Sources of Geographic Variation in Health Care:

Evidence from Patient Migration

Online Appendix

Amy Finkelstein
MIT and NBER

Matthew Gentzkow
Chicago Booth and NBER

Heidi Williams
MIT and NBER

December 2014

1 Definition of Utilization Measures

1.1 Overall Utilization

Following standard practice in the literature, we construct our health care utilization measure by aggregating care provided to Medicare beneficiaries as recorded in the inpatient, outpatient, and carrier claims data. The inpatient file records payments to inpatient hospital providers (such as hospitals), the outpatient file records payments to institutional outpatient providers (such as hospital outpatient departments), and the carrier file records payments to physicians and other non-institutional providers (such as independent ambulance providers). Following the methodology of Gottlieb et al. (2010)[1], we transform the claims (spending) data in these files into a quantity-based measure of utilization that is stripped of geographic variation in Medicare prices. This section describes our approach. The price adjustment procedures of Gottlieb et al. (2010) are specific to the types of claims examined, so we here separately describe our price adjustment procedure for inpatient, outpatient, and carrier claims. Our measure of health care utilization excludes several dimensions of care, including durable medical equipment, home health agency care, hospice care, skilled nursing facility care, inpatient rehabilitation facility care, and claims filed through Medicare Part D (prescription drug coverage). Recent work (Newhouse and Garber 2013) has suggested there is substantial variation in these additional measures of care.

1.1.1 Inpatient Claims

Inpatient payments are determined by an algorithm that we now describe.[2] First, Medicare sets so-called "standardized payment amounts" per discharge. These base payment amounts are meant to capture "the operating and capital costs that efficient facilities would be expected to incur in furnishing covered inpatient services." For example, for the fiscal year 2010, the operating base rate was $5,223, and the capital rate was $430.

Second, this base payment is adjusted by an area wage index to "reflect the expected differences in local market prices for labor." The wage index is revised annually. The wage index is applied to the labor-related portion of the base payment, where the labor-related portion is defined differently across hospitals as a function of the wage index: for hospitals with a wage index above 1.0, CMS applies a labor share of 68.8 percent; for hospitals with a wage index less than or equal to 1.0, CMS applies a labor share of 62 percent.

[1] See http://www.dartmouthatlas.org/downloads/papers/std_prc_tech_report.pdf.

[2] This section is based on details from various MedPAC reports describing reimbursement rules for inpatient services.

Third, the wage-adjusted base payment is adjusted for case mix using the DRG relative weights. To understand these DRG relative weights, note that over the time period of our data, inpatient payments are covered by a prospective payment system. Inpatient claims are centered around diagnosis related groups (DRGs) for specific services. Each DRG has a relative weight that aims to reflect "the expected relative costliness of inpatient treatment for patients in that group." DRG weights are set annually.

Finally, several factors are added to this wage- and case-adjusted base payment, including an adjustment for facilities that operate a resident training program (indirect medical education payment, IME), an adjustment for facilities that treat a disproportionate share of low-income patients (DSP), adjustments for bad debts (non-payments of deductibles and copayments by beneficiaries), new technology payments, and outlier payments for particularly expensive cases. Payments are reduced in cases of transfers, and critical access hospitals are paid separately on a cost basis.

Taken together, this description suggests that the key geographic price variation that we want to strip out comes from the area wage index, IME payments to residency programs, DSP payments to disproportionate share hospitals, bad debt adjustments, new technology payments, and outlier payments. In practice, our data do not include IME payments to residency programs, DSP payments to disproportionate share hospitals, bad debt adjustments, or new technology payments, so only the area wage index and outlier payments are relevant. Following Gottlieb et al. (2010), we define the price-adjusted inpatient level utilization for individual i's receipt of procedure k in region j at time t is estimated as:

$$U_{ikjt} = P_t \times DRG_{kt} + OP_{ikt}/WI_{jt}$$

where P_t is the national-level base payment rate at time t (not wage-adjusted), DRG_{ikjt} is the DRG weight used to determine payment for procedure k, OP_{ikt} is the outlier payment (if any), and WI_{jt} is the wage index factor defined as

$$WI_{jt} = 0.25 + 0.75 \times (\text{wage index for region } j \text{ at time } t)\,.$$

Gottlieb et al. (2010) clarify that they wage-adjust the outlier payments to account for differences in price level costs across regions, where region is defined as the provider's CBSA. If a provider is not located in a CBSA we use the state's rural wage index. For the few cases in which a provider's CBSA was uncertain and it was located within a state that does not have a rural wage index (MA, NJ, RI, DC, PR), we used the median of all the urban wage indexes in that state for that year.

To ensure that price-adjusted hospital expenditures add up to aggregated actual expenditures, we follow Gottlieb et al. (2010) and make a further adjustment (λ) to ensure the adding up constraint, where λ is defined implicitly by:

$$\sum\sum\sum\sum\nolimits_{ikjt} \text{Total Hospital Expenditures} = \lambda \sum\sum\sum\sum\nolimits_{ikjt} U_{ikjt}$$

where the sum is taken over all age groups (including the under 65 population) after randomly dropping 75% of the non-movers. Note that Gottlieb et al. (2010) further adjust for age, sex, and race, which we do not do.

1.1.2 Carrier Claims

Carrier claim based reimbursements are centered around Healthcare Common Procedure Coding System (HCPCS) codes for specific services.[3] Payments at the HCPCS code level—with the caveat that HCPCS codes are sometimes more specific when "modifier" codes are included—are determined as follows.[4]

First, CMS estimates the "amount of work required to provide a service, expenses related to maintaining a practice, and liability costs." Each of these three components—work (W), practice expense (PE), and professional liability insurance (PLI)—are assigned a relative value unit (RVU) weight.

Second, each of the three RVU components (W, PE, PLI) are adjusted by separate geographic practice cost indices (GPCIs).

Third, the GPCI-weighted sum of the three RVU components is then multiplied by a conversion factor of 0.8, reflecting that beneficiaries pay 20% of carrier costs directly through their coinsurance.

Finally, several payment modifiers are applied, including adjustments for different types of providers (physicians versus non-physicians, participating versus nonparticipating physicians); geographic bonuses paid to providers in

[3]This section is based on details from various MedPAC reports describing reimbursement rules for carrier claims.

[4]We follow Gottlieb et al.'s (2010) treatment of claims associated with multiple modifier codes, and use only the first modifier code.

designated "health provider shortage areas" (HPSAs), and service-specific adjustments for primary care and major surgical procedures.

Taken together, this description suggests that the key geographic price variation that we want to strip out comes from the three RVU-specific geographic practice cost indices (GPCIs) and the geographic specific HPSA bonuses. Gottlieb et al. (2010) estimate carrier-specific utilization by—for each HCPCS code (and HCPCS-modifier code combination, if applicable)—merging on national (that is, not area-specific) RVU weight as documented in a CMS-provided fee schedule. Some HCPCS codes have an RVU of 0 in the fee schedule (mainly due to statutory exclusions) or do not merge to the fee schedule. For such codes, we follow Gottlieb et al. (2010) and assign the RVU weight to be the median carrier payment by HCPCS code-modifier-year, divided by a year-specific price conversion factor.

As with our inpatient claims calculation, we make an adjustment to ensure that price-adjusted hospital expenditures add up to aggregated actual expenditures. Gottlieb et al. (2010) uses a different standard price adjustment for ambulatory surgery centers, anesthesia, and certified nurse anesthetists which we do not do for simplicity.

1.1.3 Outpatient Claims

Like inpatient services, outpatient payments are covered by a prospective payment system over the time period of our data.[5] Outpatient claims are centered around ambulatory payment classifications (APCs) for specific services. Each APC has a relative weight that aims to reflect "resource requirements of services." APC weights are set annually, and payments are determined as follows.

First, APC weights are multiplied by a wage-adjusted conversion factor. Specifically, the labor share—set at 60 percent for all institutions—is adjusted by a hospital wage index, while the remaining (40%) non-labor share is unadjusted.[6] Second, adjustments are made for cancer hospitals, children's hospitals, rural hospitals with 100 or fewer beds, and sole community hospitals. Finally, outlier payments can be made for particularly expensive cases.

Taken together, this description suggests that the key geographic price variation that we want to strip out comes from the wage index and the hospital type adjustments. Gottlieb et al. (2010) simplify this to focus on the wage index. Specifically, they construct

$$WI_{jt} = 0.4 + 0.6 \times (\text{wage index for region } j \text{ at time } t)$$

and divide payments to providers by this wage adjustment factor. As with inpatient and carrier claims, we also make an adjustment to ensure that price-adjusted outpatient expenditures add up to aggregated actual expenditures.

1.2 Components of Overall Utilization

To explore our aggregate utilization estimates in more detail, we construct a number of disaggregated measures. The first four measures (inpatient, outpatient, emergency room and other) are mutually exclusive and exhaustive.

- **Inpatient utilization.** This measures utilization recorded in the inpatient claims data excluding claims with revenue center codes for emergency room services. We construct a quantity-based measure of inpatient utilization that is stripped of geographic variation in Medicare prices (as described above), as well as an indicator for any hospitalization (defined as any inpatient utilization).

- **Outpatient utilization.** This measures utilization recorded in the outpatient claims data, including office visits recorded in the carrier claims data, excluding claims with revenue center codes for emergency room services. We construct a quantity-based measure of outpatient utilization that is stripped of geographic variation in Medicare prices (as described above).

- **Emergency room utilization.** This measures utilization recorded in inpatient or outpatient claims with revenue center codes for emergency room services, together with carrier claims that take place in an emergency room. We construct a quantity-based measure of emergency room utilization that is stripped of geographic variation in Medicare prices (as described above), in addition to an indicator for any emergency room utilization.

- **Other utilization.** This includes carrier claims not included in our inpatient, outpatient, and emergency room utilization measures described above; this includes laboratory claims, ambulance claims, nursing facility claims,

[5] This section is based on a 2010 MedPAC report summarizing the reimbursement rule for outpatient claims. See http://www.medpac.gov/documents/medpac_payment_basics_10_opd.pdf.

[6] New technology APCs are reimbursed differently, but as best we can tell are not addressed by Gottlieb et al. (2010).

and skilled nursing facility claims.

- **Whether a patient has seen a primary care physician or (separately) a specialist physician**. Our definition of primary care physicians and specialists follows the Dartmouth Atlas.[7] Specifically, we crosswalk the primary care and specialist definitions in the Dartmouth Atlas to the list of physician categories included in the provider specialty table provided by the Centers for Medicare and Medicaid Services (CMS) in order to replicate the categories listed in the Dartmouth Atlas.[8] When we examine physicians in the inpatient claims data, we rely on the attending physician ID; when we examine physicians in the carrier claims data, we rely on the referring physician ID; when we examine physicians in the outpatient claims data, we rely on the attending physician ID.

- **Diagnostic and imaging tests**. Our definition of diagnostic and imaging tests follows Song et al. (2010), and is based on BETOS codes: codes beginning with T are diagnostic tests, and codes beginning with I are imaging tests.

- **Preventive care procedures**. We measure a count of "preventive care" procedures following those measured in the Dartmouth Atlas and the Centers for Medicare and Medicaid:

  - Mammogram is defined following the Dartmouth Atlas.[9] Specifically, we define this based on CPT codes 76090-76092 and 76083; ICD-9 codes 87.36 and 87.37; V codes 76.11 and 76.12; and revenue center code 0403.
  - Hemoglobin A1c testing, blood lipids testing, negative retinal exam, and negative retinal or dilated eye exam are defined following the Dartmouth Atlas.[10]
  - Ambulatory visits are defined following the Dartmouth Atlas.[11]
  - Cardiovascular screening blood testing, seasonal influenza virus vaccine, diabetes self-management training, bone mass measurements, colorectal cancer screening, pap smears, pelvic examinations, and prostate cancer screening are defined following CMS's preventive care definitions.[12] Note that for colorectal cancer screening, we only use CPT code 82270 and HCPCS code G0328.

- **Number of different doctors seen**. We identify unique physicians by linking time-varying physician identifiers. The count of "different doctors seen" is constructed by summing the number of unique physicians that billed for a given patient in that year.[13]

[7] See http://www.dartmouthatlas.org/downloads/methods/research_methods.pdf, page 6.

[8] The provider specialty table used is available at http://www.resdac.org/sites/resdac.org/files/HCFA%20Provider%20Specialty%20Table.txt. Crosswalking from the Dartmouth Atlas definitions to the CMS-provided table is straightforward as the two sources enable an exact match.

[9] See http://www.dartmouthatlas.org/data/table.aspx?ind=169.

[10] See http://www.dartmouthatlas.org/data/map.aspx?ind=160.

[11] See http://www.dartmouthatlas.org/data/table.aspx?ind=170.

[12] See http://www.cms.gov/Medicare/Prevention/PrevntionGenInfo/Downloads/MPS_QuickReferenceChart_1.pdf.

[13] Any provider with a unique set of identifiers is considered a different "doctor", though these identifiers can be used for billing groups and non-physician practitioners as well.

2 Additional Results

2.1 Descriptive Evidence on Movers

We present additional descriptive evidence on movers and their moves. Online Appendix Figure 1 plots the distribution of the distance from a mover's origin to their destination HRR. Online Appendix Figure 2 shows the distribution of movers across destination HRRs. Online Appendix Table 1 shows the distribution of movers by census division.

Online Appendix Figure 3 plots mean log utilization in each relative year for movers. Since log utilization will naturally trend upward due to aging and the passage of time, for comparison we also plot the mean log utilization of a matched sample of non-movers. The figure shows that moving is associated with an increase in utilization in general, and with an upward spike in utilization in the year of the move.

Table 1 provided some comparison of the characteristics of movers and non-movers. For additional information, we also report results from the Health and Retirement Study (HRS). The HRS is a nationally representative longitudinal survey of Americans over the age of 50.[14] Since 1992, the HRS has been administered in (approximately) two-year cycles, following individuals and their spouses from their time of entry to the survey sample until their death. We use data through 2008. Data without individual identifiers and zipcode-level geographic information can be downloaded from the HRS website.[15] Our analysis uses the restricted-access HRS data, which contains zipcode-level geographic information, in order to identify movers. We define movers as individuals who move between HRRs. In order to define HRRs, we merge the HRS data with a zipcode-HRR crosswalk downloaded from the Dartmouth Atlas website.

Our analysis uses a version of the HRS data prepared by the RAND Corporation (RAND HRS). The RAND HRS contains most measures that are surveyed in the HRS, and aims to create variables consistent across the waves of the survey. We merged in the "reasons for move" question which is not included in the RAND HRS. Waves are defined as follows: Wave 1 (1992), Wave 2 (1993 and 1994), Wave 3 (1995 and 1996), Wave 4 (1998), Wave 5 (2000), Wave 6 (2002), Wave 7 (2004), Wave 8 (2006), and Wave 9 (2008).

We limit the sample to individuals aged 65+ to match our Medicare data, and define a mover as an individual who moves between HRRs exactly once. This gives us a sample of about 22,000 individuals, observed on average for about 4 waves (i.e., 8 years); about 10 percent of the sample moved during this time. In addition, about 1,100 of the 2,000 movers answered a question about why they moved.

Online Appendix Table 2 reports summary statistics for movers and non-movers in the HRS data. It indicates that movers are more likely to be white, not married, and higher educated than non-movers.

Online Appendix Figure 4 shows the most common reasons given for moving.

We also used the HRS to investigate time-varying correlates of moving. Specifically we estimated the following panel-wave level regression:

$$\text{MOVE}_{it} = \alpha_i + \tau_t + \delta X_{it} + \varepsilon_{it}. \tag{1}$$

The dependent variable MOVE_{it} is an indicator variable if person i is living in a different HRR in wave t than in wave $t-1$. Conditional on individual fixed effects (α_i) and wave fixed effects (τ_t), we analyze the bivariate association between various time varying covariates (X_{it}) and the probability of moving. Online Appendix Table 3 shows the results for different indicator variables X_{it}. In row (1) we consider an indicator variable for whether the individual is not married; in row (2), whether they are separated or divorced, in row (3) whether they are widowed, in row (4) whether they are retired or partly retired, and in row (5) an indicator variable for self-reported health being "fair" or "poor" (instead of "good", "very good" or "excellent").

2.2 Analysis of Utilization in Levels (Instead of Logs)

For the econometric and economic reasons discussed in the main text, our preferred outcome measure is the log of utilization. However, as discussed in Section 4.2 we can use our baseline estimated (log) model to ask about how the geographic variation in level utilization would change if either the place component γ_j or the distribution of the patient component c_{it} were equalized across areas.

Starting from our estimated baseline model, we first predict expected log utilization for each patient year as:

[14] The HRS (Health and Retirement Study) is sponsored by the National Institute on Aging (grant number NIA U01AG009740) and is conducted by the University of Michigan.

[15] See http://hrsonline.isr.umich.edu/.

$$\text{(2)} \qquad \hat{y}_{ijt} = \hat{c}_{it} + \hat{\gamma}_j + \hat{\tau}_t.$$

Next, we compute counterfactual log utilization when $\hat{\gamma}_j$ is set for all j to its average $\overline{\hat{\gamma}}$ as:

$$\text{(3)} \qquad \hat{y}_{ijt}^{pleq} = \hat{c}_{it} + \overline{\hat{\gamma}} + \hat{\tau}_t.$$

Finally, we compute counterfactual log utilization when the distribution of c_{it} is equalized across areas as:

$$\text{(4)} \qquad \hat{y}_{ijt}^{pateq} = \hat{c}_{it}^{'} + \hat{\tau} + \hat{\gamma}_j.$$

where $\hat{c}_{it}^{'}$ is a random draw without replacement from the set C_t of all patient components in year t.

To compute baseline and counterfactual utilization in levels, we exponentiate each of these terms, and average them by HRR, weighting non-movers up by four as usual to account for our sampling procedure. Let $\bar{\hat{y}}_j$, $\bar{\hat{y}}_j^{pleq}$, and $\bar{\hat{y}}_j^{pateq}$ be these HRR-level means. For a group of HRRs R, let $\bar{\hat{y}}_R$, $\bar{\hat{y}}_R^{pleq}$, and $\bar{\hat{y}}_R^{pateq}$ denote their simple averages across HRRs in R. Define the share of the level utilization difference between R and R' that would be eliminated if we equalized place effects as:

$$\text{(5)} \qquad \hat{S}_{place}^{level}\left(R,R'\right) = 1 - \frac{\bar{\hat{y}}_R^{pleq} - \bar{\hat{y}}_{R'}^{pleq}}{\bar{\hat{y}}_R - \bar{\hat{y}}_{R'}}$$

Define the share that would be eliminated if we equalized the distribution of patient effects as:

$$\text{(6)} \qquad \hat{S}_{pat}^{level}\left(R,R'\right) = 1 - \frac{\bar{\hat{y}}_R^{pateq} - \bar{\hat{y}}_{R'}^{pateq}}{\bar{\hat{y}}_R - \bar{\hat{y}}_{R'}}$$

Note that unlike the log versions $\hat{S}_{place}$ and $\hat{S}_{pat}$, $\hat{S}_{place}^{level}$ and $\hat{S}_{pat}^{level}$ need not sum to one.

Results are presented in Online Appendix Table 4.

2.3 Alternative Definitions of Movers

Online Appendix Table 5 shows robustness of our main results to alternative ways of defining who is a mover. The first row repeats our baseline results. As described in Section 3, our baseline definition of a mover is someone whose HRR of residence (based on their address on file for Social Security payments) changes exactly once, and their average share of claims in the destination HRR increases by at least 0.75 in years after the move relative to years before the move. Rows (2) and (3) show that our estimate of the role of patients is not sensitive to making this claim share threshold looser (0.6) or tighter (0.9). Row (4) shows results when we eliminate the claim share criterion entirely, including all individuals whose address on file changes exactly once. This increases our number of movers by about 50%, and, without further adjustment, also increases the patient share to 0.592.

However, as shown in Online Appendix Figure 5 this definition of movers based solely on address change includes a substantial number of mismeasured moves; Figure 2 shows our baseline definition. Some moves “begin” prior to the move year. In particular, there is a 6 percentage point drop in the share of claims in one’s origin HRR between relative year -2 and relative year -1 (both of which are supposed to be pre-move years). There is also a more gradual but still noticeable downward trajectory in the share of claims in the origin in all years prior to the move year; in total, about a 6 percentage point drop in the share between relative years -10 and -2. Likewise, some moves seem to happen “after” the move year, as evidenced by the continued upward trajectory of claim share in destination relative to origin in years after the move.

In addition, we make an adjustment to our basic estimating equation to allow for the possibility that we observe the timing of moves with error. Let $\hat{J}(i,t)$ be i’s current observed HRR ($o(i)$ for $r(i,t) < 0$ and $d(i)$ for $r(i,t) > 0$), and let $\hat{J}'(i,t)$be i’s other observed HRR ($d(i)$ for $r < 0$ and $o(i)$ for $r > 0$). We assume that in relative year r the current residence is reported correctly with probability λ_r, and misreported with probability $1-\lambda_r$, independent of all other variables in the model:

$$J(i,t) = \begin{cases} \hat{J}(i,t) & \text{with probability } \lambda_r \\ \hat{J}'(i,t) & \text{with probability } 1-\lambda_r. \end{cases}$$

Our model now becomes:

$$y_{it} = \alpha_i + \lambda_{r(i,t)}\gamma_{\hat{J}(i,t)} + \left(1-\lambda_{r(i,t)}\right)\gamma_{\hat{J}'(i,t)} + \tau_t + x_{it}\beta + \tilde{\varepsilon}_{it} \tag{7}$$

where $\tilde{\varepsilon}_{it} = \varepsilon_{it} + \gamma_{J(i,t)} - \lambda_{r(i,t)}\gamma_{\hat{J}(i,t)} - \left(1-\lambda_{r(i,t)}\right)\gamma_{\hat{J}'(i,t)}$ remains conditionally mean zero. We estimate λ_r from the analysis in Online Appendix Figure 2, then estimate Online Appendix equation (7) plugging in these estimates. Note that for non-movers, $\hat{J}(i,t) = \hat{J}'(i,t)$ by definition.

When we make this adjustment to our baseline definition of movers and the claim share pattern, we find that that, not surprisingly given our ability to measure move timing much more cleanly, this makes relatively little difference. As shown in row (5), our estimate of the role for patients changes from 0.465 in the baseline to 0.444 with this adjustment.

2.4 Empirical Bayes Adjustment for Event Study

Figure 6 in the main text presents event study estimates of equation (5). To account for noise in estimating $\hat{\delta}_i$, in Online Appendix Figure 6 we apply an Empirical Bayes (EB) adjustment procedure to these estimates as in Morris (1983). Specifically, we compute for each HRR j a convex combination of the estimated $\hat{y}_j$ terms and the overall mean of log utilization across HRRs, which we denote $\bar{y}$. The EB-adjusted estimates are given as

$$\hat{y}_j^{EB} = \left(1-\hat{B}_j\right)\hat{y}_j + \hat{B}_j\bar{y} \tag{8}$$

where

$$\hat{B}_j = \left(\frac{N_H-1-2}{N_H-1}\right)\frac{\hat{\sigma}_j^2}{\hat{\sigma}_j^2+\hat{\sigma}^2}. \tag{9}$$

$\hat{\sigma}_j^2$ is the standard error of the HRR mean $\hat{y}_j$, which is the within-HRR standard deviation divided by the square root of the number of observations in the HRR. $\hat{\sigma}^2$ is the weighted average of squared deviations of the $\hat{y}_j$ terms from $\bar{y}$ less the weighted average of the $\hat{\sigma}_j^2$ terms. $\hat{\sigma}^2$ is computed from the following equations through an iterative procedure.

$$\hat{\sigma}^2 = \max\left\{0, \frac{\sum_j W_j\left\{\left(\frac{N_H}{N_H-1}\right)(\hat{y}_j-\bar{y})^2 - \hat{\sigma}_j^2\right\}}{\sum_j W_j}\right\} \tag{10}$$

$$W_j = \frac{1}{\hat{\sigma}_j^2+\hat{\sigma}^2} \tag{11}$$

We iterate the weight W until we obtain a stable $\hat{\sigma}^2$ term. $N_H = 306$, the number of HRRs.

Once we have the estimates $\hat{y}_j^{EB}$ for each HRR j, we can compute $\hat{\delta}_i^{EB}$ for each patient i as follows,

$$\hat{\delta}_i^{EB} = \hat{\delta}_i = \hat{y}_{d(i)}^{EB} - \hat{y}_{o(i)}^{EB} \tag{12}$$

Then we estimate the event study regression shown in equation (13), and plot the coefficients on the $\theta_{r(i,t)}\hat{\delta}_i^{EB}$ terms.

$$y_{it} = \tilde{\alpha}_i + \theta_{r(i,t)}\hat{\delta}_i^{EB} + \tau_t + \rho_{r(i,t)} + x_{it}\beta + \varepsilon_{it} \tag{13}$$

We use the following posterior distribution for the $\hat{y}_j^{EB}$ terms when computing bootstrapped standard errors:

$$N\left(\hat{y}_j^{EB}, \hat{\sigma}_j^2\left(1-\hat{B}_j\right)\right). \tag{14}$$

2.5 Stability of Geographic Variation

Online Appendix Figure 7 confirms that geographic variation in log utilization is stable over time. We first divide HRRs into quintiles based on average log utilization across the whole sample ($\bar{y}_j$), then plot average utilization by quintile in each year.

2.6 Additional Specifications

Online Appendix Table 6 presents results for components of utilization, narrowing in on relative years -1 to 1, in the spirit of regression discontinuity.

Online Appendix Table 7 presents results for alternative measures of utilization.

Online Appendix Table 8 presents additional robustness checks.

Online Appendix Table 9 presents results separately by age quartile.
Online Appendix Table 10 shows additive decomposition results for health outcomes.
Online Appendix Table 11 shows the share of patient-years with a value of zero for various components of utilization and health measures.

2.7 Additional Event Study Figures

Online Appendix Figure 8 shows event study results for components of utilization, corresponding to the results presented in Table 4.

Online Appendix Figure 9 shows event study figures for our main robustness checks: using only part of the sample (first third, second third, third third), excluding patients who attrit for various reasons (death, entering an HMO, missing outcomes), using alternative geographic units (states and Hospital Service Areas (HSAs)), and excluding patients based on move geography (only cross-state moves, only cross-region moves), corresponding to the results presented in Table 5.

Online Appendix Figure 10 shows event studies with level utilization and percentile of utilization as the outcome variable, corresponding to the results presented in Online Appendix Table 7.
Online Appendix Figure 11 shows event studies for health outcomes.

Online Appendix Figure 12 shows event study results for the full sample, including non-movers.

References

Gottlieb, Daniel J., Weiping Zhou, Yunjie Song, Kathryn G. Andrews, Jonathan S. Skinner, and Jason M. Sutherland. 2010. “Prices Don’t Drive Regional Medicare Spending Variations.” *Health Affairs*, 29(3): 537–543.

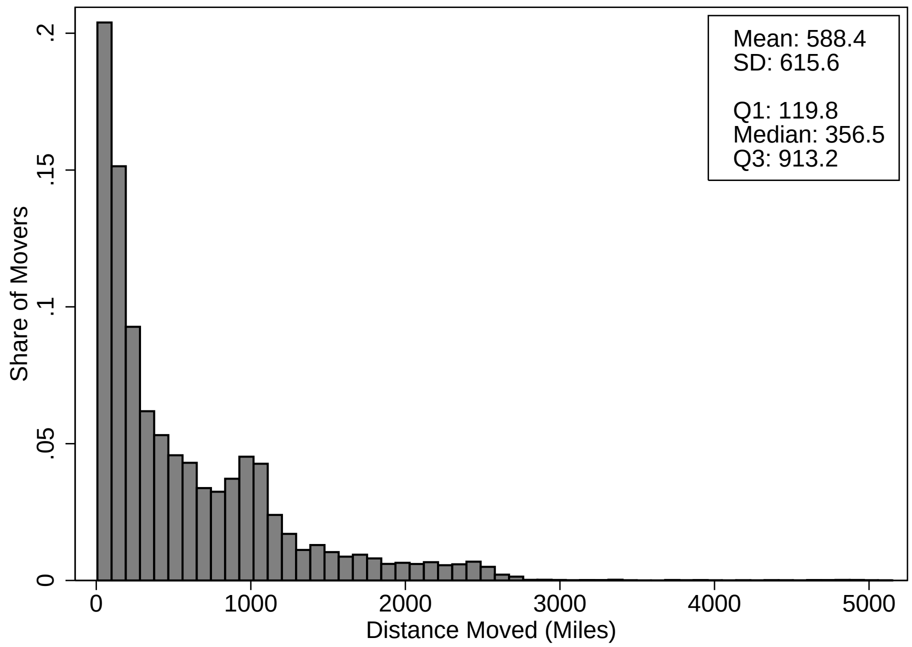
Health and Retirement Study. 2014. Public Use Datasets Produced and distributed by the University of Michigan with funding from the National Institute on Aging (grant number NIA U01AG009740).

Morris, Carl N. 1983. “Parametric Empirical Bayes Inference: Theory and Applications.” *Journal of the American Statistical Association*, 78(391): 47–55.

Newhouse, Joseph P., and Alan M. Garber. 2013. “Geographic Variation in Medicare Services.” *New England Journal of Medicine*, 368(16): 1465–1468.

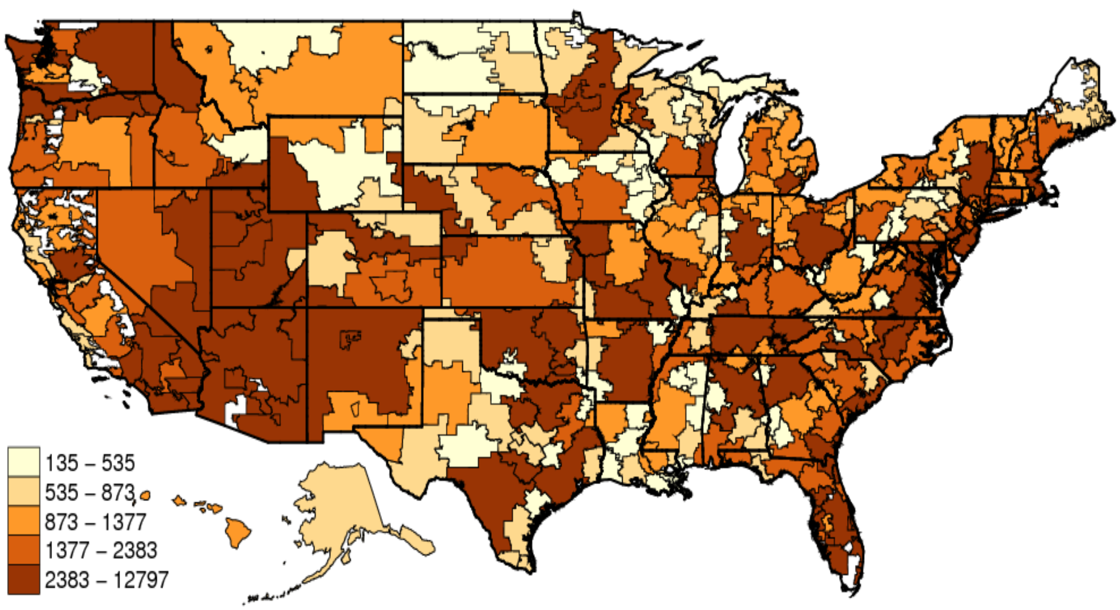
Song, Yunjie, Jonathan Skinner, Julie Bynam, Jason Sutherland, and Elliott Fisher. 2010. “Regional Variations in Diagnostic Practices.” *New England Journal of Medicine*, 363(1): 45–53.

Online Appendix Figure 1: Distribution of Distance Moved



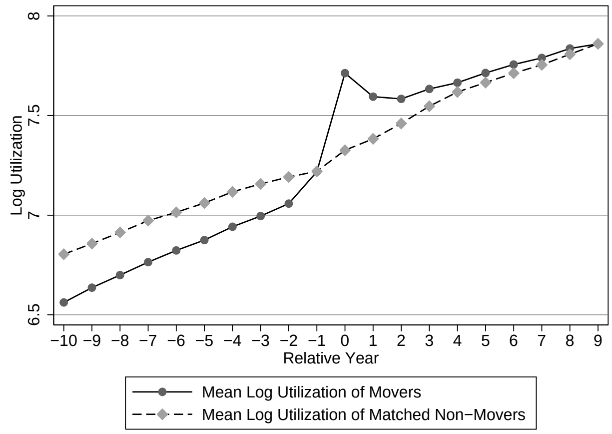
Notes: Figure shows the distribution of distances moved. Distance is measured between the population-weighted centroids of HRRs. The sample is all movers ($N = 497,097$ patients).

Online Appendix Figure 2: Distribution of the Number of Movers Across Destination HRRs



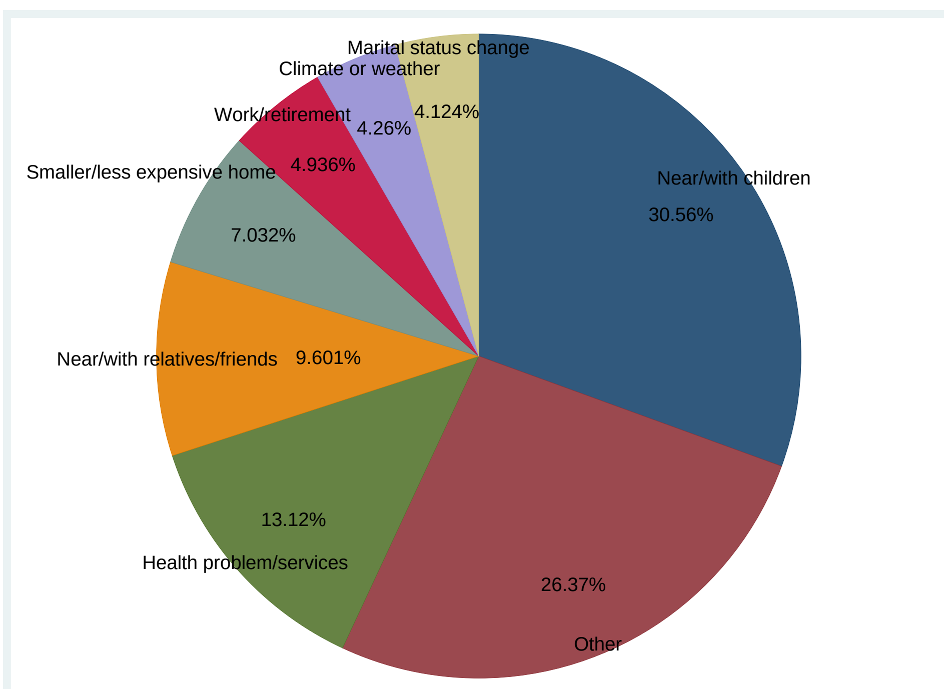
Notes: Map shows the distribution of the number of movers in different destinations in quintiles. Lower and upper limits of each quintile are displayed in the legend. The sample is all movers ($N = 497,097$ patients).

Online Appendix Figure 3: Log Utilization Over Relative Years



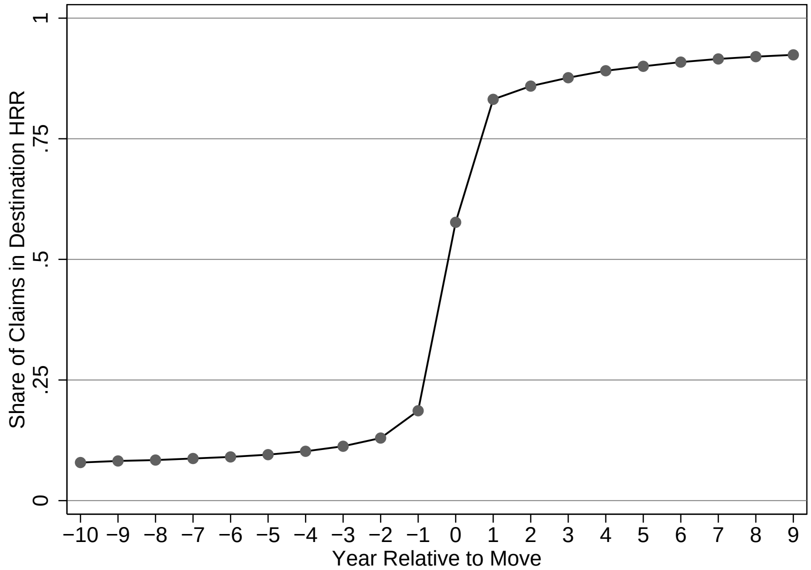
Notes: Figure shows the mean log utilization by relative year for movers and a matched sample of non-movers. In Figure 4, we describe the construction of a matched sample of non-movers from the mover's origin HRR; we construct an analogous sample of non-movers in the mover's destination HRR. For each relative year, we compute the mean of log utilization each matched sample of nonmovers, and take the average of the two. The sample is all movers ($N = 3,702,189$ patient-years) and the same number of non-mover patient-years.

Online Appendix Figure 4: HRS Top Reasons for Move



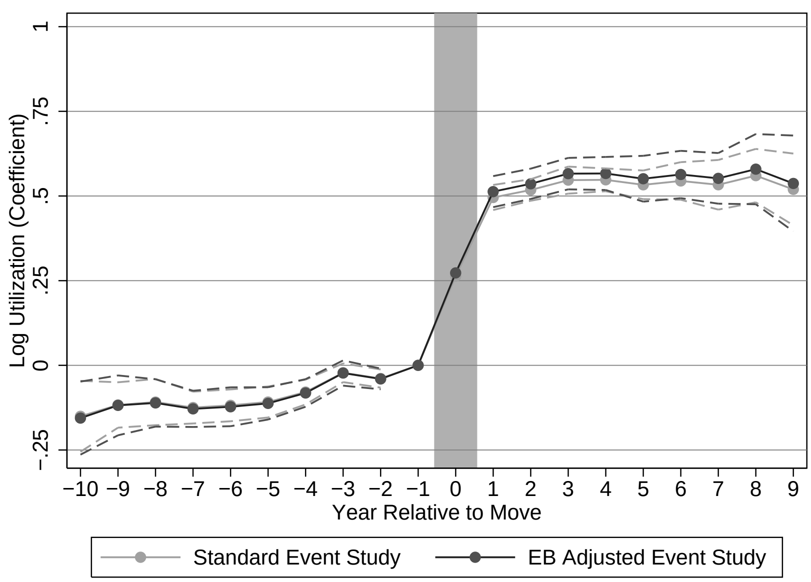
Notes: Pie chart shows the most common reasons for moving, based on the HRS. Reasons mentioned fewer than 50 times are grouped under the "Other" category. The most common reasons are to live near/with children, relatives, or friends, or for health reasons. Of the 2,025 movers in the data, 1,144 provide reasons; some provide multiple. The sample is all reasons given ($N = 1,479$ observations).

Online Appendix Figure 5: Share of Claims in Destination (All Address Changers)



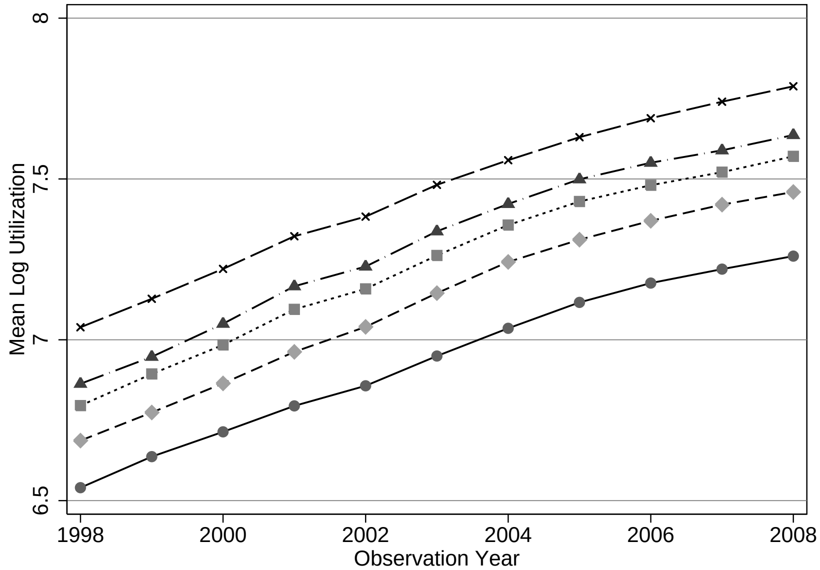
Notes: Figure displays the mean share of claims in the destination HRR by relative year. Here, we categorize someone as a mover if their HRR of residence changes exactly once ($N = 5,698,027$ patient-years). By contrast, in our baseline definition we apply the additional sample restriction that movers must also increase the share of claims in their destination HRR, among claims in either their origin or destination HRR, by at least 0.75 in the post-move years relative to the pre-move years.

Online Appendix Figure 6: Event-Study with Empirical Bayes Adjustment



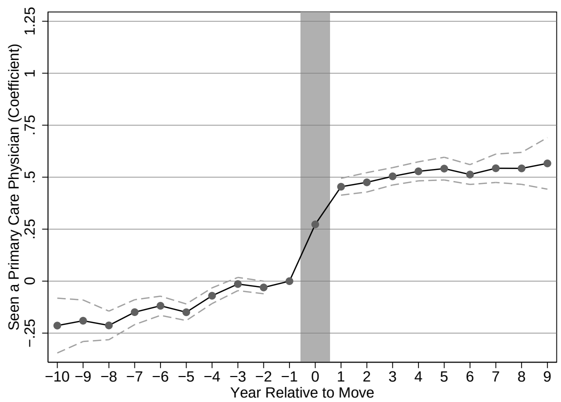
Notes: The EB adjusted event study is shown superimposed over the standard event study from Figure 6. The EB adjusted event study is constructed in the same manner as Figure 6, except the estimates of $\hat{\hat{\delta}}_i$ are adjusted using the empirical Bayes (EB) procedure. The dashed lines show the 95% confidence interval, constructed using the same bootstrap approach as in Figure 6. The sample is all movers ($N = 3,702,189$ patient-years).

Online Appendix Figure 7: Time Series of Mean Log Utilization of HRRs by Quintile

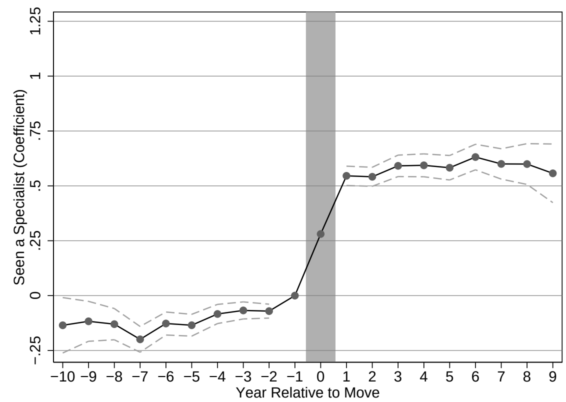


Notes: Figure displays a time series plot of the mean log utilization for each HRR quintile and year. HRR quintiles are defined by taking the average across individuals within each HRR-year, upweighting non-movers by four, and then taking the simple average within HRR across years. The sample is all movers ($N = 3,702,189$ patient-years).

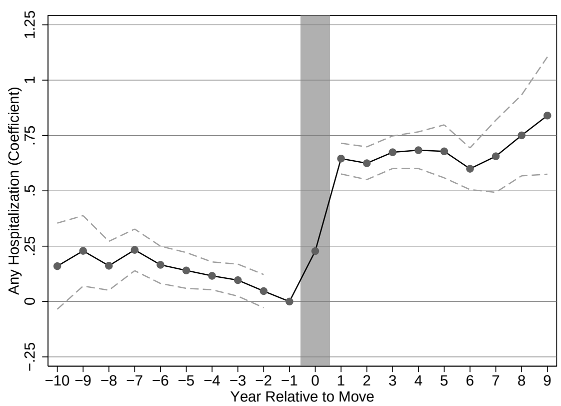
Online Appendix Figure 8: Event Study Results for Various Components of Utilization



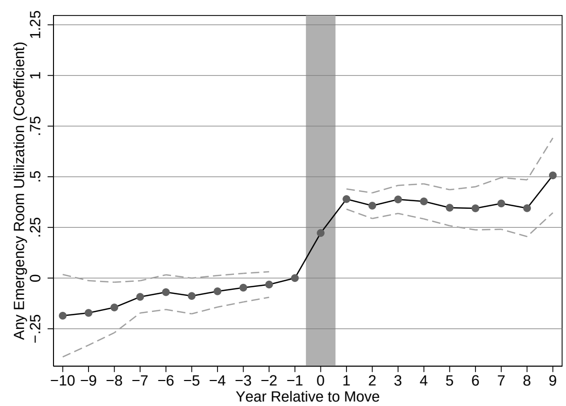
(a) Seen a Primary Care Physician



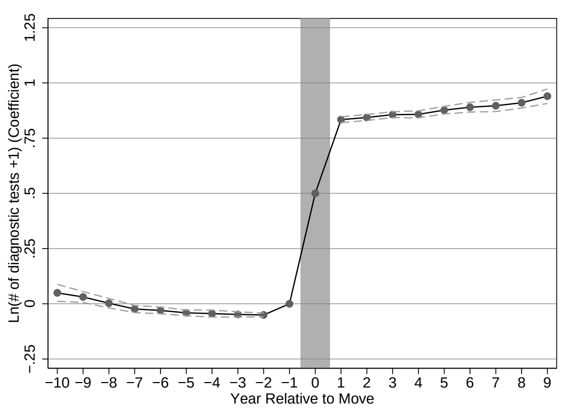
(b) Seen a Specialist



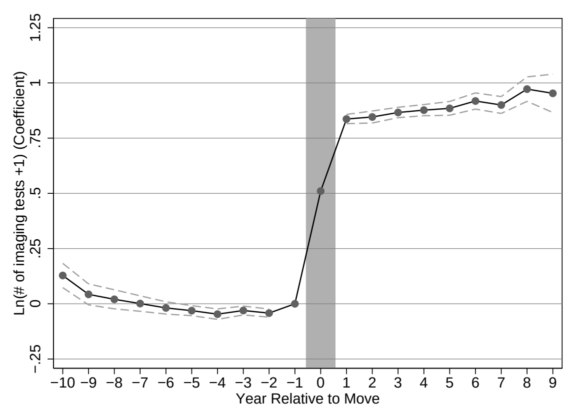
(c) Any Hospitalization



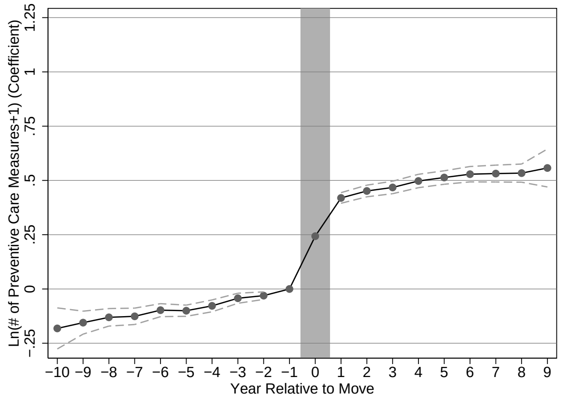
(d) Any Emergency Room Utilization



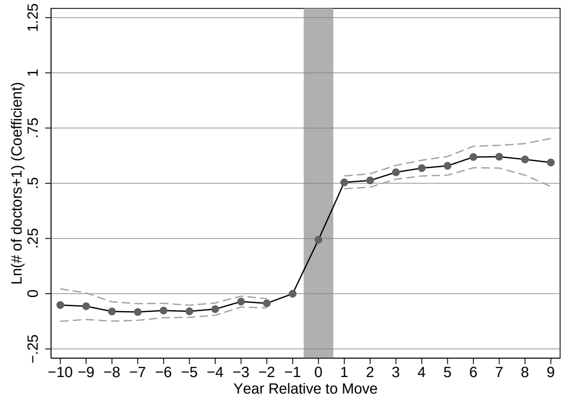
(e) Log Number of Diagnostic Tests



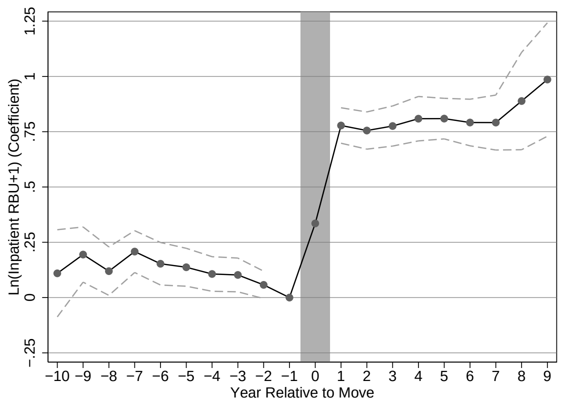
(f) Log Number of Imaging Tests



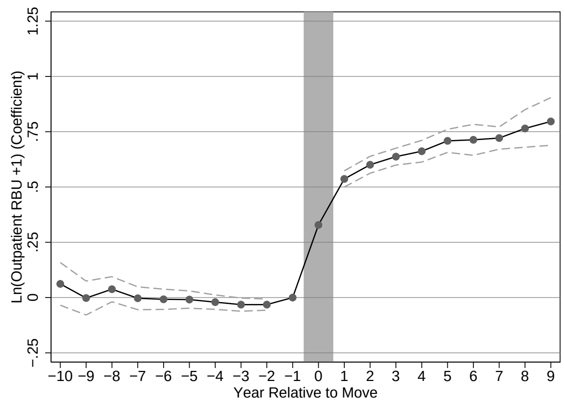
(g) Log Number of Preventive Care Measures



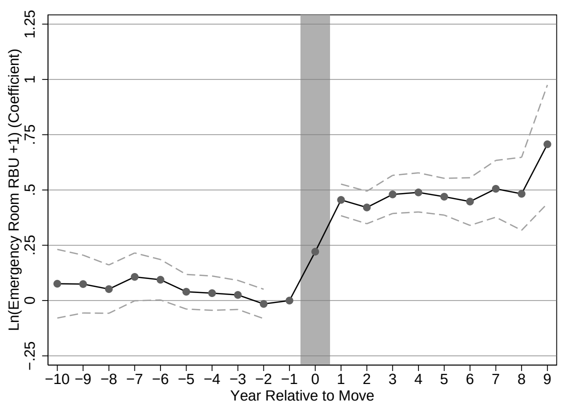
(h) Log Number of Different Doctors Seen



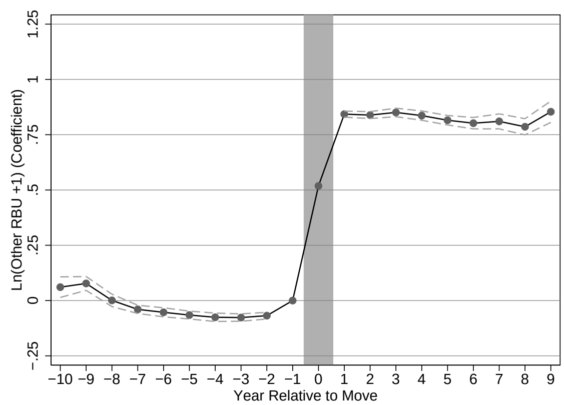
(i) Log Inpatient Utilization



(j) Log Outpatient Utilization



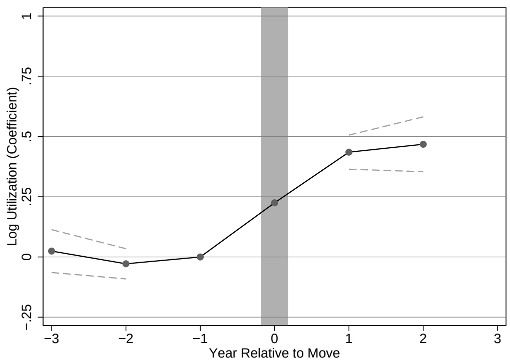
(k) Log Emergency Room Utilization



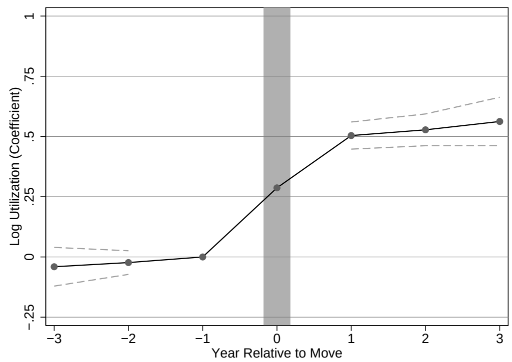
(l) Log Other Utilization

Notes: These figures are constructed in the same manner as Figure 6, except the dependent variable is now an alternate utilization measure. The dashed lines show the 95% confidence interval, constructed using the same bootstrap approach as in Figure 6. All log outcome measures are the log of the outcome plus 1. Online Appendix Table 11 shows the percent with zero for each of these outcomes. The sample is all movers ($N = 3,702,189$ patient-years).

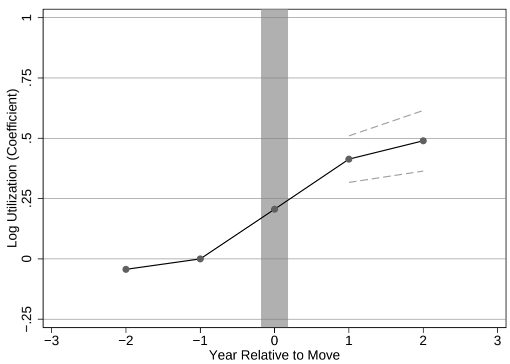
Online Appendix Figure 9: Event Studies for Robustness Specifications



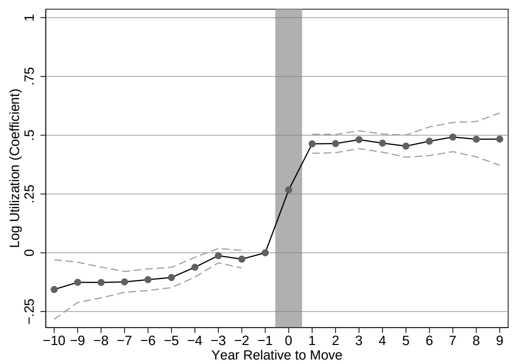
(a) First Third of Sample Only (1998-2001)



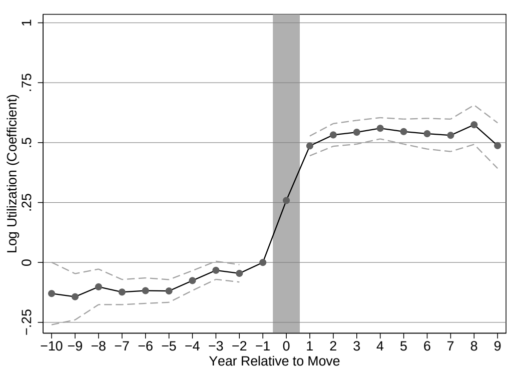
(b) Second Third of Sample Only (2002-2005)



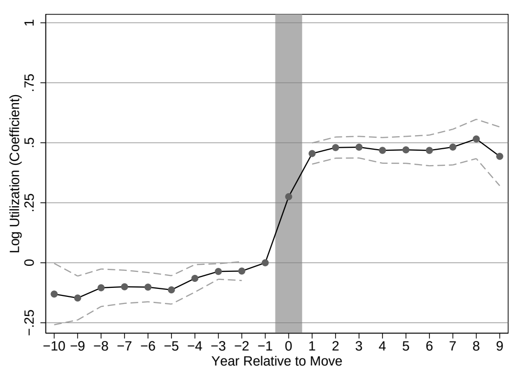
(c) Third Third of Sample Only (2006-2008)



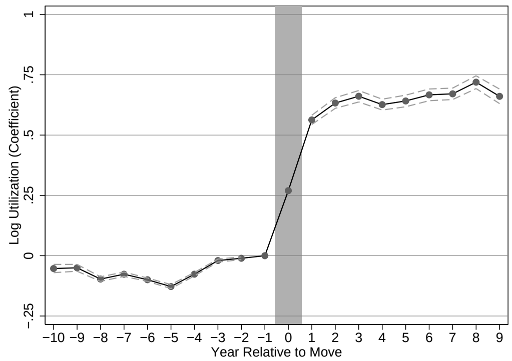
(d) Patients Who Never Die



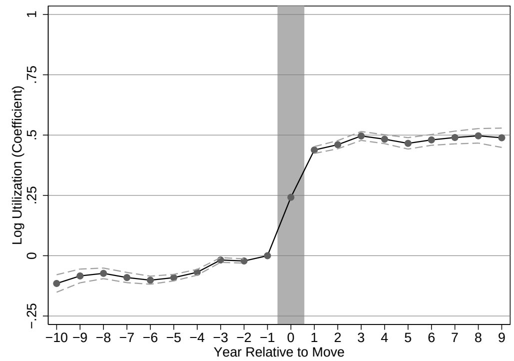
(e) Patients Never in an HMO



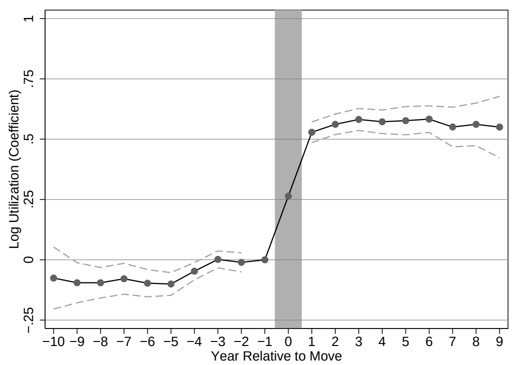
(f) Patients Never Missing Outcomes



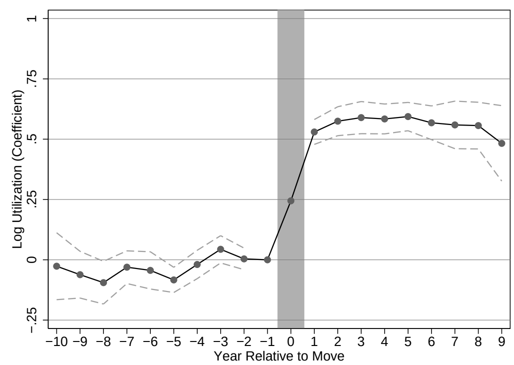
(g) Analysis at State Level



(h) Analysis at Hospital Service Area (HSA) Level



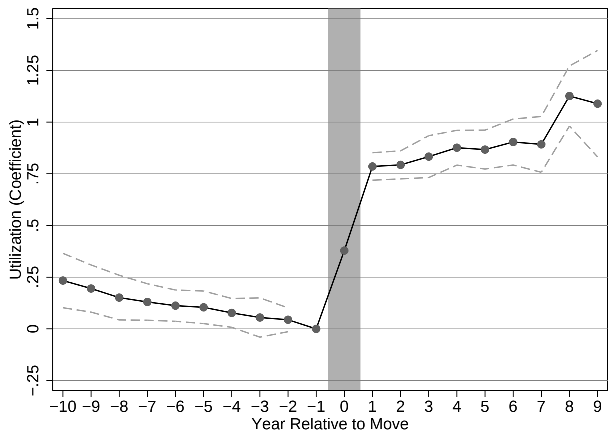
(i) Limit to Cross State Movers



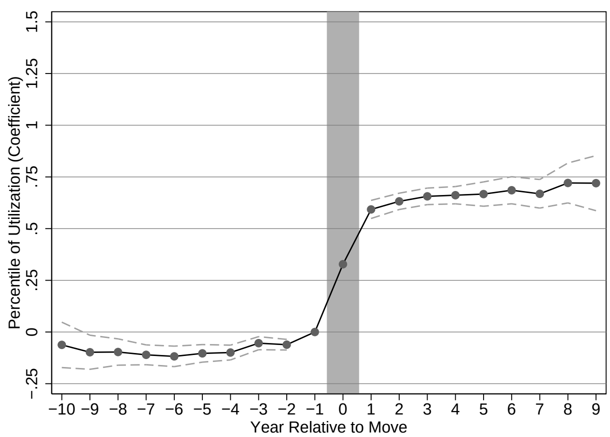
(j) Limit to Cross Census Region Movers

Notes: These figures are constructed in the same manner as Figure 6, except equation (5) and $\hat{\delta}_i$ are estimated based on only the subsample of patient-years specified. In addition, in panels (g) and (h), the definition of movers is unchanged, but the $\hat{\delta}_i$ are estimated between the origin and destination states or HSAs. The dashed lines show the 95% confidence interval, constructed using the same bootstrap approach as in Figure 6. In panel (a), the sample includes only years 1998-2001 and movers whose move-year is in that range ($N = 418,788$ patient-years); in panel (b) the sample includes only years 2002-2005 and movers whose move-year is in that range ($N = 637,041$ patient-years); in panel (c) the sample includes only years 2006-2008 and movers whose move-year is in that range ($N = 309,770$ patient-years). In panel (d) the sample includes movers who never die during the course of our study ($N = 2,505,640$ patient-years); in panel (e) the sample includes movers who never enter an HMO during the course of our study ($N = 2,938,784$ patient-years); in panel (f) the sample includes movers who are never missing an outcome for any reason (including death or entering an HMO) ($N = 1,718,427$ patient-years); in panel (g) the sample includes movers whose origin and destination state is known ($N = 3,700,363$ patient-years); in panel (h) the sample includes movers whose origin and destination HSA is known ($N = 3,702,189$ patient-years). In panel (i) the sample includes cross-state movers ($N = 2,514,160$ patient-years); in panel (j) the sample includes cross-census region movers ($N = 1,385,419$ patient-years).

Online Appendix Figure 10: Event-Study Analysis of Level Utilization and Percentile of Utilization



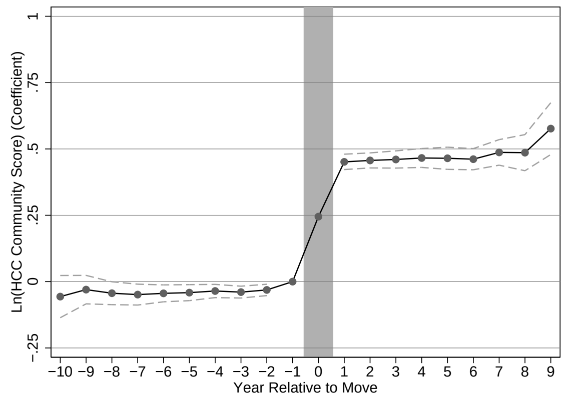
(a) Level Utilization



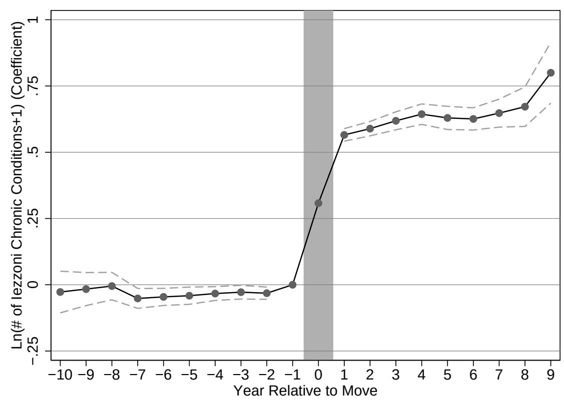
(b) Percentile of Utilization

Notes: These figures are constructed in the same manner as Figure 6 , except in panel (a) the dependent variable is level utilization and in panel (b) the dependent variable is percentile of utilization. The dashed lines show the 95% confidence interval, constructed using the same bootstrap approach as in Figure 6. The sample is all movers ($N = 3,702,189$ patient-years).

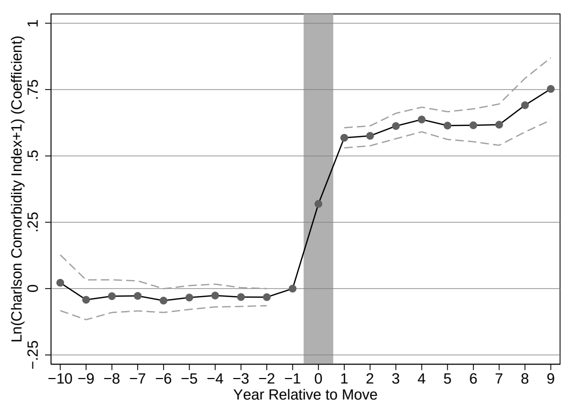
Online Appendix Figure 11: Event Study Results for Health Measures



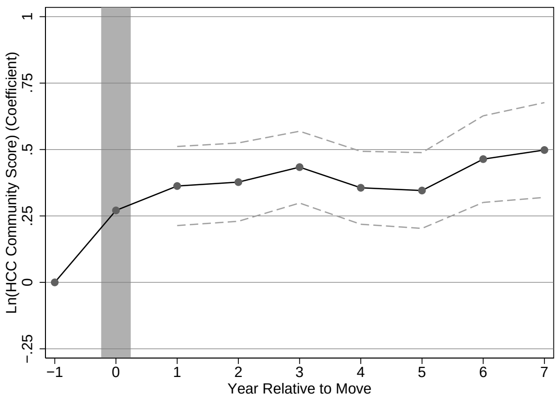
(a) Log HCC Score. All Moves



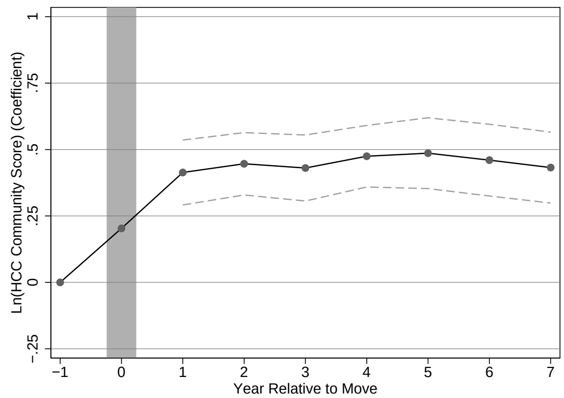
(b) Log Iezzoni Chronic Conditions. All Moves



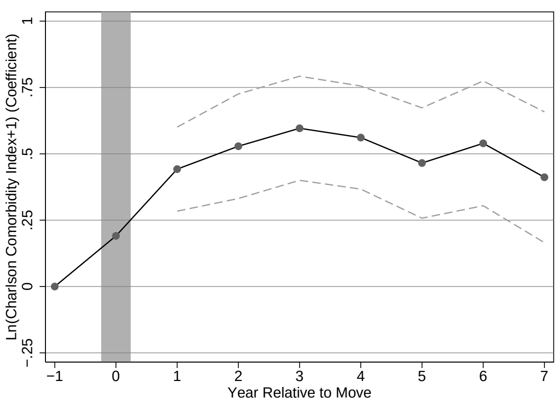
(c) Log Charlson Comorbidity Index. All Moves



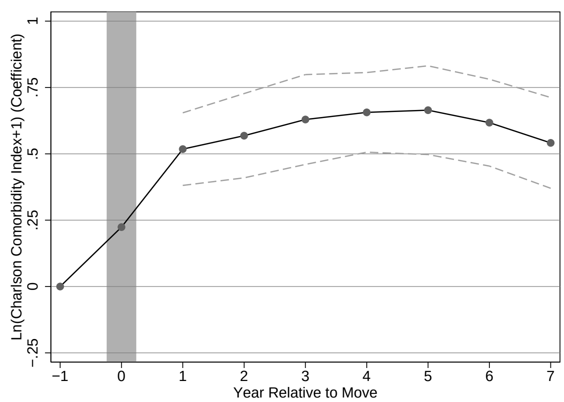
(d) Log HCC Score. Early Moves. Moves Up.



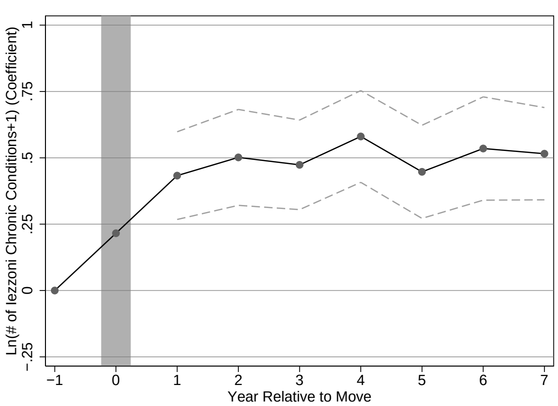
(e) Log HCC Score. Early Moves. Moves Down.



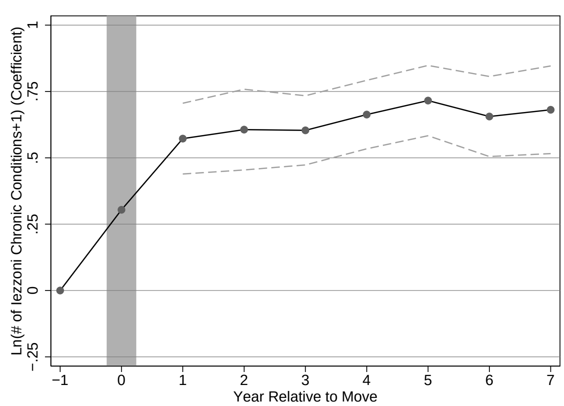
(f) Log Charlson Comorbidity Index. Early Moves. Moves Up.



(g) Log Charlson Comorbidity Index. Early Moves. Moves Down.



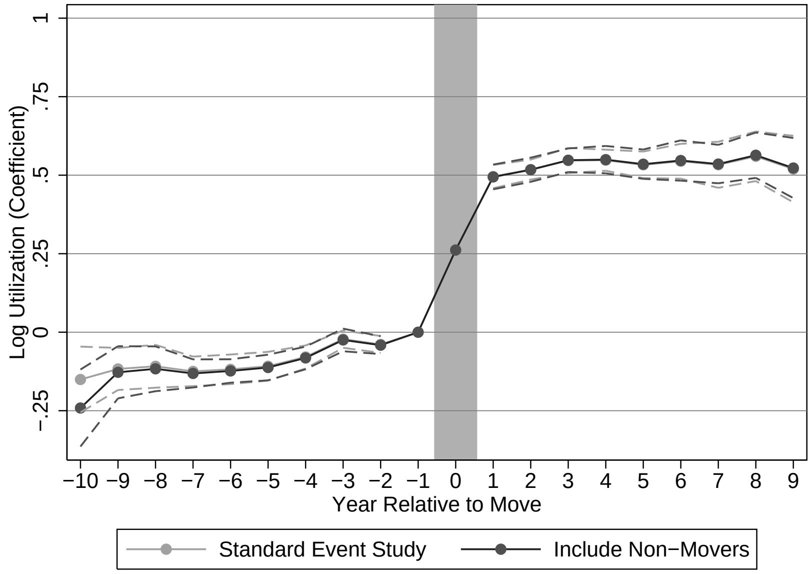
(h) Log Iezzoni Chronic Conditions. Early Moves. Moves Up.



(i) Log Iezzoni Chronic Conditions. Early Moves. Moves Down.

Notes: These figures are constructed in the same manner as Figure 6, except the dependent variables are various health measures and they are estimated on balanced-panel subsamples of movers whom we observe in each of a given set of relative years. The dashed lines show the 95% confidence interval, constructed using the same bootstrap approach as in Figure 6. All log outcome measures are the log of the outcome plus 1, except the HCC score which is bounded away from 0. Online Appendix Table 11 shows the percent with zero for each of these outcomes. In panels (a)-(c) the sample is all movers ($N = 3,702,189$ patient-years). In panels (d), (f), and (h), the sample is movers whom we observe in every relative year in [-1,7] and who move to higher utilization areas ($N = 212,958$ patient-years). In panels (e), (g), and (i), the sample is movers whom we observe in every relative year in [-1,7] and who move to lower utilization areas ($N = 209,268$ patient-years).

Online Appendix Figure 12: Event-Study with Non-Movers Included



Notes: The event study including non-movers is shown superimposed over the standard event study from Figure 6. The event study with non-movers is constructed in the same manner as Figure 6 except for non-movers we adapt equation 4, setting δ_i to zero and $o(i)$ to patient i's area of residence. This yields an event study equation similar to equation 5, with δ_i equal to zero for nonmovers. The dashed lines show the 95% confidence interval, constructed using the same bootstrap approach as in Figure 6. In the standard event study, the sample is all movers ($N = 3,702,189$ patient-years). When we include nonmovers, the sample is all movers and non-movers ($N = 16,432,955$ patient-years).

Online Appendix Table 1: Movements Between Census Divisions (Movers Only, as Percentage of All Moves)

		Destination									
		ENC	ESC	M-A	M	NE	P	SA	WNC	WSC	Total
Origin	East North Central	7.00	0.94	0.32	0.91	0.14	0.65	2.53	0.63	0.73	13.86
	East South Central	0.65	1.56	0.10	0.14	0.04	0.15	1.13	0.14	0.49	4.40
	Mid-Atlantic	0.58	0.28	6.26	0.53	0.92	0.56	5.57	0.15	0.36	15.21
	Mountain	0.62	0.17	0.20	2.64	0.10	1.58	0.57	0.65	0.75	7.29
	New England	0.15	0.08	0.40	0.19	1.56	0.21	1.54	0.05	0.10	4.26
	Pacific	0.53	0.28	0.26	2.74	0.14	9.04	0.93	0.51	1.00	15.43
	South Atlantic	2.51	1.59	2.71	0.77	1.06	0.84	13.12	0.50	1.00	24.10
	West North Central	0.52	0.15	0.07	0.63	0.04	0.36	0.47	2.44	0.66	5.35
	West South Central	0.54	0.55	0.14	0.65	0.07	0.59	0.85	0.62	6.08	10.10
	Total	13.12	5.60	10.47	9.21	4.08	13.97	26.73	5.69	11.18	100.00

Notes: Table shows the percentage of moves that take place between each of the 81 origin-destination pairs of census divisions. For example, 2.7% of all moves originate from the South Atlantic census division and have their destination in the Mid-Atlantic census division. The denominator is all movers ($N = 497,097$ patients).

Online Appendix Table 2: HRS Summary Statistics

	(1)	(2)
	Non-movers	Movers
Average age over observed waves	74.93	77.03
Average age in first observed wave	70.65	72.26
Average age in 1992	64.60	68.23
Female	0.55	0.60
White	0.81	0.89
Hispanic[a]	0.08	0.04
Education		
Less than high school	0.33	0.26
GED	0.04	0.04
High school	0.30	0.31
Some college	0.18	0.20
College	0.15	0.19
Retirement status (in first observed wave)[b]		
Retired	0.51	0.51
Partly retired	0.16	0.15
Retired or partly retired	0.66	0.66
Earnings (in first observed wave)[c]		
Average	$5,803	$5,115
Average conditional on positive	$22,641	$21,505
Median conditional on positive	$13,000	$13,000
Share with zero	0.67	0.79
Marital status (in first observed wave)		
Married or partnered	0.66	0.63
Separated or divorced	0.08	0.09
Widowed	0.23	0.27
Never married	0.03	0.01
# of patients	20,998	2,025
# of patient-years	83,202	11,068

[a]There are three race categories in the data: White/Caucasian, Black/African-American, and Other. Hispanic is a separate variable, so someone can appear as White and Hispanic, Black and Hispanic, White and Non-Hispanic, etc.

[b]Respondents are asked whether they consider themselves retired. They can have the following non-missing responses: not retired, completely retired, and partly retired.

[c]The sum of respondent's wage/salary income, bonuses/overtime pay/commissions/tips, second job or military reserve earnings, professional practice or trade income.

Online Appendix Table 3: HRS Summary of Regression Results

Row	X_{it}	(1) # patient-waves	(2) % moves in the sample	(3) Estimated coefficient	(standard error)
(1)	Unmarried & unpartnered	81,613	0.0192	0.0122	(0.0026)
(2)	Separated/divorced	81,613	0.0192	0.0033	(0.0050)
(3)	Widowed	81,613	0.0192	0.0094	(0.0025)
(4)	Retired or partly retired	66,472	0.0204	0.0035	(0.0019)
(5)	Poor/fair Health	94,270	0.0215	-0.0021	(0.0015)

Notes: Table shows the coefficients and standard errors from estimating the regression in Online Appendix equation (1) for each of the indicator variables in the rows of the table. For each row, column (1) shows the number of patient-waves for which the indicator variable is not missing. Column (2) shows the fraction of patient-waves counted in column (1) in which a move happens. Column (3) shows results from estimating a linear regression with wave fixed effects and person fixed effects. Because of missing data, rows have different sample sizes (column 1) and slightly different percentages moving (column 2).

Online Appendix Table 4: Additive Decomposition (Estimation in Logs, Decomposition in Levels)

	(1) Above / below median	(2) Top & bottom 25%	(3) Top & bottom 10%	(4) Top & bottom 5%	(5) McAllen & El Paso	(6) Miami & Minneapolis
Difference in mean predicted utilization						
Overall	807	1,294	1,768	2,074	2,053	3,692
If we equalize patient effects	592	898	1,159	1,319	1,340	1,790
If we equalize place effects	228	410	605	749	551	1,493
$\hat{S}^{level}_{pat}(R,R')$	0.267	0.306	0.345	0.364	0.347	0.515
	(0.041)	(0.036)	(0.041)	(0.045)	(0.214)	(0.054)
$\hat{S}^{level}_{place}(R,R')$	0.718	0.683	0.658	0.639	0.732	0.596
	(0.041)	(0.038)	(0.045)	(0.049)	(0.183)	(0.060)

Notes: Table based on estimation of equation (1); the estimates are transformed based on Online Appendix equations (2), (3), and (4). The measures reported in this table are explained in Online Appendix Section 2.2. Each column defines a set of areas R and R'. The first row reports the difference in average predicted level utilization overall between the two sets ($\bar{\hat{y}}_R - \bar{\hat{y}}_{R'}$); the second row reports the difference in predicted level utilization that would remain if we counterfactually equalized patient effects ($\bar{\hat{y}}_{R,pateq} - \bar{\hat{y}}_{R',pateq}$); the third row reports the difference in predicted level utilization that would remain if we counterfactually equalized place effects ($\bar{\hat{y}}_{R,pleq} - \bar{\hat{y}}_{R',pleq}$). The fourth row reports the share of difference in predicted level utilization due to patients. The fifth row reports the share of the difference in predicted level utilization due to place. Standard errors (in parentheses) are calculated using a bootstrap procedure with 50 repetitions at the patient level. In columns (1)-(4), the partitions of places shown in the columns are defined based on average utilization in each HRR. The sample size is the same as in Table 2.

Online Appendix Table 5: Alternative Definitions of Movers

	Mover definition	(1) Mean of log utilization	(2) Above / below median utilization difference	(3) Share due to patients	(4) N (% of movers retained)	
(1)	Baseline	7.193	0.283	0.465	16,031,875	(100%)
(2)	Looser claim share criterion (0.6)	7.195	0.282	0.489	16,250,644	(106%)
(3)	Stricter claim share criterion (0.9)	7.190	0.284	0.443	15,659,995	(90.5%)
(4)	No claim share criterion	7.199	0.282	0.592	17,792,048	(153%)
(5)	Baseline, adjusted for move timing measurement error	7.193	0.283	0.444	16,031,875	(100%)

Notes: Table reports the share of the difference in utilization between above and below median HRRs due to patients, analogous to column (1) of Table 2, for alternative samples and specifications. Columns report the mean of log utilization, difference in average utilization between above and below median utilization HRRs , patient share ($\hat{S}_{pat}\left(R,R'\right)$), and sample size and percent of movers retained. Row (1) repeats our baseline results.

Row (2): We modify the baseline definition to categorize someone as a mover if their HRR of residence changes and their average claim share in the destination HRR increases by at least 0.6 instead of 0.75; the remainder of the definition remains unchanged.

Row (3): We modify the baseline definition to categorize someone as a mover if their HRR of residence changes and their average claim share in the destination HRR increases by at least 0.9 instead of 0.75; the remainder of the definition remains unchanged.

Row (4): We categorize someone as a mover if their HRR of residence changes.

Row (5): Same as Row (1), but we adjust for measurement error in move timing by estimating Online Appendix equation (7).

Online Appendix Table 6: Narrow Window Specifications for Components of Utilization

	Utilization Measure	(1) Share due to patients	(2) Share due to Patients, Relative Years -1 to 1
(1)	Baseline: Log(utilization)	0.465	0.557
(2)	Seen a primary care physician	0.452	0.547
(3)	Seen a specialist	0.322	0.390
(4)	Any hospitalization	0.410	0.376
(5)	Any emergency room visit	0.714	0.679
(6)	Log (# of diagnostic tests)	0.092	0.129
(7)	Log(# of imaging tests)	0.142	0.176
(8)	Log(# of preventive care measures)[a]	0.611	0.652
(9)	Log(# of different doctors seen)	0.392	0.467
(10)	Log(inpatient utilization)[b]	0.242	0.195
(11)	Log(outpatient utilization)[b]	0.358	0.406
(12)	Log(emergency room utilization)[b]	0.639	0.662
(13)	Log(other utilization)[b]	0.124	0.145

Notes: Table reports the share of the difference in utilization between above and below median HRRs due to patients, analogous to column (1) of Table 2, for alternative samples and specifications. Column (1) reports the share of the difference in the average utilization measure between above and below median HRRs that is due to patients ($\hat{S}_{pat}(R,R')$). Column (2) reports this same share, but narrows the sample of movers used for estimation to relative years -1 and +1. All log outcome measures are the log of the outcome plus 1. The partition of HRRs into above and below median groups is based on the utilization of individuals in the baseline sample and differs in each row according to the definition of utilization used; it is computed separately for each column. Online Appendix Table 11 shows the percent with zero for each of these outcomes. The sample size is the same as in Table 2 in column (1). In column (2), the sample is all non-movers and mover relative years -1 and +1 ($N = 13,511,698$ patient-years).

[a]"# of preventive care measures" includes how many of the following treatments the patient received in the past year: Ambulatory Care, Eye Screening, Hemoglobin Test, Lipid Screen, Cardio Screen, Diabetes Management, Pelvic Screen, Bone Mass Test, Colorectal Cancer Screening, Flu Shot, or in the past two years: Mammogram, Pap Test, Prostate Cancer Screening.

Online Appendix Table 7: Alternative Measures of Utilization

	Outcome	(1) Mean of outcome	(2) Above / below median difference in outcome	(3) Share due to patients
(1)	Utilization (in levels)	6629.120	1231.389	0.228
(2)	Percentile in national distribution	49.791	3.694	0.295
(3)	In top 80% of utilization	0.791	0.048	0.505
(4)	In top 50% of utilization	0.484	0.059	0.252
(5)	In top 20% of utilization	0.195	0.032	0.317
(6)	In top 10% of utilization	0.097	0.022	0.165
(7)	In top 5% of utilization	0.048	0.014	0.225

Notes: Table reports the share of the difference in utilization between above and below median HRRs due to patients, analogous to column (1) of Table 2, for alternative samples and specifications. Columns report the mean of log utilization, difference in average utilization between above and below median utilization HRRs ($\hat{y}_R - \hat{y}_{R'}$) and patient share ($\hat{S}_{pat}\,(R,R')$). In row (1), the outcome is level utilization. In row (2), the outcome is the percentile of the national distribution of utilization that a patient is in. In rows (3) to (7), the outcome is an indicator for being in the top 80%, 50%, 20%, 10%, and 5%, respectively, of the national distribution of utilization. The partition of HRRs into above and below median groups is based on the utilization of individuals in the baseline sample and differs in each row according to the definition of utilization used. The sample size is the same as in Table 2.

Online Appendix Table 8: Additional Robustness Checks

	Specification	(1) N	(2) Mean of log utilization	(3) Above / below median utilization difference	(4) Share due to patients
(1)	Baseline	16,031,875	7.193	0.283	0.465
(2)	Movers only	3,301,109	7.252	0.287	0.481
(3)	Drop age as a covariate	16,031,875	7.193	0.283	0.446
(4)	Drop relative year as a covariate	16,031,875	7.193	0.283	0.485
(5)	Log(total expenditure)	16,031,875	7.156	0.291	0.453
(6)	Log(utilization+0.1)	16,031,875	7.062	0.326	0.508
(7)	Log(utilization+10)	16,031,875	7.339	0.243	0.360
(8)	Drop moves to Florida	15,640,033	7.193	0.282	0.447
(9)	Drop moves to Florida, Arizona, and California	15,250,903	7.192	0.283	0.428
(10)	Early moves	13,106,078	7.190	0.284	0.453
(11)	Middle moves	13,152,510	7.190	0.284	0.456
(12)	Late moves	13,214,518	7.190	0.284	0.532

Notes: Table reports the share of the difference in utilization between above and below median HRRs due to patients, analogous to column (1) of Table 2, for alternative samples and specifications. Columns report the sample size, mean of log utilization, difference in average utilization between above and below median utilization HRRs ($\hat{y}_R - \hat{y}_{R'}$) and patient share ($\hat{S}_{pat}(R, R')$). Row (2) only uses movers when estimating equation (1). Row (3) drops age as a covariate when estimating equation (1) and row (4) drops relative year when estimating equation (1). Rows (5) to (7) change the outcome variable to the log of total expenditures, the log of utilization + 0.1, and the log of utilization + 10, respectively. Rows (8) and (9) drop moves to Florida and to Florida, Arizona, and California, respectively. Row (10) includes movers only if they are observed continuously in each relative year in [-1,7] and all non-movers; row (11) includes movers only if they are observed continuously in each relative year [-4,4] and all non-movers; row (12) includes movers only if they are observed continuously in each relative year [-7,1] and all non-movers.

Online Appendix Table 9: Results by Age Quartile

		(1) Mean of log utilization	(2) Above / below median utilization difference	(3) Share due to patients	(4) N
(1)	Baseline	7.193	0.283	0.465	16,031,875
	Age quartiles				
(2)	First	6.674	0.362	0.552	4,295,051
(3)	Second	7.152	0.307	0.494	3,859,244
(4)	Third	7.442	0.263	0.454	3,980,690
(5)	Fourth	7.585	0.271	0.358	3,896,890

Notes: Table reports the share of the difference in utilization between above and below median HRRs due to patients, analogous to column (1) of Table 2, for alternative samples and specifications. Columns report the mean of log utilization, difference in average utilization between above and below median utilization HRRs ($\hat{y}_R - \hat{y}_{R'}$), patient share ($\hat{S}_{pat}(R,R')$), and sample size. In rows (2) to (5), stratification groups were determined by patient-level averages of age over all observed years. Row (2) provides estimates for the first quartile of age (mean age 68.1), row (3) provides estimates for the second quartile of age (mean age 72.8), row (4) provides estimates for the third quartile of age (mean age 78.3), and row (5) provides estimates for the fourth quartile of age (mean age 86.3). The partition of HRRs into above and below median utilization is based on the age quartile sample specified in the row.

Online Appendix Table 10: Additive Decomposition of Health Outcomes

	Outcome	(1) Mean of Outcome	(2) Difference in average outcome: Overall	(3) Difference in average outcome: Due to patients	(4) Difference in average outcome: Due to place	(5) Share of difference due to patients	(6) Share of difference due to place
(1)	Log(HCC score)	-0.132	0.115	0.056	0.059	0.485	0.515
(2)	Log(Charlson Comorbidity Index)	0.672	0.102	0.044	0.058	0.430	0.570
(3)	Log(Iezzoni chronic conditions)	0.495	0.100	0.042	0.058	0.418	0.582
(4)	Log(# of chronic conditions)	1.181	0.138	0.059	0.079	0.430	0.570

Notes: Table reports the share of the difference in utilization between above and below median HRRs due to patients and places, analogous to column (1) of Table 2, for alternative dependent variables. Columns report the mean of the dependent variable, difference in the average outcome between above and below median utilization HRRs, the difference due to place ($\hat{\gamma}_R - \hat{\gamma}_{R'}$), the difference due to patients ($\hat{c}_R - \hat{c}_{R'}$), the patient share ($\hat{S}_{pat}(R,R')$), and the place share ($\hat{S}_{place}(R,R')$). HRRs are partitioned into above and below median places based on the average of the given outcome (by row) in each HRR. All log outcome measures are the log of the outcome plus 1, except the HCC score which is bounded away from 0. Online Appendix Table 11 shows the percent with zero for each of these outcomes. The sample size is the same as in Table 2 in rows (1)-(3). In row (4), the sample also excludes the year 1998, as chronic conditions are not observed in that year ($N = 14,598,443$ patient-years).

Online Appendix Table 11: Share Zero for Components of Utilization and Health Measures

		(1) Non-Movers	(2) Movers
(1)	# of diagnostic tests	0.33	0.30
(2)	# of imaging tests	0.45	0.42
(3)	# of preventive care measures	0.10	0.09
(4)	# of different doctors seen	0.07	0.07
(5)	Inpatient utilization	0.75	0.74
(6)	Outpatient utilization	0.07	0.07
(7)	Emergency room utilization	0.71	0.69
(8)	Other utilization	0.35	0.31
(9)	Charlson Comorbity Index	0.16	0.14
(10)	# of Iezzoni Chronic Conditions	0.41	0.40
(11)	# of Chronic Conditions	0.49	0.47
	# of patient-years	12,730,766	3,702,189

Notes: We report the share of patient-years for which the utilization component or health measure has a value of zero. We use the log of these measures plus 1 in Tables 4 and 7, Online Appendix Tables 6 and 10, and Online Appendix Figures 8 and 11. The sample is all movers and nonmovers ($N = 16,432,955$ patient-years).