. Reduced travel costs for Paducah residents improves their market access but reduces the quality of care produced in Paducah itself. The increase in imports by Paducah residents causes service quality in neighboring regions to increase because of scale effects. This higher quality in turn attracts additional patients from the ring surrounding them, reducing quality slightly in that distant ring.

Table 1: Aggregate medical services exhibit a strong home-market effect

Estimation method:	(1)	(2)	(3)	(4)
	PPML	PPML	PPML	IV
Provider-market population (log)	0.638 (0.0634)	0.643 (0.0610)	0.645 (0.0455)	0.597 (0.0732)
Patient-market population (log)	0.377 (0.0615)	0.376 (0.0587)	0.406 (0.0423)	0.360 (0.0521)
Distance (log)	-1.664 (0.0501)	0.0996 (0.307)		0.0796 (0.270)
Distance (log, squared)		-0.178 (0.0299)		-0.177 (0.0265)
Observations	93,636	93,636	93,636	93,636
Distance elasticity at mean		-2.46		-2.46
Distance deciles			Yes	

*Notes:* This table reports estimates of equation (8), which estimates the presence of weak or strong home-market effects. The sample is all HRR pairs ( $N = 306^2$ ), and the dependent variable in all regressions is the value of trade. The independent variables are patient- and provider-market log population, log distance between HRRs, and an indicator for same-HRR observations ( $i = j$ ). The positive coefficient on provider-market log population implies a weak home-market effect, and the fact that this coefficient exceeds that on patient-market population implies a strong home-market effect. Column 2 makes the distance coefficient more flexible by adding a control for the square of log distance. Column 3 replaces parametric distance specifications with fixed effects for each decile of the distance distribution. Column 4 uses the provider-market and patient-market log populations in 1940 as instruments for the contemporaneous log populations when estimating by generalized method of moments (GMM). Trade flows are computed from the Medicare 20% carrier Research Identifiable Files, using the dollar value of physician services, excluding emergency-room care and assigning each procedure its national average price. HRR definitions are from the Dartmouth Atlas Project. Standard errors (in parentheses) are two-way clustered by patient market and provider market.



Table 2: The home-market effect is stronger for rare procedures

	(1)	(2)	(3)	(4)	(5)	(6)
Provider-market population (log)	0.638 (0.0634)	0.624 (0.0613)	0.623 (0.0614)		0.630 (0.0598)	
Patient-market population (log)	0.377 (0.0615)	0.379 (0.0590)	0.380 (0.0591)		0.379 (0.0572)	
Provider-market population (log) $\times$ rare			0.306 (0.0472)	0.291 (0.0455)	0.316 (0.0480)	0.287 (0.0458)
Patient-market population (log) $\times$ rare			-0.229 (0.0698)	-0.219 (0.0671)	-0.232 (0.0704)	-0.211 (0.0658)
Observations	187,272	113,468	113,468	113,468	113,468	113,468
Distance controls	Yes	Yes	Yes	Yes		
Distance [quadratic] controls					Yes	Yes
Patient-provider-market-pair FEs				Yes		Yes

*Notes:* This table reports estimates of equation (11), which introduces interactions with an indicator for whether a procedure is “rare” (provided less often than the median procedure, when adding up all procedures provided nationally). The interactions with patient- and provider-market population reveal whether the home-market effect is larger for rare procedures. The unit of observation is {rare indicator, exporting HRR, importing HRR} so the number of observations is  $2 \times 306^2$  in column 1, and the dependent variable in all regressions is the value of trade. Columns 2 onwards drop HRR pairs with zero trade in both procedure groups, and column 2 shows that this restriction has a negligible impact on the estimated log population coefficients. Columns 3 onwards include the rare indicator interacted with patient- and provider-market populations and distance covariates. Columns 1–4 control for distance using the log of distance between HRRs. Columns 5 and 6 add a control for the square of log distance. Columns 4 and 6 introduce a fixed effect for each  $ij$  pair of patient market and provider market, so these omit all covariates that are not interacted with the rare indicator. The positive coefficient on provider-market population  $\times$  rare across all columns indicates that the home-market effect is stronger for rare than for common services. The negative coefficient on patient-market population  $\times$  rare across all columns indicates that the *strong* home-market effect has a larger magnitude for rare services. Trade flows are computed from the Medicare 20% carrier Research Identifiable Files, using the dollar value of physician services, excluding emergency-room care and assigning each procedure its national average price. HRR definitions are from the Dartmouth Atlas Project. Standard errors (in parentheses) are two-way clustered by patient market and provider market.

Table 3: The stronger home-market effect for rare procedures is robust to instrumenting for population

	(1)	(2)	(3)	(4)	(5)	(6)
Geography:	HRR	HRR	CBSA	CBSA	CBSA	CBSA
Instrument:	1940 pop	1940 pop	1940 pop	1940 pop	Bedrock	Bedrock
Procedure Sample:	Common	Rare	Common	Rare	Common	Rare
Provider-market population (log)	0.595 (0.0733)	1.080 (0.0913)	0.716 (0.0249)	0.895 (0.0388)	1.157 (0.307)	1.753 (0.524)
Patient-market population (log)	0.361 (0.0522)	0.0476 (0.114)	0.396 (0.0261)	0.328 (0.0344)	0.182 (0.373)	-0.582 (0.580)
Distance (log)	0.0756 (0.270)	0.973 (0.449)	-3.412 (0.294)	-1.378 (0.989)	-4.678 (1.049)	-4.631 (2.520)
Distance (log, squared)	-0.177 (0.0265)	-0.261 (0.0503)	0.105 (0.0287)	-0.0742 (0.0935)	0.210 (0.0845)	0.181 (0.199)
Observations	93,636	93,636	857,476	857,476	781,456	781,456
Distance elasticity at mean	-2.45	-2.76	-1.91	-2.43	-1.68	-2.05

*Notes:* This table reports estimates of equation (8), when separating procedures into those above- and below-median frequency and instrumenting for log population. The dependent variable in all regressions is the value of trade. Trade flows are computed from the Medicare 20% carrier Research Identifiable Files, using the dollar value of physician services, excluding emergency-room care and assigning each procedure its national average price. We report coefficients on provider market population, patient market population, log distance, and log distance squared. Every specification also includes a same-market ( $i = j$ ) indicator variable. The odd-numbered columns are trade in above-median-frequency procedures; the even-numbered columns are trade in below-median-frequency procedures. In columns 1 and 2, the sample is all HRR pairs ( $N = 306^2$ ). In columns 3 and 4, the sample is all CBSA pairs ( $N = 926^2$ ). In columns 5 and 6, the sample is all CBSA pairs for which the bedrock-depth instrumental variable is available ( $N = 844^2$ ). We use 1940 population counts to produce two instrumental variables: 1940 population in the patient market and 1940 population in the provider market are instruments for log population in the patient market and log population in the provider market, respectively. Similarly, we use bedrock depth to produce two instrumental variables for CBSAs. Both the strong home-market effect and its larger magnitude for rare procedures are robust to instrumenting for population, estimating by GMM. Standard errors (in parentheses) are two-way clustered by patient market and provider market.

Table 4: The home-market effect is stronger for rarer diagnoses

	(1)	(2)	(3)	(4)	(5)	(6)
Provider-market population (log)	0.638 (0.0638)	0.624 (0.0616)	0.619 (0.0600)		0.625 (0.0584)	
Patient-market population (log)	0.379 (0.0616)	0.381 (0.0591)	0.384 (0.0579)		0.383 (0.0561)	
Provider-market population (log) $\times$ rare			0.0726 (0.0557)	0.0687 (0.0520)	0.0762 (0.0564)	0.0681 (0.0508)
Patient-market population (log) $\times$ rare			-0.0430 (0.0429)	-0.0409 (0.0404)	-0.0427 (0.0443)	-0.0378 (0.0396)
Observations	187,272	112,626	112,626	112,626	112,626	112,626
Distance controls	Yes	Yes	Yes	Yes		
Distance [quadratic] controls					Yes	Yes
Patient-provider-market-pair FEs				Yes		Yes

*Notes:* This table augments equation (8) by adding interactions with an indicator for whether a diagnosis is “rare” (provided less often than the median diagnosis, when adding up all patients receiving the diagnosis nationally) or “common” (more often than median). The interactions with patient- and provider-market population reveal whether the home-market effect is larger for rare diagnoses. The unit of observation is {rare indicator, exporting HRR, importing HRR} so the number of observations is  $2 \times 306^2$  in column 1, and the dependent variable in all regressions is the value of trade. Valid primary diagnoses observed in 1,000 distinct claims or more nationally in the professional fees 20% sample are included. Columns 2 onwards drop HRR pairs with zero trade, and column 2 shows that this restriction has a negligible impact on the estimated log population coefficients. Columns 1–4 control for distance using the log of distance between HRRs. Columns 5 and 6 add a control for the square of log distance. Columns 4 and 6 introduce a fixed effect for each  $ij$  pair of patient market and provider market, so these omit the patient- and provider-market population covariates. The positive coefficient on provider-market population  $\times$  rare across all columns indicates that the home-market effect is stronger for rare than for common diagnoses. The negative coefficient on patient-market population  $\times$  rare across all columns indicates that the *strong* home-market effect is especially true for rare diagnoses. Trade flows are computed from the Medicare 20% carrier Research Identifiable Files, using the dollar value of physician services, excluding emergency-room care and assigning each procedure its national average price. HRR definitions are from the Dartmouth Atlas Project. Standard errors (in parentheses) are two-way clustered by patient market and provider market.

Table 5: Scale elasticity estimates

Panel A: All services	No Controls		Controls	
	No Diag	Diag	No Diag	Diag
OLS	0.804 (0.044)	0.778 (0.030)	0.884 (0.046)	0.793 (0.038)
2SLS: population (log)	0.799 (0.049)	0.716 (0.030)	0.871 (0.052)	0.721 (0.037)
2SLS: population (1940, log)	0.660 (0.093)	0.550 (0.069)	0.640 (0.082)	0.559 (0.059)
Panel B: Rare services				
OLS	1.089 (0.045)	0.945 (0.030)	1.124 (0.046)	0.956 (0.033)
2SLS: population (log)	1.033 (0.049)	0.910 (0.037)	1.072 (0.051)	0.920 (0.041)
2SLS: population (1940, log)	0.951 (0.081)	0.832 (0.061)	0.940 (0.073)	0.832 (0.055)

*Notes:* This table reports estimates of  $\hat{\alpha}$  from ordinary least squares (OLS) or two-stage least squares (2SLS) regressions of the form  $\widehat{\ln \delta}_i = \alpha \ln Q_i + \ln R_i + u_i$ , where  $\widehat{\ln \delta}_i$  is estimated in equation (12),  $Q_i$  is region  $i$ 's total production of non-emergency-room physician services for Medicare beneficiaries,  $R_i$  is Medicare's Geographic Adjustment Factor, and  $u_i$  is an error term. In the rows labeled "2SLS" we instrument for  $\ln Q_i$  using the specified instruments. The  $\ln R_i$  control is omitted in the columns labeled "no controls". Appendix Table D.11 reports analogous estimates at the CBSA level, which allows us to also control for input costs (as input cost data are more reliable for CBSAs than for HRRs). In the columns labeled "no diag",  $Q_{ii}$  observations were omitted when estimating  $\widehat{\ln \delta}_i$  in equation (12). Standard errors are robust to heteroskedasticity. Across all of the permutations of our method, we estimate substantial scale economies.

Table 6: Regression of  $\hat{\Phi}_{j\mathfrak{t}}$  on tercile dummies and trade shares

	Nationwide Reimbursement Increase			Boston Reimbursement Increase		
	(1)	(2)	(3)	(4)	(5)	(6)
Income tercile = 2	1.085 (0.0559)	1.090 (0.0560)	-0.144 (0.0325)	3.53e-05 (0.00938)	0.00300 (0.00966)	-0.00496 (0.00215)
Income tercile = 3	1.579 (0.0718)	1.563 (0.0704)	-0.242 (0.0561)	0.0698 (0.0349)	0.0649 (0.0347)	-0.00730 (0.00267)
Imported share ( $1 - m_{0j\kappa} - m_{jj\kappa}$ )			-0.520 (0.131)			
$m_{0j\kappa}$			-12.66 (0.429)			
$m_{\text{Boston},j\kappa}$						36.93 (0.438)
Constant	8.759 (0.0598)	8.762 (0.0406)	11.11 (0.0781)	0.0983 (0.105)	0.0989 (0.0127)	-0.0690 (0.00266)
Observations	885	885	885	885	885	885
R-squared	0.498	0.673	0.988	0.000	0.980	1.000
HRR fixed effects	No	Yes	Yes	No	Yes	Yes

*Notes:* This table uses linear regressions to summarize how market access changes across HRRs  $j$  and income terciles  $\kappa$  in response to two different counterfactual policies. The dependent variable in all columns is the percentage change in market access,  $100 \times (\hat{\Phi}_{j\mathfrak{t}} - 1)$ . Standard errors (in parentheses) are clustered by market. Columns 1, 2, and 3 consider a 10% reimbursement increase nationwide. Columns 4, 5, and 6 consider a 30% reimbursement increase in Boston only. The constant in the first regression reports the percentage change for the lowest income terciles, and the coefficients on the other terciles are the additional percentage point gain for those terciles relative to the lowest. Other controls include the outside option market share  $m_{0j\kappa}$ , imported share  $1 - m_{0j\kappa} - m_{jj\kappa}$  (where  $m_{jj\kappa}$  is local production), and Boston's market share  $m_{\text{Boston},j\kappa}$ . The coefficients are much smaller in columns 4, 5, and 6 because only Boston is treated, so most of the country is hardly affected. When we add market share controls, the coefficients indicating tercile differences become much smaller, indicating that baseline trade patterns drive the distributional impacts.

# Appendix – For Online Publication

Market Size and Trade in Medical Services

Jonathan I. Dingel, Joshua D. Gottlieb, Maya Lozinski, and Pauline Mourot  
March 2023

## A Theory appendix

### A.1 Model with multiple types of patients

This section extends the model to feature multiple types of patients who face different trade costs. There is a finite set of patient types, which are indexed by  $\kappa$ . A patient type is defined by the trade costs  $\rho_{ij(k)} = \rho_{ij}^\kappa \forall k \in \kappa$ . Qualities  $\delta_i$ , including the outside option  $\delta_0$ , are the same for all patient types. The demand by patients of type  $\kappa$  residing in location  $j$  for procedures performed by providers in location  $i$  is now given by

$$Q_{ij}^\kappa = \frac{\delta_i N_j^\kappa}{\Phi_j^\kappa} \rho_{ij}^\kappa.$$

The aggregate gravity equation is the sum of type-specific gravity equations:

$$Q_{ij} = \sum_{\kappa} Q_{ij}^\kappa = \delta_i \sum_{\kappa} \frac{N_j^\kappa}{\Phi_j^\kappa} \rho_{ij}^\kappa. \quad (\text{A.1})$$

The free-entry condition (4) remains unchanged with the introduction of multiple patient types:

$$R_i = \frac{w_i \delta_i}{A_i Q_i^\alpha}.$$

In equilibrium, market clearing requires that

$$Q_i = \left( \frac{w_i \delta_i}{A_i R_i} \right)^{1/\alpha} = \delta_i \sum_j \sum_{\kappa} \frac{N_j^\kappa}{\Phi_j^\kappa} \rho_{ij}^\kappa \implies \delta_i = \left( \frac{A_i R_i}{w_i} \right)^{1/(1-\alpha)} \left( \sum_j \sum_{\kappa} \frac{N_j^\kappa}{\Phi_j^\kappa} \rho_{ij}^\kappa \right)^{\alpha/(1-\alpha)}.$$

## A.2 Derivations of results in Section 1.5

Abusing notation so that  $\mathcal{I}$  is both the set and number of regions, equations (2) and (3) together constitute  $2\mathcal{I}$  equations with  $2\mathcal{I}$  unknowns. For the special case of  $H(Q_i) = Q_i^\alpha$  and  $K(\delta_i) = \delta_i$ , this reduces to the following  $\mathcal{I}$  equations with the unknowns  $\{\delta_i\}_{i=1}^{\mathcal{I}}$ :

$$\delta_i = \left( \frac{\bar{R}A_i}{w_i} \right)^{\frac{1}{1-\alpha}} \left( \sum_{j \in \mathcal{I}} \frac{\rho_{ij}}{\sum_{i' \in 0 \cup \mathcal{I}} \delta_{i'} \rho_{i'j}} N_j \right)^{\frac{\alpha}{1-\alpha}}$$

Following Costinot et al. (2019), we examine the home-market effect in the neighborhood of a symmetric equilibrium. For brevity, assume  $\frac{\bar{R}A_i}{w_i} = 1 \forall i$ . Note that at the symmetric equilibrium:

$$\bar{\delta}^{\frac{1-\alpha}{\alpha}} = \frac{1}{1 + \bar{\delta} + \sum_{i' \neq i} \bar{\delta} \rho} \bar{N} + \sum_{j \neq i} \frac{\rho}{1 + \bar{\delta} + \sum_{i' \neq j} \bar{\delta} \rho} \bar{N} = \frac{1 + (\mathcal{I} - 1)\rho}{\bar{\Phi}} \bar{N} = \frac{\bar{\Phi} - 1}{\bar{\Phi}} \frac{\bar{N}}{\bar{\delta}}. \quad (\text{A.2})$$

Given  $\alpha > 0$ , totally differentiating the above system of equations in terms of  $\{d\delta_i, dN_i\}_{i=1}^{\mathcal{I}}$  and evaluating it at the symmetric equilibrium yields the following expression:

$$\frac{\bar{\Phi}^2}{\bar{N}} \frac{1 - \alpha}{\alpha} \bar{\delta}^{\frac{1-2\alpha}{\alpha}} d\delta_i = - \left[ d\delta_i + \rho \sum_{i' \neq i} d\delta_{i'} \right] + \bar{\Phi} \frac{dN_i}{\bar{N}} + \sum_{j \neq i} -\rho \left[ d\delta_j + \rho \sum_{i' \neq j} d\delta_{i'} \right] + \sum_{j \neq i} \rho \bar{\Phi} \frac{dN_j}{\bar{N}}.$$

Given  $dN_1 > 0$  and  $dN_j = 0 \forall j \neq 1$ , we obtain the following expression for  $d \ln \delta_1$ :

$$d \ln \delta_1 = \frac{\frac{\bar{\Phi}}{\bar{\delta}} d \ln N_1 - (\mathcal{I} - 1)(2\rho + ((\mathcal{I} - 2)\rho^2)) d \ln \delta_{j \neq 1}}{\frac{\bar{\Phi}^2}{\bar{N}} \frac{(1-\alpha)}{\alpha} \bar{\delta}^{\frac{1-2\alpha}{\alpha}} + 1 + (\mathcal{I} - 1)\rho^2}. \quad (\text{A.3})$$

Further tedious algebra delivers the following expression for quality changes:

$$d \ln \delta_1 - d \ln \delta_{j \neq 1} = \frac{(1 - \rho)}{\frac{\bar{\Phi}^2}{\bar{N}} \frac{(1-\alpha)}{\alpha} \bar{\delta}^{\frac{1-2\alpha}{\alpha}} + (1 - \rho)^2} \frac{\bar{\Phi}}{\bar{\delta}} d \ln N_1. \quad (\text{A.4})$$

Equation (A.2) implies that  $\frac{\Phi^2 (1-\alpha)}{N} \bar{\delta}^{-\frac{1-2\alpha}{\alpha}} = \left(\frac{1-\alpha}{\alpha}\right) \frac{\Phi(\Phi-1)}{\bar{\delta}}$  and therefore

$$d \ln \delta_1 - d \ln \delta_{j \neq 1} = \frac{(1-\rho)}{\left(\frac{1-\alpha}{\alpha}\right) \frac{\Phi(\Phi-1)}{\bar{\delta}} + (1-\rho)^2 \frac{\bar{\Phi}}{\bar{\delta}}} d \ln N_1 = \left[ \frac{1-\alpha}{\alpha} \frac{(\bar{\Phi}-1)}{(1-\rho)\bar{\delta}} + \frac{(1-\rho)\bar{\delta}}{\bar{\Phi}} \right]^{-1} d \ln N_1 > 0.$$

The last expression above is reported in Section 1.5.

Prior to deriving the weak and strong home-market effects, we obtain an expression for  $\frac{d \ln \delta_j}{d \ln N_1}$  for  $j \neq 1$  around the symmetric equilibrium. Define  $\bar{Q} \equiv \frac{\Phi^2 (1-\alpha)}{N} \bar{\delta}^{-\frac{1-2\alpha}{\alpha}} > 0$ . Combining the expressions for  $d \ln \delta_1$  from equation (A.3) and for  $d \ln \delta_1 - d \ln \delta_{j \neq 1}$  from equation (A.4) yields the following:

$$\frac{d \ln \delta_{j \neq 1}}{d \ln N_1} = \frac{\bar{\Phi}}{\bar{\delta}} \frac{\bar{Q}\rho + \rho^3(\mathcal{I}-1) - \rho^2(\mathcal{I}-2) - \rho}{(\bar{Q} + (1-\rho)^2)(\bar{Q} + 1 + \rho^2 + 2\rho(\mathcal{I}-1) + \mathcal{I}\rho^2(\mathcal{I}-2))}$$

The weak home-market effect is derived as follows:

$$\begin{aligned} \ln Q_{1,j \neq 1} &= \alpha \ln Q_1 + \ln \rho - \ln \Phi_j + \ln N_j \\ \frac{d \ln Q_{1,j \neq 1}}{d \ln N_1} &= \alpha \frac{d \ln Q_1}{d \ln N_1} - \frac{\alpha}{\Phi_j} \left( \rho Q_1^{\alpha-1} \frac{d Q_1}{d \ln N_1} + Q_j^{\alpha-1} \frac{d Q_j}{d \ln N_1} + \rho \sum_{i' \neq 1,j} Q_{i'}^{\alpha-1} \frac{d Q_{i'}}{d \ln N_1} \right) \\ &= \frac{d \ln \delta_1}{d \ln N_1} - \frac{1}{\Phi_j} \left( \rho \delta_1 \frac{d \ln \delta_1}{d \ln N_1} + \delta_j \frac{d \ln \delta_j}{d \ln N_1} + \rho \sum_{i' \neq 1,j} \delta_{i'} \frac{d \ln \delta_{i'}}{d \ln N_1} \right) \\ &= \left( \frac{\bar{N} - Q_{1j}}{\bar{N}} \right) \frac{d \ln \delta_1}{d \ln N_1} - \left( \frac{\bar{N} - Q_{0j} - Q_{1j}}{\bar{N}} \right) \frac{d \ln \delta_j}{d \ln N_1} \\ &= \left( \frac{\bar{N} - Q_{1j}}{\bar{N}} \right) \left[ \frac{d \ln \delta_1}{d \ln N_1} - \frac{d \ln \delta_j}{d \ln N_1} \right] + \frac{Q_{0j}}{\bar{N}} \frac{d \ln \delta_j}{d \ln N_1} \\ &= \frac{\bar{\Phi}}{\bar{\delta} \bar{N}} \frac{1}{\bar{Q} + (1-\rho)^2} \left[ (Q_{jj} + (\mathcal{I}-2)Q_{1j})(1-\rho) + \frac{Q_{0j}}{\bar{Q} + 1 + \rho^2 + 2\rho(\mathcal{I}-1) + \mathcal{I}\rho^2(\mathcal{I}-2)} \right. \\ &\quad \left. \times \{ \bar{Q} + (\rho-1)^2 + 2(\mathcal{I}-1)(\rho-\rho^2) + (\mathcal{I}-1)(\mathcal{I}-2)[\rho^2 - \rho^3] \} \right] \\ &> 0. \end{aligned}$$



The condition for the strong home-market effect is derived as follows:

$$\begin{aligned}
Q_{1,j \neq 1} - Q_{j \neq 1,1} &= \frac{Q_1^\alpha \rho}{1 + Q_1^\alpha \rho + Q_j^\alpha + \sum_{i \neq 1,j} Q_i^\alpha \rho} N_j - \frac{Q_j^\alpha \rho}{1 + Q_1^\alpha + Q_j^\alpha \rho + \sum_{i \neq 1,j} Q_i^\alpha \rho} N_1 \\
d \ln Q_{1,j \neq 1} - d \ln Q_{j \neq 1,1} &= d \ln N_j - d \ln N_1 + \alpha \left[ 1 + (1 - \rho) \frac{\bar{Q}^\alpha}{\bar{\Phi}} \right] (d \ln Q_1 - d \ln Q_j) \\
&= -d \ln N_1 + \left[ 1 + (1 - \rho) \frac{\bar{\delta}}{\bar{\Phi}} \right] (d \ln \delta_1 - d \ln \delta_j) \\
&= \left[ \frac{1 - \frac{1-\alpha}{\alpha} \frac{1+(\mathcal{I}-1)\rho}{1-\rho}}{\frac{1-\alpha}{\alpha} \frac{1+(\mathcal{I}-1)\rho}{(1-\rho)} + \frac{(1-\rho)\bar{\delta}}{1+(1+(\mathcal{I}-1)\rho)\bar{\delta}}} \right] d \ln N_1.
\end{aligned}$$

There is a strong home-market effect in the neighborhood of the symmetric equilibrium if and only if  $d \ln Q_{1,j \neq 1} - d \ln Q_{j \neq 1,1} > 0$ .

$$\left[ \frac{1 - \frac{1-\alpha}{\alpha} \frac{1+(\mathcal{I}-1)\rho}{1-\rho}}{\frac{1-\alpha}{\alpha} \frac{1+(\mathcal{I}-1)\rho}{(1-\rho)} + \frac{(1-\rho)\bar{\delta}}{1+(1+(\mathcal{I}-1)\rho)\bar{\delta}}} \right] d \ln N_1 > 0 \iff \frac{\alpha}{1-\alpha} > \frac{1 + (\mathcal{I} - 1)\rho}{1 - \rho}$$

This is true if  $\alpha$  is large enough and  $\rho$  is small enough.

Our difference-in-differences prediction concerns how the effect of market size on net exports varies with the number of potential patients  $\bar{N}$ . Given the scale elasticity  $\alpha$  and (inverse) trade costs  $\rho$ , the denominator of the right side of equation (6) is increasing in the symmetric-equilibrium quality  $\bar{\delta}$ . For two procedures that both exhibit a strong home-market effect because they have the same scale elasticity and trade costs, the effect of population size on net exports will be larger for the procedure with lower service quality. The symmetric-equilibrium service quality is increasing in the number of potential patients  $\bar{N}$  because there are increasing returns (see equation (A.2)). Thus, in the neighborhood of the symmetric equilibrium, the strength of a strong home-market effect is decreasing in the number of potential patients.

## B Data appendix

### B.1 Procedure frequency in main sample compared with aggregate and private data

Medicare provides two public-use files based on 100 percent claims. The first one contains the complete count of procedures billed by HCPCS code but does not have information about providers. We use it to confirm that procedure counts based on the confidential data do not suffer substantial sampling bias. In Figure D.9, we split procedure codes into 10 deciles based on their national frequencies, separately in the confidential and public datasets. This generates a 100-cell matrix by decile pair. We plot the share of procedures in each cell in this matrix to determine how well the two datasets align. The vast majority of the codes are on the diagonal, with almost all of the remainder adjacent to the diagonal. This suggests that sampling error is not causing us to mischaracterize procedure frequency.

Medicare provides a second public file at the level of physician-by-procedure (HCPCS code). This summary does not contain any patient-level information so cannot be used to study trade flows, but we can use it to replicate analyses based on the location of production and physician experience. This file is censored such that physician-by-procedure pairs with 10 or fewer observations per year are suppressed, which makes for a more complicated bias than simple 20 percent random sampling. Nevertheless, all of the results that can be tested on this sample confirm those found in the 20 percent sample.

Since our procedure frequency measures rely on Medicare data, we would mismeasure frequency if the Medicare population uses a substantially different composition of care from the broader population. For example, childbirth is less common among Medicare beneficiaries. So our frequency measures may not capture the true national frequency of a procedure.

We address this by comparing procedure frequencies between the Medicare public data and private data from the Health Care Cost Institute (HCCI). The HCCI data contain claims for about 55 millions of privately insured patients (about 35% of individuals with employer-

based insurance). We only consider HCPCS codes performed on at least eleven patients in the HCCI data. Note that frequencies are computed for all providers here, not only MDs and DOs. The authors acknowledge the assistance of the Health Care Cost Institute (HCCI) and its data contributors, Aetna, Humana, and Blue Health Intelligence, in providing the claims data analyzed in this section.

We examine whether procedures classified as above median frequency in one dataset are above median frequency in the other dataset. Table D.12 shows that 88% of the services above median frequency in Medicare are also as above median frequency in the HCCI data. Similarly, 82% of the services below median frequency in Medicare are also below median frequency in the HCCI data.

We next compare classifications of procedures' frequency deciles in Figure D.10. Analogous to Figure D.9, this plot visualizes the share of procedures which fall into each of pair of frequency decile bins in HCCI and Medicare data. The two classifications appear to coincide relatively well, with slightly stronger agreement for very frequent procedures compared to rarer procedures in the Medicare public-use data. Overall, the frequency classifications of procedures coincide well between Medicare public-use data and HCCI data.

## B.2 Additional details on data sources

**Physician earnings.** The Gottlieb et al. (2020) earnings data depicted in Appendix Figure D.2 are only available for 111 commuting zones. The American Community Survey (ACS) covers far more CBSAs, but this source top-codes income for a substantial share of doctors.

**U.S. News and World Report.** The publication produces an overall ranking and rankings for 12 particular specialties. We count the number of times each HRR's hospitals appear on any of these 13 lists.<sup>53</sup> Thus, higher ranking on the horizontal axis indicates a region

---

<sup>53</sup>Results are similar when we use other methods to aggregate the rankings information, including when we account for the ordered nature of the lists.

has some combination of more top-ranked hospitals, or each of its hospitals performs well in many specialty areas.

### **B.3 Geographic price adjustments**

**Professional fees.** To adjust for geographic price variation in the professional fees, we compute a national average price per Healthcare Common Procedure Coding System (HCPCS) code as the sum of the line allowed amount, which includes the line item’s Medicare-paid and beneficiary-paid amounts (i.e., deductible, copayment, and coinsurance), divided by the sum of the line service count per HCPCS code nationally. We then apply this average price to all billing for the HCPCS code when computing total spending across services.

**Hospital inpatient fees.** We use the field “final standard payment amount” in the MedPAR file, which is computed as described in Finkelstein, Gentzkow, and Williams (2016) and Gottlieb et al. (2010). This represents “a standard Medicare payment amount, without the geographical payment adjustments and some of the other add-on payments that go to the hospitals” according to the data documentation.

**Hospital outpatient fees.** To adjust for geographic price variation in hospital outpatient fees, we compute a national average price per Healthcare Common Procedure Coding System (HCPCS) code, Ambulatory Payment Classifications (APC) code, and revenue center code. HCPCS codes reflect the procedure performed and APC codes reflect a prospective payment system applicable to outpatient analogous to Diagnosis Related Groups (DRGs) for inpatient claims. Revenue center contains information on the place of service, e.g. rehabilitation or acute care, so we consider two procedures performed in different revenue centers as different procedures for price adjustment purposes.

The total amount per claim line is calculated as the sum of the claim (Medicare) payment amount, the primary payer amount, the Part B beneficiary co-insurance amount, the beneficiary Part B deductible amount, and the beneficiary blood deductible amount. These

amounts are summed nationally for each {HCPCS code, APC code, revenue center code} triplet, and divided by the frequency of that triplet to obtain a national average price. We then apply this average price to all instances of that {HCPCS code, APC code, revenue center code} combination when computing total spending across services.

## **B.4 Residential measurement error**

This appendix uses two methods to investigate potential measurement error in patients' residential location. The first source of potential error is “snowbird” patients, who have multiple residences and therefore may appear to travel farther than they actually do. They may need medical care while spending months in a warmer HRR that is not the one listed as their main residence (or vice versa). Our results are robust to two methods of removing potential snowbirds: excluding Arizona, California, and Florida, following Finkelstein, Gentzkow, and Williams (2016), and excluding the 10% of HRRs with the highest share of second homes in American Community Survey data. These results are in Tables [D.4](#) and [D.5](#). The results are little changed by these sample restrictions.

We test for more general location measurement error by examining how far patients appear to travel for dialysis. Since Medicare patients requiring dialysis must generally visit a dialysis center thrice weekly, they are unlikely to go substantial distances for this service. Table [D.6](#) compares travel distances for dialysis with other care. Dialysis patients appear to travel less than one-quarter as often as other patients—and even less when excluding snowbird states—suggesting that our residential location assignment is largely accurate.

## **B.5 Scale elasticity estimation with unobserved market segments**

Our data only contain procedure-level production and consumption in Traditional Medicare (TM), not for Medicare Advantage (MA) or non-Medicare (NM) patients. We quantify how this biases our estimate of the scale elasticity,  $\alpha$ , based on geographic variation. Suppose

the production function is

$$\ln \delta_i = \alpha \ln Q_i + u_i,$$

where  $Q_i = Q_i^{\text{TM}} + Q_i^{\text{MA}} + Q_i^{\text{NM}}$  is the total quantity produced in region  $i$ , of which we only observe  $Q_i^{\text{TM}}$ . When we estimate the scale elasticity  $\alpha$  using  $Q_i^{\text{TM}}$  as a proxy for  $Q_i$ , our regression coefficient may be biased:

$$\frac{\text{Cov}(\ln \delta_i, \ln Q_i^{\text{TM}})}{\text{Var}(\ln Q_i^{\text{TM}})} = \frac{\text{Cov}(\alpha \ln Q_i, \ln Q_i^{\text{TM}})}{\text{Var}(\ln Q_i^{\text{TM}})} + \frac{\text{Cov}(u_i, \ln Q_i^{\text{TM}})}{\text{Var}(\ln Q_i^{\text{TM}})} = \alpha \zeta,$$

where  $\zeta$ , which governs the bias, is the regression coefficient from  $\ln Q_i = \zeta \ln Q_i^{\text{TM}} + u_i$ .

To compute  $\zeta$  we differentiate the identity  $Q_i = Q_i^{\text{TM}} \left( 1 + \frac{Q_i^{\text{MA}}}{Q_i^{\text{TM}}} + \frac{Q_i^{\text{NM}}}{Q_i^{\text{TM}}} \right)$  with respect to  $Q_i^{\text{TM}}$ , which we observe:

$$\frac{d \ln Q_i}{d \ln Q_i^{\text{TM}}} = 1 + s_i^{\text{MA}} \varrho_i^{\text{MA}} + s_i^{\text{NM}} \varrho_i^{\text{NM}},$$

where  $s_i^{\text{MA}} \equiv \frac{Q_i^{\text{MA}}}{Q_i^{\text{TM}} + Q_i^{\text{MA}} + Q_i^{\text{NM}}}$  is the MA share of production in region  $i$ ,  $\varrho_i^{\text{MA}} \equiv \frac{d \ln \frac{Q_i^{\text{MA}}}{Q_i^{\text{TM}}}}{d \ln Q_i^{\text{TM}}}$  is the TM production elasticity of relative production, and  $s_i^{\text{NM}}$  and  $\varrho_i^{\text{NM}}$  are similarly defined for non-Medicare (NM) insurance. To make it feasible to estimate these elasticities, we assume that they are constant across regions. If relative quantities produced are uncorrelated with the Traditional Medicare quantity produced ( $\varrho^{\text{MA}} = \varrho^{\text{NM}} = 0$ ), then  $\zeta = 1$  and  $\alpha \zeta$  is an unbiased estimate of the scale elasticity  $\alpha$ .<sup>54</sup> Otherwise, we need estimates of the average production shares  $\bar{s}^{\text{MA}}$  and  $\bar{s}^{\text{NM}}$  and the regression coefficients  $\varrho^{\text{MA}}$  and  $\varrho^{\text{NM}}$  to compute  $\zeta$ .

We compute the production shares using data on aggregate expenditures and price deflators from prior research. Medicare, including both TM and MA, paid for \$153 billion of the \$525 billion spent nationally on physician services in 2017 (Centers for Medicare and Medicaid Services, 2022). Per capita spending and prices are similar between the two parts of

---

<sup>54</sup>A special case would be if the quantity of care produced outside of TM is perfectly correlated with volume inside TM, so the shares  $s_i^{\text{MA}}$  and  $s_i^{\text{NM}}$  are constant.

Medicare (Berenson et al., 2015; Gupta, Navathe, and Schwartz, 2022). Given this similarity, we apportion Medicare’s production between TM and MA based on relative enrollment and obtain  $\bar{s}^{\text{MA}} = 0.111$ . Next we consider Non-Medicare (NM) production. Private insurance spent \$226 billion, which we deflate by a factor of 1.43 to account for the higher prices private insurance pays to make quantities comparable to Medicare (Lopez and Jacobson, 2020). Medicaid spent roughly \$41 billion, which we deflate by its relative price of 0.72 (Zuckerman, Skopec, and Aarons, 2021). We incorporate other residual categories of production without price adjustments.<sup>55</sup> Combining these, we obtain an average  $\bar{s}^{\text{NM}} = 0.676$ .

To estimate  $\varrho^{\text{MA}}$  and  $\varrho^{\text{NM}}$ , we assume that relative production is proportionate to relative resident beneficiaries. We obtain the number of TM beneficiaries and number of MA beneficiaries by HRR from Medicare enrollment data and compute the number of NM patients as total population minus Medicare enrollees.<sup>56</sup> Regressing the respective beneficiary ratios on log TM production yields  $\hat{\varrho}^{\text{MA}} = 0.073$  and  $\hat{\varrho}^{\text{NM}} = 0.069$ . Putting these together means  $\hat{\zeta} = 1.055$ , so our estimated  $\alpha\zeta = 0.66$  from Table 5 implies a scale elasticity of  $\alpha = \frac{0.66}{1.055} = 0.63$ .

## C Details of counterfactual calculations

Section C.1 describes how we compute counterfactual equilibrium outcomes relative to baseline equilibrium outcomes in the model. Section C.2 describes the assumptions we make to infer the number of potential patients  $N_j$  and hence import shares  $m_{ij}$ , which are inputs into these calculations. Section C.3 describes how to compute counterfactual outcomes in the

---

<sup>55</sup>These other categories in the National Health Expenditure data are labeled Other Health Insurance Programs and Other Third Party Payers, along with out-of-pocket spending. Our simplifying approach here amounts to assuming Medicare prices for these residual categories.

<sup>56</sup>Ideally we would like to use the quantity of production in NM and MA markets, but we do not have this available at the HRR level. Beneficiaries might seem like a problematic proxy because the composition of NM beneficiaries varies widely across space, with some regions having a high Medicaid share and others a high private share. In aggregate, these two markets turn out to have similar per capita quantities of physician service spending: while private spending is \$1,118 per capita and Medicaid spending is \$550 per capita, the price adjustments mentioned above the quantities are relatively similar at \$782 and \$764, respectively, when valued at Medicare prices.

model when there are multiple (observed) types of patients who differ in their trade costs. Section C.4 describes how we infer the number of potential patients of each type.

## C.1 Computing equilibrium outcomes in counterfactual scenarios

We compute counterfactual equilibrium outcomes relative to baseline equilibrium outcomes by rewriting the equilibrium system of equations in terms of the initial allocation, constant elasticities, relative exogenous parameters, and relative endogenous equilibrium outcomes, a technique known as “exact hat algebra” in the trade literature.

If  $K(\delta) = \delta$  and  $H(Q) = Q^\alpha$ , an equilibrium is a set of quantities and qualities  $\{Q_i, \delta_i\}_{i \in \mathcal{I}}$  that simultaneously satisfy equations (4) and (1) and  $Q_i = \sum_j Q_{ij}$ . Consider two equilibria: the baseline equilibrium and the counterfactual equilibrium. Define export shares  $x_{ij} \equiv \frac{Q_{ij}}{\sum_{j'} Q_{ij'}}$  and import shares  $m_{ij} \equiv \frac{Q_{ij}}{N_j}$  in the baseline equilibrium. Denote the counterfactual parameters and equilibrium outcomes by primes. Plugging  $Q_i = \sum_j Q_{ij}$  into equation (4), we can write the system of equations for each equilibrium as

$$\begin{aligned} \delta'_i &= \left( \frac{R'_i \hat{A}'_i}{w'_i} \right) \left( \sum_j Q'_{ij} \right)^\alpha & Q'_{ij} &= \delta'_i \frac{\rho'_{ij}}{\sum_{i' \in 0 \cup \mathcal{I}} \delta'_{i'} \rho'_{i'j}} N'_j \\ \delta_i &= \left( \frac{R_i \hat{A}_i}{w_i} \right) \left( \sum_j Q_{ij} \right)^\alpha & Q_{ij} &= \delta_i \frac{\rho_{ij}}{\sum_{i' \in 0 \cup \mathcal{I}} \delta_{i'} \rho_{i'j}} N_j \end{aligned}$$

Define  $\hat{y} \equiv \frac{y'}{y}$  for every variable  $y$ . For example,  $\hat{\delta}_i \equiv \frac{\delta'_i}{\delta_i}$ .

We now rewrite the counterfactual equilibrium equations in terms of baseline equilibrium shares  $x_{ij}, m_{ij}$ , the scale elasticity  $\alpha$ , (relative) counterfactual exogenous parameters  $\hat{A}, \hat{R}, \hat{w}, \hat{\rho}, \hat{N}$ , and (relative) counterfactual endogenous qualities  $\hat{\delta}$ .

First, divide the counterfactual free-entry condition by the baseline free-entry condition to obtain an expression for relative quality:

$$\frac{\delta'_i}{\delta_i} = \frac{\hat{R}_i \hat{A}_i}{\hat{w}_i} \left( \frac{\sum_{j \in \mathcal{I}} Q'_{ij}}{\sum_{j \in \mathcal{I}} Q_{ij}} \right)^\alpha = \frac{\hat{R}_i \hat{A}_i}{\hat{w}_i} \left( \sum_{j \in \mathcal{I}} \frac{Q_{ij}}{\sum_{j \in \mathcal{I}} Q_{ij}} \frac{Q'_{ij}}{Q_{ij}} \right)^\alpha = \frac{\hat{R}_i \hat{A}_i}{\hat{w}_i} \left( \sum_{j \in \mathcal{I}} x_{ij} \frac{Q'_{ij}}{Q_{ij}} \right)^\alpha \quad (\text{C.1})$$

Second, divide the counterfactual gravity equation by the baseline gravity equation to obtain



an expression for relative bilateral flows:

$$\begin{aligned} \frac{Q'_{ij}}{Q_{ij}} &= \frac{\delta'_i}{\delta_i} \left( \frac{\frac{\rho'_{ij}}{\sum_{i' \in 0 \cup \mathcal{I}} \delta_{i'} \rho'_{i'j}} N'_j}{\frac{\rho_{ij}}{\sum_{i' \in 0 \cup \mathcal{I}} \delta_{i'} \rho_{i'j}} N_j} \right) = \frac{\frac{\delta'_i \rho'_{ij} N'_j}{\delta_i \rho_{ij} N_j}}{\sum_{i' \in 0 \cup \mathcal{I}} \frac{\delta_{i'} \rho_{i'j}}{\sum_{i' \in 0 \cup \mathcal{I}} \delta_{i'} \rho_{i'j}} \frac{\delta'_{i'} \rho'_{i'j}}{\delta_{i'} \rho_{i'j}}} \\ &= \frac{\hat{\delta}_i \hat{\rho}_{ij} \hat{N}_j}{\sum_{i' \in 0 \cup \mathcal{I}} \frac{Q_{i'j}}{N_j} \hat{\delta}_{i'} \hat{\rho}_{i'j}} = \frac{\hat{\delta}_i \hat{\rho}_{ij} \hat{N}_j}{m_{0j} + \sum_{i' \in \mathcal{I}} m_{i'j} \hat{\delta}_{i'} \hat{\rho}_{i'j}} \end{aligned}$$

Plug this expression for relative bilateral flows into equation (C.1) and rearrange terms to obtain the following system of  $\mathcal{I}$  equations with unknowns  $\{\hat{\delta}_i\}_{i=1}^{\mathcal{I}}$ :

$$\hat{\delta}_i = \left( \hat{R}_i \hat{A}_i / \hat{w}_i \right)^{\frac{1}{1-\alpha}} \left( \sum_{j \in \mathcal{I}} \frac{x_{ij} \hat{\rho}_{ij} \hat{N}_j}{m_{0j} + \sum_{i' \in \mathcal{I}} m_{i'j} \hat{\delta}_{i'} \hat{\rho}_{i'j}} \right)^{\frac{\alpha}{1-\alpha}}. \quad (\text{C.2})$$

## C.2 Inferring the number of potential patients

A baseline calibration of our model requires  $\alpha$ ,  $x_{ij}$ , and  $m_{ij}$  in order to use equation (C.2) to compute relative counterfactual outcomes. We have estimated  $\alpha$ . The export shares  $x_{ij} \equiv \frac{Q_{ij}}{\sum_j Q_{ij}}$  are easily computed using the observed trade matrix.<sup>57</sup> The challenge is computing import shares  $m_{ij} \equiv \frac{Q_{ij}}{N_j}$  because we do not observe  $N_j$ ; while we observe the number of Medicare beneficiaries in region  $j$ , not all beneficiaries are in the market for all services. This section describes the assumptions we make in order to infer the values of the relevant market size  $N_j \forall j \in \mathcal{I}$ . Specifically, we assume per capita demand is uniform, outside-option quality is constant across regions, and the average outside-option share is 10%, as described below.

We have estimated  $\theta_j = N_j / \Phi_j$  in equation (12). We observe the number of beneficiaries enrolled in Traditional Medicare in region  $j$ , which we denote  $S_j^{\text{TM}}$ . By definition,  $m_{0j} = \frac{\delta_{0j}}{\Phi_j}$ .

---

<sup>57</sup>Dingel and Tintelnot (2021) document overfitting problems when calibrating gravity models using noisy observed shares. We obtain similar counterfactual outcomes when calibrating our model using gravity-predicted shares.

We assume  $\delta_{0j} = \delta_0 \forall j$  and  $N_j \propto S_j^{\text{TM}}$ . This implies

$$m_{0j} = \frac{\delta_{0j}}{\Phi_j} = \frac{\delta_0 \theta_j}{N_j} = \frac{\delta_0 \theta_j}{\mathfrak{s} S_j^{\text{TM}}},$$

where  $\mathfrak{s}$  is a constant of proportionality. We set  $\frac{\delta_0}{\mathfrak{s}}$  such that the average outside-option share is 10%,  $\frac{1}{\mathcal{I}} \sum_j m_{0j} = 0.1$ . This requires  $\frac{\delta_0}{\mathfrak{s}} = \frac{0.1 \times \mathcal{I}}{\sum_j \theta_j / S_j^{\text{TM}}}$ . With  $m_{0j}$  in hand, we can infer  $N_j$ :

$$m_{0j} = 1 - \sum_{i \in \mathcal{I}} m_{ij} = 1 - \frac{1}{N_j} \sum_{i \in \mathcal{I}} Q_{ij} \implies N_j = \frac{1}{1 - m_{0j}} \sum_{i \in \mathcal{I}} Q_{ij}.$$

With  $N_j$  in hand, we can compute all import shares,  $m_{ij} = \frac{Q_{ij}}{N_j} \forall i \in 0 \cup \mathcal{I}, \forall j \in \mathcal{I}$ .

We exclude the Anchorage, Alaska HRR from our counterfactual computations. The entire state of Alaska is one (geographically isolated and very large) HRR. The average within-Alaska-HRR procedure incurs more than 60 kilometers of travel. In the gravity regression, Alaska has the smallest exporter fixed effect: very few patients travel to Alaska for care. Alaska's importer fixed effect is quite large because Alaskans import about 15% of their services and the average import traverses more than 3,600 kilometers. As a result, the implied outside-option share would exceed one when we set the nationwide average to 10%. We therefore exclude the Alaska HRR from the economy when computing counterfactual outcomes. Given its considerable geographic isolation, Alaska would have little influence on outcomes in other regions.

The qualitative and spatial patterns of counterfactual outcomes are the same if we assume the average outside-option share is 20% rather than 10%.

### C.3 Counterfactual outcomes with multiple patient types

This section describes how to compute counterfactual equilibrium outcomes relative to baseline equilibrium outcomes when there are multiple patient types who face heterogeneous trade costs. The derivation is very similar to that of Section C.1. Define import shares  $m_{ij\kappa} \equiv \frac{Q_{ij\kappa}}{N_{j\kappa}}$  in the baseline equilibrium. Define patient-type shares  $n_{j\kappa} \equiv \frac{N_{j\kappa}}{N_j}$ . We rewrite

the system of baseline and counterfactual gravity equations (A.1) and free-entry condition (4) as follows:

$$\begin{aligned}\delta'_i &= \left(\frac{R'_i A'_i}{w'_i}\right) \left(\sum_j Q'_{ij}\right)^\alpha & Q'_{ij} &= \delta'_i \sum_\kappa \frac{\rho'_{ij\kappa}}{\sum_{i' \in 0 \cup \mathcal{I}} \delta'_{i'} \rho'_{i'j\kappa}} N'_{j\kappa} \\ \delta_i &= \left(\frac{R_i A_i}{w_i}\right) \left(\sum_j Q_{ij}\right)^\alpha & Q_{ij} &= \delta_i \sum_\kappa \frac{\rho_{ij\kappa}}{\sum_{i' \in 0 \cup \mathcal{I}} \delta_{i'} \rho_{i'j\kappa}} N_{j\kappa}.\end{aligned}$$

As above, dividing the counterfactual free-entry condition by the baseline free-entry condition yields the expression for relative quality in equation (C.1). Second, divide the counterfactual gravity equation by the baseline gravity equation to obtain an expression for relative bilateral flows:

$$\frac{Q'_{ij}}{Q_{ij}} = \frac{\delta'_i}{\delta_i} \left( \frac{\sum_\kappa \frac{\rho'_{ij\kappa}}{\sum_{i' \in 0 \cup \mathcal{I}} \delta'_{i'} \rho'_{i'j\kappa}} N'_{j\kappa}}{\sum_\kappa \frac{\rho_{ij\kappa}}{\sum_{i' \in 0 \cup \mathcal{I}} \delta_{i'} \rho_{i'j\kappa}} N_{j\kappa}} \right) = \frac{\delta'_i}{\delta_i} \left( \frac{\sum_\kappa \frac{\rho'_{ij\kappa}}{\Phi'_{j\kappa}} N'_{j\kappa}}{\sum_\kappa \frac{\rho_{ij\kappa}}{\Phi_{j\kappa}} N_{j\kappa}} \right) = \hat{\delta}_i \sum_\kappa \frac{n_{j\kappa}}{m_{ij\kappa}} \frac{\hat{\rho}_{ij\kappa}}{\hat{\Phi}_{j\kappa}} \hat{N}_{j\kappa}$$

Plugging this expression for relative bilateral flows into equation (C.1) and then rearranging terms yields the following system of  $\mathcal{I}$  equations with unknowns  $\{\hat{\delta}_i\}_{i=1}^{\mathcal{I}}$ :

$$\hat{\delta}_i = \left(\hat{R}_i \hat{A}_i / \hat{w}_i\right)^{\frac{1}{1-\alpha}} \left( \sum_{j \in \mathcal{I}} x_{ij} \left( \sum_\kappa \frac{n_{j\kappa}}{m_{ij\kappa}} \frac{m_{ij\kappa} \hat{\rho}_{ij\kappa}}{m_{0j\kappa} + \sum_{i' \in \mathcal{I}} m_{i'j\kappa} \hat{\delta}_{i'} \hat{\rho}_{i'j\kappa}} \hat{N}_{j\kappa} \right) \right)^{\frac{\alpha}{1-\alpha}}. \quad (\text{C.3})$$

## C.4 Inferring the number of potential patients of each type

Because we do not observe patients who select the outside option, we make assumptions that allow us to infer  $N_{j\kappa}$  and thus  $n_{j\kappa}$  and  $m_{ij\kappa}$ , which are needed to compute counterfactual outcomes using equation (C.3). We start from a type-specific variant of the gravity equation (7) with fixed effects, as in the single-type equation (12). The estimating equation is

$$\ln \mathbb{E}(\bar{R} Q_{ij\kappa}) = \ln \delta_i + \ln \left( \frac{N_{j\kappa}}{\Phi_{j\kappa}} \right) + \gamma^\kappa X_{ij} = \ln \delta_i + \ln \theta_{j\kappa} + \gamma^\kappa X_{ij}.$$

This yields an estimate of  $\theta_{j\kappa} = N_{j\kappa} / \Phi_{j\kappa}$ .

As in the single-type case above, we assume per capita demand is uniform and outside-

option quality is constant across regions. We observe the number of beneficiaries of type  $\kappa$  enrolled in Traditional Medicare in region  $j$ , which we denote  $S_{j\kappa}^{\text{TM}}$ . We assume  $\delta_{0j} = \delta_0 \forall j$  and  $N_{j\kappa} = \mathfrak{s}S_{j\kappa}^{\text{TM}}$ , where  $\mathfrak{s}$  is a constant of proportionality that is common across types. This implies

$$m_{0j\kappa} = \frac{\delta_0}{\Phi_{j\kappa}} = \frac{\delta_0 \theta_{j\kappa}}{N_{j\kappa}} = \frac{\delta_0 \theta_{j\kappa}}{\mathfrak{s}S_{j\kappa}^{\text{TM}}}.$$

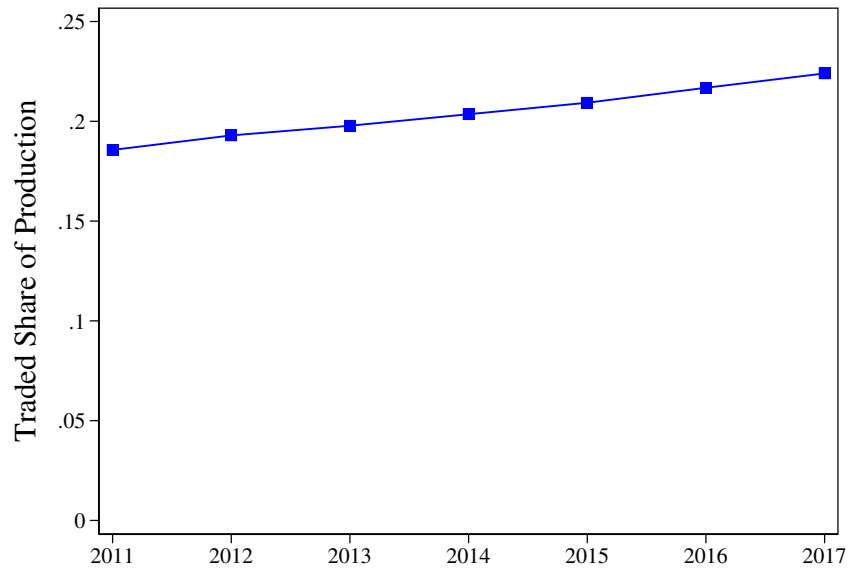
Let  $\mathcal{K} = \sum_{\kappa} 1$  denote the number of patient types. We set  $\frac{\delta_0}{\mathfrak{s}}$  such that the average outside-option share, across all types, is 10%,  $\frac{1}{\mathcal{IK}} \sum_{j\kappa} m_{0j\kappa} = 0.1$ . This implies

$$m_{0j\kappa} = 0.1 \times \frac{\theta_{j\kappa}/S_{j\kappa}^{\text{TM}}}{\frac{1}{\mathcal{IK}} \sum_{j'\kappa'} \theta_{j'\kappa'}/S_{j'\kappa'}^{\text{TM}}}.$$

Using the resulting  $N_{j\kappa} = \frac{1}{1-m_{0j\kappa}} \sum_{i \in \mathcal{I}} Q_{ij\kappa}$  allows us to compute all import shares.

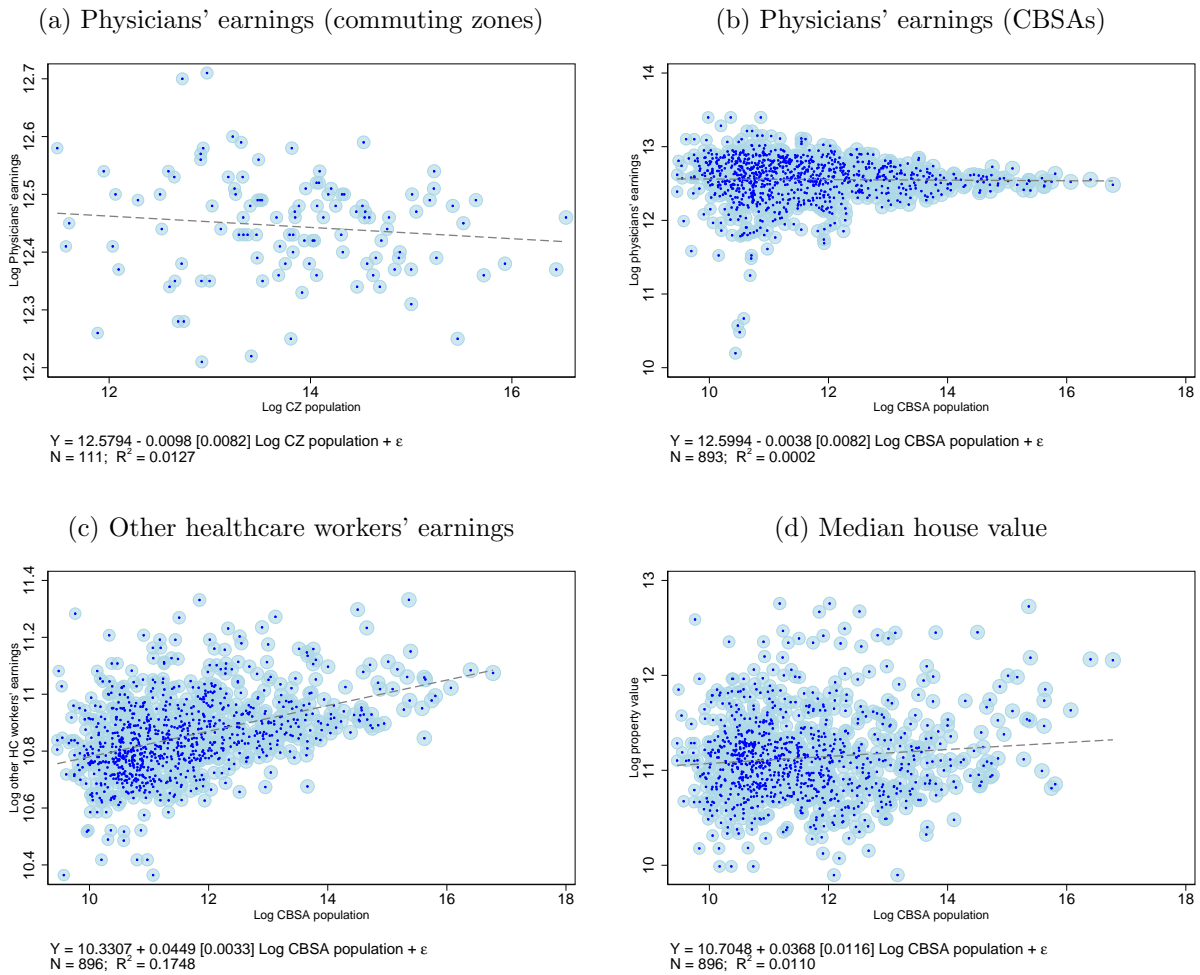
## D Additional exhibits

Figure D.1: Trade in medical services has increased over time



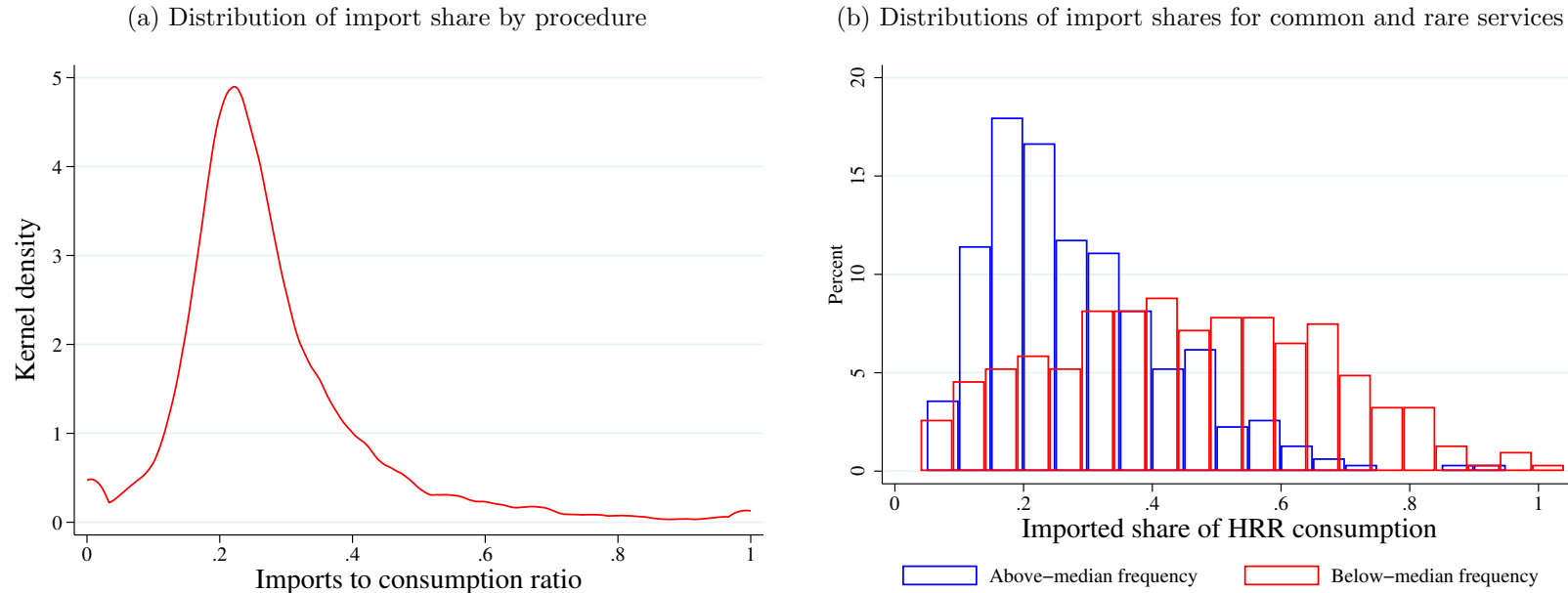
*Note:* This figure shows the annual exported share of production from 2011 to 2017. Production and trade are computed using the Medicare 20% carrier Research Identifiable Files for the relevant years. Production is exported when the patient's address and the service location are in different HRRs. HRR definitions are from the Dartmouth Atlas Project.

Figure D.2: Population elasticities of input costs



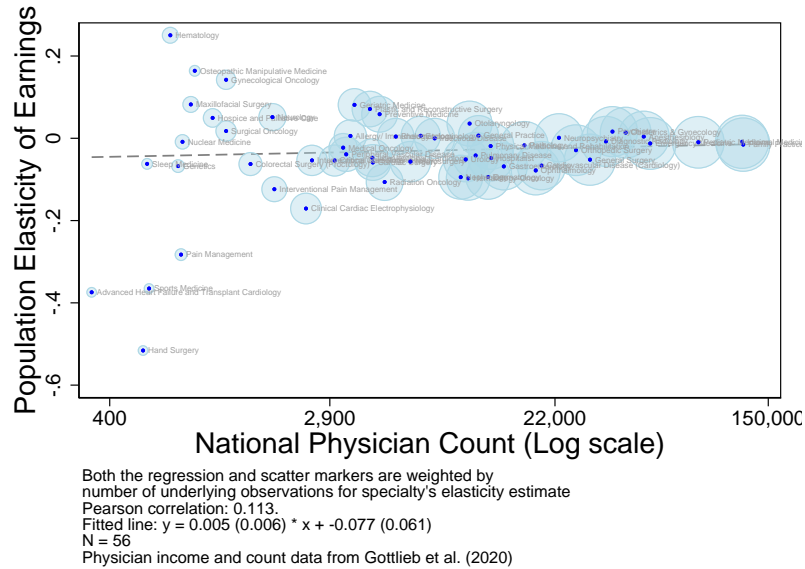
*Notes:* This figure depicts relationships between input costs and population sizes. Panel (a) shows physicians' earnings across 111 commuting zones using data from Gottlieb et al. (2020). Panels (b), (c), and (d) show variation across CBSAs in physicians' earnings, other healthcare workers' earnings, and median house values (a proxy for real estate and other locally priced inputs) using data from the 2015–2019 American Community Survey.

Figure D.3: Variation in trade shares across procedures and regions



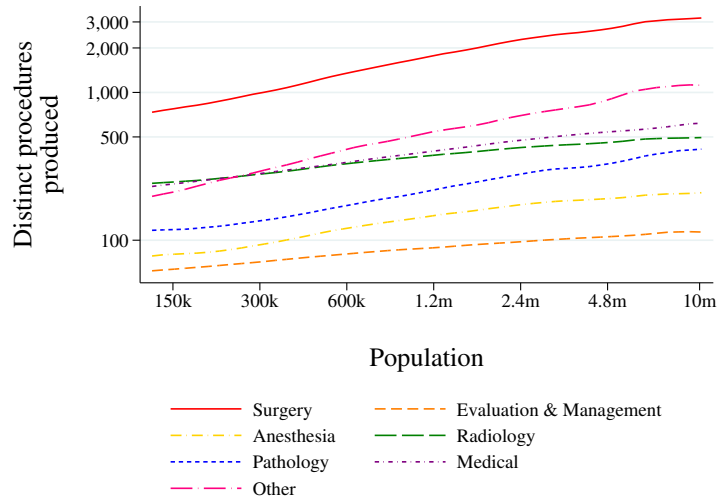
*Notes:* Panel (a) shows the distribution of the imported consumption share across procedures for procedures performed at least 20 times (in our 20% sample of Medicare claims). Imports are defined as care provided to a patient who lives in one HRR at a service location in a different HRR. Panel (b) splits all services into two groups based on how often they are performed nationally. Those performed less often than the median are shown in red, and those performed more often than the median service are shown in blue. Import shares are substantially higher for the rarer services.

Figure D.4: Specialists' income patterns do not explain the output-population gradient



Notes: This figure shows the population elasticity of income for different medical specialties against the total number of physicians in those specialties. For each specialty, we estimate the elasticity of income with respect to population across commuting zones, using data from Gottlieb et al. (2020). The graph shows that these elasticities are unrelated to the total national count of physicians in those specialties.

Figure D.5: Larger markets produce a greater variety of procedures

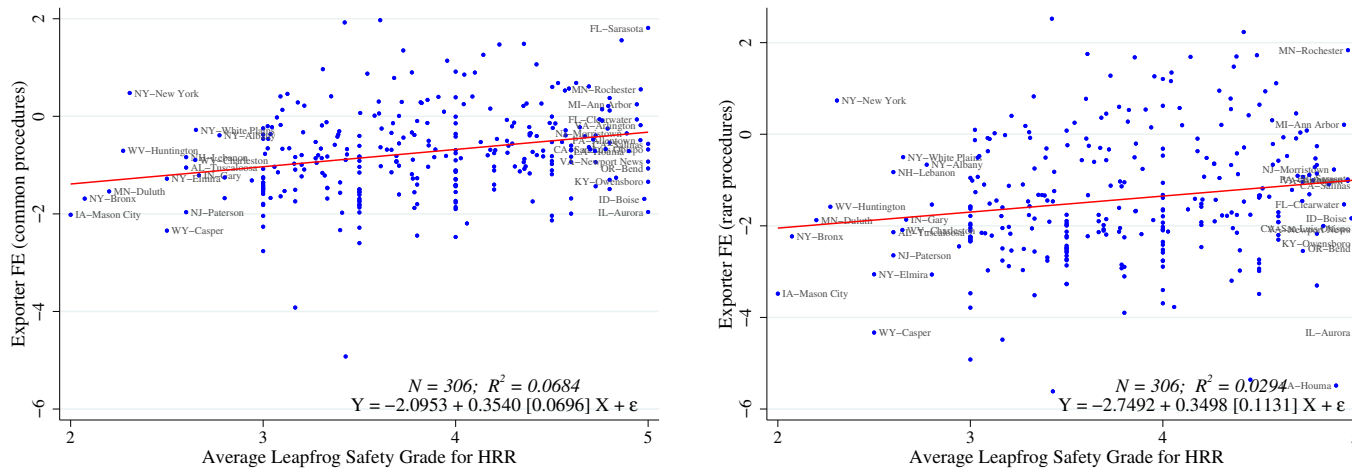


Notes: This figure shows the local relationship between the number of distinct services performed in the Medicare data in a given HRR and that HRR's population. More populous HRRs perform more unique services; Table D.14 reports the population elasticities. We use procedure classifications from the American Academy of Professional Coders, which groups codes into surgeries, anesthesia, radiology, pathology, medical, and evaluation & management services (AAPC, 2021). We combine Category II codes, Category III codes and Multianalyte Assays into "other."



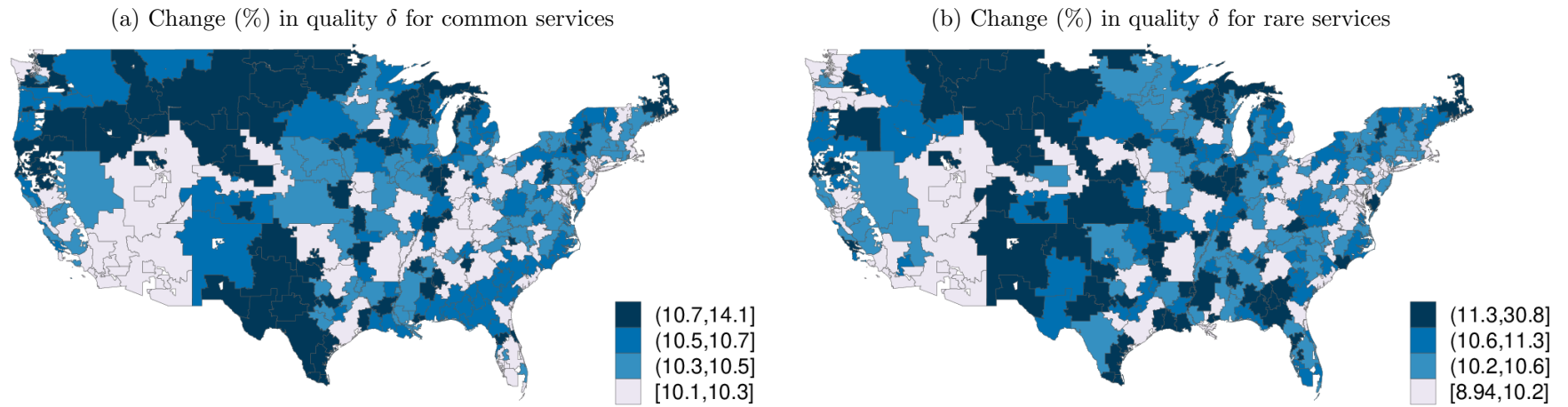
Figure D.6: Leapfrog Safety Grade vs. estimated quality: common and rare

(a) Leapfrog Safety Grade vs. quality for common services      (b) Leapfrog Safety Grade vs. quality for rare services



*Notes:* This figure shows the relationship between exporter fixed effects, estimated separately for common and rare services, and the Leapfrog Safety Grade. The vertical axis shows the exporter fixed effects for each HRR estimated from equation (12), in Panel (a) using trade in common services, and in Panel (b) using trade in rare services. The horizontal axis in both panels is the average safety grade for hospitals in an HRR, determined by the Leapfrog Group. The Leapfrog Safety Grades range from A to F, which we scale as integers from 1 (for F) to 5 (for A). We then compute the mean score for all hospitals in the HRR. The Safety Grades are positively associated with the exporter fixed effects for both rare and common procedures.

Figure D.7: Counterfactual change in quality  $\delta$  for rare vs. common services when increasing reimbursement by 10% everywhere

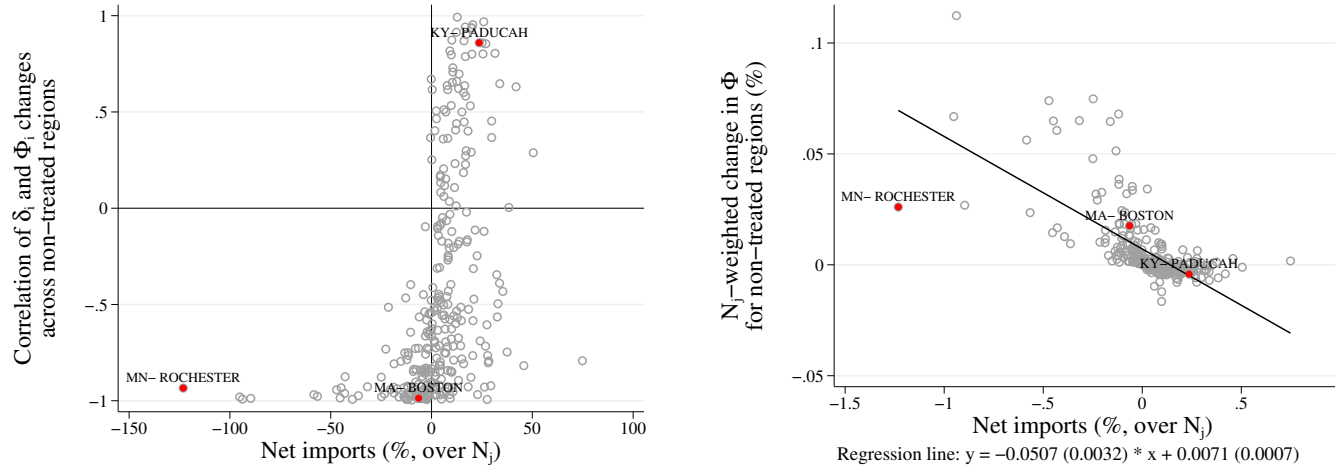


Cutoffs: Percentiles 25, 50, 75.

Cutoffs: Percentiles 25, 50, 75.

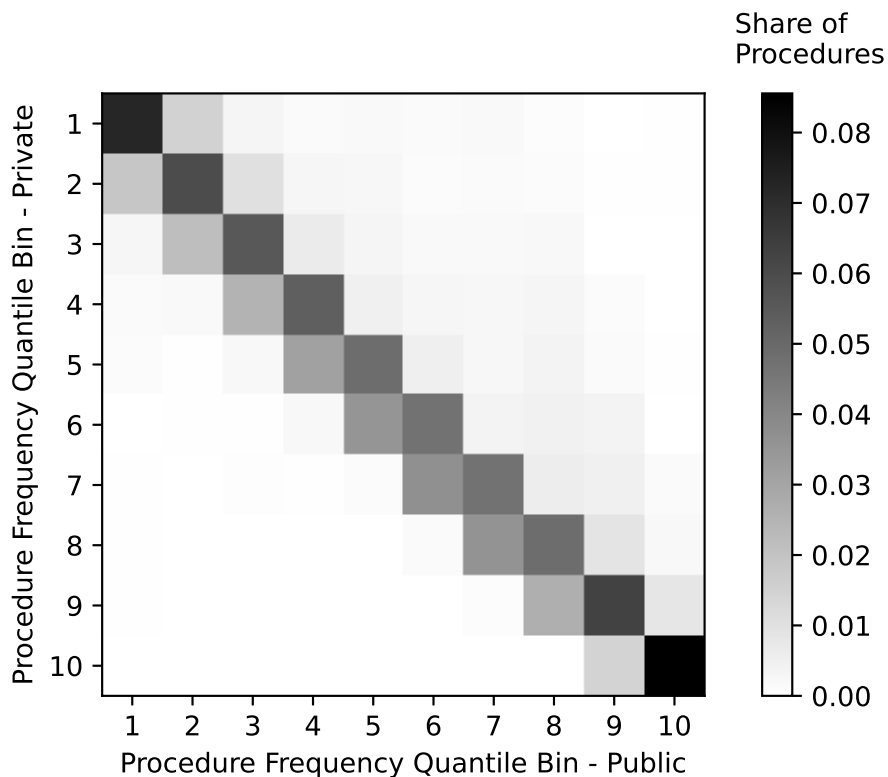
*Notes:* Both panels show the impacts of increasing reimbursements by 10% everywhere ( $\hat{R}_i = 1.1$  for all  $i$ ) on the quality of production in each region,  $\delta_i$ . Panel (a) illustrates the change for common services, and Panel (b) for rare services. Each panel is based on the baseline trade matrix for the respective set of services. Panel (a) uses an agglomeration elasticity of  $\alpha = 0.6$  and Panel (b) uses  $\alpha = 0.9$ . The common-services scenario excludes the Alaska HRR and the rare-services scenario excludes four HRRs (Alaska, Hawaii, Houma, La., and Minot, N.D.). The pattern of outcomes is qualitatively similar but the magnitudes vary more for rare services.

Figure D.8: Spillovers from higher reimbursements in one region depend on that region's net imports

(a) Correlation of  $\hat{\delta}_i$  and  $\hat{\Phi}_i$  across non-treated regions      (b) Change in non-treated regions' aggregate market access

*Notes:* This figure characterizes counterfactual outcomes when raising reimbursements by 30 percent in one HRR. We conduct this exercise for each region, one at a time, and each observation in each panel represents one such counterfactual scenario. Panel (a) illustrates the contrast in spillovers as a function of net imports of the treated region. The vertical-axis value for each observation reports the correlation—across *all regions other than the treated one* for the exercise in question—between the counterfactual changes  $\hat{\delta}_i$  and  $\hat{\Phi}_i$ . The scatterplot relates these correlations to the *treated* region's net import share, which is plotted on the horizontal axis. When the treated region is a net exporter, changes in quality  $\delta_i$  and in market access  $\Phi_i$  for non-treated regions move in opposite directions: a region whose output quality declines experiences an increase in market access through imports from the treated region. However, increasing reimbursements in a net-importing region often has the opposite effect: neighboring regions with quality reductions also experience lower market access, (changes in  $\delta_i$  and  $\Phi_i$  are positively correlated). For each counterfactual, the vertical-axis value in Panel (b) shows the aggregate impact on patient market access *excluding the treated region*. The panel relates this impact to the *treated* region's net imports, shown on the horizontal axis. When the treated region is a net importer, the aggregate impact on market access for non-treated regions tends to be smaller or even negative.

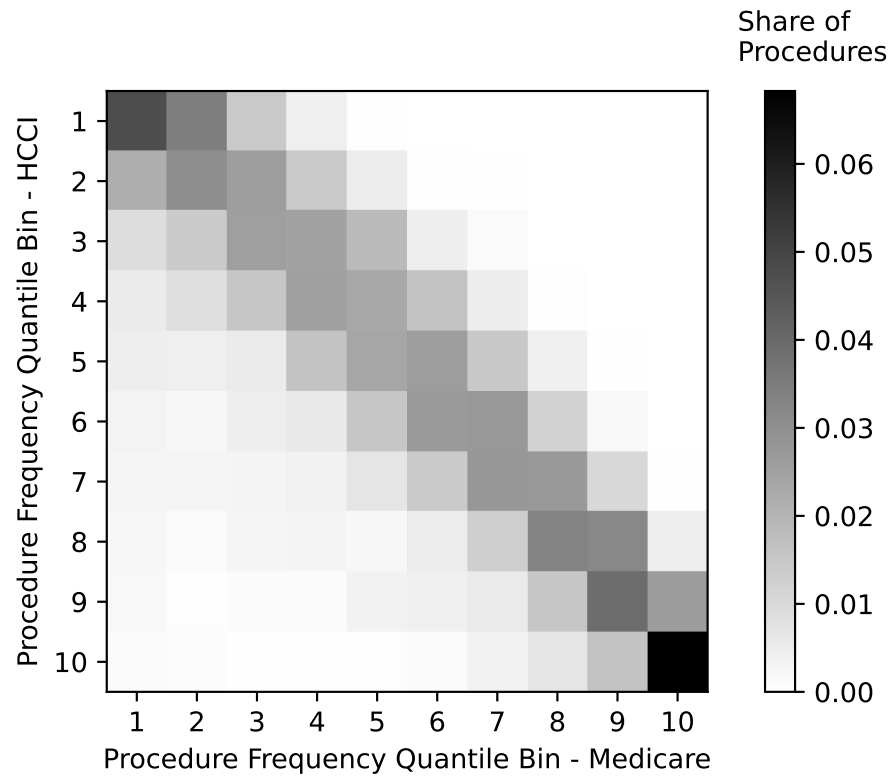
Figure D.9: Deciles of Procedure Frequency in Confidential and Public Medicare Data



The simple correlation between quantile bins is 0.9068

*Notes:* This figure shows the share of procedures in each frequency decile in the Medicare public data compared to the Medicare confidential data. The classification of procedures by frequency deciles appears largely consistent between the two data sources for Medicare patients.

Figure D.10: Deciles of Procedure Frequency in Medicare and Private Insurance Data



The simple correlation between quantile bins is 0.8287

*Notes:* This figure shows the share of procedures in each frequency decile in the Medicare versus privately insured data. The classification of procedures by frequency deciles appears largely consistent when comparing public Medicare data with data on privately insured patients from the Health Care Cost Institute (HCCI).

Table D.1: Higher-income patients are less sensitive to distance: Procedure-level estimates

	(1)	(2)	(3)	(4)	(5)
	25min visit	cataract removal	knee joint repair	heart artery bypass	gallblader removal
Distance (log)	-2.082 (0.0755)	-2.283 (0.0823)	-2.262 (0.0908)	-2.247 (0.0875)	-2.141 (0.0821)
Distance (log) $\times$ income tercile 2	0.0939 (0.0604)	0.143 (0.0822)	0.172 (0.0741)	0.0980 (0.0823)	0.205 (0.0686)
Distance (log) $\times$ income tercile 3	0.211 (0.0756)	0.290 (0.0909)	0.233 (0.0915)	0.403 (0.0927)	0.319 (0.0891)
Observations	271,728	268,400	262,352	240,352	250,800
Patient market-income FE & Provider market FE	Yes	Yes	Yes	Yes	Yes

*Notes:* This table reports the coefficient on log distance for each income tercile from gravity regressions estimated separately for five procedures varying in frequency: 25 min office visit (HCPCS 99214), cataract removal (66984), knee joint repair (27447), heart artery bypass (33533), and gallblader removal (47562). The dependent variable in all regressions is the number of procedures traded. Each regression includes log distance interacted with an income tercile indicator, an indicator for same-HRR observations ( $i = j$ ), an exporting HRR fixed effect, and an income-tercile-importing-HRR fixed effect. The coefficients for higher income terciles are positive, indicating that patients residing in higher-income ZIP codes are less sensitive to distance. Trade flows are computed from the Medicare 20% carrier Research Identifiable Files. HRR definitions are from the Dartmouth Atlas Project. Standard errors (in parentheses) are two-way clustered by patient market and provider market.

Table D.2: Estimates of a strong home-market effect by CBSA

	(1)	(2)	(3)	(4)	(5)	(6)
Estimation method:	PPML	PPML	PPML	IV	PPML	IV
Instrument:				1940 pop		Bedrock
Provider-market population (log)	0.734 (0.0232)	0.739 (0.0234)	0.703 (0.0205)	0.716 (0.0249)	0.739 (0.0259)	1.161 (0.307)
Patient-market population (log)	0.395 (0.0290)	0.393 (0.0292)	0.417 (0.0264)	0.396 (0.0261)	0.394 (0.0311)	0.178 (0.373)
Distance (log)	-2.311 (0.0493)	-3.464 (0.324)		-3.403 (0.295)	-3.400 (0.347)	-4.677 (1.056)
Distance (log, squared)		0.110 (0.0323)		0.104 (0.0288)	0.105 (0.0346)	0.210 (0.0850)
Observations	857,476	857,476	857,476	857,476	781,456	781,456
Sample:	All CBSAs	All CBSAs	All CBSAs	All CBSAs	Bedrock data	Bedrock data
Distance elasticity at mean		-1.90		-1.92	-1.90	-1.68
Distance deciles			Yes			

*Notes:* This table reports estimates of equation (8), which estimates the presence of weak or strong home-market effects. The dependent variable in all regressions is the value of trade computed by assigning each procedure its national average price. The independent variables are patient- and provider-market log population, log distance between CBSAs, and an indicator for same-CBSA observations ( $i = j$ ). The positive coefficient on provider-market log population implies a weak home-market effect, and the fact that this coefficient exceeds that on patient-market population implies a strong home-market effect. Column 2 makes the distance coefficient more flexible by adding a control for the square of log distance. Column 3 replaces parametric distance specifications with fixed effects for each decile of the distance distribution. Column 4 uses the provider-market and patient-market log populations in 1940 as instruments for the contemporaneous log populations when estimating by generalized method of moments. Column 5 reports the PPML estimate on the subsample of regions for which we have data on depth to bedrock available ( $N = 884^2$ ). Column 6 uses depth to bedrock in the importing and exporting regions as instruments for current log population in those regions, respectively. Trade flows are computed from the Medicare 20% carrier Research Identifiable Files. HRR definitions are from the Dartmouth Atlas Project. Standard errors (in parentheses) are two-way clustered by patient market and provider market.

Table D.3: Estimates of a strong home-market effect including facility spending

Estimation method:	(1) PPML	(2) PPML	(3) PPML	(4) IV
Provider-market population (log)	0.692 (0.0573)	0.700 (0.0526)	0.690 (0.0383)	0.828 (0.0585)
Patient-market population (log)	0.222 (0.0564)	0.218 (0.0508)	0.255 (0.0345)	0.266 (0.0470)
Distance (log)	-1.645 (0.0489)	0.865 (0.327)		0.921 (0.262)
Distance (log, squared)		-0.253 (0.0330)		-0.257 (0.0259)
Same hrr	0.403 (0.176)	1.781 (0.246)	4.685 (0.0637)	
Observations	93,636	93,636	93,636	93,636
Distance elasticity at mean		-2.76		-2.76
Distance deciles			Yes	

*Notes:* This table reports estimates of equation (8), which estimates the presence of weak or strong home-market effects, when including professional and facility fees. The sample is all HRR pairs ( $N = 306^2$ ). The dependent variable in all regressions is the value of trade when including professional and facility (inpatient and outpatient) fees at national average prices. The independent variables are patient- and provider-market log population, log distance between HRRs, and an indicator for same-HRR observations ( $i = j$ ). The positive coefficient on provider-market log population implies a weak home-market effect, and the fact that this coefficient exceeds that on patient-market population implies a strong home-market effect. Column 2 makes the distance coefficient more flexible by adding a control for the square of log distance. Column 3 replaces parametric distance specifications with fixed effects for each decile of the distance distribution. Column 4 uses the provider-market and patient-market log populations in 1940 as instruments for the contemporaneous log populations when estimating by generalized method of moments. Trade flows are computed from the Medicare 20% carrier, MedPAR, and outpatient claims Research Identifiable Files, excluding emergency-room care and skilled nursing facilities. HRR definitions are from the Dartmouth Atlas Project. Standard errors (in parentheses) are two-way clustered by patient market and provider market.



Table D.4: Estimates of a strong home-market effect excluding AZ, FL, CA

Estimation method:	(1)	(2)	(3)	(4)
	PPML	PPML	PPML	IV
Provider-market population (log)	0.647 (0.0812)	0.647 (0.0701)	0.649 (0.0425)	0.663 (0.0625)
Patient-market population (log)	0.376 (0.0810)	0.385 (0.0694)	0.415 (0.0421)	0.401 (0.0570)
Distance (log)	-1.748 (0.0608)	1.688 (0.427)		1.705 (0.397)
Distance (log, squared)		-0.360 (0.0429)		-0.361 (0.0396)
Observations	67,600	67,600	67,600	67,600
Distance elasticity at mean		-3.35		-3.35
Distance deciles			Yes	

*Notes:* This table reports estimates of equation (8), which estimates the presence of weak or strong home-market effects, excluding snowbird states. The sample is all HRR pairs, excluding those in Arizona, Florida, or California. The dependent variable in all regressions is the value of trade computed by assigning each procedure its national average price. The independent variables are patient- and provider-market log population, log distance between HRRs, and an indicator for same-HRR observations ( $i = j$ ). The positive coefficient on provider-market log population implies a weak home-market effect, and the fact that this coefficient exceeds that on patient-market population implies a strong home-market effect. Column 2 makes the distance coefficient more flexible by adding a control for the square of log distance. Column 3 replaces parametric distance specifications with fixed effects for each decile of the distance distribution. Column 4 uses the provider-market and patient-market log populations in 1940 as instruments for the contemporaneous log populations when estimating by generalized method of moments. Trade flows are computed from the Medicare 20% carrier Research Identifiable Files. HRR definitions are from the Dartmouth Atlas Project. Standard errors (in parentheses) are two-way clustered by patient market and provider market.

Table D.5: Estimates of a strong home-market effect excluding HRRs with high second-home share

Estimation method:	(1) PPML	(2) PPML	(3) PPML	(4) IV
Provider-market population (log)	0.657 (0.0666)	0.664 (0.0646)	0.662 (0.0455)	0.679 (0.0572)
Patient-market population (log)	0.366 (0.0650)	0.362 (0.0625)	0.393 (0.0426)	0.381 (0.0567)
Distance (log)	-1.686 (0.0511)	0.346 (0.327)		0.354 (0.298)
Distance (log, squared)		-0.208 (0.0319)		-0.209 (0.0290)
Observations	76,176	76,176	76,176	76,176
Distance elasticity at mean		-2.63		-2.63
Distance deciles			Yes	

*Notes:* This table reports estimates of equation (8), which estimates the presence of weak or strong home-market effects, excluding HRRs with a high second-home share. The sample is all HRR pairs excluding those in the top 10% based on the share of housing units that are vacant for seasonal/recreational purposes in the 2013–2017 American Community Survey. See Table D.4 notes on the variables, instruments, geographic units, and standard errors.

Table D.6: Travel for dialysis

Distance (km)	Share of output				
	All (Professional)	All (Facility)	All (Dialysis)	No snowbird states (Dialysis)	Snowbird states (Dialysis)
[0, 50)	0.77	0.77	0.94	0.94	0.93
[50, 100)	0.12	0.12	0.03	0.03	0.03
[100, .)	0.11	0.11	0.03	0.02	0.04

*Notes:* For the care described in each column and the distance intervals in each row, the entries in this table report the share of patients traveling that distance from their residential ZIP code to the service location's ZIP code. The first column shows professional claims (from Medicare's "carrier" file), the second column shows facility (hospital) claims, and the third column shows dialysis claims. The remaining columns split dialysis claims between "snowbird" states (AZ, CA, and FL, following Finkelstein, Gentzkow, and Williams 2016) and other states. In non-snowbird states, the table shows that 94% of patients travel less than 50 km from their home for dialysis, and only 2% more than 100 km. This is less than one-fifth as much as for other facility or professional care, suggesting that residential location is recorded correctly for almost all patients.

Table D.7: Contrasting geographies of colonoscopies and LVAD insertions

	Colonoscopy	LVAD Insertion
Code	G0121	33979
N	58,798	333
Physicians	13,475	177
$\hat{\beta}_p^{\text{production}}$	0.00	0.71
$\hat{\beta}_p^{\text{consumption}}$	-0.01	0.03
Share traded (HRR)	0.15	0.50
Share traded (CBSA)	0.15	0.48
Median distance traveled (km)	18.44	65.50
Share > 100km	0.06	0.37

*Notes:* This table reports statistics for two HCPCS codes: screening colonoscopy (G0121) and LVAD insertion (33979). We report the number of times the procedure is performed in 2017 in our 20% sample of Medicare patients and the number of distinct physicians performing it. The population elasticities of production and consumption are estimated using the Poisson models in equations (9) and (10) based on production HRR and patients' residential HRR, respectively. We also report the shares of procedures in which the patient and service location are in different HRRs or CBSAs, the median distance traveled for all care, and the share in which the patient and service location are more than 100 kilometers apart.

Table D.8: Estimates of a stronger home-market effect for rare diagnoses including facility spending

	(1)	(2)	(3)	(4)	(5)	(6)
Provider-market population (log)	0.669 (0.0570)	0.663 (0.0558)	0.653 (0.0555)		0.662 (0.0517)	
Patient-market population (log)	0.236 (0.0561)	0.237 (0.0546)	0.243 (0.0544)		0.239 (0.0500)	
Provider-market population (log) $\times$ rare			0.225 (0.0245)	0.209 (0.0206)	0.231 (0.0236)	0.207 (0.0204)
Patient-market population (log) $\times$ rare			-0.0819 (0.0216)	-0.0753 (0.0158)	-0.0837 (0.0211)	-0.0698 (0.0162)
Observations	187,272	147,814	147,814	147,814	147,814	147,814
Distance controls	Yes	Yes	Yes	Yes		
Distance [quadratic] controls					Yes	Yes
Patient-provider-market-pair FEs				Yes		Yes

*Notes:* This table reports estimates of equation (11), which introduces interactions with an indicator for whether a diagnosis is “rare” (provided to less patients than the median diagnosis, when adding up all diagnoses nationally). The dependent variable in all regressions is the value of trade when including professional and facility (inpatient and outpatient) fees at national average prices. The interactions with patient- and provider-market population reveal whether the home-market effect is larger for rare diagnoses. The unit of observation is {rare indicator, exporting HRR, importing HRR} so the number of observations is  $2 \times 306^2$  in column 1. All diagnoses are included. Columns 2 onwards drop HRR pairs with zero trade in both diagnosis groups, which leads to a larger sample than in Table 4 because trade in facility fees is included in addition to professional fees for all diagnoses. Column 2 shows that this restriction has a negligible impact on the estimated log population coefficients. Columns 1–4 control for distance using the log of distance between HRRs. Columns 5 and 6 add a control for the square of log distance. Columns 4 and 6 introduce a fixed effect for each  $ij$  pair of patient market and provider market, so these omit all covariates that are not interacted with the rare indicator. The positive coefficient on provider-market population  $\times$  rare across all columns indicates that the home-market effect is stronger for rare than for common services. The negative coefficient on patient-market population  $\times$  rare across all columns indicates that the *strong* home-market effect has a larger magnitude for rare services. Trade flows are computed from the Medicare 20% carrier, MedPAR, and outpatient claims Research Identifiable Files, excluding emergency-room care and skilled nursing facilities. HRR definitions are from the Dartmouth Atlas Project. Standard errors (in parentheses) are two-way clustered by patient market and provider market.

Table D.9: Home-market effect is stronger for rare services controlling for patient engagement

	(1)	(2)
Provider-market population (log) $\times$ common $\times$ high engagement	-0.0355 (0.0351)	-0.0355 (0.0348)
Provider-market population (log) $\times$ rare $\times$ low engagement	0.231 (0.0483)	0.244 (0.0370)
Provider-market population (log) $\times$ rare $\times$ high engagement	0.481 (0.0812)	0.358 (0.143)
Patient-market population (log) $\times$ common $\times$ high engagement	0.0439 (0.0258)	0.0450 (0.0257)
Patient-market population (log) $\times$ rare $\times$ low engagement	-0.145 (0.0376)	-0.124 (0.0244)
Patient-market population (log) $\times$ rare $\times$ high engagement	-0.477 (0.0926)	-0.573 (0.272)
Distance (log) $\times$ common $\times$ high engagement	-0.0550 (0.0210)	0.151 (0.121)
Distance (log) $\times$ rare $\times$ low engagement	0.0485 (0.0376)	0.726 (0.174)
Distance (log) $\times$ rare $\times$ high engagement	-0.130 (0.0821)	2.476 (2.784)
Distance (log, squared) $\times$ common $\times$ high engagement		-0.0197 (0.0103)
Distance (log, squared) $\times$ rare $\times$ low engagement		-0.0624 (0.0155)
Distance (log, squared) $\times$ rare $\times$ high engagement		-0.278 (0.326)
Observations	226,936	226,936
Distance controls	Linear	Quadratic
Patient-provider-market-pair FEs	Yes	Yes
Additional distance elasticity at mean for high engagement: common procedures	-0.05	-0.13
Additional distance elasticity at mean for high engagement: rare procedures	-0.18	-1.34

*Notes:* This table reports estimates of a variant of equation (11), which adds interactions with indicators for whether a procedure is “rare” (provided less often than the median procedure) and for whether a procedure is “high engagement” (median number of distinct claims per patient for the procedure in a given year is above one) or low engagement. The unit of observation is {rare indicator, high-engagement indicator, exporting HRR, importing HRR}, and the dependent variable is the value of trade. Each column includes fixed effects for each  $ij$  pair of patient market and provider market, rare versus common procedures, and high- versus low-engagement procedures, plus indicators for three categories (common  $\times$  high-engagement, rare  $\times$  low-engagement, and rare  $\times$  high-engagement) interacted with patient- and provider-market populations and distance covariates. Covariates for common  $\times$  low-engagement procedures are omitted, since they would lead to collinearity with the  $ij$  fixed effects. Column 2 adds a control for the square of log distance and its interactions. The negative coefficient on provider-market population and the positive coefficient on patient-market population for common and high-engagement procedures indicate that the home-market effect is slightly less *strong* compared to common and low-engagement procedures, even though these effects are not all statistically different from zero. The positive coefficient on provider-market population  $\times$  rare and the negative coefficient on patient-market population  $\times$  rare for both high- and low-engagement procedures indicates that the *strong* home-market effect is stronger for rare services, whether they are high- or low-engagement. The distance elasticity is more negative for high-engagement procedures (both rare and common). Trade flows are computed from the Medicare 20% carrier Research Identifiable Files. HRR definitions are from the Dartmouth Atlas Project. Standard errors (in parentheses) are two-way clustered by patient market and provider market.

Table D.10: Gravity regression by procedure: individual procedures exhibit a strong home-market effect

	(1)	(2)	(3)	(4)	(5)	(6)
Procedure:	Colonoscopy	Cataract surgery	Brain tumor	Brain radiosurgery	LVAD	Colon removal
HCPCS code:	G0121	66982	61510	61798	33979	44155
Provider-market population (log)	0.515 (0.0692)	0.466 (0.0730)	0.928 (0.0885)	1.149 (0.119)	1.251 (0.168)	0.998 (0.164)
Patient-market population (log)	0.351 (0.0694)	0.437 (0.0691)	0.192 (0.0726)	0.166 (0.0816)	0.182 (0.141)	-0.146 (0.146)
Distance (log)	0.436 (0.413)	0.948 (0.508)	0.997 (0.548)	1.518 (0.701)	2.168 (0.920)	3.090 (1.651)
Distance (log, squared)	-0.216 (0.0410)	-0.268 (0.0503)	-0.266 (0.0577)	-0.307 (0.0712)	-0.365 (0.0930)	-0.499 (0.173)
Observations	93,636	93,636	93,636	93,636	93,636	93,636
Distance elasticity at mean	-2.66	-2.89	-2.81	-2.89	-3.06	-4.06
Total count	58,798	43,604	1,922	752	333	112

*Notes:* This table reports estimates of equation (8) for procedure-level trade for six selected HCPCS codes, which vary in how common they are. For all procedures, the sample is all HRR pairs ( $N = 306^2$ ). The dependent variable in all regressions is the value of trade in the procedure (computed using each procedure's national average price). The independent variables are patient- and provider-market log population, log distance and square of log distance between HRRs, and an indicator for same-HRR observations ( $i = j$ ). The positive coefficient on provider-market log population implies a weak home-market effect, and the fact that this coefficient exceeds that on patient-market population implies a strong home-market effect. Trade flows are computed from the Medicare 20% carrier Research Identifiable Files. HRR definitions are from the Dartmouth Atlas Project. Standard errors (in parentheses) are two-way clustered by patient market and provider market. The bottom row reports the total national count of the procedure in our sample. Common procedures include screening colonoscopy (column 1) and cataract surgery (column 2). In a screening colonoscopy, the physician visualizes the large bowel with a camera to look for cancer. In a cataract surgery, the surgeon removes a cloudy lens from the eye to improve vision. Relatively rare procedures include brain radiosurgery (column 3), brain tumor removal (column 4), left ventricular assist device (LVAD) implantation (column 5) and colon removal (column 6). In brain radiosurgery, an area of the brain is irradiated, often to kill a tumor. In an LVAD implantation, a pump is implanted in the chest to assist a failing heart in pumping blood. Brain tumor and colon removals involve surgical removal of the respective structures.

Table D.11: Scale elasticity estimates for CBSAs

Panel A: All services	No Controls		Controls	
	No Diag	Diag	No Diag	Diag
OLS	1.052 (0.017)	0.888 (0.009)	1.054 (0.021)	0.904 (0.010)
2SLS: population (log)	1.023 (0.016)	0.845 (0.010)	1.006 (0.023)	0.849 (0.013)
2SLS: population (1940, log)	0.928 (0.025)	0.848 (0.014)	0.909 (0.033)	0.833 (0.017)
2SLS: bedrock depth	0.762 (0.099)	0.810 (0.038)	0.699 (0.109)	0.833 (0.040)
Panel B: Rare services				
OLS	1.108 (0.028)	0.941 (0.010)	1.071 (0.035)	0.936 (0.011)
2SLS: population (log)	1.106 (0.026)	0.914 (0.013)	1.043 (0.040)	0.893 (0.017)
2SLS: population (1940, log)	1.019 (0.044)	0.942 (0.017)	0.960 (0.062)	0.919 (0.022)
2SLS: bedrock depth	0.095 (0.393)	0.814 (0.063)	-0.361 (0.627)	0.818 (0.074)

*Notes:* This table reports estimates of  $\alpha$  from ordinary least squares (OLS) or two-stage least squares (2SLS) regressions of the form  $\widehat{\ln \delta}_i = \alpha \ln Q_i + \ln R_i + \ln w_i + u_i$  using core-based statistical areas (CBSAs) as the geographic units. The dependent variable  $\widehat{\ln \delta}_i$  is estimated in equation (12),  $Q_i$  is region  $i$ 's total production for Medicare beneficiaries,  $R_i$  is Medicare's Geographic Adjustment Factor, the  $w_i$  covariate includes mean two-bedroom property value and mean annual earnings for non-healthcare workers. and  $u_i$  is an error term. In the rows labeled "2SLS" we instrument for  $\ln Q_i$  using the specified instruments. The  $\ln R_i$  and  $\ln w_i$  controls are omitted in the columns labeled "no controls". In the columns labeled "no diag",  $Q_{ii}$  observations were omitted when estimating  $\widehat{\ln \delta}_i$  in equation (12). Standard errors (in parentheses) are robust to heteroskedasticity.



Table D.12: Classification of rare and common procedures in Medicare vs. private insurance data

Above median HCCI	0	1	total
Above median CMS			
0	82	18	100
1	12	88	100

*Notes:* This table compares the percentage of procedures classified as rare (above median frequency equals one) or common (above median frequency equals zero) in the public Medicare data versus the private insurance data from the Health Care Cost Institute (HCCI). Classifying procedures as rare versus common is consistent when using Medicare or privately insured data.

Table D.13: Specialization earnings and frequency

	(1)	(2)	(3)
Dependent variable: Per capita population elasticity			
Number of physicians in specialization (log, national)	-0.0716 (0.0139)		-0.0677 (0.0137)
Mean earnings (log)		-0.245 (0.0697)	-0.174 (0.0543)
Observations	209	209	209
R-squared	0.199	0.050	0.223

*Notes:* This table reports estimates of a regression of per capita population elasticity of physician count on the national count of physicians and mean earnings. Each observation is an NPPES taxonomy code. Earnings (wage and business income) data from Gottlieb et al. (2020) are reported by Medicare specialty groups. We use a crosswalk to map Medicare specialty groups to NPPES taxonomy codes. The estimation sample excludes 11 taxonomy codes that are not mapped to any Medicare specialty. Standard errors (in parentheses) are robust to heteroskedasticity.

Table D.14: Larger markets produce a greater variety of procedures

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	All	Anesthesia	E&M	Medical	Other	Pathology	Radiology	Surgery
Population (log)	0.357 (0.00737)	0.292 (0.0133)	0.169 (0.00529)	0.294 (0.00665)	0.428 (0.0116)	0.359 (0.0204)	0.202 (0.00611)	0.400 (0.00960)
Observations	306	306	306	306	306	306	306	306

*Notes:* This table reports the population elasticity of the number of distinct service codes produced in a region, estimated using Poisson pseudo-maximum likelihood (PPML). Column 1 shows the coefficient including all service types. The remaining columns show the coefficients for specific categories of service types. We use procedure classifications from the American Academy of Professional Coders, which groups codes into surgeries, anesthesia, radiology, pathology, medical, and evaluation & management (“E&M”) services (AAPC, 2021). We combine Category II codes, Category III codes and Multianalyte Assays into “other.”