Sources of Geographic Variation in Health Care:
Evidence from Patient Migration

Amy Finkelstein, *MIT and NBER*

Matthew Gentzkow, *Chicago Booth and NBER*

Heidi Williams, *MIT and NBER*[*]

April 2014

PRELIMINARY. Do not cite or circulate. Comments welcome.

Abstract

We analyze the role of demand and supply factors in explaining geographic variation in US healthcare utilization, using an empirical strategy that exploits migration of Medicare patients. Our approach allows us to account for demand differences driven by both observable and unobservable patient characteristics, and to separate the role of patient health from place-specific diagnosis propensities. We find that roughly 40% of geographic variation is attributable to patient demand, with the remainder due to supply-side factors. Demand differences do not appear to result from differences in past treatment, and are explained to a significant degree by differences in patient health.

Keywords: Healthcare spending, regional variation, Dartmouth Atlas

JEL: H51, I1, I11

[*]We are grateful to Lizi Chen, Grant Graziani, Sara Kwasnick, Tamar Oostrom, Daniel Prinz, and Tony Zhang for excellent research assistance.

1 Introduction

Health care utilization varies widely across the United States (Fisher et al., 2003a, 2003b). Adjusting for regional differences in age, sex, and race, the average Medicare enrollee in Miami, FL consumed $14,423 of care in 2010, while the average enrollee in Minneapolis, MN consumed just $7,819. The average enrollee in McAllen, TX consumed $13,648, compared to $8,714 in nearby and demographically similar El Paso, TX.[1] Similar variation is observed in the frequency of specific treatments (Chandra et al., 2012) and in measures of total health care utilization that adjust for regional variation in prices (Gottlieb et al., 2010). High utilization is not generally correlated with better patient outcomes.[2]

Understanding what drives geographic variation has first-order implications for policy. If high-spending areas like McAllen and Miami are different mainly because their doctors' incentives or beliefs lead them to order excessive treatments with low return, policies that change those incentives or beliefs could result in savings on the order of several percentage points of GDP (Congressional Budget Office 2008; Gawande 2009; Skinner 2012). If, on the other hand, patients in high-spending areas are simply sicker and efficiently demand a high level of care, policies aimed at eliminating geographic variation could be ineffective or counter-productive (Zuckerman et al., 2010).

In this paper, we use an empirical strategy that exploits patient migration to separate variation due to patient characteristics such as health or preferences from variation due to place-specific variables such as doctors' incentives and beliefs, endowments of physical capital, and characteristics of hospital markets. As a shorthand, we refer to the former as "demand" factors and the latter as "supply" factors. To see the intuition for our approach, imagine a patient who moves from high-spending Miami to low-spending Minneapolis. If all of the spending difference between these cities arises from supply-side differences like doctor incentives, we would expect the migrant's spending to drop immediately following the move, to a level similar to other patients of the low-spending doctors in Minneapolis. If all of the spending difference reflects the demand-side reality

[1]Authors' tabulations based on total Medicare Parts A and B reimbursements per enrollee, from Dartmouth Atlas of Health Care, http://www.dartmouthatlas.org/downloads/tables/pa_reimb_hrr_2010.xls.

[2]See Skinner (2012) for an extensive discussion. The Congressional Budget Office (2008) concludes that high-spending areas "tend to score no better and, in some cases, score worse than other areas do on process-based measures of quality and on some measures of health outcomes," and that more intensive treatment in high-spending areas "appear[s] to improve health outcomes for some types of patients, but worsen outcomes for others."

that residents of Miami are sicker, we would expect the migrant's spending to remain constant after the move, at a level similar to the typical person in Miami. Where the observed spending change falls between these two extremes identifies the relative importance of demand and supply factors.

We implement this strategy using claims data for a 20 percent sample of Medicare beneficiaries from 1998 to 2008.[3] In our baseline model, a patient's average annual healthcare utilization arises from a combination of a patient fixed effect, a location fixed effect, and a vector of time-varying controls. Our main outcome measure is total utilization adjusted for price differences as in Gottlieb et al. (2010). The model allows for the possibility that migrants have systematically different utilization levels from non-migrants, and that these levels are correlated with the migrant's origin and destination region. It also allows for arbitrary differences in utilization trends of migrants relative to non-migrants. The key identifying assumption is that such differential trends do not vary systematically with the migrant's origin and destination.

We begin with a descriptive event-study analysis of changes in utilization around moves. We observe sharp changes in the year of a move, equal to about 60 percent of the utilization difference between origin and destination. There is little systematic trend in utilization pre-move, and no systematic adjustment in utilization post-move. The on-impact effect is similar for moves from low-to-high and high-to-low utilization regions, and for moves where the absolute value of the origin-destination difference in utilization is either small or large.

Our estimated model exploits this variation to infer that about 40 percent of the difference in average utilization between above- and below-median utilization areas is due to demand-side factors. The share is similar for differences between the top and bottom quartiles, deciles, or ventiles. The share is similar when we isolate differences between the the very highest-spending areas, such as McAllen, TX and Miami, FL, and the very lowest-spending areas, such as El Paso, TX or Minneapolis, MN.

[3]We follow the predominant tendency of the literature on geographic variation in health care utilization and focus on health care utilization among Medicare beneficiaries. The Medicare setting is appealing due to the availability of high quality, rich data on large numbers of Medicare beneficiaries, and the relatively geographically uniform insurance environment. The literature has explored - to the extent feasible given existing data - whether the geographic variation in health care utilization that has been extensively documented in the Medicare setting exists in other settings as well. Regional variation appears to be the norm; geographic variation has been documented in the US Veterans Affairs system (Ashton et al., 1999, CBO, 2008, Subramanian et al., 2002), in private insurance markets (Baker et al., 2008, Chernew et al., 2010, Dunn et al., 2011, Philipson et al., 2010, Rettenmaier and Saving, 2009), and in other countries including the UK and Canada (McPherson et al., 1981), although the magnitudes of this variation and the correlation with Medicare variation is the subject of some debate in these studies.

An important question in interpreting these findings is to what extent differences in patients' demand for care today result from differences in the care they consumed in the past. In other words, to what extent is health utilization subject to habit formation in the sense of Becker and Murphy (1988)? If patients in Miami are diagnosed with more conditions than patients elsewhere, for example, they may demand continued treatment for these conditions even after moving to Minneapolis. Our estimates in this case would still be an accurate decomposition of current variation, but it would be important in interpreting them that the demand-side component today could ultimately be caused by supply-side factors in the past.

We present several pieces of evidence that suggest habit-formation is limited, at least over the 10-year horizon of our data. First, as already noted, we observe similar changes for patients moving from high-to-low utilization areas as low-to-high. Stickiness arising from the difficulty of stopping treatments once begun would suggest more flexibility upwards than downwards. Second, we find that patients with high average utilization, in poor health, and who are older actually change their utilization more when they move. This also argues against persistence driven by past diagnoses, as well as many habit models that would predict older patients have a larger stock of past consumption and so would change by less. Finally, we find no evidence that utilization continues to adjust in the years after a move. Such adjustment is a robust prediction of habit models, and is the key feature of the data that identifies empirical studies of habit formation such as Bronnenberg et al. (2012).

A second important question is how much of the demand-side effect is due to differences in patient health, as opposed to other patient characteristics such as preferences or information. Separating the two would be straightforward if the exogenous component of patient health were observable. The challenge is that observable health measures such as risk scores or indicators for chronic conditions are based on diagnoses that are themselves endogenous to supply-side factors (Song et al., 2010). If areas with high patient fixed effects also have doctors who diagnose more aggressively, a naive estimate would exaggerate the role of patient health. To address this, we extend our empirical strategy using changes in the observable health measures around moves to separate the exogenous (patient) and endogenous (place) components of health. We then examine the role of these different components of health in explaining geographic variation in utilization.

Consistent with (Song et al., 2010) our examination of movers suggests the effect of endogenous diagnosis is quite large. Once we adjust for this, we find that about 35 percent of the difference in average utilization between above- and below-median utilization areas can be explained by the patient component of health (independent of the place-specific diagnostic propensity). The remaining role for patients presumably reflects other factors such as unmeasured patient risk factors, information or preferences. We also find that about 40 percent of the difference in average utilization between above- and below-median utilization areas can be explained by the "endogenous" component of health, i.e., the place-specific diagnostic propensity.

Our work contributes to a large existing literature seeking to separate the role of demand-side and supply-side factors in driving geographic variation. All of these studies infer the role of demand-side factors from the explanatory power of patient observables. They find little role for patient demographics or other exogenous characteristics. Health status measures explain somewhat more variation (Zuckerman et al., 2010), but the interpretation of this is unclear given the problem of endogenous diagnosis (Song et al., 2010). Overall, the tentative conclusion has been that the role of patients is limited, and that most of the variation likely originates on the supply side. A review by Chandra et al. (2012), for example, concludes: "In general, the literature points to the importance of supply-side incentives over demand-side factors in driving treatment choice" (p. 425) and "most of the literature agrees that patient characteristics and preferences do not explain much of the differences across areas" (p. 402).[4]

Our strategy has two important advantages relative to this literature. First, we can capture the effect of both observed and unobserved patient characteristics, since both will be reflected in persistence of utilization following a move. Second, our approach correctly incorporates variation arising from differences in patient health, even if observable measures of these factors are endogenous to supply-side differences in diagnosis rates, and in fact allows us to separate the effect of health from other patient characteristics.

Our bottom-line conclusion that patients account for 40 percent of geographic variation suggests an upward revision relative to the conventional wisdom. Moreover, the lack of evidence of any role for habit formation, in which prior health care utilization affects subsequent patient de-

[4]This is a typical characterization of the existing literature. For example, in his recent handbook chapter reviewing the literature on regional variations in health care, Skinner (2012) writes "It seems likely... that regional practice norms trump patient preferences." (p. 74) Likewise, the Congressional Budget Office (2008) concludes that "Income and the preferences of individuals for specific types of care appear to explain little of the variation in spending." (p. 1) though they note that there is less consensus about the role of variation in health status per se. A recent study by Cutler et al. (2013) confirms this impression, finding "Patient demand is relatively unimportant in explaining variations. Physician organizational factors... may matter, but the single most important factor is physician beliefs about treatment."

mand, suggests that patient demand may be relatively “fixed”, at least in the 65+ population (which accounts for about a third of total health care spending) over our approximately 10 year study period. These findings imply potential limits to the impact of supply-side policies aimed at reducing geographic variation in health care spending.

Like past decompositions, ours is not sufficient to draw conclusions about the efficiency of observed geographic variation. A natural interpretation might be that geographic variation driven by patient demand may likely be efficient, but the variation driven by supply may not be. However, supply-driven heterogeneity could reflect different allocations of physical or human capital, and so be consistent with efficiency (Chandra and Staiger, 2007). Demand-driven heterogeneity could reflect patient misinformation, and so contribute to inefficiency. We present suggestive evidence from the cross section that moves to higher utilization areas are associated with reductions in mortality, but caution that this evidence is, by design, less compelling than the panel analysis in the rest of the paper. We view our findings as both a first step toward a more welfare-relevant understanding and a clarification of an influential body of existing evidence.

Our work is also related to the broader literature on geographic variation in health practice, including the original “Dartmouth Atlas” studies of Fisher et al. (2003a, 2003b), and subsequent work reviewed by Skinner (2012) and Chandra et al. (2012). Most closely related are two studies that use migration as a key source of identification. As previously mentioned,Song et al. (2010) exploit patient migration to estimate the extent to which cross-region differences in Medicare risk scores are confounded by variation in diagnosis rates. Molitor (2011) looks at physician migration, estimating how cardiologist behavior changes around moves. Ours is the first study to exploit migration to decompose aggregate variation in health spending across regions.

Finally, our paper relates to work outside of the health care sector that uses changes in residence or employment to separate effects of individual characteristics from geographic or institutional factors. A number of papers beginning with Abowd et al. (1999) use matched worker-firm data to separately identify worker and firm fixed effects. In this vein, we draw especially on Card et al.’s (2013) recent study of German workers and firms. A second strand of work including Fernandez and Fogli (2006) and Luttmer and Singhal (2011) uses data on immigrants to identify effects of culture. Early work includes Aaronson’s (1998) study of neighborhood effects on children and Ichino and Maggi’s (2000) study of shirking at an Italian bank. Recently, researchers have used moves to identify determinants of brand preferences (Bronnenberg et al., 2012), tax reporting (Chetty et al. 2013b), teacher value added (Chetty et al. 2013a), and retirement savings (Chetty et al., 2012).

Section 2 introduces our model and estimation strategy. Section 3 describes our data and presents summary statistics. Section 4 presents descriptive evidence on the effects of moving across different areas. Section 5 presents our main analysis of the role of patient vs. place in explaining geographic variation in health care utilization; it also explores the validity of our identifying assumptions and the robustness of our finding. Section 6 explores some potential mechanisms for the role of patients. Section 7 concludes.

2 Model and Identification

2.1 Model

We index patients by i, geographic areas by j, and years by t. Some patients are "non-movers" who live in one area throughout the sample, while others are "movers" whose area changes exactly once. For a mover i who moves during year t_i^*, we define the year relative to move to be $r(i,t) = t - t_i^*$. The outcome of interest is y, which in our main specifications will be the log of total healthcare utilization. We discuss the precise definition of these variables in the context of our data in section 3 below.

We assume y for a patient i who lives in area j throughout year t is given by:

$$y_{ijt} = \alpha_i + \rho_{r(i,t)} + x_{it}\beta + \gamma_j + \tau_t + \varepsilon_{ijt}. \tag{1}$$

Here α_i, γ_j, and τ_t are fixed effects for patient, area, and time respectively, and x_{it} is a vector of time-varying patient characteristics, which in our baseline specification is simply a series of dummies for five-year age bins. The term $\rho_{r(i,t)}$ is a fixed effect for movers in relative year $r(i,t)$, which we normalize to zero for non-movers. We do not model outcomes for movers in year t_i^*, when (as we show below) they spend part of the year in their origin area and part of the year in their destination; when we estimate equation (1), we omit these observations. We let $c_{it} = \alpha_i + \rho_{r(i,t)} + x_{it}\beta$ denote the combined effect of patient characteristics. Naturally, γ_j denotes the effect of place. We normalize $\mathrm{E}(\alpha) = 0$, $\tau_0 = 0$, and $\rho_0 = 0$, and we assume that the error term ε_{it} satisfies $\mathrm{E}\left(\varepsilon_{ijt} | \alpha_i, \rho_{r(i,t)}, x_{it}, \gamma_j, \tau_t\right) = 0$.

Our main goal is to decompose variation in average utilization across regions into a demand-side component attributable to patients and a supply-side component attributable to place. To define this decomposition formally, let $\overline{y}_{jt}$ denote the expectation of y_{it} across patients living in area j in year t, and let $\overline{y}_j$ denote the average of $\overline{y}_{jt}$ across t. Let $\overline{c}_{jt}$ and $\overline{c}_j$ denote the analogous expectation of c_{it}. Then the difference in average utilization between any two areas j and j' is the sum of the difference of a place component and the difference of a patient component: $\overline{y}_j - \overline{y}_{j'} = \left(\gamma_j - \gamma_{j'}\right) + \left(\overline{c}_j - \overline{c}_{j'}\right)$. When we talk about larger geographic regions R that consist of multiple areas j, we abuse notation by letting $\overline{y}_R$, $\overline{c}_R$, and $\overline{\gamma}_R$ denote the simple averages of $\overline{y}_j$, $\overline{c}_j$, and γ_j across areas in R.

We define the share of the difference between j and j' attributable to place to be

$$S_{place}\left(j, j'\right) = \frac{\left(\gamma_j - \gamma_{j'}\right)}{\left(\overline{y}_j - \overline{y}_{j'}\right)} \tag{2}$$

and we define the share attributable to patients to be

$$S_{pat}\left(j, j'\right) = 1 - S_{place}\left(j, j'\right).$$

Note that although $S_{pat}(j, j')$ and $S_{place}(j, j')$ mechanically sum to 1, neither need be between 0 and 1, since it is possible that $\left(\gamma_j - \gamma_{j'}\right)$ and $\left(\overline{c}_j - \overline{c}_{j'}\right)$ have opposite signs. We define $S_{pat}(R, R')$ and $S_{place}(R, R')$ to be the analogous shares for sets of of areas R and R'. [5]

We let $\hat{y}_j$ denote the sample analogue of $\overline{y}_j$. Given estimates $\hat{\gamma}_j$ of the γ_j, we form a consistent estimate $\hat{c}_j = \hat{y}_j - \hat{\gamma}_j$ of $\overline{c}_j$.

2.2 Identification

The model in equation (1) is only identified if the data include movers. If all patients were non-movers, there would be no way to separate differences in the area fixed effects γ_j from differences in the average patient characteristics $\overline{c}_j$. The key to separate identification of these two components

[5] That is, $S_{place}(R, R') = \left(\overline{\gamma}_R - \overline{\gamma}_{R'}\right) / \left(\overline{y}_R - \overline{y}_{R'}\right)$ and $S_{pat}(R, R') = 1 - S_{place}(R, R')$.

is the observed changes in utilization when patients move.[6]

To make this more precise, consider a simplified version of our model in which the τ_t, x_{it}, and $\rho_{r(i,t)}$ are all set to zero, and so utilization depends only on patient and place fixed effects. Suppose we observe a large number of patients who move from area j' to area j. Then the difference $\Delta_{j'}^{j}$ between their average y_{it} in the years after the move and the years before the move is a consistent estimator of $\left(\gamma_j - \gamma_{j'}\right)$. If we observe similar samples of patients moving between the other areas in the sample, along with the overall mean of utilization $\bar{y}$, we can form consistent estimates $\hat{\gamma}_j$ of each γ_j. The $\bar{c}_j$ would then be consistently estimated by $\hat{y}_j - \hat{\gamma}_j$.

Identification in the full model is similar. Identifying the τ_t and β is standard and does not rely on movers. Adding the $\rho_{r(i,t)}$ has a more substantial effect. It allows for arbitrary changes in utilization for movers pre- and post-move, with the restriction that these changes are the same regardless of the origin and destination. In the full model, therefore, observing only movers from j' to j is not enough to identify $\left(\gamma_j - \gamma_{j'}\right)$, because $\Delta_{j'}^{j}$ would also depend on the difference between the post-move and pre-move $\rho_{r(i,t)}$. Identification in this case comes from the differences in the changes across movers with different origins and destinations. If we have movers from j' to j and also movers from j to j', for example, we can estimate $\left(\gamma_j - \gamma_{j'}\right)$ consistently as $\left(\Delta_{j'}^{j} - \Delta_{j}^{j'}\right)/2$.

Importantly, our model permits movers to differ arbitrarily from non-movers in both levels of utilization (via the α_i) and trends in utilization around their moves (via the $\rho_{r(i,t)}$). We can in principle allow substantially more flexibility, including different calendar year effects τ_t for movers, area or individual-specific trends, and interactions between τ_t or $\rho_{t(i,t)}$ and characteristics of a mover's origin and destination area. We can also add flexibility by using data for movers only in the years just before or after their move, in the spirit of a regression discontinuity. We explore robustness to specifications along these lines below.

Our model is nevertheless restrictive in several important ways. First, we cannot allow for shocks to utilization that coincide exactly with the timing of the move and that are correlated with spending in the origin and destination. In the example above, suppose that for movers from j' to j the conditional expectation of ε_{ijt} in years just after the move is strictly greater than for movers from j to j'. This would inflate $\Delta_{j'}^{j}$ relative to $\Delta_{j}^{j'}$, and lead $\left(\Delta_{j'}^{j} - \Delta_{j}^{j'}\right)/2$ to be an over-estimate

[6] A sufficient condition for identification is that the number of movers between any pair of areas j and j' grows large as the total sample size approaches infinity. Abowd et al. (2002) discuss weaker conditions for identification.

of $(\gamma_j - \gamma_{j'})$. As a concrete example, this could occur if a subset of movers move in order to seek intensive treatment in their destination, and they are differentially likely to move to relatively high or low-spending areas. Evidence discussed below argues against this specific story, but we cannot in general rule out such correlated shocks.

Second, our model assumes that α_i and γ_j are additively separable. It is easy to imagine violations of this assumption. For example, if α_i primarily captures patient health, and γ_j primarily captures doctor incentives to order excessive treatments, it might well be that the effect of a given change in α_i on utilization is greater where γ_j is large than where it is small. We argue below that additive separability appears to be a reasonable approximation, but it is almost surely not literally correct and this should be borne in mind in interpreting the results.

Finally, our model does not allow for the possibility that α_i in a given period is a function of past values of y_{it}. If, for example, patients in high-utilization areas become accustomed to visiting the doctor frequently and receiving a large number of tests when they do, they might continue to demand these services post-move. In this case, variation across areas in current α_i could partly be caused by the influence of γ_j in the past. We discuss this possibility and evidence that bears on it at length below.

2.3 Event Study Representation

To visualize the way utilization changes when patients move, we define an alternative "event study" representation of equation (1).

To build intuition, it helps to start with the simple case where τ_t, x_{it}, and $\rho_{r(i,t)}$ are all set to zero and where our panel of movers is balanced in the sense that each mover is observed for the same number of years pre- and post-move.

Note that even in this case we cannot simply plot the average of y for movers by relative year $r(i,t)$. This plot would be informative if all movers had the same origin j' and destination j—in this case, we could estimate $(\gamma_j - \gamma_{j'})$ directly from the discontinuity in movers' y_{ijt} on move, and estimate $S_{place}(j,j')$ and $S_{pat}(j,j')$ by comparing this jump to $(\bar{y}_j - \bar{y}_{j'})$. In reality, however, there are a large number of areas and for any pair there are moves in both directions. If for any pair j' and j there are an equal number of movers going from j to j' as j' to j, we would expect the graph of average y_{ijt} across all movers to show *no* change around the move, since positive changes for some would be exactly offset by negative changes for others. Depending on the relative flows, the plot could show a jump up, no change, or a jump down, and the magnitude of the jump would be uninformative about the relative importance of patient and place.

To produce a more informative plot, we would like to scale y so that the direction and magnitude of the jump on move are informative regardless of the origin and destination. For a mover i whose origin and destination areas are $o(i)$ and $d(i)$ respectively, we denote by δ_i the difference in average utilization in the mover's destination and origin:

$$\delta_i = \overline{y}_{d(i)} - \overline{y}_{o(i)}. \tag{3}$$

Similarly, let $S^i_{place} = S_{place}(d(i), o(i))$ and $S^i_{pat} = S_{pat}(d(i), o(i))$. Following Bronnenberg et al. (2012), we could define for movers i:

$$y^{scaled}_{it} = \frac{y_{it} - \overline{y}_{o(i)}}{\delta_i}.$$

Note that y^{scaled}_{it} will be 0 if the mover's utilization is the same as the average utilization in his origin, 1 if it is the same as the average utilization in his destination, and in between if his utilization is partway between the average utilization in origin and destination. If the model is correct, the expectation of y^{scaled}_{it} should be flat both before and after move and the jump on move will be equal to the average value of S^i_{place}. Plotting the averages of y^{scaled}_{it} by relative year would thus produce an event study figure with a direct interpretation in terms of the model quantities of interest. The larger the jump in y^{scaled}_{it} on move, the greater the share of geographic variation we would attribute to place, and the smaller the share we would attribute to patients.

To implement this in the full model, we must deal with three additional complications. First, we need to allow for the controls τ_t, x_{it}, and $\rho_{r(i,t)}$. Second, our panel is not balanced and so changes in the composition of movers could introduce pre- or post-trends into the event study figure. To avoid this, we need to control for the individual fixed effects α_i explicitly. Third, the difference δ_i can be very small in some cases, which would make the simple average of y^{scaled}_{it} poorly behaved. This leads us to prefer a regression implementation that avoids dividing by δ_i.

Observe that we can rewrite equation (1) as:

$$y_{it} = \alpha_i + \gamma_{o(i)} + I_{r(i,t)>0} S^i_{place} \delta_i + \rho_{r(i,t)} + \tau_t + x_{it}\beta + \varepsilon_{it}, \tag{4}$$

where for non-movers we set δ_i to zero and $o(i)$ to the patient's area of residence. Combining $\alpha_i + \gamma_{o(i)}$ into a single patient fixed effect $\tilde{\alpha}_i$, replacing δ_i with its sample analogue $\hat{\delta}_i$, and parameterizing the interaction with $\left(\bar{y}_{d(i)} - \bar{y}_{o(i)}\right)$ as a flexible function of relative year yields

$$y_{it} = \tilde{\alpha}_i + \theta_{r(i,t)} \hat{\delta}_i + \rho_{r(i,t)} + \tau_t + x_{it}\beta + \varepsilon_{it}. \tag{5}$$

This is the event study equation we take to the data. The $\theta_{r(i,t)}$ are the parameters of interest: they measure changes in y_{it} years around the move scaled relative to δ_i. If the sampling error in $\hat{\delta}_i$ is ignorable, and heterogeneity in S^i_{place} is orthogonal to the other variables in the model, the plot of the $\theta_{r(i,t)}$ will have a precise interpretation similar to that of the average y^{scaled}_{it} in the simple case: the plot should be flat before and after move, and jump on move by the appropriate average of S^i_{place}.

3 Data and Summary Statistics

3.1 Data and variable definitions

Our primary data source is a 20 percent random sample of Medicare beneficiaries ("patients") from 1998 through 2008.[7] These data contain approximately 13 million patients. For each patient, we observe information on all Medicare claims for inpatient care, outpatient care, and physician services. For each claim, the data include information on the diagnosis, the type and quantity of care provided, and the dollar value reimbursed by Medicare. We also observe demographic information for each patient, including age, gender, race, and zip code of residence, defined as the address on file for Social Security payments as of March 31st of each year. To match the timing with which we observe patients' residence, we define all outcome variables for year t to be

[7] The sample is a panel defined by taking all Medicare beneficiaries in each year whose social security number ends in either "0" or "5." The sample thus varies from year to year, but a given patient remains in the sample as long as they are enrolled in Medicare.

aggregates of claims from April 1 of year t through March 31 of year $t+1$.

Our primary outcome is an index of overall health care utilization by individual by year, constructed by adjusting an individual's total annual expenditure for regional variation in prices, following Gottlieb et al. (2010). We refer to this throughout as simply "utilization." Appendix A describes the construction of this measure in detail.

We follow the literature in using as our geographic unit of analysis a Hospital Referral Region (HRR), as defined by the Dartmouth Atlas of Healthcare.[8] HRRs are collections of zip codes designed to approximate hospital markets, and are determined through an algorithm that reflects both commuting patterns and the location of major referral hospitals.[9] We map zip codes to 306 HRRs using the 1998 Dartmouth Atlas definitions.[10] About 17% of claims in any year are outside of a one's HRR, reflecting the fact that the HRR is, not surprisingly, an imperfect measure of the "market", and that some share of medical services are constructed outside of one's HRR. [11]

We define patients to be "non-movers" if (based on their address on file for Social Security payments) they live in the same HRR throughout our sample period. We define patients to be "movers" if (based on their address on file for Social Security payments) their HRR of residence 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. About 5 percent of patients are classified as movers. We drop about 3 percent of patients whose HRR of residence changes two or more times or whose address changes exactly once but for whom their average share of claims in the destination increases by less than 0.75 after the move compared to before the move.[12]

[8] www.dartmouthatlas.org/downloads/geography/ziphsahrr98.xls

[9] Each HRR consists of a collection of zip codes that contain at least one hospital that performs major cardiovascular procedures and neurosurgeries. Zip codes are grouped into an HRR based on where the highest proportion of cardiovascular procedures are referred. Each HRR must have a population of at least 120,000.

[10] We drop roughly 3 percent of patient-years whose zip codes do not match the 1998 definitions.

[11] We follow the "Dartmouth" literature and define "utilization in HRR j" to mean "utilization of patients living in HRR j", not "claims located in HRR j".Our analysis examines how a beneficiary's utilization changes when she moves across HRRs with different utilization patterns; to the extent that people within a given HRR consume some of their care outside of that HRR, we are still capturing the impact of a change in HRR on utilization.

[12] About a third of patients whose address changes exactly once appear to not satisfy this second criterion. Examination of the data reveal a variety of alternative explanations. In many cases, the claim share in destination doesn't change at all after the move, suggesting that despite the change in address to where the Social Security check is mailed, the patient did not actually move. We show below that our results are robust to other definitions of movers, such as using other cut-points for the increase in average claim share in destination after move (like 0.6 or 0.9) instead of 0.75.

3.2 Sample Restrictions and Summary Statistics

For computational tractability, we retain a random 25 percent of non-movers, along with all movers. We further limit our analysis to patient-years aged 65 and over, so that we exclude individuals who are not old-age eligible for Medicare. We then sequentially exclude the 5.1% of patients whose HRR of residence changes two or more times during our sample period, the 0.4% of patient-years older than 99, and the 18.7% of patient-years who are in Medicare Advantage (and for whom we do not therefore observe utilization), and the 7.1% of patient-years who do not have Medicare Part A or B coverage in all months. We then drop 9.3% patients who satisfy the change of address criterion for being categorized a mover but do not satisfy the criterion for increase in claim share in destination.

Our final sample includes 2.6 million patients, of whom approximately $504,000$ are movers. We note that individuals may enter or exit our sample due to age or being in a Medicare Advantage plan; they may also exit due to death. Below we explore the sensitivity of our results to potential attrition bias.

Figure 1 shows the distribution of average utilization across HRRs. We show results for both our baseline utilization measure (top panel) and for spending (bottom panel). Both show substantial variation across HRRs. Mean utilization is about $6,500 per person. The across-HRR standard deviation in average utilization is about $750; this somewhat lower than the approximately $1,000 standard deviation for spending, which reflects the fact that higher utilization areas also tend to have higher prices (Gottlieb et al. (2010)).

Table 1 presents some summary statistics of our patients. Column 1 ("original sample") shows the full sample of Medicare beneficiaries aged 65+ (limited to a random 25 percent sample of non-movers). Column 2 shows our baseline sample after the additional restrictions described above. In both cases, we up-weight the non-movers by four to account for our random sampling procedure.

Our baseline sample is about 57 percent female and 86 percent white; the average age at which we first observe someone is 72. On average we observe a person for just over 6 years. About one-third of our sample dies during our sample period. Average utilization is about $7,200 per beneficiary (standard deviation is about $10,800). Health care utilization is notoriously right skewed. About 7 percent of our sample has zero utilization, median utilization is about $4,000, and the 90th percentile of utilization is over $16,000.

In our baseline specification, we define our output measure y_{it} as the log of our utilization measure plus 1 ("log utilization"). We do this because of large differences in average utilization across areas and across people and the large secular (national) increases in utilization; this makes it undesirable to constrain each market or person to grow by the same (level) amount each year. We explore other functional forms in the sensitivity analysis below.

When we compute the sample analogue $\hat{y}_j$ of $\bar{y}_j$, we weight non-movers by 4 to account for our sampling procedure, and we omit movers in relative year 0.

3.3 Movers and Their Moves

Our empirical strategy relies heavily on movers; the fact that we observe the same individual in different places allows us to identify the role of place (i.e. to estimate the place fixed effects in equation (1)) separately from the that of person. In this sub-section, therefore, we provide some descriptive information on movers and their moves.

3.3.1 Movers vs. non-movers

Naturally people who move are not a randomly selected sub-set of the population. As shown in columns 3 and 4 of Table 1, movers are more likely to be female (60 percent vs 57 percent) and white (88 percent vs. 86 percent), and are older when first observed.[13] They are also more likely to be initially in the West (21 percent relative to 16 percent) than in other regions (particularly the Midwest). Additional information on the characteristics of elderly, cross-HRR movers can be gleaned from the panel data in the Health and Retirement Study. As in the Medicare data, we found that movers tend to be older, more likely to be female, and more likely to be white than non-movers. We also find that movers are more educated (39 percent have some college or more, compared to 32 percent of non-movers). In the first wave they are observed, movers are less likely to be married, but similarly likely to be retired as non movers. Details of this analysis and results

[13]We observe movers on average for about an extra 1.5 years; this is by construction since we must observe someone for at least two years to classify them as a mover. Despite being somewhat older, they are slightly less likely to die during our sample. This is also somewhat by construction since once has to survive long enough to change location in order to be classified as a mover.

can be found in Appendix B. Of course, any fixed differences between movers and non-movers will be captured by the patient fixed effects in equation (1).

3.3.2 Geography of moves

The average move is 589 miles, the median is 357, and the interquartile range is 119 to 914 miles. Appendix Figure 13 shows the distribution of distances moved (measured from the population-weighted centroids of the HRRs). About 60 percent of moves are cross state, and about 40 percent are across the 9 census divisions. Appendix Table 15 shows the distribution of moves by census division. Movers are (relative to the population in the census division) disproportionately likely to move out of the Mountain region and disproportionately unlikely to move out of the West South Central; they are disproportionately likely to move into the Mountain Region and into the South Atlantic or East South Central. Appendix Table 16 provides more detail on the share of moves to and from specific census divisions; it shows that the South Atlantic is the most likely destination. Indeed, moves with Florida as a destination account for 12.2% of all moves (and 9.6% of moves whose origin is outside of Florida). Moves to Florida, Arizona and California account for 24.9% of moves (and 18.1% of moves that originate outside of those states).

3.3.3 Timing of move

Our analysis requires that we be able to identify the timing of the move relatively precisely in the data. Figure 20 graphs average mover claims in their destination HRR, as a share of claims in either their origin or destination HRR, for different years relative to the year in which the move occurred. For movers, let $r(i,t)$ be the year relative to the move, which we define as occurring when $r(i,t) = 0$.[14] We refer to relative year 0 as the "move year" because the move takes place in between the end of relative year -1 and the end of relative year 0. The figure shows a clear "switch" in the location of the majority of claims from origin HRR in relative year -1 to destination HRR in relative year 1; the share of claims in the origin and destination is roughly equal in relative year 0, which is to be expected if moves are roughly uniformly distributed throughout the year. The overall picture therefore gives us confidence that we are able to identify a set of movers and their

[14]Specifically, relative year 0 is the last year in which the March 31 address is the origin HRR, and relative year 1 is the first year in which the March 31 address is in the destination HRR.

move year.

The HRS data provides some insight into the reasons for moving. In response to a question about the reasons why they moved, the most common reason given was to be near or with children (about one-third of respondents), followed by "medical or health problems or services; for health reasons" (13%), or to be near relatives and friends (10%). We also directly investigated what was correlated with moving and found that losing one's marital partner between waves increased the probability of moving between waves by 1.2 percentage points (off of a base of 1.9 percent move rate); this seemed predominantly to reflect the effects of becoming widowed. Retiring was associated with an increased probability of move of 0.4 percentage points, while a decline in self-reported health status from excellent,very good, or good to fair or poor was not associated with a statistically or substantively significant change in move probability (point estimate of -0.02 percentage point change). Of course, any direct effect of moving will be captured by the relative year fixed effects in equation (1) and will not affect our analysis, which is identified off of differential effects of moving across pairs of HRRs.

3.3.4 "Size" of moves

A key measure for our descriptive analysis will be the "size" of the move, δ_i, which we previously defined in equation (3) as the average difference in utilization between the mover's destination and origin..

Figure 2 shows the distribution of the sample analogue $\hat{\delta}_i$ across movers in our sample. It shows substantial heterogeneity in the size of the move; a one standard deviation move represents a change of 0.25 in average log utilization between origin and destination. Moves appear roughly symmetric; moves "up" to higher utilization areas are about as common as moves "down" to lower utilization areas. Indeed, the average move has a $\hat{\delta}_i$ of 0.

4 Descriptive Evidence

Before estimating the economic model, we present some initial, descriptive results designed to give a feel for the basic patterns in the data driving our identification of the role of patient and place.

Our first look at the data in Figure 3 is a binned scatter plot of how the change in utilization that occurs upon moving varies with the "size" of the move. We divide our sample of movers into 20 equal-sized bins by the "size" of their move $(\hat{\delta}_i)$. For each bin, we plot the difference in average utilization for movers from two to five years after the move relative to two the five years before the move. We also graph the linear best fit of the 20 points.

Several interesting findings emerge from Figure 3. It shows that the change in outcome is systematically, positively correlated with the size of the move. Moreover, the relationship between size of move and change in outcome seems roughly symmetric; similarly sized moves "up" and "down" have similarly sized impacts; this has economic content that we return to below in discussing the assumptions of our model. Finally, as a point of comparison, the figure also shows the average change in utilization over the same period for a matched (on covariates) non-mover sample (to whom we assign a "move size" of 0 for graphical purposes). The positive values for the average change in utilization after the move for all bins reflects the positive aggregate time trend in utilization. The fact that the non-mover point lies quite close to the impact of moving for a mover with an approximately 0-sized moves suggests little direct effect of moving per se on changes in utilization.

We can use Figure 3 to form an initial rough estimate of the role of patients relative to place. If the variation in health care utilization is mainly due to differences in individual preferences regarding treatment, then people who change areas will not necessarily experience much systematic changes in treatment and we would expect a slope or "pass-through" rate of about 0. If, on the other hand, providers in the area are the primary determinants of treatment patterns with the patient in a relatively passive role, then individuals who move across areas with different practice styles will experience, on average, a change in utilization that would bring them quite close to the new area's average norm, and we would expect a slope, or "pass through" rate of about 1. Figure 3 has a slope of about 0.67. In other words, this initial look at the raw data suggests that when an individual moves to a place that is 1 unit different in terms of utilization, the individual's utilization changes by about 0.67, or that about one-third of the patient's utilization is explained by the patient, rather than the place.

To better exploit the information we have on the timing of the move, we also present a graphical, event-study analysis that visualizes the time path of an individual's utilization changes as he moves across areas with different utilization patterns. This allows us to verify that our empirical strategy based on individual moves captures something that happens at the time of the move.

Figure 4 plots estimates of equation (5). Since the level of the graph is not separately identified, we normalize it to 0 in the year prior to the move (relative year -1). Mean utilization jumps up sharply when individuals move to higher utilization places (i.e. between relative year -1 and relative year 1); the effect is fairly linear through the move year (relative year 0), consistent with the fact that this is roughly an "in between" year (see Figure 20). There is some evidence of a gradual pre-trend upward in utilization prior the move; below we will show that the quantitative effect of "narrowing in" our analysis around the timing of move is fairly small. We view the results as broadly reassuring on the concern of the timing of moving being correlated with shocks to the demand for care. There is no evidence of any systematic trend in utilization after the move, suggesting that the impact of place is relatively "immediate". This is suggestive evidence against models of gradual adaptation or habit formation in this settings. We will return to both these points in more detail below when we discuss the assumptions of the model and the potential sources of patient heterogeneity.

Recall that the size of the jump tells us about the relative role of patient and place. If there was no jump at the time of move, this would suggest that utilization behavior was entirely determined by patient. If at move the individual jumped immediately to the destination level (i.e. size of jump of 1), this would suggest that utilization was determined entirely by the place. Indeed, as noted above, the size of the jump can be interpreted as an average of $S^i_{place} = 1 - S^i_{pat}$. The event study just suggests that about 46 percent of the difference across HRRs is due to patients. We now turn to a more formal quantification.

5 Main Results: Patient vs. Place

5.1 Decompositions

The event study analysis provides visual evidence that utilization changes sharply at the time of the move across area. We use the identifying variation provided by patients moving across area to estimate the role of patients versus place using the basic framework provided by equation (1). Using the estimates from equation (1), we report two difference decompositions of the role of patient and place.

We first ask: what share of the difference in average utilization across two sets of HRRs (R and R') is due to the patient (as opposed to the place). This is the sample analog of $S_{pat}(R, R')$.

Table 2 reports this share, as well as the components $(\hat{y}_R - \hat{y}_{R'})$, $(\hat{\gamma}_R - \hat{\gamma}_{R'})$, and $(\hat{c}_R - \hat{c}_{R'})$ for a number of different choices of R and R'. In each case, we choose R to be a group of HRRs with relatively high values of $\hat{y}_j$ and R' to be a group of HRRs with relatively low values of $\hat{y}_j$.

Column 1 decomposes the share of the difference in average utilization between above and below median HRRs into the share due to patients and the share due to place. We find that about 43 percent of the difference in average utilization between above and below median HRRs is due to patients, with the other 57 percent due to place. This estimate is fairly precise; we can reject a role for patients of more than about 49 percent or less than about 37 percent.

Results for other partitions of HRRs suggest a similar share of differences in cross-HRR average utilization that is due to patients. Columns 2 through 4 show that a similar estimate for the share of variation due to patients for other cuts of the distribution of HRRs - patients account for between 37 and 43 percent of the variation in utilization between the top and bottom 25% of HRRs, top and bottom 10%, and top and bottom 5%. In the last two columns we look at the two case studies highlighted in the introduction that have received a lot of popular attention: McAllen relative to El Paso and Miami relative to Minneapolis. Here, we find that patients account respectively for 43 and 31 percent of the difference in average utilization, although naturally the standard errors increase with these smaller samples.

The fact that our estimate of the share of the differences across groups of HRRs that due to patients is relatively stable across different cuts of the HRR distribution is related to our finding of a relatively linear relationship between the size of the move and the average impact of the move in Figure 3. These results suggest that $S_{pat}(j, j')$ is not strongly correlated with $\bar{y}_j - \bar{y}_{j'}$. This also relates to our finding that our estimates in Table 2 of the share of difference in cross-HRR average utilization that is due to patients ranging from 31 percent to 43 percent are broadly in line with the descriptive analysis, which suggested a 46 percent role for patients implied by the magnitude of the jump in utilization from relative year -1 to 1 in the event study (Figure 4) or a one-third role for patients implied by the initial pre-post analysis in Figure 3. These estimates differ for two reasons. First, the descriptive analysis produces an estimate of the weighted average role of patients across pairs of moves, with the weights inversely proportional to the variance of the estimates. Second, the slight pre-trends in Figure 4 are handled differently when we look at the immediate jump in Figure 4 relative to the other two exercises.

We also report results from a second exercise for gauging the contribution of patient and place that is not specific to a particular partition of HRRs. We ask what share of the cross-HRR variance in outcome would be eliminated if we (counterfactually) equalized patient effects (or, equivalently, if we randomly distributed patients across HRRs). We define this to be

$$S_{pat}^{var} = 1 - \frac{\text{Var}\left(\gamma_j\right)}{\text{Var}\left(\bar{y}_j\right)}. \tag{6}$$

Similarly, we can define the share of variance that would be eliminated if we equalized area fixed effects to be

$$S_{place}^{var} = 1 - \frac{\text{Var}\left(\bar{c}_j\right)}{\text{Var}\left(\bar{y}_j\right)}$$

Note that unlike S_{pat} and S_{place}, this is not an additive decomposition; the sum of S_{pat}^{var} and S_{place}^{var} will not sum to one so long as $Cov\left(\bar{c}_j, \gamma_j\right)$ is non-zero.

Table 3 reports the estimated values of the relevant quantities $\text{Var}\left(\hat{y}_j\right)$, $\text{Var}\left(\hat{\gamma}_j\right)$, $\text{Var}\left(\hat{c}_j\right)$, and the implied estimates of S_{pat}^{var} and S_{place}^{var}. We find that about 36 percent of the variation in outcomes across HRRs would be eliminated if patient fixed effects were equalized. We also find that about 55 percent of the variation in outcomes across HRRs would be eliminated if place fixed effects were equalized. The fact that these two shares sum to less than 1 is indicative of a negative covariance between patients and areas, suggesting that patients with high fixed effects for utilization sort disproportionately to areas with low utilization fixed effects; this is confirmed by our estimate of a negative covariance of patient and HRR fixed effects as shown in the table.

5.2 Investigation of key assumptions

The foregoing results for the role of patient and place are based on estimating the model in equation (1) using individuals who move across areas to identify the role of patient separately from that of place. In this section we investigate the validity of three key assumptions of the economic model and estimation strategy.

5.2.1 Exogenous moving

Moving may be correlated with changes in outcomes, and the difference between destination and origin utilization may be non-random with respect to the mover. Indeed, Figure 3 already suggested a small, positive direct effect of moving on changes in utilization. Similarly, Figure 5 suggests that, prior to their move, movers' utilization appears about 10 percent closer to the average utilization in their subsequent destination than non-movers in their origin, suggesting that there is some selection of the direction of the move with the prior level of utilization. However, as is evident from equation (1), and as discussed in section 3.3, our empirical strategy can handle any direct effects (or correlates) of moving per se with utilization through relative year fixed effects, as well as any fixed differences in utilization between movers and non-movers through patient fixed effects.

Rather, a key assumption of our empirical strategy is that where a given person lives is not correlated with changes in the error term. In particular, it would violate the model if there were a shock to utilization at the time of the move that is correlated with where the individual was moving to or from; for example, if people moving from low to high utilization places experienced positive shocks to their expected utilization at the time of move.

One testable implication of this assumption is that we should not see pronounced trends in utilization prior to the move correlated with the difference between average utilization in the mover's origin and destination. A natural way to assess this is to look for trends in utilization prior to the move and how they correlate with the size of the move. As we saw previously in Figure 4, there is a slight pre-trend, suggest that individuals who move to higher utilization places are trending upward in their utilization prior to their move.[15] This pre-trend persists when we limit the sample to a "balanced panel" of movers whom we observe in every year from relative year -7 to 1 (Figure 6(c)). Although this slight pre-trend suggests a strict violation of the model, we show in the robustness below that it is not quantitatively influential in our findings.

[15] Indeed, if we re-estimate the event study regression from equation (5), but imposing a linear trend in relative year interacted with delta for relative years less than or equal to minus 1, we estimate a statistically significant pre-trend of -0.015 (standard error = 0.0035).

5.2.2 Additivity

The framework in equation (1) assumes that the role of patient and place is additively separable. In other words, there are no interaction terms between patient and place in determining utilization. While unlikely to be strictly true, several pieces of evidence suggest that this assumption is a reasonable approximation of reality. First, an implication of the additive separability assumption is that the impact of moving across areas should be symmetric (i.e. a given move "up" should have the same sized effect (in absolute value) as a given move "down"). As noted, Figure 3 provided evidence of just this symmetry.

Second, another implication of the additively separable model is the absence of habit formation. In other words, utilization is a function of current location and a person effect, but not of prior location. One implication of an alternative model with habit formation (such as in Bronnenberg et al., 2012) is that we would expect some gradual adjustment post move as one's habits adjust to the new area. Figure 4 and the balanced panel analysis in Figures 6(a) and **6**(b) provide one way to look at this and did not suggest such habit formation.

Third, we estimated an augmented version of the model in equation (1) in which we allow for a complete set of interactions between patient and place. The explanatory power of the model improves slightly - the adjusted R-squared rises from 0.430 to 0.445. This suggests the presence of a relatively small interaction effect.

Finally, we examined the pattern of mean residuals from our estimation of equation (1). Violations of the additively separable model might be expected to cause relatively large mean residuals for particular types of interactions - for example, high utilizing people at high utilizing places. To examine this possibility, Figure 7 plots the mean residuals for the 100 cells formed by the deciles of the estimated HRR effect and the deciles of the estimated person fixed effect. These results show mean residuals that are very close to zero in virtually all cases, and no systematic patterns except at the very highest (and lowest) decile of the person effects where it looks like mean residuals are increasing (decreasing) with the HRR decile. Whether or not this violation of additivity is quantitatively important to our findings is something we still need to investigate.

5.2.3 Patient and place effects are constant over time

Finally, our model assumes that each patient and place has a fixed (constant over time) component that determines utilization. The baseline model rules out any time-varying patient or place effects, except as they manifest themselves for patients through age or relative year. Again, while unlikely to be strictly true, several pieces of evidence suggest that these assumptions are a reasonable approximations of reality.

To examine whether the assumption of constant place effects seemed reasonable, we first divided our data into two time periods (1998-2003 and 2004-2008) and ranked HRRs by mean utilization within each time period. The rank correlation between the two time periods is 0.91. This is consistent with the findings of the prior literature that patterns of geographic variation in health care utilization has been relatively stable since data became available in the early 1990s (Chandra et al. 2009, Rettenmaier and Saving 2009, Weinstein et al. 2004). We also show in Appendix Figure 14 that quintiles of HRRs are stable over time; there is a secular trend of rising utilization over time, but there is no differential trend for different quintiles of HRRs.

Even with rank stability, we could be concerned that the effect of place is changing over time - for example high utilization places are exerting a systematically larger effect on utilization over time. To assess this, Figure 8 plots the mean residuals of the 110 cells formed by the deciles of the estimated HRR effect and by years. The results are reasonably reassuring; there is no systematic time pattern in the residuals overall or separately by HRR decile (as various violations of constant place effects might produce), and the mean residuals tend to be close to zero.

As a final check on the assumption of constant place fixed effects, we estimated an augmented version of the baseline model from equation (1) in which we now allow place effects to vary across the two time periods:

$$y_{ijt} = \alpha_i + \rho_{r(i,t)} + x_{it}\beta + \gamma_j + \gamma_j^{post} I_t^{post} + \tau_t + \varepsilon_{ijt}, \tag{7}$$

where I_t^{post} is an indicator variable for years 2004 and later. Calculation of the share of the differences due to patient and place is analogous to our main specification, with the estimated value of $\left(\gamma_j + \gamma_j^{post} I_t^{post}\right)$. Table 4 shows that this has no discernible effect on the results (compare rows 1 and 2).

In a similar manner, we investigated the validity of the assumption that patients exert a constant (time invariant) contribution to utilization. Figure 9 plots the mean residuals of the 110 cells formed by the decile of the estimated person effects and by years. Once again, we would be concerned by either systematic patterns of residuals by patient decile and time or by substantial deviations from zero mean by cell; there are a handful of large outliers, concentrated in low deciles and late years. Once again, whether this violation is quantitatively important to our findings is something we still need to investigate.

We likewise investigated the sensitivity of our results to allow patient effects to vary over time. To do this, in row 3 of Table 4 we consider the data in two time periods (1998-2003 and 2004-2008) and only include data on a mover during the time period in which they moved. This effectively allows patients to have a constant effect over only about half the length of time as in the baseline specification. In the next two rows, we subdivide the analysis in Row 3 so that we separately examine only movers who moved in the first period and movers who moved in the second period; this effectively requires both place and patient effects to be constant over only half the length of time as in the baseline specification. Finally, the last four rows of the table repeat this exercise but now dividing the data into three periods (1998-2001, 2002-2005, and 2006-2008) instead of two. The estimate of the share due to patients is relatively stable across these alternatives. For example, relative to the baseline estimate that about 43 percent of the difference in utilization between above and below median HRRs would be eliminated if patient fixed effects were equalized, dividing the data into two time periods (row 3) or three (row 6) raises this estimate to 46 and 45 percent respectively. Estimating the data separately by periods (rows 4,5, and 7-9) produces estimates of the reduction in utilization differences within a given period ranging from 0.39 to 0.59 for different periods.

5.3 Robustness

We investigated the sensitivity of our results to potential attrition bias and to alternative specifications. The results appear reasonably robust, with our baseline estimate of a 0.43 patient share in explaining the average difference in utilization between above and below median utilization HRRs ranging from 0.39 to 0.54 across the various alternative specifications we examined.

5.3.1 Attrition

On average, about 30 percent of patients who are ever in our baseline sample are absent in a given year (see Table 1). This is due in roughly equal parts to death (an absorbing exit state) and being in an HMO in that year.[16]

Such "attrition" poses a threat to our identification strategy if it is correlated with shocks to outcome, the timing of the move, and where the individual is moving to or from. For example, if at the time of their move individuals who experience negative shocks to utilization are more likely to enter an HMO if they are moving from high to low utilization places. One way to shed some light on whether this is likely a concern is to re-estimate the event study in equation (5), using a measure of attrition for the dependent variable. Specifically, the dependent variable is a binary variable that is 1 if the individual is absent from the sample in that person-year, and 0 otherwise; the right hand side is defined as before.

Figure 10 shows the results. They suggest that moving to an HRR with an average of one log utilization more than your origin is associated with a 3.5 percent decrease in the probability of attrition; given a cross-HRR standard deviation of log utilization of about 0.2, this indicates that a move up of 1 standard deviation of log utilization is associated with a 0.6 percent decrease in the probability of attrition. [17]

As another informal check on the sensitivity of our results to attrition, we examine the robustness of our core results to re-estimating the basic economic model in equation (1) limiting the sample to patients who never die, to patients who are never in an HMO, or to both. We also try estimating the model limiting the sample to a balanced panel of early movers, middle movers, or late movers (analogously to Figures 6(a)-(c)). The results are shown in Table 5; the Appendix Figure 15 shows the corresponding event study figures. In general the results seem fairly stable across these alternative cuts, with the patient share in explaining the average difference in utilization between above and below median utilization HRRs ranging from 0.41 to 0.54.

[16] 3 percent of patients are absent from the baseline sample in a given year due to not being enrolled in Parts A and B for all months eligible that year.

[17] Of course, "attrition due to death" and its correlates with the change in average utilization in the destination relative to the origin HRR may potentially be of substantive interest. We explore this in Section 6.3.

5.3.2 Alternative specifications

We also explored the sensitivity of our results to a number of our empirical choices. Table 6 shows the robustness of our results to some of our specification choices; the Appendix Figure 16 shows the corresponding event study figures. Row 2 shows the results estimated on movers only. Row 3 drops age as a time-varying covariate. And in row 4 we change the dependent variable to the log of expenditure (plus 1) rather than the log of utilization (plus 1).[18] The share of the difference in above vs. below median utilization HRRs due to patients varies from 0.422 to 0.450 across these specifications.

In rows 5 through 8, we re-estimate the baseline model "narrowing in" on relative years closer to the move year. Not surprisingly, given the evidence of a slight upward pre-trend in the event study (see Figure 4), limiting the data to years closer to the move year increases the estimated role of the patient. In particular, relative to the 0.43 estimate in our baseline specification, the role of the patient increases to 0.47 when we limit the data to relative years -1 and 1.

In rows 9 and 10 we re-estimate the model with a dependent variable of $ln(utilization+0.1)$and $ln(utilization+10)$ instead of our baseline dependent variable $ln(utilization+1)$. In row 11 we estimate the model in first difference. In row 12 we add patient and HRR fixed effects to the first-differenced model to allow for separate (linear) trends for each patient and each HRR. Our estimates of the patient share range from 0.38 to 0.52 across these specifications.

Finally, we explored the robustness of our results to limiting to different subsamples of movers. Table 7 shows the results are robust. In particular, we show results limiting to the approximately 2/3 of movers who move across states (row 2), or to the approximately 1/3 of movers who move across the 9 census regions (row 3). We also show that the results are robust to dropping the 12% of movers whose destination is in Florida (row 4), or to dropping the 25% of movers whose destination is destination is Florida, Arizona or California (row 5). Appendix Figure 17 shows the corresponding event studies. In Appendix C.4, we also show that the results are robust to alternative definitions of movers.

[18]We would expect the patient share to be lower for spending $(p*q)$ than utilization (q). We find, however, that the patient share is slightly higher for spending (0.450 compared to 0.426 for utilization). We suspect this is because Medicare prices vary within areas over time (e.g. are functions of local wage indices), and some of the geographic variation in price therefore does not load onto the place fixed effects.

6 Mechanisms

We have presented robust evidence of a non-trivial role for patients in explaining geographic variation in health care utilization. Our central estimate is that about 40 percent of the difference in average utilization between above and below median utilization HRRs would be eliminated if patient effects were equalized across those HRRs. As noted in the introduction, an estimate of the relative role of patients and places in explaining geographic variation has potential implications both for thinking about the likely impact of policies that might reduce geographic variation and for the efficiency implications of this geographic variation. In this final section of the paper, we briefly explore what we can learn - or rule out - about why (and how) patients contribute to the observed geographic variation in health care utilization.

6.1 Mechanical effects: patients encoding place

We start by considering a potential alternative interpretation of our patient effects: that they are really capturing the role of place prior to move. One reason we might observe a patient who moves from high-spending Miami to low-spending Minneapolis to continue consuming a lot of care is that high consumption pre-move creates high demand for care post-move. Several mechanisms might lead to this kind of state dependence. It may be that high spending is associated with high rates of diagnosis (Song et al. 2010), and once conditions are diagnosed patients continue to receive treatment for them wherever they go. It may be that behaviors such as regular doctor visits or flu shots are subject to habit formation as in Becker and Murphy (1988). Or, it may be that social learning leads patients who have lived around high-spending individuals to conclude that high treatment levels are optimal (or vice versa). These types of stories suggest that we might be attributing to the person what is really a function of their past location. And thus mis-allocating to the "person" and effect that is really about their prior "place."

However, three separate results point away from this type of story. First, we can rule out that utilization is "downward sticky" (e.g. once treated intensely, always treated intensely even if move to a low intensity place. This implies asymmetry in the nature of the move: moves to more intensive places should increase utilization more than moves to less intensive places should decrease it. As noted already in discussing the evidence for the model's assumption of additively separability, Figure 3 suggests symmetric effects of moves "up" and "down" in utilization intensity.

Second, this type of "downward stickiness" also implies that patients should play less of a role if they have already had substantial prior experience with the health care system. In other words, the role of patients in explaining geographic variation in utilization should be lower for sicker and older patients. We examined this by re-estimating the economic model in equation (10) separately for individuals stratified by their experience with the health care system (i.e. above vs below median on number of chronic conditions, number of hospital days, or utilization), as well as by age. Table 8 reports the results, and Appendix Figure 18 shows the corresponding event study figures for these different strata. They suggest, if anything, more of a role of healthy patients in explaining cross-area variation in utilization than for sicker patients.[19] These findings also mitigate against mechanical serial correlation.

Finally, as noted previously, models of habit formation (what I do now depends in part on what I did in the past) would tend to imply a gradual adaption to the new place. However, as shown in e.g. Figure 6(a), the effect of the move appears to be relatively "immediate". This is suggestive evidence against models of gradual adaptation or habit formation in this settings. The fact that the role for patients is declining with age (Table 8) - at least among a 65+ year old sample - also mitigates against models of habit formation. Of course, these results do not speak to the role of experience at younger (pre 65) ages. Still, they suggest that patient demand is relatively "fixed" among the elderly, who account for about one third of total annual healthcare spending (Moses et al. 2013).

6.2 Patient health

An advantage of our mover-based empirical strategy is that it does not require us to infer the role of patients from observable patient characteristics. As noted in the Introduction, an observables-based approach to inferring the role of patients is hampered both by the potential existence of important unobservable patient characteristics (e.g. unmeasured health or preferences) and by the potential endogeneity of observed patient characteristics to place (Song et al., 2010). We therefore use our empirical strategy to investigate both of these issues.

[19] Results are similar if we limit the sample to movers and stratify on these characteristics in the 1 to 2 years prior to the move (not shown).

Having estimated the share of the difference between above and below-median utilization places that would be eliminated if we equalized all observed and unobserved patient characteristics across areas, we can also ask how much of the gap would be eliminated if we only equalized a specific vector of patient observables z_{it}. This could include both time-varying characteristics such as health status and time-invariant characteristics such as race and sex whose effect is absorbed in the patient fixed effects in equation (1).

To answer this question for a given z_{it}, we first run an auxiliary regression

$$y_{ijt} = \gamma_j + \tau_t + z_{it}\beta_z + \varepsilon_i \tag{8}$$

using only non-movers. The coefficient $\hat{\beta}_z$ gives us a predictive relationship between z_{it} and y_{ijt} based on only within-area variation. Naturally, β_z does not represent the "causal" effect of some observable on utilization, but merely a conditional correlation in the spirit of prior literature examining the role of certain patient observables in "explaining" geographic variation. We then define the patient component due to z to be $c_{it}^{obs} = z_{it}\beta_z$, and we define the averages $\overline{c}_{jt}^{obs}$, $\overline{c}_{j}^{obs}$, and $\overline{c}_{R}^{obs}$ (for a set of areas R) analogously to $\overline{c}_{jt}$, $\overline{c}_{j}$, and $\overline{c}_{R}$. Finally, we define $S_{pat}^{obs}(R,R') = \left(\overline{c}_{R}^{obs} - \overline{c}_{R'}^{obs}\right) / \left(\overline{y}_R - \overline{y}_{R'}\right)$. Note that like S_{pat}, S_{pat}^{obs} need not lie between 0 and 1, and that S_{pat}^{obs} may be larger than S_{pat}.

Table 9 shows the results. We find that age, race and sex explain little of the geographic variation in health care utilization; row 2 indicates that equalizing age race and sex would eliminate only about 6 percent of the difference in average utilization between above and below median utilization HRRs. The next two rows show the results of equalizing two different measures of patient health. The first, Hierarchical Condition Categories (HCC) risk scores, are CMS measures of predicted spending (used by Medicare for program payment) based on a coding algorithm that is designed to predict spending as a function of demographics (including age, gender, and Medicaid eligibility) and that year's inpatient and outpatient diagnoses (conditions) which it aggregates into a hierarchy of condition categories; this measure therefore includes both health measures and other demographics. The second measure - the number of chronic conditions (CCs) - is purely a health measure, indicating how many of 27 possible chronic conditions the patient has been diagnosed with (within a 1-3 year window) by the end of that year.

Equalizing differences in patient health - as measured by HCC scores - would eliminate about

41 percent of the difference in average utilization (row 3); equalizing patient health as measured by the number of CCs would eliminate 75% percent of the difference in average utilization (row 4). Equalizing patient health (as measured by the number of CCs) and demographics would eliminate 79% (row 5). These results are consistent with the prior literature which has tended to find that controlling for basic demographics has little effect in reducing geographic variation in utilization, but controlling for measured health has non-trivial effect (for a recent review, see Skinner 2012). For example, Zuckerman et al. (2010) document that the gap between the highest and lowest spending quintiles is 52 percent in the raw data, 48 percent controlling for demographics, and 33 percent controlling for measured health.

At the same time, the literature has recently recognized that measured patient health may itself be partly endogeneous to place, as it partly captures differential propensities to diagnosis. Song et al. (2010) examined individuals who moved across quintiles of the HRR distribution and found that measured health increased as individuals moved to HRRs in higher quintile of end of life spending (a standard proxy for an intensive area). This suggests that measures of patient health like "CCs" and "HCCs" may not reflect purely patient-based characteristics. Indeed, the fact that we estimate that eliminating variation across patients in the number of CCs would reduce substantially more of the difference in average utilization across HRRs than eliminating the entire "patient effect" (see row 4) is suggestive of the fact that part of what the health measures are capturing is actually a function of place rather than of person.

We can use our same mover strategy to quantify the component of measured health that reflects endogenous diagnosis ("place component of health") and the component of measured health that reflects the patient's underlying health (i.e. "patient component of health"). Specifically, we can re-estimate our baseline model in equation (1) using a health measure as the left hand side variable y and then re-calculate S_{pat}. We do this for both HCCs and the number of CCs. The results are shown in Table 10, while Appendix Figure 19 show the analogous event studies. Note that in columns 5 and 6 the share of difference due to patients and due to place now reflects the share of the difference in the average outcome in column 1 between HRRs that are above and below median for that outcome; thus the partition of HRRs differs in each row. The results suggest in rows 1 and 2 suggest that, by either measure, about 50 percent of measured geographic health variation is in fact endogenous to place..

We can then ask how much of the mean difference in utilization between above and below median utilization places would go away if we equalized the *patient* component of health or if we equalized the *place* component of health. Table 11 shows the results. The results show that equalizing the patient component of health as measured by number of chronic conditions (CCs) would eliminate about 36 percent of the geographic variation in utilization. In other words, the exogenous component of patient health can explain about 85 percent of the patient effect, leaving about 15 percent to be explained by unmeasured risk factors or preferences. Likewise, we see that equalizing the place-based component of health (again as measured by number of chronic conditions) would eliminate about 40 percent of geographic variation in utilization. In other words, the diagnosis propensity of a place can explain about 70 percent of the place effect. Of course, as discussed previously, these are not decompositions that must sum to 1; it is possible there are other important patient components that offset, and likewise for place.

6.3 Mortality

We have presented a decomposition of the relative roles of demand and supply in contributing to geographic variation in health care utilization, and to some of the mechanisms behind the patient demand. Our findings suggest that about 40 percent of the geographic variation in health care utilization is due to patient demand, and that most of this in turn reflects variation in underlying patient health. A natural question is the efficiency of the observed variation. A full examination of this issue is beyond the scope of the current paper.

However, one way that previous researchers have tried to shed light on this question is by examining the impact of higher utilization areas on patient outcomes, particularly mortality. In the cross-section, it has been widely noted that higher utilization is not generally correlated with better patient outcomes (Skinner (2012) provides a recent review). Using tourists who are get ill while on vacation in Florida, Doyle (2011) finds that becoming ill in a higher-spending area is associated with lower mortality.

We can likewise present some suggestive evidence by examining the association between post-move mortality and average utilization in the destination. Unlike the analyses thus far which have relied on panel data, our mortality analysis is by necessity cross sectional. This strongly suggests that these mortality estimates should be interpreted more cautiously. With this important caveat in mind, Figure 11 plots mean mortality post-move for four sets of patients: patients moving to above-median utilization HRRs, patients moving to below-median utilization HRRs, patients moving to HRRs in the top decile of utilization, and patients moving to HRRs in the bottom decile of utilization. Comparing the above and below median plots reveals that patients moving to above-median utilization HRRs tend to have slightly lower average annual mortality rates; this pattern is much more pronounced in a comparison of patients moving to the top and bottom decile HRRs, where patients moving to top decile HRRs appear to have substantially different annual mortality rates.

While this figure is suggestive that moves to higher utilization HRRs are associated with lower post-move mortality rates, the concern with this comparison is that patients moving to higher utilization HRRs may have lower baseline mortality probabilities than do patients moving to lower utilization HRRs. To attempt to address this concern, Table 12 presents regression results for the relationship between post-move mortality and average utilization in the destination ($\bar{y}_d$) that successively condition on various covariates. Column 1 shows results with no controls. Column 2 add controls for origin HRR dummies, so that all analysis is among movers from the same HRR who move to destinations with different average utilization rates. Column 3 adds additional controls for the mover's year of move, age, sex and race. Column 4 adds additional controls for the pre-move health conditions of the mover. The panels display estimates for all available post-move years, as well as estimates from relative year 0 (the year of the move; recall that in relative year 0 only about half of our sample has moved), relative year 1 (the year after the move) and relative year 5 (five years post-move).

Overall, our estimates are sensitive to the choice of controls, which suggests that destination utilization is correlated with underlying characteristics of the mover. Nonetheless, with this important caveat in mind, our preliminary regression estimates suggest that moves to higher utilization areas are associated with reductions in mortality. Conditional on all controls (Column 4), our preliminary estimates thus far suggest that a one-standard deviation (0.2 log point) increase in average utilization in the destination is associated with about a 0.13 percentage point (1.6 percent) reduction in the annual mortality rate post move. This effect appears to be relatively constant over the approximately 8 years we can observe post move. We plan on exploring the robustness of this finding further in future drafts.

7 Conclusion

We investigated the relative roles of demand-side and supply-side factors in the well-documented geographic variation in health care utilization. Our strategy exploits patient migration in panel data of a 20% random sample of Medicare beneficiaries from 1998 to 2011. The descriptive analysis shows a clear jump in utilization at the time of move as individuals move from lower to higher utilization areas (and conversely when they move from higher to lower utilization areas). Our baseline estimate suggests that about 40 percent of geographic variation in healthcare utilization is due to demand-side (i.e. patient) factors rather than supply-side (i.e. place-based) factors. This estimate is relatively stable across a range of alternative specifications, as well as across the exact definition of geographic variance used.

To better understand what is behind the role for patients that we found, in the last part of the paper we briefly explored some potential mechanisms. We found no evidence of habit formation in which prior health care utilization affects subsequent patient demand. We find a large role for patient health in contributing to geographic variation in health care utilization, estimating that approximately 35 percent of the geographic variation in health care utilization can be explained by the "exogenous" component of patient health that we estimate. The remaining role for patients presumably reflects unmeasured health and/or differences in preferences for care.

These findings suggest potential limits to the impact of supply-side policies aimed at reducing geographic variation in health care, at least in the near term. Patient demand appears to play a larger role than the conventional wisdom may have assumed. Moreover, the lack of habit formation suggests that patient effects are relatively "fixed", at least among the 65 and over population, a population that accounts for about a third of total annual health care spending (Moses et al. 2013). It is possible that very long-term forms of habit formation are operative, such as models in which health care preferences are strongly influenced by early life experiences, and this would be a natural idea to explore in future work with data on younger individuals.

Our findings may also be relevant for thinking about the efficiency of this geographic variation. To the extent that variation in patient demand reflects variation in marginal impact of treatment or marginal utility from a given impact, that portion of the variation may well be efficient. Of course, it is also possible that it reflects mis-information on the part of patients. We presented some preliminary, suggestive evidence that moves to higher utilization areas are associated with lower mortality in the cross section. This suggests there may be some "return" to higher utilization. But a more thorough and careful examination of the efficiency implications of the geographic variation is an important direction for further work.

Finally, we note that in the same way we have begun to decompose the mechanisms behind patient effects in this paper, we hope in future work to begin to decompose the mechanisms behind the place effects we estimate. Particularly interesting questions concern the role of physicians and the role of health care organizations. We suspect that similar strategies exploiting *physician* migration across geographic areas and across health care organizations within a geographic area may prove useful for gaining further insights.

References

Aaronson, D. (1998). Using sibling data to estimate the impact of neighborhoods on children's educational outcomes. *Journal of Human Resources*, 33(4):915–946.

Abowd, J. M., Creecy, R. H., and Kramarz, F. (2002). Computing person and firm effects using linked longitudinal employer-employee data. *U.S. Census Bureau Technical Paper*.

Abowd, J. M., Kramarz, F., and Margolis, D. N. (1999). High wage workers and high wage firms. *Econometrica*, 67(2):251–334.

Ashton, C. M., Peterson, N. J., Souchek, J., Menke, T. J., Yu, H.-J., Pietz, K., Eigenbrodt, M. L., Barbour, G., Kizer, K., and Wray, N. P. (1999). Geographic variations in utilization rates in veterans affairs hospitals and clinics. *The New England Journal of Medicine*, 340(1):32–39.

Baker, L. C., Fisher, E. S., and Wennberg, J. E. (2008). Variations in hospital resource use for medicare and privately insured populations in California. *Health Affairs*, 27(2):w123–134.

Becker, G. S. and Murphy, K. M. (1988). A theory of rational addiction. *Journal of Political Economy*, 96(4):675–700.

Bronnenberg, B. J., Dube, J.-P. H., and Gentzkow, M. (2012). The evolution of brand preferences: evidence from consumer migration. *American Economic Review*, 102(6):2472–2508.

Card, D., Heining, J., and Kline, P. M. (2013). Workplace heterogeneity and the rise of West German wage inequality. *Quarterly Journal of Economics*, 128(3):967–1015.

Chandra, A., Cutler, D., and Song, Z. (2012). Who ordered that? The economics of treatment choices in medical care. In Mark V. Pauly, T. G. M. and Barros, P. P., editors, *Handbook of Health Economics*, volume 2 of *Handbook of Health Economics*, pages 397–432. Elsevier.

Chandra, A., Sabik, L., and Skinner, J. S. (2009). Cost growth in Medicare: 1992 to 2006. In Wise, D. A., editor, *Explorations in the Economics of Aging*, pages 133–157. University of Chicago Press.

Chandra, A. and Staiger, D. O. (2007). Productivity spillovers in health care: Evidence from the treatment of heart attacks. *Journal of Political Economy*, 115(1):103–140.

Chernew, M., Sabik, L., Chandra, A., Gibson, T., and Newhouse, J. (2010). Geographic correlation between large-firm commercial spending and Medicare spending. *American Journal of Managed Care*, 16(2):131–138.

Chetty, R., Friedman, J. N., Leth-Petersen, S., Nielsen, T., and Olsen, T. (2012). Active vs. passive decision and crowd-out in retirement savings accounts: evidence from Denmark. Working Paper 18565, NBER.

Chetty, R., Friedman, J. N., and Rockoff, J. E. (2013a). Measuring the impacts of teachers ii: teacher value-added and student outcomes in adulthood. Working Paper 19424, NBER.

Chetty, R., Friedman, J. N., and Saez, E. (2013b). Using differences in knowledge across neighborhoods to uncover the impacts of the eitc on earnings. *American Economic Review*.

Congressional Budget Office (2008). Geographic variation in health care spending. Washington: Government Printing Press.

Cutler, D., Skinner, J., Stern, A. D., and Wennberg, D. (2013). Physician beliefs and patient preferences: A new look at regional variation in spending. Working Paper 19320, NBER.

Doyle, J. J. (2011). Returns to local-area health care spending: Evidence from health shocks to patients far from home. *American Economic Journal: Applied Economics*, 3(3):221–43.

Dunn, A. C., Shapiro, A., and Liebman, E. (2011). Geographic variation in commercial medical care expenditures: A decomposition between price and utilization. BEA Working Papers 0075, Bureau of Economic Analysis.

Fernandez, R. and Fogli, A. (2006). Fertility: the role of culture and family experience. *Journal of the European Economic Association*, 4(2-3):552–561.

Fisher, E. S., Wennberg, D. E., Stukel, T. A., Gottlieb, D. J., Lucas, F. L., and Pinder, E. L. (2003a). The implications of regional variations in medicare spending: the content, quality, and accessibility of care. Part 1. *Annals of Internal Medicine*, 138(4):273–287.

Fisher, E. S., Wennberg, D. E., Stukel, T. A., Gottlieb, D. J., Lucas, F. L., and Pinder, E. L. (2003b). The implications of regional variations in medicare spending: the content, quality, and accessibility of care. Part 2. *Annals of Internal Medicine*, 138(4):288–299.

Gawande, A. (2009). The cost conundrum. *New Yorker*.

Gottlieb, D. J., Zhou, W., Song, Y., Andrews, K. G., Skinner, J. S., and Sutherland, J. M. (2010). Prices don't drive regional medicare spending variations. *Health Affairs*, 29(3):537–543.

Ichino, A. and Maggi, G. (2000). Work environment and individual background: explaining regional shirking differentials in a large Italian firm. *Quarterly Journal of Economics*, 115(3):1057–1090.

Luttmer, E. F. P. and Singhal, M. (2011). Culture, context, and the taste for redistribution. *American Economic Journal: Economic Policy*, 3(1):157–179.

McPherson, K., Strong, P., Epstein, A., and Jones, L. (1981). Regional variations in the use of common surgical procedures: Within and between England and Wales, Canada and the United States of America. *Social science & medicine. Part A, Medical sociology*, 15(3 pt 1):273–288.

Molitor, D. (2011). The evolution of physician practice styles: Evidence from cardiologist migration. MIT Working Paper: http://economics.mit.edu/files/7301.

Moses, H., Matheson, D. H. M., Dorsey, E. R., George, B. P., Sadoff, D., and Yoshimura, S. (2013). The anatomy of health care in the United States. *Journal of the American Medical Association*, 310(18):1947–1964.

Philipson, T. J., Seabury, S. A., Lockwood, L. M., Goldman, D. P., Lakdawalla, D. N., and Cutler, D. M. (2010). Geographic variation in health care: The role of private markets. *Brookings Papers on Economic Activity*, Spring:325–361.

Rettenmaier, A. J. and Saving, T. R. (2009). Perspectives on the geographic variation in health care spending. Special Publication National Center for Policy Analysis.

Skinner, J. (2012). Causes and consequences of regional variations in health care. In Mark V. Pauly, T. G. M. and Barros, P. P., editors, *Handbook of Health Economics*, volume 2 of *Handbook of Health Economics*, pages 45–93. Elsevier.

Song, Y., Skinner, J., Bynam, J., Sutherland, J., and Fisher, E. (2010). Regional variations in diagnostic practices. *New England Journal of Medicine*, 363(1):45–53.

Subramanian, U., Weinberger, M., Eckert, G. J., L'Italien, G. J., Lapuerta, P., and Tierney, W. (2002). Geographic variation in health care utilization and outcomes in veterans with acute myocardial infarction. *Journal of General Internal Medicine*, 17(8):604–611.

Weinstein, J. N., Bronner, K. K., Morgan, T. S., and Wennberg, J. E. (2004). Trends and geographic variations in major surgery for degenerative diseases of the hip, knee, and spine. *Health Affairs*, 9:w81.

Zuckerman, S., Waldman, T., Berenson, R., and Hadley, J. (2010). Clarifying sources of geographic differences in Medicare spending. *New England Journal of Medicine*, 363(1):54–62.

Table 1: Summary Statistics

	(1)	(2)	(3)	(4)
	Original Sample[1]	Baseline Sample[1]	Non-Movers	Movers
	Share of Patients			
Female	0.57	0.57	0.57	0.60
White	0.85	0.86	0.86	0.88
Age First Observed:				
65 – 74	0.69	0.67	0.67	0.59
75 – 84	0.22	0.24	0.24	0.31
$\geq$ 85	0.09	0.09	0.09	0.09
Average Age First Observed	71.7	72.3	72.2	73.4
Average # of Years Observed	6.94	6.31	6.24	7.51
First Observed Residence:				
Northeast	0.20	0.20	0.20	0.17
South	0.36	0.39	0.38	0.41
Midwest	0.24	0.26	0.26	0.19
West	0.20	0.16	0.16	0.23
Have Chronic Condition[2]	0.41	0.50	0.50	0.51
Average # of Chronic Conditions[2]	1.35	1.65	1.65	1.62
Utilization[3]:				
Mean	$6,728	$7,211	$7,217	$7,117
S.D.	$10,407	$10,788	$10,883	$9,054
Share with Zero	0.12	0.07	0.07	0.07
Die During Sample	0.34	0.35	0.35	0.33
Years Missing Outcomes[4]	0.41	0.30	0.30	0.24
# of Patients	3,661,927	2,600,540	2,096,774	503,766
# of Patient Years	26,915,203	16,871,545	13,090,585	3,780,960

[1]Includes all movers and 25% random sample of non-movers aged 65 and over. Nonmovers receive a weight of 4.
[2]Measured in first year patient is observed in 1999 or later, since chronic conditions are first measurable in 1999.
[3]Patient-years enrolled in Medicare Advantage are included in the Original Sample but excluded from the utilization statistics. Calculated at patient-year level. (All other statistics report at patient-level).

[4]Years missing outcomes are all post-sample entry years through 2008 for which the patient is not in the sample and any years between 1998 and 2008 that the patient was enrolled in Medicare Advantage or not enrolled in Medicare Part A & B for all months eligible.

Table 2: Additive Decomposition of Log Utilization

	(1)	(2)	(3)	(4)	(5)	(6)
	Above / Below Median	Top & Bottom 25%	Top & Bottom 10%	Top & Bottom 5%	McAllen & El Paso	Miami & Minneapolis
Diff. in Average Log Utilization:						
Overall	0.279	0.458	0.667	0.819	0.590	0.556
Due to Place	0.161	0.283	0.422	0.468	0.339	0.386
Due to Patients	0.119	0.175	0.245	0.351	0.251	0.170
Share of Difference Due to						
Patients	0.426	0.382	0.368	0.429	0.425	0.306
	(0.031)	(0.026)	(0.029)	(0.029)	(0.159)	(0.091)
Place	0.575	0.618	0.632	0.571	0.575	0.694

Notes: This table reports the additive decomposition of the share of the difference in average log utilization across two sets of areas due to patients and due to place. Results are based on estimating equation (1); the dependent variable is log utilization, x_{it} includes indicator variables for age in five year bins, and individuals in the year of their move are omitted from the estimation. Each column defines a set of areas R and R'. The first row reports the difference in average utilization overall between the two areas $(\hat{y}_R - \hat{y}_{R'})$; the second row reports the difference due to place $(\hat{\gamma}_R - \hat{\gamma}_{R'})$; the third row reports the difference due to patients $(\hat{c}_R - \hat{c}_{R'})$. The fourth row reports the share of the difference in average utilization between the two areas due to patient $\left(\hat{S}_{pat}\left(R,R'\right)\right)$ which is simply the ratio of the third row to the first row. The last row reports the share of the difference in average utilization between the two areas due to place $\left(\hat{S}_{place}\left(R,R'\right)\right)$ which is the ratio of the second row to the first row. Standard errors (in parentheses) are calculated using a bootstrap procedure with 50 repetitions. In columns 1-4, the partitions of places shown in the columns are defined based on average utilization in each HRR; see text for more details. $N = 16,464,297$.

Table 3: Variance Decomposition of Log Utilization

	(1)	(2)
Cross-HRR Variance of Average:		
Log Utilization	0.034	
HRR Effects	0.022	
Patient Effects	0.015	
Covariance of Average		
HRR and Patient Effects	-0.002	(0.001)
Share Variance would be Reduced if:		
HRR Effects were Made Equal	0.552	(0.029)
Patient Effects were Made Equal	0.357	(0.040)

Notes: All results are based on estimating equation (1), using the same specification as in Table (2). The first three rows report variances of $\hat{y}_j$, $\hat{\gamma}_j$ and $\hat{c}_j$, respectively. The fourth row reports the covariance of $\hat{\gamma}_j$ and $\hat{c}_j$. The last two rows of the table report the share of the variance in cross-HRR utilization that would be reduced if HRR effects were made equal across areas $\left(\hat{S}^{var}_{place}\right)$ and if patient effects were made equal across areas $\left(\hat{S}^{var}_{pat}\right)$. Standard errors (in parentheses) are calculated using a bootstrap procedure with 50 repetitions. $N = 16,464,297$.

Table 4: Sensitivity of Additive Decomposition To Potentially Time Varying Place or Person Effects

	Specification	(1) Mean of log utilization	(2) Above / below median utilization difference	(3) Share due to patients	(4) N
(1)	Baseline	7.105	0.279	0.426	16,464,297
(2)	Baseline with second period interaction	7.105	0.279	0.428	16,464,297
(3)	Movers included only in sample half of move[a]	7.102	0.280	0.464	14,630,538
(4)	First half of sample only (1998-2003)	6.930	0.279	0.502	7,915,918
(5)	Second half of sample only (2004-2008)	7.304	0.294	0.532	6,714,620
(6)	Movers included only in sample third of move[b]	7.097	0.280	0.446	14,081,978
(7)	First third of sample only (1998-2001)	6.842	0.279	0.556	4,958,756
(8)	Second third of sample only (2002-2005)	7.163	0.286	0.392	5,370,059
(9)	Third third of sample only (2006-2008)	7.342	0.302	0.594	3,753,163

Notes: This table reports our estimate of the share of the difference in utilization between above and below median HRRs due to patients, based on estimating equation (1) (or a variant on it) on the sample shown in the different rows. Column (1) reports the mean of log utilization for the sample analyzed in a given row. Column (2) reports the difference in average utilization between above and below median utilization HRRs $(\hat{y}_R - \hat{y}_{R'})$. Column (3) reports the share of the difference in column (2) that is due to patients $\left(\hat{S}_{pat}(R, R')\right)$. Column (4) reports the sample size in patient-years. Row 1 contains our baseline estimate (previously shown in the first column of table 2) based on estimating equation (1). Row 2 is based on estimating equation (7) which includes a second period (2004-2008) interaction term with the HRR fixed effects. All other rows are based on estimating equation (1). Rows 4 and 5 are mutually exclusive and exhaustive subsets of Row 3: row 4 includes only the first time period - 1998-2003 - and movers who moved during this period; row 5 includes only the second time period (2004-2008) and patients who move during that time period. For these three rows, the partition of HRRs into above and below median utilization is based on the utilization of movers and non-movers in the relevant row. Rows 7 through 9 are estimated in an analogous fashion to rows 4 and 5, except now with three time periods (1998-2001, 2002-2005, 2006-2008).

[a]We partition the sample into two time periods and limit observations on movers to the move year and all other years for that mover that fall within the same time period (1998-2003) or (2004-2008). For example, if a patient moved in 2000, only years 1998-2003 (if available) would be included for this mover.

[b]Sample limits observations on movers to the the move year for movers and all other years for that mover that fall within the same time period (1998-2001) or (2002-2005) or (2006-2008). For example, if a patient moved in 2000, only years 1998-2001 (if available) would be included for this mover.

Table 5: Sensitivity of Additive Decomposition to Handling of Attrition

	Specification	(1) Mean of log utilization	(2) Above / below median utilization difference	(3) Share due to patients	(4) N
(1)	Baseline	7.105	0.279	0.426	16,464,297
(2)	Patients who never die	6.979	0.292	0.521	11,023,865
(3)	Patients never in an HMO	7.129	0.276	0.437	13,889,323
(4)	Patients never in an HMO and never die	7.004	0.290	0.543	9,072,339
(5)	Patients never missing outcomes	7.040	0.281	0.525	8,323,365
(6)	Early Moves	7.093	0.280	0.544	13,482,761
(7)	Middle Moves	7.095	0.280	0.418	13,533,057
(8)	Late Moves	7.094	0.280	0.513	13,605,281

Notes: This table reports our estimate of the share of the difference in utilization between above and below median HRRs due to patients, based on estimating equation (1) for different restrictions of our baseline sample, as shown in the different rows. The partition of HRRs into above and below median utilization is based on the utilization of individuals in the sample specified in the row. Column (1) reports the mean of log utilization for the given sample. Column (2) reports the difference in average utilization between above and below median utilization HRRs $(\hat{y}_R - \hat{y}_{R'})$. Column (3) reports the share of the difference in column (2) that is due to patients $\left(\hat{S}_{pat}\left(R,R'\right)\right)$. Column (4) reports the sample size in patient-years. Row 1 contains our baseline estimate (previously shown in the first column of table 2). Row 2 restricts the baseline sample to patients who are never observed to die during the course of our study. Row 3 restricts the baseline sample to patients who are never in an HMO. Row 4 restricts the baseline sample to patients who are both never in an HMO and never die during the course of our study. Row 5 includes movers only if they are observed continuously in each relative year in [-1,7] and all non-movers; Row 6 includes movers only if they are observed continuously in each relative year [-4,4] and all non-movers; Row 7 includes movers only if they are observed continuously in each relative year [-7.1] and all non-movers. The event study analogues of rows 6 through 8 can be found in Figure 6.

Table 6: Sensitivity of Additive Decomposition to Alternate Specifications

	Specification	(1) Mean of log utilization	(2) Above / below median utilization difference	(3) Share due to patients	(4) N
(1)	Baseline	7.105	0.279	0.426	16,464,297
(2)	Movers Only	7.182	0.287	0.422	3,373,712
(3)	Drop age as a covariate	7.105	0.279	0.384	16,464,297
(4)	Ln(Total Expenditure)	7.080	0.286	0.450	16,464,297
(5)	Relative years -5 to 5	7.108	0.280	0.424	15,850,600
(6)	Relative years -3 to 3	7.105	0.280	0.447	15,093,787
(7)	Relative years -2 to 2	7.102	0.280	0.449	14,551,943
(8)	Relative years -1 to 1	7.097	0.281	0.474	13,888,046
(9)	Ln(Utilization+0.1)	6.937	0.323	0.438	16,464,297
(10)	Ln(Utilization+10)	7.288	0.239	0.376	16,464,297
(11)	First differences	7.119	0.278	0.453	16,871,545
(12)	First differences with fixed effects	7.119	0.278	0.517	16,871,545

Notes: This table reports our estimate of the share of the difference in utilization between above and below median HRRs due to patients, based on estimating equation (1) for alternative specifications. The partition of HRRs into above and below median utilization is based on the utilization of individuals in the sample specified in the row (which differs from the baseline in row 2 and rows 5-8) and the definition of utilization used (which differs from the baseline in row 4). Rows 11 and 12 use the baseline partition and utilization definition. Column (1) reports the mean of the outcome for the given sample. Column (2) reports the difference in average utilization between above and below median utilization HRRs $(\hat{y}_R - \hat{y}_{R'})$. Column (3) reports the share of the difference in column (2) that is due to patients $\left(\hat{S}_{pat}(R,R')\right)$. Column (4) reports the sample size in patient-years. The outcome is log utilization in every row except for row 4 where it is log expenditures. Rows 5 to 8 narrow the sample of movers used for estimation to relative years -5 to 5, relative years -3 to 3, relative years -2 to 2, and relative years -1 to 1, respectively. Row 11 estimates the equation (1) in first differences, and row 12 estimates equation (1) in first differences adding patient and HRR fixed effects to allow for differential trends. In rows 11 and 12 we do not drop the year of the move (relative year 0), but use the adjustment technique described in C.4.1, assuming that there is no misreporting in move timing other than a 50% chance of a patient being in their origin in relative year 0 and a 50% chance a patient being in their destination in relative year 0.

Table 7: Sensitivity of additive decomposition to restrictions on mover geography

	Sample	(1) Mean of log utilization	(2) Above / below median utilization difference	(3) Share due to patients	(4) N (% of movers retained)
(1)	Baseline	7.105	0.279	0.426	16,464,297 (100%)
(2)	Cross state movers only	7.103	0.280	0.411	15,381,739 (68%)
(3)	Cross census region movers only	7.098	0.281	0.436	14,354,315 (37.9%)
(4)	Drop moves to FL	7.100	0.279	0.433	16,064,008 (87.8%)
(5)	Drop moves to FL, AZ, and CA	7.102	0.280	0.392	15,665,458 (75.1%)

Notes: This table reports our estimate of the share of the difference in utilization between above and below median HRRs due to patients, based on estimating equation (1) for alternative specifications. The partition of HRRs into above and below median utilization is based on the utilization of individuals in the sample specified in the row. We retain all non-movers in every row. The baseline sample has $503,766$ movers. Column (1) reports the mean of the outcome for the given sample. Column (2) reports the difference in average utilization between above and below median utilization HRRs $(\hat{y}_R - \hat{y}_{R'})$. Column (3) reports the share of the difference in column (2) that is due to patients $\left(\hat{S}_{pat}(R,R')\right)$. Column (4) reports the sample size in patient-years.

Table 8: Additive Decomposition Results by Sub-Samples

	(1) Mean of log utilization	(2) Mean of stratification variable	(3) Above / below median utilization difference	(4) Share due to patients	(5) N
Baseline	7.105	.	0.279	0.426	16,464,297
# of Chronic Conditions					
Above Median	8.021	4.80	0.135	-0.027	8,142,854
Below Median	6.215	1.63	0.376	0.607	8,230,964
# of Hospital Days					
Above Median	7.845	4.84	0.171	0.095	8,197,162
Below Median	6.372	0.21	0.338	0.597	8,267,135
Age Quartiles					
First	6.733	68.1	0.383	0.583	4,399,076
Second	7.079	72.8	0.288	0.409	3,932,299
Third	7.315	78.2	0.271	0.562	4,068,121
Fourth	7.325	86.3	0.249	0.271	4,064,801
Utilization Group					
Above Median	8.070	11301	0.137	-0.075	8,213,922
Below Median	6.145	1985	0.322	0.561	8,250,375

Notes: This table reports our estimate of the share of the difference in utilization between above and below median HRRs due to patients, based on estimating equation (1) for different stratifications of our baseline sample, as shown in the different rows. Stratification groups were determined by patient-level averages over all observed years. The partition of HRRs into above and below median utilization is based on the utilization of individuals in the sample specified in the row. Column (1) reports the mean of log utilization and column (2) reports the mean of the stratification variable for a given sample. Column (3) reports the difference in average log utilization across the two sets of HRRS $(\hat{y}_R - \hat{y}_{R'})$, and column (4) reports the difference in average utilization across the two sets of HRRS that is due to patients $\left(\hat{S}_{pat}\left(R, R'\right)\right)$. Row 1 replicates our baseline results (previously shown in the first column of table 2). The subsequent rows are all stratifications of the baseline sample and include both movers and non-movers.

Table 9: Equalizing Patient Observables

	Patient Observable	(1) Share due to patient observables	(2) Share due to patients	(3) Fraction of patient share
(1)	Age	0.035	0.426	0.082
(2)	Age, Race, Sex	0.057	0.426	0.134
(3)	Ln(HCC Score)	0.405	0.426	0.953
(4)	Ln(# of Chronic Conditions+1)	0.753	0.434	1.737
(5)	Ln(# of Chronic Conditions+1), Age, Race, Sex	0.789	0.434	1.819

Notes: The areas R and R' are defined as above/below median HRRs in terms of mean log utilization. In Column 1, we report the share of the difference in average log utilization between these areas due to patient observables $\left(\hat{S}_{pat}^{obs}(R,R')\right)$ and in column 2 we report the share due to patients $\left(\hat{S}_{pat}(R,R')\right)$. The rows contain various combinations of patient observables that we want to equalize across HRRs. Ln(HCC Score) refers to the natural log of Hierarchical Condition Category (HCC) risk score for a given patient and year. Ln(# of Chronic Conditions +1) refers the natural log of how many of CMS's 27 chronic conditions a patient has been diagnosed with (within a 1-3 year lookback window) by the end of a year. These data are only available from 1999, so estimates in Rows (4) and (5) are based on years 1999-2008 only. The age, race and sex dummies are fully interacted in rows (2) and (5). If multiple patient characteristics are listed, each characteristic is included as a regressor in Equation (8) and $\left(\hat{S}_{pat}^{obs}(R,R')\right)$ refers to the share of the difference between high and low utilization places that would go away if we equalized all given patient observables.

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)	Ln(HCC Score)	-0.123	0.114	0.056	0.058	0.492	0.508
(2)	Ln(# of Chronic Conditions+1)	1.204	0.139	0.060	0.079	0.433	0.567

Notes: This table reports the additive decomposition of the share of the difference in the average outcome across two sets of areas due to patients and due to place. HRRs are partitioned into above and below median places based on the average of the given outcome (by row) in each HRR. Results are based on estimating equation (1); the dependent variable is the given outcome, x_{it} includes indicator variables for age in five year bins, and individuals in the year of their move are omitted from the estimation. The second column reports the difference in average outcome overall between the two areas $(\hat{y}_R - \hat{y}_{R'})$; the third column reports the difference due to patients $(\hat{c}_R - \hat{c}_{R'})$; the fourth column reports the difference due to place $(\hat{\gamma}_R - \hat{\gamma}_{R'})$. The fifth column reports the share of the difference in average outcome between the two areas due to patient $\left(\hat{S}_{pat}(R,R')\right)$, which is simply the ratio of the third column to the second column. The last column reports the share of the difference in average outcome between the two areas due to place $\left(\hat{S}_{pat}(R,R')\right)$ which is the ratio of the fourth column to the second column. In Row (1), $N = 16,464,297$ and in Row (2) $N = 14,996,992$ as we do not observe chronic conditions in 1998.

Table 11: Geographic Variation in Log Utilization Explained by Health

	Outcome	(1) Mean of Outcome	(2) Share due to patient observables	(3) Share due to patients	(4) Fraction of patient share	Outcome	(5) Share due to place observables	(6) Share due to place	(7) Fraction of place share
	Patient Component of:					Place Component of:			
(1)	Ln(HCC Score)	-0.123	0.190	0.426	0.447	Ln(HCC Score)	0.215	0.575	0.374
(2)	Ln(# of Chronic Conditions+1)	1.204	0.363	0.434	0.836	Ln(# of Chronic Conditions+1)	0.390	0.566	0.689

Notes: The areas R and R' are defined as above/below median HRRs in terms of average utilization. The first column shows the overall patient-year mean of the whole outcome (not just the patient or place component). In Column 2, we report the share of the difference in average log utilization between these areas due to patient observables $\left(\hat{S}^{obs}_{pat}(R,R')\right)$ and in column 3 we report the share due to patients $\left(\hat{S}_{pat}(R,R')\right)$. The fourth column contains the fraction of patient share explained by patient observables, which is column (2) divided by column (3). Columns 5, 6 & 7 repeat this analysis for place. The rows contain various combinations of patient observables that we want to equalize across HRRs.

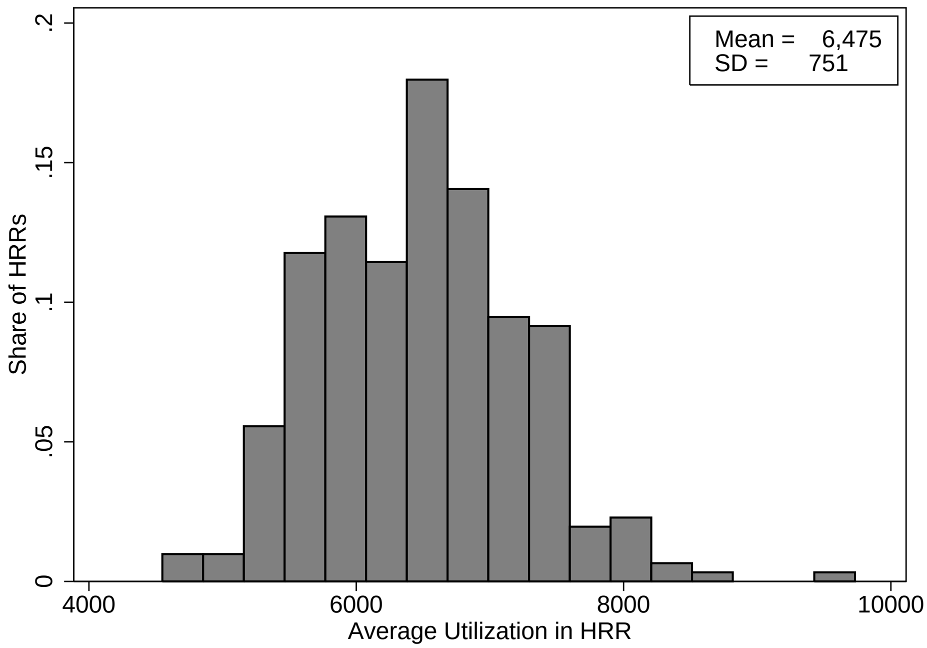
Table 12: Post-Move Mortality by Destination of Move

		(1)	(2)	(3)	(4)
	Mean Utilization in	-0.0109	-0.0182	-0.0097	-0.0067
All Post Move Years	Destination HRR	(0.0009)	(0.0010)	(0.0010)	(0.0011)
	N	1,927,616	1,927,616	1,927,616	1,624,411
	Mean Utilization in	-0.0098	-0.0163	-0.0107	-0.0084
Relative Year 0	Destination HRR	(0.0015)	(0.0017)	(0.0017)	(0.0017)
	N	407,248	407,248	407,248	368,763
	Mean Utilization in	-0.0099	-0.0198	-0.0097	-0.0058
Relative Year 1	Destination HRR	(0.0022)	(0.0025)	(0.0025)	(0.0026)
	N	382,563	382,563	382,563	342,693
	Mean Utilization in	-0.0144	-0.0207	-0.0088	-0.0096
Relative Year 5	Destination HRR	(0.0034)	(0.0039)	(0.0038)	(0.0042)
	N	143,163	143,163	143,163	113,493
Controls for:					
Origin HRR		N	Y	Y	Y
Move Year, Demographics		N	N	Y	Y
Pre-Move Health		N	N	N	Y

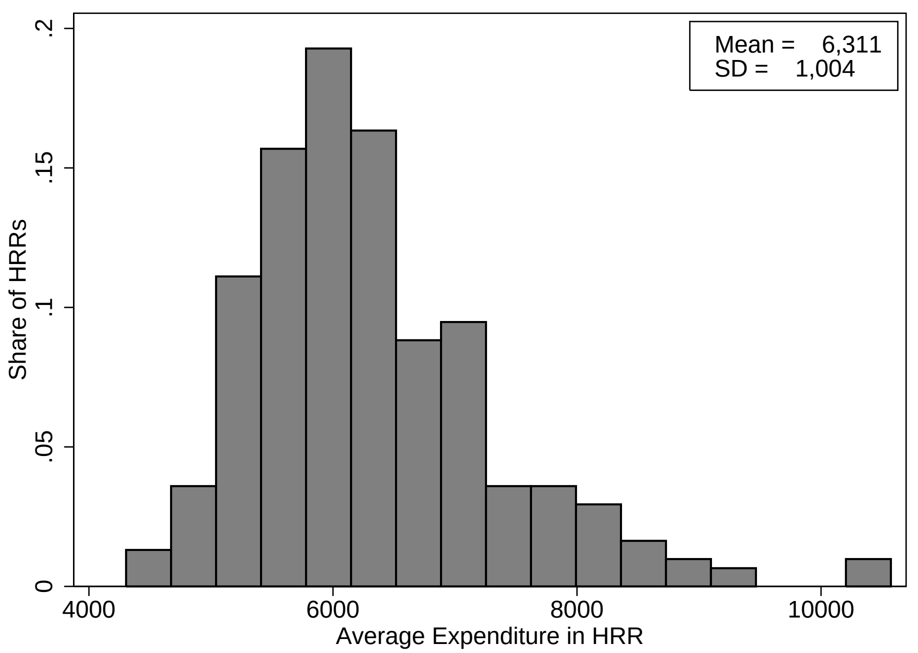
Notes: This table shows the relationship between post-move mortality and average utilization in the mover's destination HRR. The dependent variable is a binary flag for whether the mover died by the end of the relative year. We show the coefficient on the average utilization in that mover's destination HRR. Origin HRR controls are a set of dummies for each HRR of origin. Move year controls include dummies for each calendar move year. Demographic controls are fully interacted dummies for each 5-yr age bin, white, non-white, female and male. Pre-move health is the pre-move average value for each of 27 chronic condition dummy variables.

Figure 1: Variation Across HRRs in Average Utilization and Spending

(a) Average Utilization



(b) Average Total Expenditure



Figures display average utilization (panel a) and average expenditures (panel b) across HRRs. Specifically, we compute average utilization across every individual i in HRR j in year t (omitting movers in the year of their move and weighting each non-mover by 4) and then average over years to get the HRR average that we report in the figure.

Figure 2: Distribution of “size” of move ($\hat{\delta}_i$) among movers

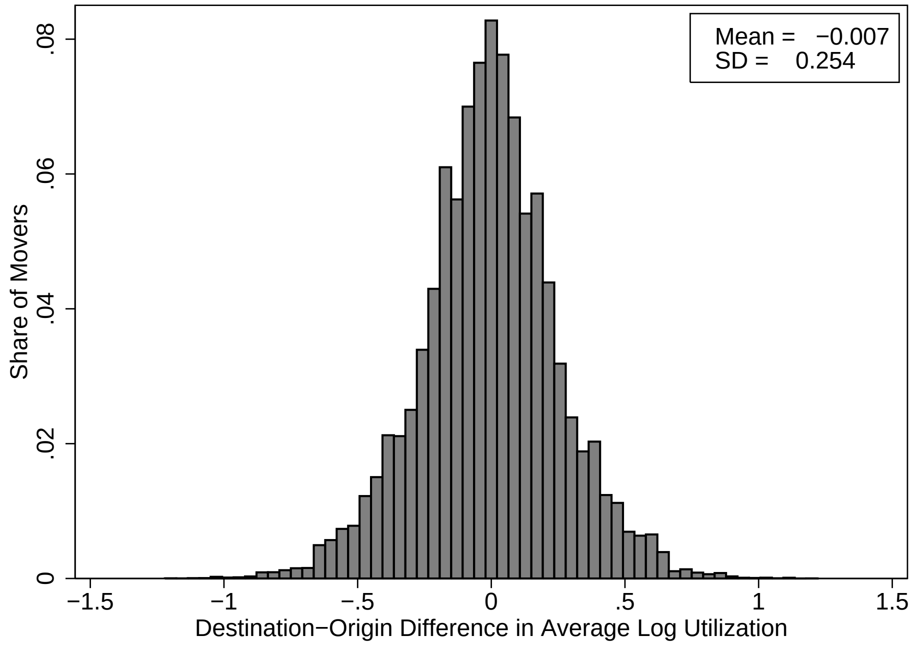
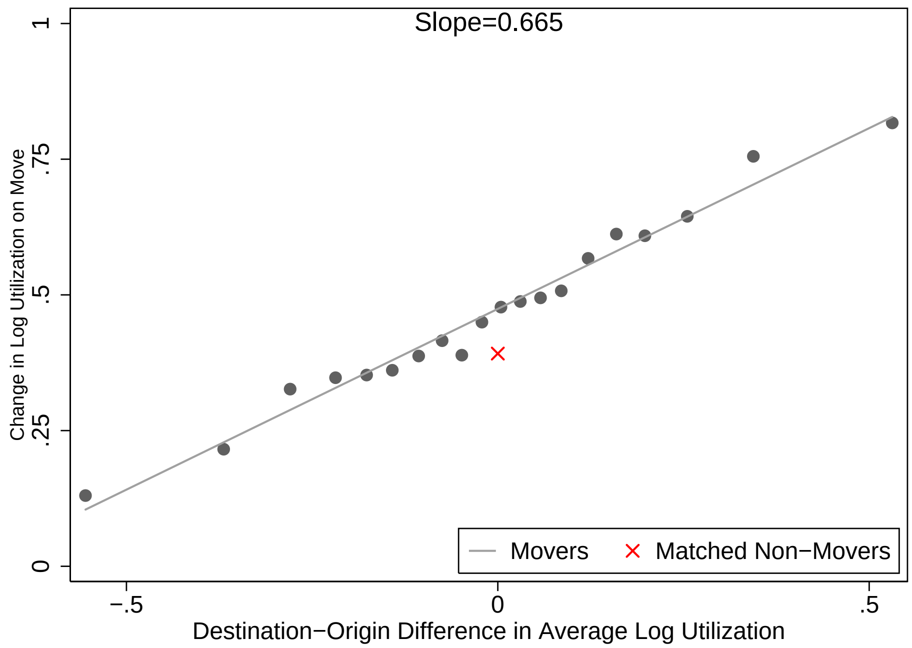


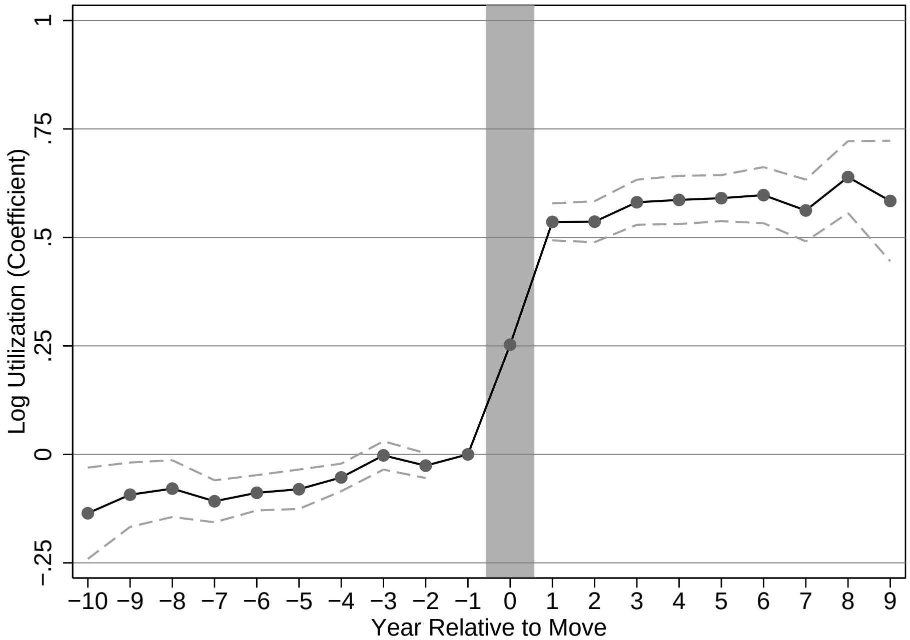
Figure shows a histogram of the average difference in log utilization between a mover’s origin and destination HRRs $\left(\hat{\delta}_i = \hat{y}_{d(i)} - \hat{y}_{o(i)}\right)$. $N = 503,766$ movers.

Figure 3: Change in Log Utilization By Size of Move.



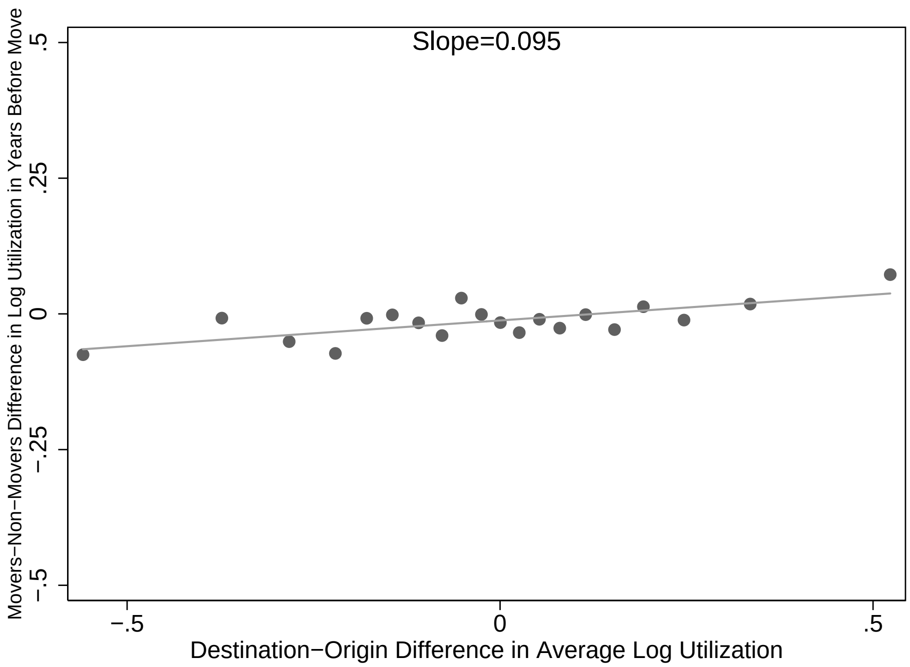
This figure displays a binned scatterplot that characterizes the change in log utilization before and after move. Sample consists of all movers for whom we observe utilization in at least one of relative years [-5,-2] as well as in at least one of relative years [2, 5]. For each mover, we calculate the size of the move,$\hat{\delta}_i = \hat{y}_{d(i)} - \hat{y}_{o(i)}$,i.e. the average difference in log utilization between a mover's destination and origin HRR; we group $\hat{\delta}_i$ into ventiles (20 equal-sized bins). The x-axis displays the mean of $\hat{\delta}_i$ for movers in each ventile. The y axis shows the difference, for each ventile, of the average log utilization for mover-years in relative years [2,5] minus average log utilization for mover-years in years [-5,-2]. The line of best fit is obtained from simple OLS regression of the y-variable on the x-variable using the 20 data points corresponding to movers. The slope of the line is reported as on the graph. For comparison, we also compute the average change in log utilization for a matched non-mover sample, which we show with the red cross on the graph. Specifically, for each mover-year [-5,-2] we randomly draw a non-mover in the mover's origin HRR who is of the same gender, race (white or not), calendar year, and five year age bin as the mover; we do the same thing for each mover-year [2,5] except now drawing from the mover's destination HRR; we then report the difference of the average log utilization for the matched non-mover sample in relative years [2,5] minus the average utilization for this matched non-mover sample in relative years [-5,-2].

Figure 4: Event-study Analysis of Log Utilization



The solid line in the figure connects the estimated coefficients on the $\tilde{\theta}_{r(i,t)}\hat{\delta}_i$ terms from the estimation of Equation (5). These are the interactions of indicator variables for year relative to move year ($\tilde{\theta}_{r(i,t)}$)and the "size of the move ($\hat{\delta}_i = \hat{y}_{d(i)} - \hat{y}_{o(i)}$), i.e. the average difference in log utilization between a mover's destination and origin HRR. We normalize the coefficient on the term corresponding to the year immediately prior to the move (relative year -1) to 0. The dependent variable y_{it} is the log of (utilization+1). The regression also controls for a series of indicator variables for year relative to move ($\tilde{\theta}_{r(i,t)}$), patient fixed effects ($\tilde{\alpha}_i$), calendar year fixed effects (τ_t), and indicator varies for five year age bins. The dashed lines capture the upper and lower bounds of the 95% Confidence Interval of the estimated coefficients on the $\tilde{\theta}_{r(i,t)}\hat{\delta}_i$ terms. We construct these confidence interval using a 2-step procedure. In the first step, for each HRR j, we construct the asymptotic distribution of $\bar{y}_j$, which is a normal distribution with mean μ_j and standard deviation σ_j calculated from the data. In the second step, we bootstrap equation (5) with 50 repetitions drawn at the patient level, making a random draw from the distribution of $\bar{y}_j$for each mover's origin and destination to construct their $\hat{\delta}_i$ for each repetition. $N = 3,780,960$ patient-years (movers only).

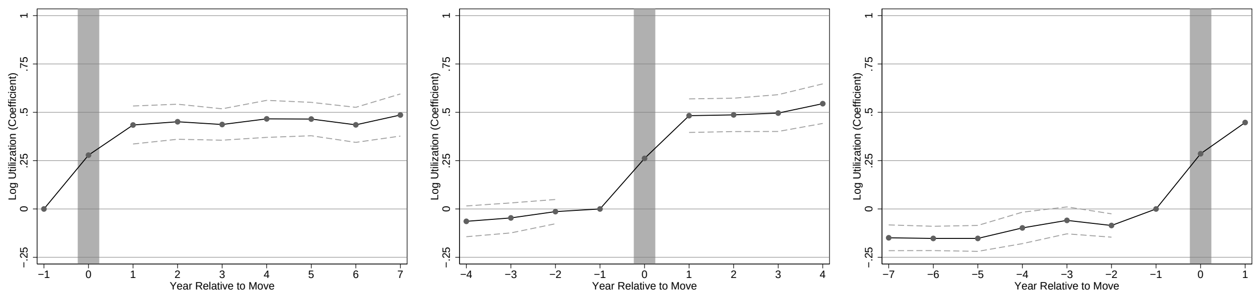
Figure 5: Pre-move Log Utilization of Movers Relative to Non-Movers, By Size of Move.



This figure displays a binned scatterplot that characterizes the pattern of movers' log utilization prior to movers by the size of their subsequent move. Sample consists of all movers for whom we observe utilization in at least one year of relative years [-5,-2]. For each mover, we calculate the size of the move ($\hat{\delta}_i = \hat{y}_{d(i)} - \hat{y}_{o(i)}$) i.e. the average difference in log utilization between a mover's destination and origin HRR; we group $\hat{\delta}_i$ into ventiles (20 equal-sized bins). The x-axis displays the mean of $\hat{\delta}_i$ for movers in each ventile. The y axis shows for each ventile the average of the difference in log utilization between mover and matched non-mover patient-years in years [-5,-2]. The matched non-mover sample is defined by selecting, for each mover, a non-mover in mover's origin HRR who is the same as the mover in terms of 5-year age bin, birth cohort, sex, and calendar year. The line of best fit is obtained from simple OLS regression of y-variable on x-variable using the 20 data points corresponding to movers . The slope of the line is reported as on the graph.

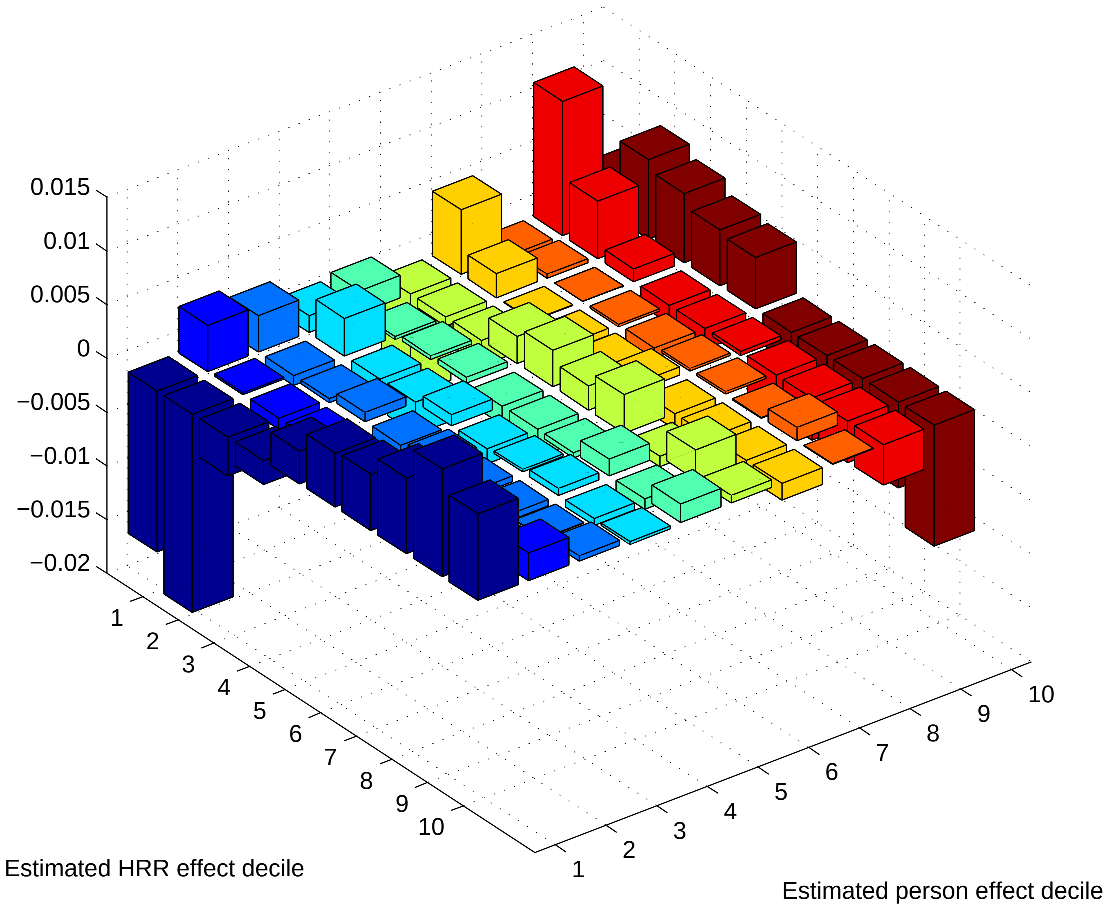
Figure 6: Balanced-Panel Event-study Analyses of Log Utilization

(a) Early Moves. (b) Middle Moves. (c) Late Moves.



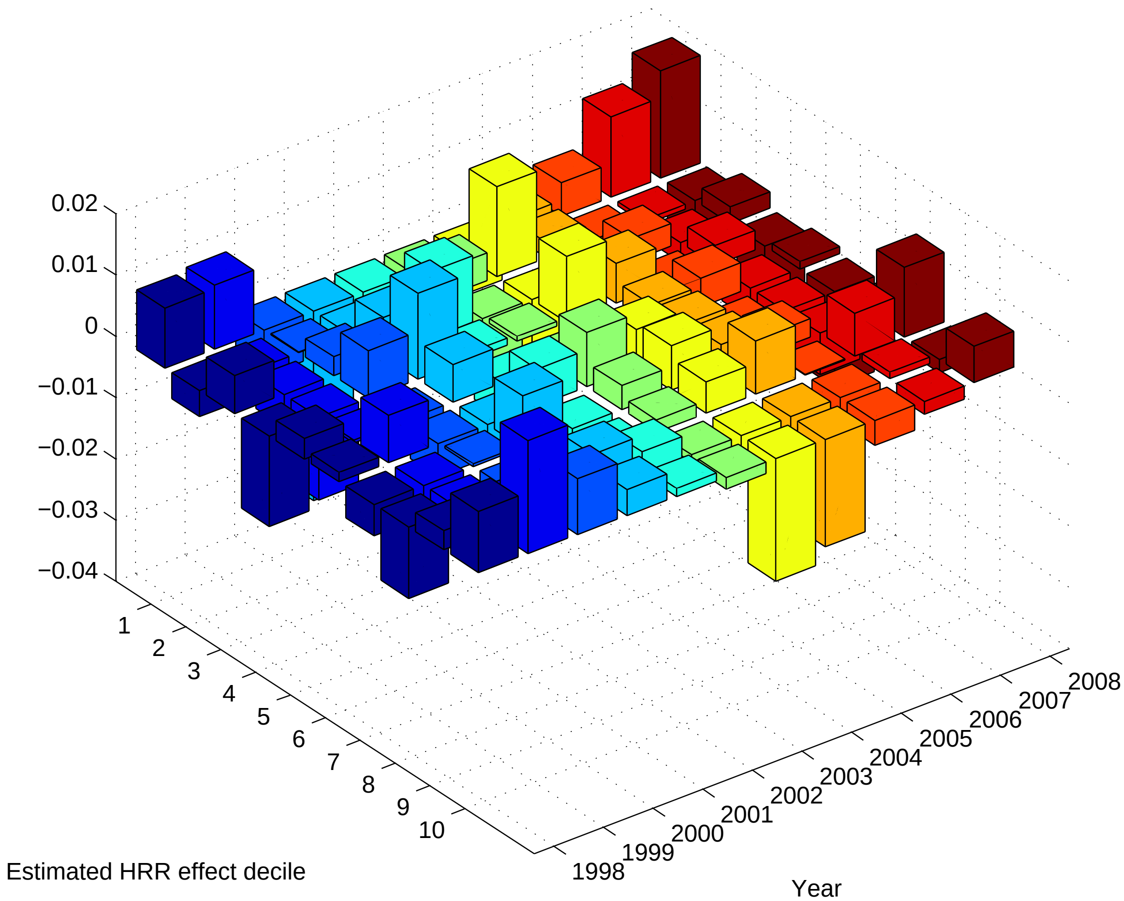
These figures are constructed in the same manner as in Figure 4 above, except they are estimated on balanced panel sub-samples of movers whom we observe in each of a given set of relative years. Panel (a) restricts to movers whom we observe in every relative year in [-1,7] (this consists of $49,022$ patients per year, so $N = 441,198$ patient-years).Panel (b) restricts to movers whom we observe in every relative year [-4,4] (this consists of $55,309$ patients for year, so $N = 497,781$ patient years). Panel (c) restricts the sample whom we observe in every relative year in [-7,1] (this consists of $64,337$ patients per year, so $N = 579,033$ patient years). The dashed lines show the upper and low 95% confidence interval of the estimated $\hat{\theta}_r$, constructed using the same bootstrap approach as in Figure 4.

Figure 7: Mean Residuals by Estimate Person Effect Decile and HRR Effect Decile



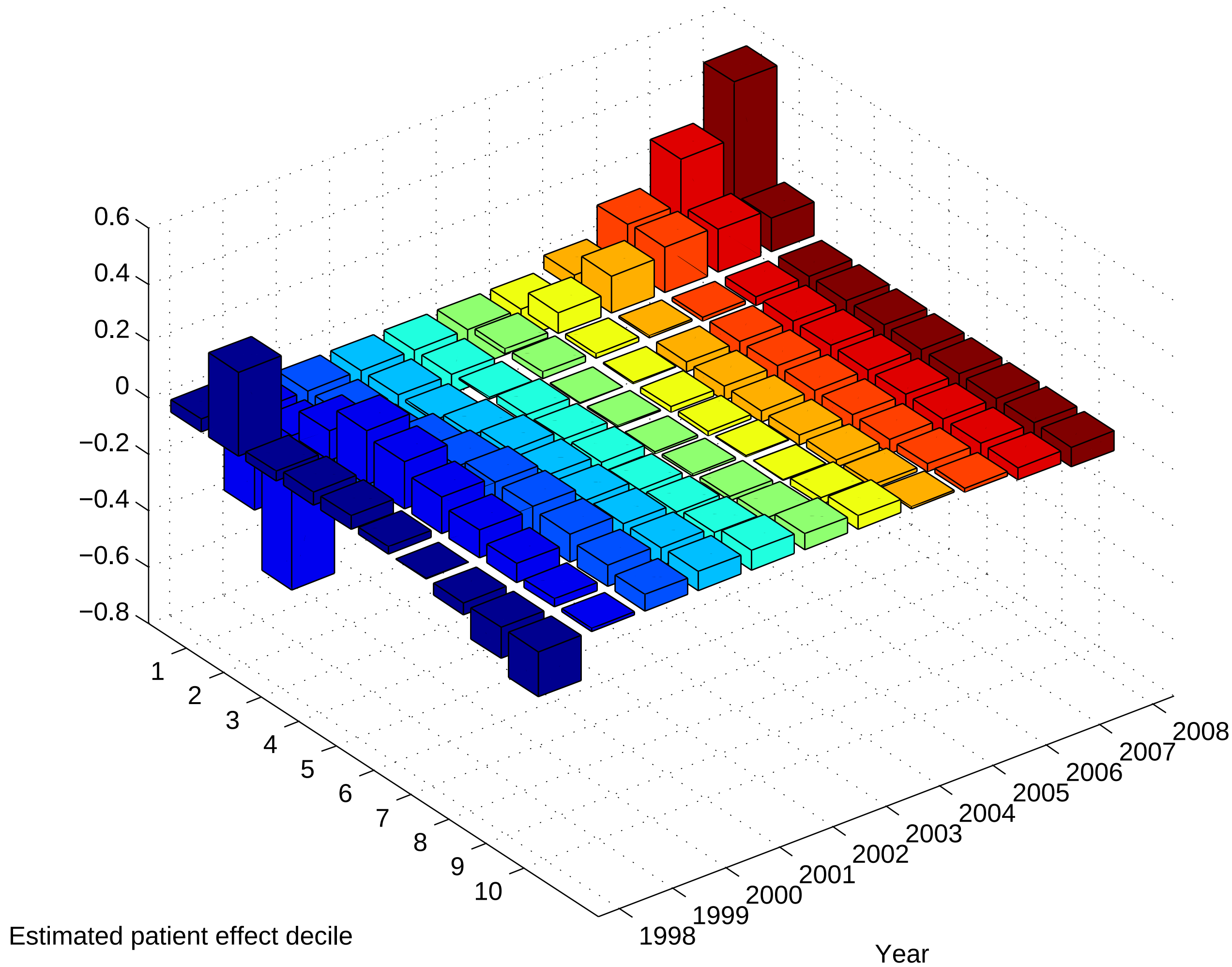
This figure displays the mean residuals from equation(1) by deciles of estimated person and HRR effects. After estimating equation (1) and predicting the residuals for each patient-year observation, we assign each observation to one of the 100 cells defined by the 10 deciles of estimated patient effects ($\hat{c}_{it}$) and the 10 deciles of estimated HRR effects($\hat{\gamma}_j$), and compute the mean residual for each cell... $N = 16,464,297$ patient-years.

Figure 8: Mean Residuals by Estimated HRR Effect Decile and Calendar Year



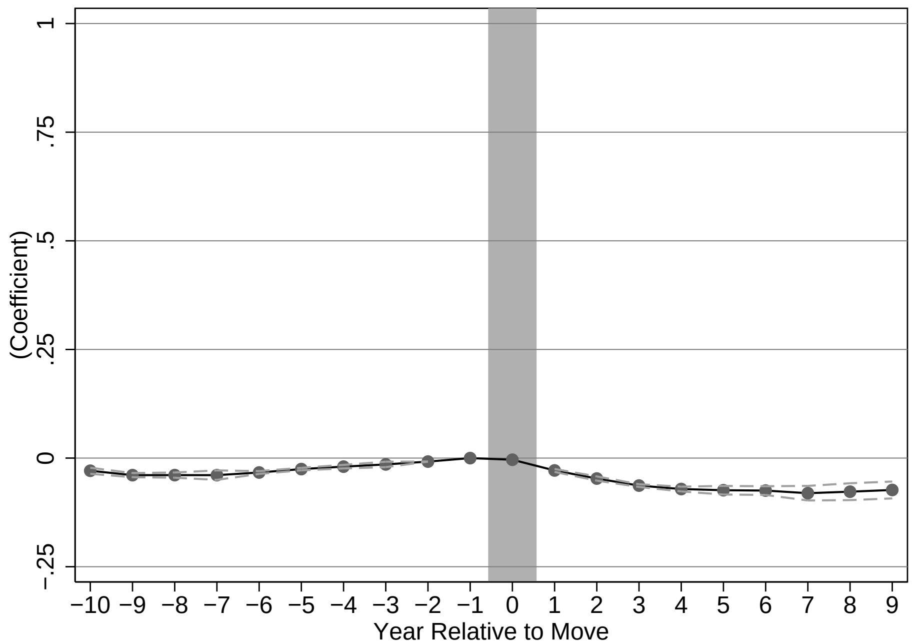
This figure displays the mean residual from estimation of equation (1) for each of the 110 cells generated by the 10 deciles of estimated HRR effects $\left(\hat{\gamma}_j\right)$ and the 11 calendar years in the data. $N = 16,464,297$ patient-years.

Figure 9: Mean Residuals by Person Effect Decile and Year



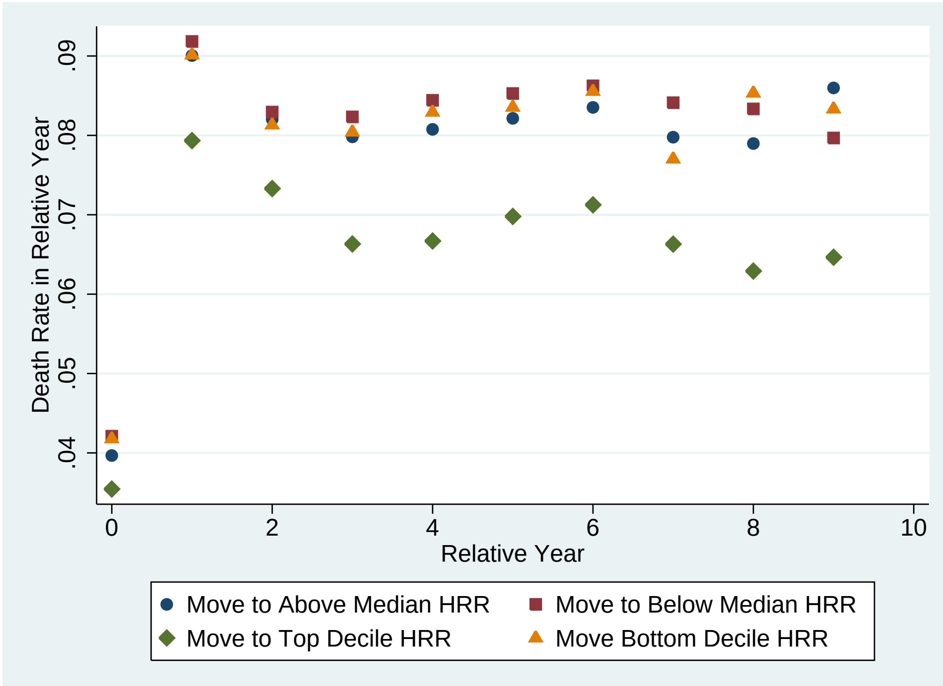
This figure displays the mean residual from estimation of equation (1) for each of the 110 cells generated by the 10 deciles of estimated patient effects ($\hat{c}_{it}$) and the 11 calendar years in the data. $N = 16,464,297$ patient-years.

Figure 10: Event-study Analysis of Missing Utilization Indicator



The solid line in the figure connects the estimated coefficients on the $\tilde{\theta}_{r(i,t)}\hat{\delta}_i$ terms from the estimation of equation (5) where the dependent variable is now an indicator for whether the patient-year observation is missing the utilization measure. The dashed lines show the 95% Confidence Interval. The sample is our baseline sample of movers, augmented to include patient-years for these movers that are missing outcomes during our 1998-2009 sample period. Patient-years with missing outcomes include all post-sample entry years through 2008 for which the patient is not in the sample (due to death, enrollment in Medicare Advantage, or not being enrolled in Medicare Part A & B for all months that they were eligible) and any years between 1998 and 2008 that the patient was over 65 and on Medicare (i.e. eligible for our sample) but was enrolled in Medicare Advantage or not enrolled in Medicare Part A & B for all months eligible. Other than the different dependent variable, the figure is constructed in the same manner as Figure 4. $N = 5,541,426$patient-years (movers only).

Figure 11: Post-Move Mortality By Destination of Move



This figure plots mean mortality rates for different sets of movers in post-move years. The sets are defined based on the average utilization in that mover's destination HRR. The death rate is defined as the fraction of movers observed in a given year who die during that year.

Appendix

A Utilization Measure

A.1 Overview

We construct our health care utilization measure by aggregating care provided to Medicare beneficiaries as recorded in the inpatient, outpatient, and carrier claims data. Loosely, 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 files record payments to physicians and other non-institutional providers (such as independent ambulance providers). Importantly, this 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). While recent work (Newhouse and Garber 2013) has suggested there is substantial variation in these additional measures of care, we follow much of the earlier literature in focusing on measuring care provided in the inpatient, outpatient, and carrier files.

To focus on a "quantity" measure of utilization - so that geographic variation in our measure of utilization will not reflect geographic variation in prices - we follow the methodology of Gottlieb et al. (2010).[20] Section A.2 describes how we construct this price-adjusted utilization measure. In addition, to explore our aggregate estimates in more detail we construct a number of disaggregated measures of utilization - measuring, e.g., primary care visits and mammograms; we describe the construction of these disaggregated measures in Section A.3.

A.2 Construction of utilization measure

The price adjustment procedures of Gottlieb et al. (2010) is specific to the types of claims examined, so we here separately describe our price adjustment procedure for inpatient, outpatient, and carrier claims.

A.2.1 Inpatient claims

A 2009 MedPac report summarizes the reimbursement rule for inpatient services.[21] 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

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

[21]See http://www.medpac.gov/documents/medpAc_payment_Basics_09_hospital.pdf.

a relative weight that aims to reflect "the expected relative costliness of inpatient treatment for patients in that group." DRG weights are set annually, and payments are determined as follows.

First, Medicare sets so-called "standardized payment amounts" per discharge. These base payment accounts are meant to capture "the operating and capital costs that efficient facilities would be expected to incur in furnishing covered inpatient services." 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.

Third, the wage-adjusted base payment is adjusted for case mix using the DRG relative weights.

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 utilization for individual i's receipt of procedure k in region j at time t is estimated as:

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

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

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

. 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:

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

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

When we examine physicians in the inpatient claims data, we rely on the attending physician ID.

A.2.2 Carrier claims

A 2012 MedPac report summarizes the reimbursement rule for carrier claims.[22] Carrier claim based reimbursements are centered around Healthcare Common Procedure Coding System (HCPCS) codes for specific services. 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.[23]

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 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.

When we examine physicians in the carrier claims data, we rely on the referring physician ID.

[22] See http://www.medpac.gov/documents/MedPAC_Payment_Basics_12_Physician.pdf.

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

A.2.3 Outpatient claims

A 2010 MedPac report summarizes the reimbursement rule for outpatient claims.[24] Like inpatient services, outpatient payments are covered by a prospective payment system over the time period of our data. 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.[25]Second, adjustments are made for cancer hospitals, children's hospitals, rural hospitals with 100 or fewer beds, and sole community hospitals. Finally, oulier 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

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

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

When we examine physicians in the outpatient claims data, we rely on the attending physician ID.

A.3 Disaggregated measures of utilization

To explore our aggregate utilization estimates in more detail, we construct a number of disaggregated measures:

- **Physician types**. Our definitions of "primary care physicians" and "specialist physicians" follow the Dartmouth Atlas.[26] Specifically, all physicians in the AMA Masterfile except primary physicians and those with unspecified specialties are classified as specialists.

- **Ambulatory care visit to a primary care physician**. Our definition of ambulatory care visits to primary care physicians follows the Dartmouth Atlas.[27] Specifically, this variable is defined as a CPT code of 99201-99205, 99211-99215, 99381-99387, 99391-99397, 99241-99245, or 99271-99275; a place of service of office (place of service code 11), outpatient

[24] See http://www.medpac.gov/documents/medpac_payment_basics_10_opd.pdf.

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

[26] See http://www.dartmouthatlas.org/data/table.aspx?ind=137.

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

hospital (22), rural health clinic (72) or federally qualified health center (50); and a physician specialty of general practice (specialty code 1), family practice (8), internal medicine (11), pediatrician (37), nurse practitioner (50), physician assistant (97) or clinic (70). This also includes any visit to a rural health center (RHC) or federally qualified health center (FQHC) recorded in the outpatient file.

- **Diagnostic and imaging tests**. Our definition of diagnostic and imaging tests follows Song et al. (2011), and is based on BETOS codes: codes beginning with T are diagnostic tests, and codes beginning with I are imaging tests.

- **Outpatient procedures**. We code several outpatient procedures which are broadly categorized as "preventive care" in our variable measuring any preventive care:

  - Mammogram is defined following the Dartmouth Atlas.[28] 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.[29]
  - Cardiovascular screening blood testing, diabetes self-management training, bone mass measurements, colorrectal cancer screening, pap smears, pelvic examinations, and prostate cancer screening are defined following CMS's preventive care definitions.[30] Note that for colorrectal cancer screening, we use CPT code 82270 and HCPCS code G0328.

[28] See http://www.dartmouthatlas.org/data/table.aspx?ind=169.
[29] See http://www.dartmouthatlas.org/data/map.aspx?ind=160.
[30] See http://www.cms.gov/Medicare/Prevention/PrevntionGenInfo/Downloads/MPS_QuickReferenceChart_1.pdf.

B Analysis of the Health and Retirement Study (HRS)

B.1 The HRS data

The Health and Retirement Study (HRS) is a nationally representative longitudinal survey of Americans over the age of 50. Since 1992, the HRS has been administered in two-year cycles, following individuals and their spouses from their time of entry to the survey sample until their death. Data without individual identifiers and zipcode-level geographic information can be downloaded from the HRS website. Our analysis also used the restricted-access HRS data which contains zipcode-level geographic information which we needed in order to identify movers. Movers are defined to be 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. The restricted-access RAND HRS data we use incorporates data from 1992, 1993, 1994, 1995, 1996, 1998, 2000, 2002, 2004, 2006, and 2008. 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). Note that Wave 2 and Wave 3 have two years, but the sample surveyed in the two years is different. We only keep patient-waves with age of at least 65 years. The reasons for move come from the public-access general (non-RAND) HRS data.

We limited the sample to individuals aged 65+ to match our Medicare data, and defined a mover in a like-fashion as an individual who moves between HRRs. This gave 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.

B.2 Summary statistics

Table 13: Summary statistics

	(1) Non-movers	(2) Movers	(3) t-statistic
Average age over observed waves	74.9286	77.0283	-26.5428
Average age in first observed wave	70.6508	72.2606	-10.3132
Average age in 1992	64.6010	68.2324	-14.7578
Female	0.5508	0.5965	-3.9595
White	0.8056	0.8908	-9.4020
Hispanic[1]	0.0810	0.0440	5.9444
Education			
Less than high school	0.3321	0.2626	6.3737
GED	0.0411	0.0396	0.3259
High school	0.3019	0.3096	-0.7208
Some college	0.1778	0.2003	-2.5137
College	0.1471	0.1879	-4.9035
Retirement status (in first observed wave)[2]			
Retired	0.5059	0.5103	-0.2566
Partly retired	0.1590	0.1535	0.4398
Retired or partly retired	0.6649	0.6638	0.0685
Earnings (in first observed wave)[3]			
Average	$5,803	$5,115	0.9866
Average conditional on positive	$22,641	$21,505	0.7718
Median conditional on positive	$13,000	$13,000	
Share with zero	0.6738	0.7896	
Marital status (in first observed wave)			
Married or partnered	0.6566	0.6274	2.6134
Separated or divorced	0.0837	0.0910	-1.1221
Widowed	0.2294	0.2677	-3.8659
Never married	0.0303	0.0139	4.2125
# of Patients	20,998	2,025	
# of Patient Years	83,202	11,068	

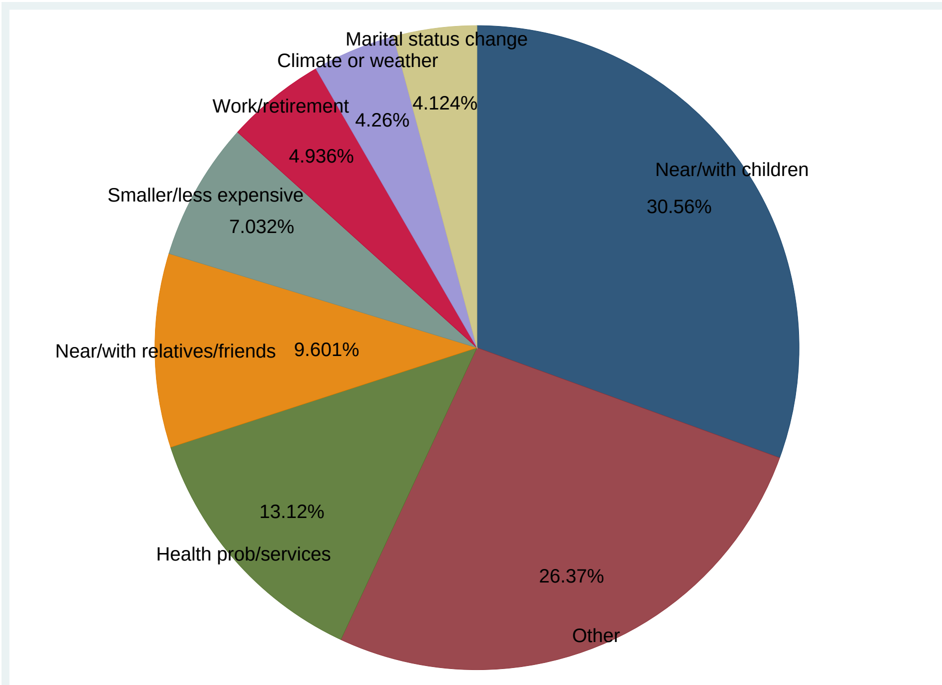
[1]There are three race categories in the data: White/Caucasian, Black/African-American, and Other. Hispanic is a separate variable, so someone can appear as White and Hispanic, Black and Hispanic, White and Non-Hispanic, etc.

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

[3]The sum of respondent's wage/salary income, bonuses/overtime pay/commissions/tips, 2nd job or military reserve earnings, professional practice or trade income.

B.3 Self-reported reasons for move

Figure 12: Top reasons for move



The pie chart shows the most common reasons for move. 1,144 of the 2,025 movers in the data provide reasons. Some provide more than one, these are counted separately. In total, the pie chart is based on 1,479 observations. Reasons mentioned fewer than 50 times are grouped under the "Other" category. The most common reasons to move people mention are wanting to live near or with children, relatives, or friends and health reasons.

B.4 Regression analysis of correlates of move

This section presents a person-wave level panel regression. Only person-waves with age of at least 65 are considered, multiple movers are also excluded.

On the left hand side, we consider a binary variable which takes value 0 if someone moves in that wave and value 1 if she does not move in that wave. On the right hand side, we consider variables representing health, marital status change and retirement status.

On the right hand side we have dummies for being unmarried and unpartnered, for being divorced or separated, for being widowed, for being retired or partly retired, and for having fair or poor health. (Health is reported on a 1 to 5 scale, higher values correspond to worse health. A score of 4 corresponds to "fair" and a score of 5 corresponds to "poor.") For each dummy, we are estimating the following regression:

$$\text{move in wave}_{it} \quad = \quad \alpha_i + \delta \text{Dummy}_{it} + \tau_t + \varepsilon_{it} \tag{9}$$

where α_i denote individual fixed effects for individual i and t indexes waves; τ_t represents wave fixed effects.

Table 14: Summary of regression results

Row	Dummy included	(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: This table shows the coefficients and standard errors from running the regression in equation (9) on each dummy. Columns (1) to (3) show results from estimating a linear regression with wave fixed effects and person fixed effects. Because of missing data, lines have different sample sizes and slightly different percentages moving.

C Additional results

C.1 Geography of moves

Figure 13: Distribution of distance moved

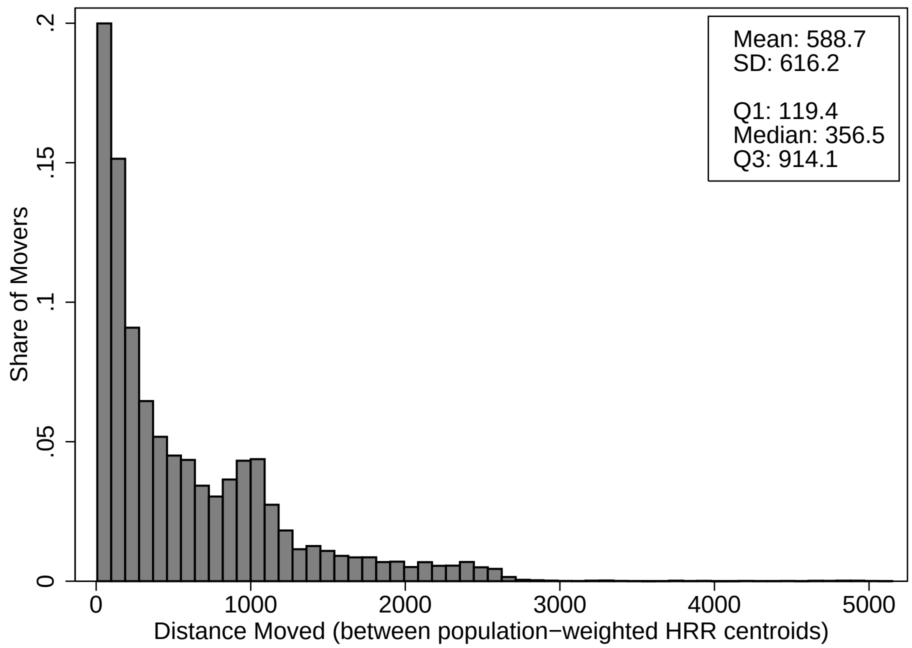


Figure shows a histogram of the distribution of distances moved. Distance is measured between the population-weighted centroids of HRRs. $N = 503,766$ movers.

Table 15: Summary of moves by census division

Census division (CD)	First observed in CD	and move	and leave CD	Last observed in CD	and move	and enter CD	Leaving relative to original	Entering relative to original
East North Central	1,577,977	69,423	35,767	1,574,609	65,722	32,399	0.0227	0.0205
East South Central	637,847	22,039	15,438	643,817	28,073	21,408	0.0242	0.0336
Mid-Atlantic	1,285,862	76,745	45,677	1,261,584	52,764	21,399	0.0355	0.0166
Mountain	467,561	36,939	24,436	477,189	46,642	34,064	0.0523	0.0729
New England	472,925	21,417	13,763	472,549	20,454	13,387	0.0291	0.0283
Pacific	965,906	78,375	33,004	958,319	70,809	25,417	0.0342	0.0263
South Atlantic	1,819,208	120,989	56,854	1,831,913	134,394	69,559	0.0313	0.0382
West North Central	699,013	26,849	15,662	700,819	28,595	17,468	0.0224	0.0250
West South Central	962,660	50,779	20,713	968,160	56,281	26,213	0.0215	0.0272

Notes: We use the baseline sample and upweight non-movers by a factor of 4.
Column (1) shows the number of patients (both movers and non-movers) who start out in each census division.
Column (2) shows the number of patients who start out in each census division and move to another HRR.
Column (3) shows the number of patients who start out in each census division and end up in another census division.
Column (4) shows the number of patients (both movers and non-movers) who end up in each census census division.
Column (5) shows the number of patients who end up in each census division and had moved to another HRR.
Column (6) shows the number of patients who end up in a census division after starting out in another census division.
Column (7) shows the number of patients who left each census division relative to the number of patients who started out in the census division, i.e., Column (7) is Column (3) divided by Column (1).

Column (8) shows the number of patients who entered each census division relative to the number of patients who started out in the census division, i.e., Column (8) is Column (6) divided by Column (1).

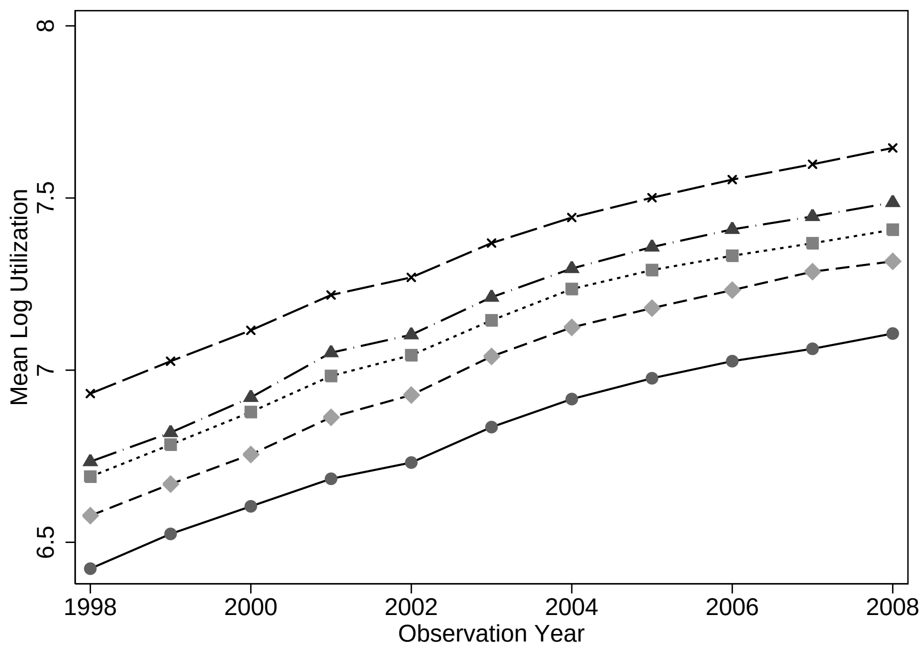
Table 16: Movements between census divisions (Movers only, as percentage of all moves)

		Destination								
		ENC	ESC	M-A	M	NE	P	SA	WNC	WSC
Origin	East North Central	6.96	0.93	0.32	0.91	0.14	0.64	2.52	0.63	0.73
	East South Central	0.64	1.56	0.10	0.14	0.04	0.15	1.12	0.14	0.48
	Mid-Atlantic	0.58	0.28	6.28	0.54	0.92	0.56	5.57	0.15	0.37
	Mountain	0.62	0.17	0.20	2.67	0.10	1.59	0.57	0.66	0.76
	New England	0.15	0.08	0.40	0.19	1.56	0.21	1.53	0.05	0.10
	Pacific	0.53	0.28	0.26	2.76	0.14	9.13	0.94	0.51	1.01
	South Atlantic	2.50	1.58	2.70	0.77	1.05	0.84	13.09	0.49	1.00
	West North Central	0.52	0.15	0.07	0.63	0.04	0.36	0.47	2.43	0.65
	West South Central	0.54	0.55	0.14	0.65	0.07	0.58	0.85	0.62	6.08

Notes: This table shows the percentage of moves that take place between each fo the 81 origin-destination pairs of census divisions. The percentages are computed over the total number of movers (503,766). For example, 2.7% of all moves originate from the South Atlantic census division and have their destination in the Mid-Atlantic census division.

C.2 Patient and place effects are constant over time

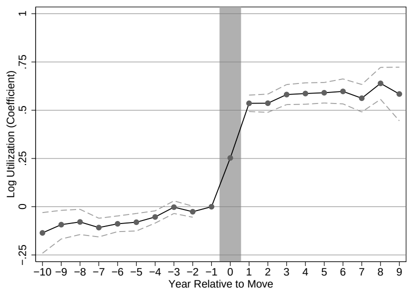
Figure 14: Time series of mean log utilization of HRRs by quintile.



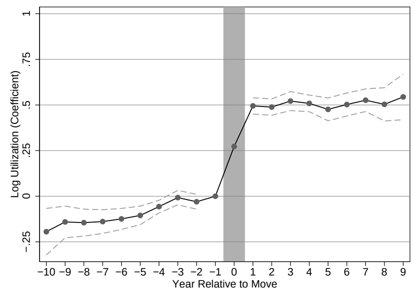
This figure shows a time series plot of the mean log utilization of HRRs by quintile. For this analysis, we computed the mean utilization within each quintile for each year. HRR Quintiles are defined by taking the average over years of the HRR-year mean utilization.

C.3 Additional event studies

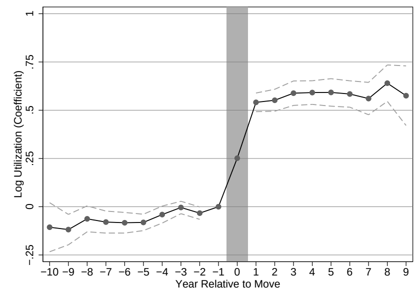
Figure 15: Sensitivity of Event Study Results on Log Utilization to Handling of Attrition



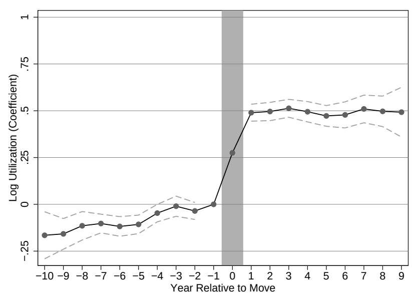
(a) Baseline Sample



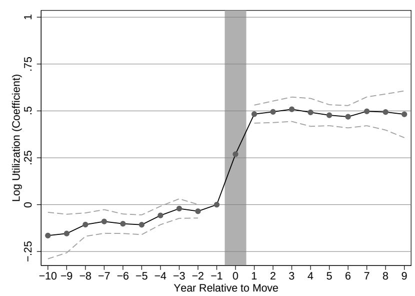
(b) Drop patient who ever die



(c) Drop patients ever in an HMO



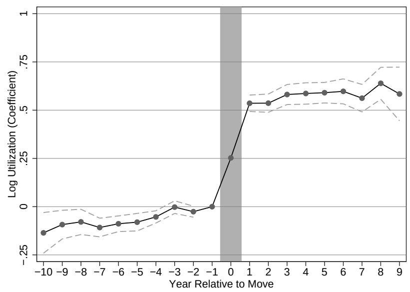
(d) Drop patients ever in an HMO or ever die



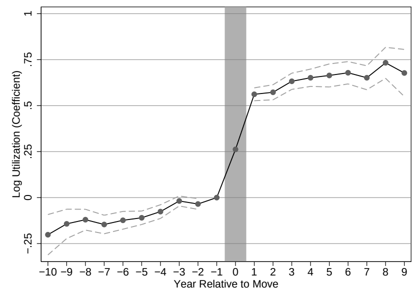
(e) Drop patients ever missing outcomes

The solid line in the figures connects the estimated $\hat{\theta}_r$ from the estimation of equation (5). Figure 15 (a), replicates our baseline specification in Figure 4 (a) ($N = 3,780,960$ patient-years). Figure 15 (b) repeats the same analysis as for Figure 15 (a), except equation (5) and $\hat{\delta}_i$ are estimated only for the subsample of movers who are never observed to die during the course of our study ($N = 2,502,399$ patient-years). In Figure 15 (c), equation (5) and $\hat{\delta}_i$ are estimated only for the subsample of movers who are never in an HMO ($N = 3,022,969$ patient-years), in Figure 15 (d), they are estimated only for movers who are both never in an HMO and never die ($N = 1,956,761$ patient-years), and in Figure 15 (e), they are estimated only for movers who never have missing outcomes for any reason (including death or entering an HMO) ($N = 1,759,509$ patient-years). Since the level of the graph is not separately identified, we normalize the coefficient on the term corresponding to the year immediately prior to the move (relative year -1) to 0. The dashed lines show the upper and low 95% confidence interval of the estimated $\hat{\theta}_r$, constructed using the same bootstrap approach as in Figure 21.

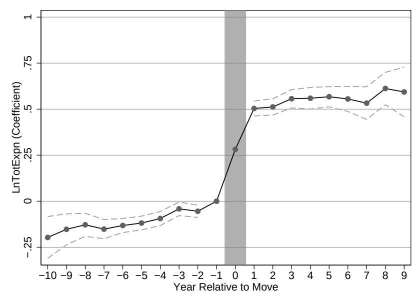
Figure 16: Sensitivity of Event Study Results to Alternate Specifications



(a) Log Utilization



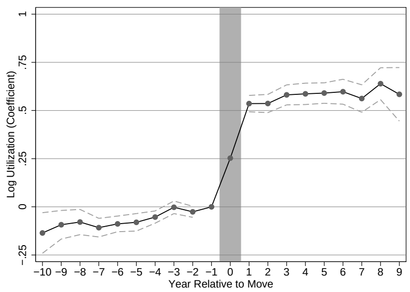
(b) Drop age as a covariate



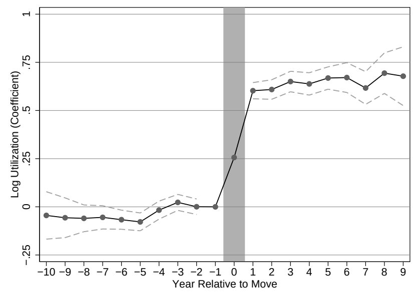
(c) Log Total Expenditure

The solid line in the figures connects the estimated $\hat{\theta}_r$ from the estimation of equation (5). Figure 15 (a), replicates our baseline specification in Figure 4. Figure 16 (b) repeats the same analysis as for Figure 16 (a) except x_{it} now doesn't include 5-year age bins. Figure 16 (c) repeats the same analysis as for Figure 16 (a), but y_{it} is now the log of expenditure (plus 1) instead of the log of utilization (plus 1). For all panels $N = 3,780,960$ patient-years. Since the level of the graph is not separately identified, we normalize the coefficient on the term corresponding to the year immediately prior to the move (relative year -1) to 0. The dashed lines show the upper and low 95% confidence interval of the estimated $\hat{\theta}_r$, constructed using the same bootstrap approach as in Figure 21.

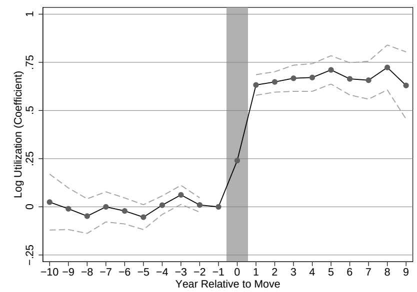
Figure 17: Sensitivity of Event Study Results on Log Utilization to Restrictions on Mover Geography



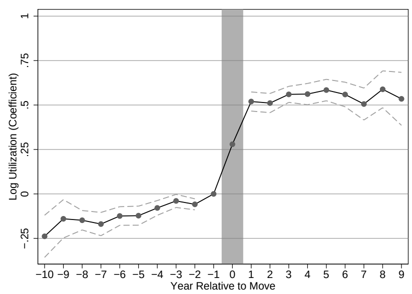
(a) Baseline Sample



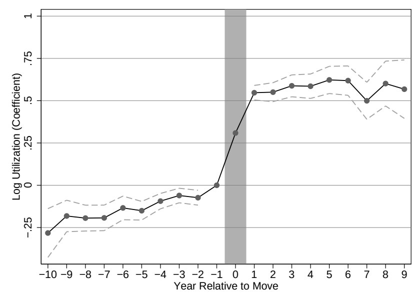
(b) Limit to cross state movers



(c) Limit to cross census region movers



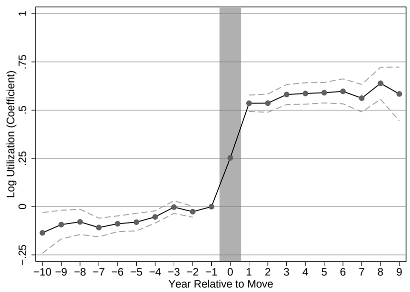
(d) Limit to movers whose destination is not Florida



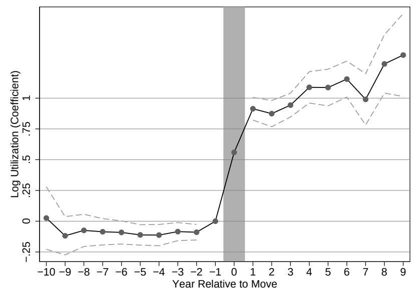
(e) Limit to movers whose destination is not Florida, Arizona, or California

The solid line in the figures connects the estimated $\hat{\theta}_r$ from the estimation of equation (5). Figure 17 (a), replicates our baseline specification in Figure 4 (a) ($N = 3,780,960$ patient-years). Figure 17 (b) repeats the same analysis as for Figure 17 (a), except equation (5) and $\hat{\delta}_i$ are estimated only for the subsample of movers who move across states ($N = 2,566,097$ patient-years). In Figure 17 (c), equation (5) and $\hat{\delta}_i$ are estimated only for the subsample of movers who move across census regions ($N = 1,414,687$ patient-years), in Figure 17 (d), they are estimated only for movers who do not move to Florida ($N = 3,332,637$ patient-years), and in Figure 17 (e), they are estimated only for movers who do not move to Florida, Arizona, or California ($N = 2,886,083$ patient-years). Since the level of the graph is not separately identified, we normalize the coefficient on the term corresponding to the year immediately prior to the move (relative year -1) to 0. The dashed lines show the upper and low 95% confidence interval of the estimated $\hat{\theta}_r$, constructed using the same bootstrap approach as in Figure 4.

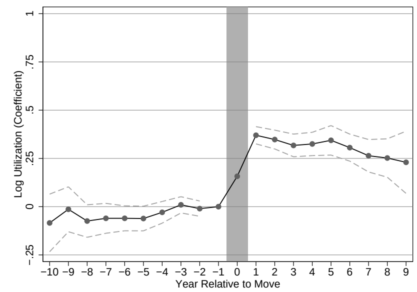
Figure 18: Event Study Results for Alternate Sub-Samples



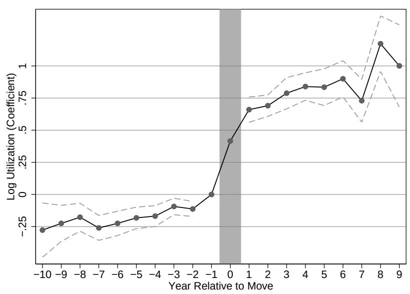
(a) Log Utilization



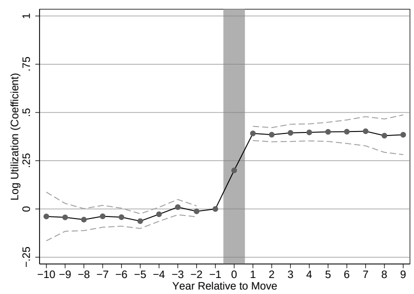
(b) Above Median - # of Chronic Conditions



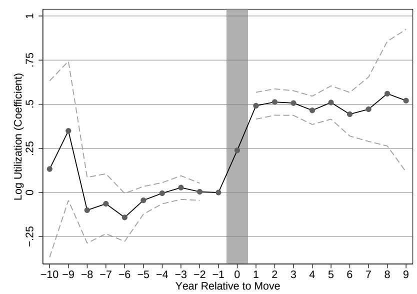
(c) Below Median - # of Chronic Conditions



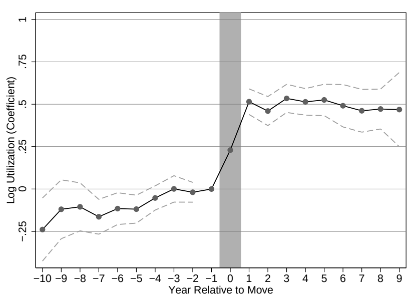
(d) Above Median - # of Hospital Days



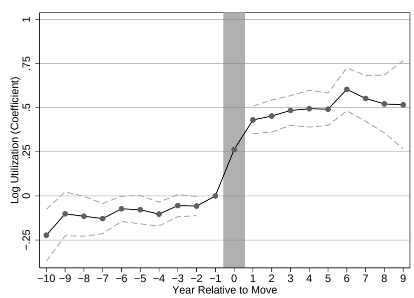
(e) Below Median - # of Hospital Days



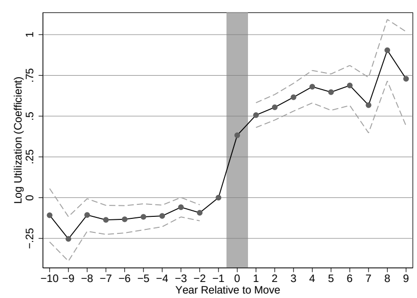
(f) Age Quartile 1



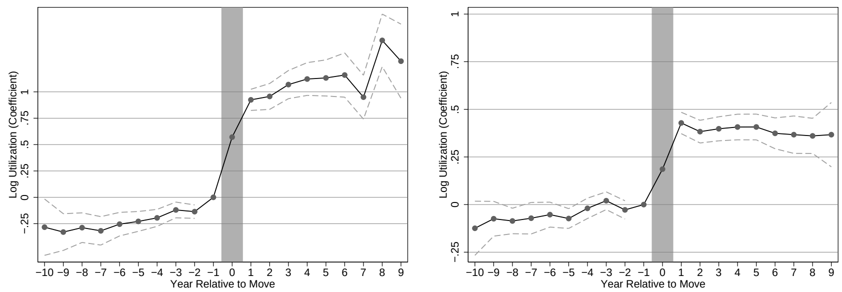
(g) Age Quartile 2



(h) Age Quartile 3



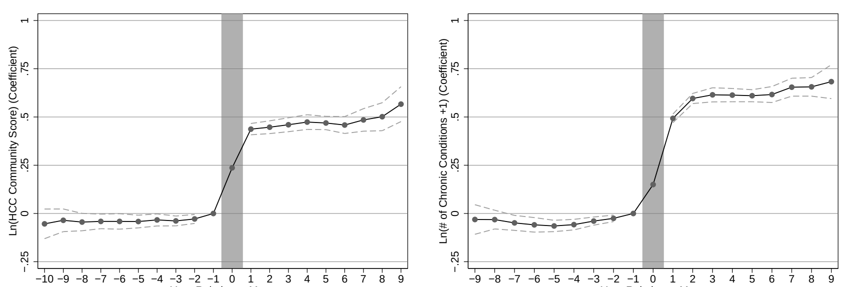
(i) Age Quartile 4



(j) Above Median Utilization (k) Below Median Utilization

The solid line in the figures connects the estimated $\hat{\theta}_r$ from the estimation of equation (5). Figure 18 (a), replicates our baseline specification in Figure 4 (a) ($N = 3,780,960$ patient-years). Figure 18 (b) repeats the same analysis as for Figure 18 (a), except equation (5) and $\hat{\delta}_i$ are estimated only for the subsample of movers who have an above median number of chronic conditions ($N = 2,075,296$ patient-years). In Figure 18 (c), equation (5) and $\hat{\delta}_i$ are estimated only for the subsample of movers who have a below median number of chronic conditions ($N = 1,702,303$ patient-years), and in Figures 18 (d)-(e), they are estimated only for movers who have an above or below median numbers of hospital days respectively ($N = 2,026,477$ and $1,754,483$ patient-years). In Figures 18 (f)-(i) equation (5) and $\hat{\delta}_i$ are estimated only for movers whose average age is in the designated quartile ($N = 758,983$, $881,598$, $994,285$, and $1,146,094$ patient-years, respectively.) In Figures 18 (j)-(k) they are estimated only for movers who have above or below median utilization, respectively ($N = 2,027,364$ and $1,753,596$ patient-years). Since the level of the graph is not separately identified, we normalize the coefficient on the term corresponding to the year immediately prior to the move (relative year -1) to 0. The dashed lines show the upper and low 95% confidence interval of the estimated $\hat{\theta}_r$, constructed using the same boostrap approach as in Figure 21.

Figure 19: Event Study Results for Alternate Outcomes



(a) Ln(HCC Score) (b) Ln(# of Chronic Conditions+1)

The solid line in the figures connects the estimated $\hat{\theta}_r$ from the estimation of equation (5). Figure 19 (a) repeats the same analysis as for Figure Figure 4., but y_{it} is now the log of HCC Score instead of the log of utilization. Similarly, in Figure 19 (b) y_{it} the estimation of equation (5) is now the log of the number of chronic conditions +1. For all panels $N = 3,780,960$ patient-years. Since the level of the graph is not separately identified, we normalize the coefficient on the term corresponding to the year immediately prior to the move (relative year -1) to 0. The dashed lines show the upper and low 95% confidence interval of the estimated $\hat{\theta}_r$, constructed using the same bootstrap approach as in Figure 4.

C.4 Robustness to alternative definitions of movers

C.4.1 Accounting for measurement error in move timing: decomposition

Table 17 shows the results. The first row shows the results for our baseline definition of movers. 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. The 0.75 threshold was arbitrary, designed to capture "real" moves as opposed to people whose address on file changes without their actual location changing. Rows 2 and 3 show that our estimate of the role of patients is not sensitive to making this threshold slightly looser (0.6) or slightly tighter (0.9) threshold for the required increase in average share of claims in the destination after the move relative to before. In row 4, we include all individuals whose address on file changes exactly once, regardless of how the average share of claims in the destination changes on move. This increases our number of movers by about 50%, and, without further adjustment, also increases the patient share to 0.55.

However, as shown in Figure 21a this definition of movers based solely on address change includes a substantial number of mis-measured moves. 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.

To handle the fact that we are mismeasuring some moves with this solely address-based definition of movers, we make an adjustment to our basic estimating equation to allow in our estimation 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:

$$y_{it} = \alpha_i + \lambda_{r(i,t)}\gamma_{\hat{J}(i,t)} + \left(1-\lambda_{r(i,t)}\right)\gamma_{\hat{J}'(i,t)} + \tau_t + x_{it}\beta + \tilde{\varepsilon}_{it} \tag{10}$$

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 Figure 20, then estimate equation (10) plugging in these estimates. Note that for non-movers, $\hat{J}(i,t) = \hat{J}'(i,t)$ by definition.

Row 5 shows that this adjustment decreases the estimated role for patients with the address-change definition of movers from 0.55 to 0.38. By contrast, when we make this adjustment to our baseline definition of movers and the claim share pattern in Figure 21c we find that that, not surprisingly given our ability to measure move timing much more cleanly, this makes relatively little difference to the baseline estimate; as shown in row 6 our estimate of the role for patients changes from 0.426 in the baseline to 0.404 with this adjustment.

C.4.2 Accounting for measurement error in move timing: event study

Summary Analogously to the adjustment to our basic estimating equation, we describe a correction of our event study figures for measurement error in move timing. We define a set of corrected coefficients:

$$\hat{\theta}_r^{corrected} = \begin{cases} \hat{\theta}_r - (1-\lambda_r)\left(1-\hat{S}_{pat}\right) & \text{for } r<0 \\ \hat{\theta}_r + (1-\lambda_r)\left(1-\hat{S}_{pat}\right) & \text{for } r>0 \end{cases}. \tag{11}$$

where $\lambda_r$is (as previously defined) the probability the current location of the mover is defined correctly, $\hat{\theta}_r$ was estmated in equation (5) and our preliminary estimate

$$\hat{S}_{pat} = 1 - \left(\overline{\theta}_{post} - \overline{\theta}_{pre}\right) / \left(\overline{\lambda}_{post} + \overline{\lambda}_{pre} - 1\right)$$

where $\overline{\theta}_{pre} = \overline{\hat{\theta}}_{r<0}$, $\overline{\theta}_{post} = \overline{\hat{\theta}}_{r>0}$, $\overline{\lambda}_{pre} = \overline{\lambda}_{r<0}$ and $\overline{\lambda}_{post} = \overline{\lambda}_{r>0}$.

Figure 21(a) shows the corrected event study, which shows a flat trend in utilization prior to the move and after the move. The lack of a systematic trend in utilization prior to the move helps alleviate concerns about moving being correlated with an underlying secular trend in demand for care. The lack of any pronounced trend in utilization after the move suggests that the impact of place is relatively "immediate". This is suggestive evidence against models of gradual adaptation or habit formation in this settings. We will return to both these points in more detail below when we discuss the assumptions of the model and the potential sources of patient heterogeneity.

Recall that the size of the jump tells us about the relative role of patient and place. If there was no jump at the time of move, this would suggest that utilization behavior was entirely determined by patient. If at move the individual jumped immediately to the destination level (i.e. size of jump of 1), this would suggest that utilization was determined entirely by the place. Indeed, as noted above, 1 - the size of the jump gives us the share due to the patients. Eyeballing the event study, it looks like about 40 percent of the difference across HRRs is due to patients.

Derivation Under this form of measurement error, the θ_{it} in equation (4) becomes:

$$\theta_{it} = \begin{cases} \left(1-\lambda_{r(i,t)}\right)\left(1-S_{pat}\left(d\left(i\right),o\left(i\right)\right)\right) & \text{for } r<0 \\ \lambda_{r(i,t)}\left(1-S_{pat}\left(d\left(i\right),o\left(i\right)\right)\right) & \text{for } r>0 \end{cases}.$$

where, recall $\lambda_r$is the probability the mover's residence is reported correctly. Let θ_{it}^{true} be the value of these parameters in the model with no measurement error (i.e., zero for $r<0$ and $1-S_{pat}\left(d\left(i\right),o\left(i\right)\right)$ for $r>0$). We can express θ_{it}^{true} as a function of θ_{it}, λ_r, and $S_{pat}\left(d\left(i\right),o\left(i\right)\right)$:

$$\theta_{it}^{true} = \begin{cases} \theta_{it}-\left(1-\lambda_{r(i,t)}\right)\left(1-S_{pat}\left(d\left(i\right),o\left(i\right)\right)\right) & \text{for } r<0 \\ \theta_{it}+\left(1-\lambda_{r(i,t)}\right)\left(1-S_{pat}\left(d\left(i\right),o\left(i\right)\right)\right) & \text{for } r>0 \end{cases}.$$

To correct our event study figures, we begin from estimates of $\hat{\theta}_r$ from equation (5). We construct a preliminary estimate $\hat{S}_{pat}$ of the average value of $S_{pat}\left(d\left(i\right),o\left(i\right)\right)$ as follows. First, let $\overline{\theta}_{pre} = \overline{\hat{\theta}}_{r<0}$ and let $\overline{\theta}_{post} = \overline{\hat{\theta}}_{r>0}$. Let $\overline{\lambda}_{pre} = \overline{\lambda}_{r<0}$ and let $\overline{\lambda}_{post} = \overline{\lambda}_{r>0}$. Then $\hat{S}_{pat} = 1-\left(\overline{\theta}_{post}-\overline{\theta}_{pre}\right)/\left(\overline{\lambda}_{post}+\overline{\lambda}_{pre}-1\right)$. (Note that if $S_{pat}\left(d\left(i\right),o\left(i\right)\right)$ is a constant S_{pat} for all movers, $\text{plim}\hat{S}_{pat} = S_{pat}$).

We define a set of corrected coefficients:

$$\hat{\theta}_r^{corrected} = \begin{cases} \hat{\theta}_r - (1-\lambda_r)\left(1-\hat{S}_{pat}\right) & \text{for } r<0 \\ \hat{\theta}_r + (1-\lambda_r)\left(1-\hat{S}_{pat}\right) & \text{for } r>0 \end{cases}.$$

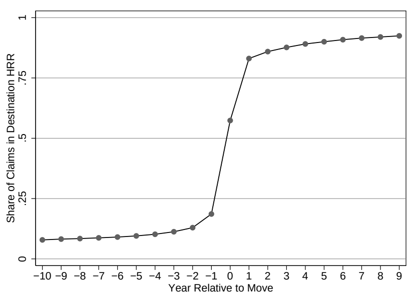
It is straightforward to show that this correction eliminated the bias due to measuring moves with error, in the sense that if $S_{pat}(d(i), o(i))$ is a constant S_{pat} for all movers, $plim\hat{\theta}_r^{corrected}$ equals 0 for $r<0$ and equals $1-S_{pat}$ for $r>0$.

C.4.3 Movers as address-changers

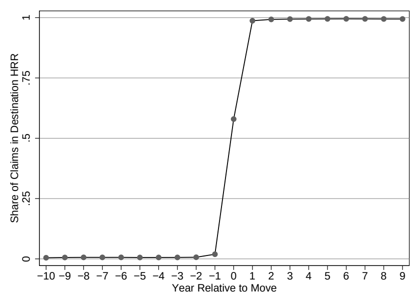
Finally, we explore an alternative definition of movers defined as individuals whose HRR of residence (based on their address on file for Social Security payments) changes exactly once, and who do not have more than 10% of claims in their destination HRR in the years before a move (i.e. relative years -1 and earlier) or more than more than 10 percent of claims in her origin HRR in the years after her move (relative years 1 and later); this approach follows in the spirit of Song et al's (2010) strategy for eliminating so-called "snowbirds" from their moving analysis. Like our baseline definition of movers, Figure 21b shows that it limits the sample to a set of people where we can cleanly identify the timing of the move. However, unlike our baseline definition of movers, which excludes approximately 1/3 of address changers, this definition excludes approximately 40% of address changers, resulting in a strict subset of our baseline sample. Row 7 shows that with this definition of movers we estimate a role for patients of 0.42.

C.4.4 Identifying timing of moves: Graphical evidence

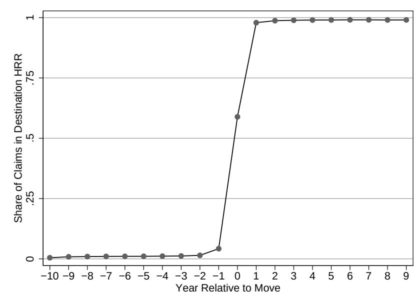
Figure 20: Identifying timing of moves.



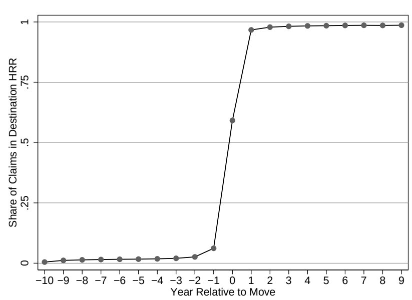
(a) Change of address criterion



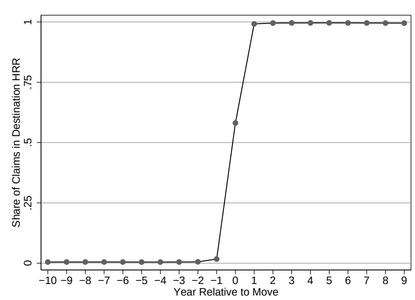
(b) Change of address criterion, snowbirds dropped



(c) Change of address + claim share criterion ($\Delta > 0.75$)



(d) Change of address + claim share criterion ($\Delta > 0.6$)



(e) Change of address + claim share criterion ($\Delta > 0.9$)

C.4.5 Additive decomposition results

Table 17: Additive decomposition results: alternative definitions of movers

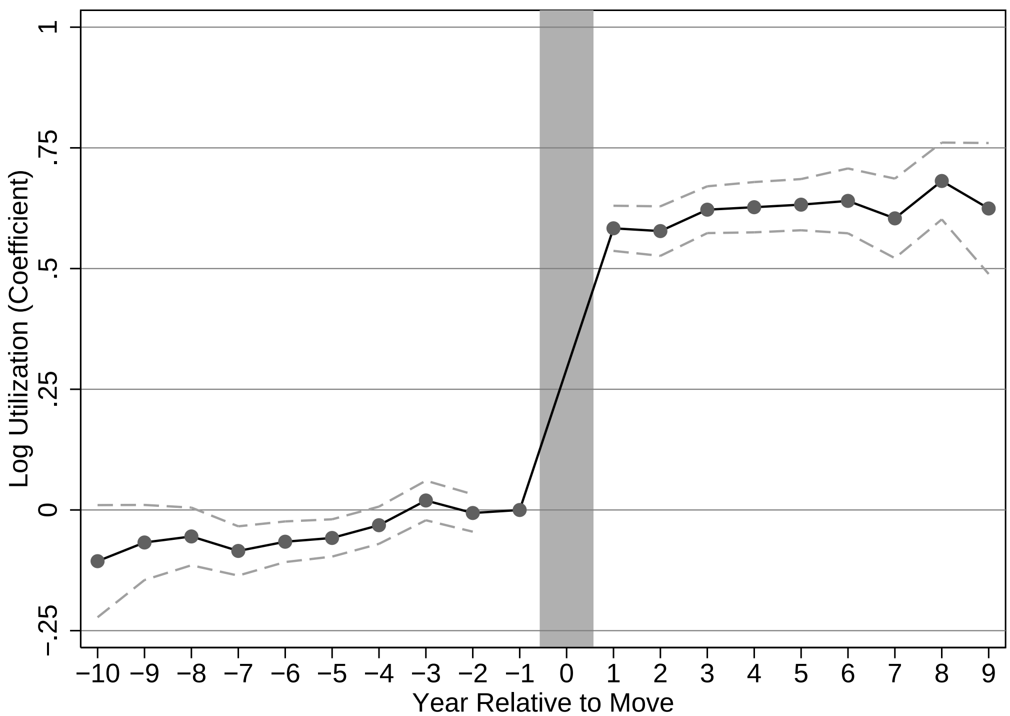
	Mover definition	(1) Mean of log utilization	(2) Above / below median utilization difference	(3) Share due to patients	(4) N (% of movers retained)
(1)	Baseline	7.105	0.279	0.426	16,464,297 (100%)
(2)	Looser claim share change	7.112	0.279	0.452	16,688,958 (106%)
(3)	Stricter claim share change	7.094	0.281	0.401	16,082,219 (90.5%)
(4)	Address change	7.123	0.278	0.547	18,263,890 (153%)
(5)	Address change, adjusted for move timing measurement error	7.123	0.278	0.376	18,263,890 (153%)
(6)	Baseline, adjusted for move timing measurement error	7.105	0.279	0.404	16,464,297 (100%)
(7)	Address change, excluding snowbirds	7.097	0.280	0.416	16,164,310 (92.5%)

Notes: This table reports our estimate of the share of the difference in utilization between above and below median HRRs due to patients, based on estimating equation (1) (or a variant on it) on the sample shown in the different rows. In rows 4 and 6, our estimating equation is Equation (10). We retain all non-movers in every row. The baseline sample has $503,766$ movers. Column (1) reports the mean of the outcome for the given sample. Column (2) reports the difference in average utilization between above and below median utilization HRRs $(\hat{y}_R - \hat{y}_{R'})$. Column (3) reports the share of the difference in column (2) that is due to patients $\left(\hat{S}_{pat}\left(R,R'\right)\right)$. Column (4) reports the sample size in patient-years.
Row 1: We categorize someone as a mover if their HRR of residence changes and (i) they have claims both before and after the residence change, with $\Delta > 0.75$, (ii) they have claims only after the address change, with a claim share in destination of over 0.95, (iii) they have claims only before the address change, with a claim share in destination of under 0.05, or (iv) they don't have any claims. Our sample includes non-movers and movers, but excludes patients who satisfy the first (change of address) criterion but not the second (claim share) criterion.
Row 2: We categorize someone as a mover if their HRR of residence changes and (i) they have claims both before and after the residence change, with $\Delta > 0.6$, (ii) they have claims only after the address change, with a claim share in destination of over 0.95, (iii) they have claims only before the address change, with a claim share in destination of under 0.05, or (iv) they don't have any claims. Our sample includes non-movers and movers, but excludes patients who satisfy the first (change of address) criterion but not the second (claim share) criterion.
Row 3: We categorize someone as a mover if their HRR of residence changes and (i) they have claims both before and after the residence change, with $\Delta > 0.9$, (ii) they have claims only after the address change, with a claim share in destination of over 0.95, (iii) they have claims only before the address change, with a claim share in destination of under 0.05, or (iv) they don't have any claims. Our sample includes non-movers and movers, but excludes patients who satisfy the first (change of address) criterion but not the second (claim share) criterion.
Row 4: We categorize someone as a mover if their HRR of residence changes.
Row 5: Same as Row 4, but we adjust for measurement error in move timing by estimating Equation (10).
Row 6: Same as Row 1, but we adjust for measurement error in move timing by estimating Equation (10).

Row 7: Same as Row 4, but we exclude patients who have at least 10% of their claims in the destination pre-move or at least 10% of their claims in the origin post-move.

C.4.6 Event study adjusted for measurement error in move timing

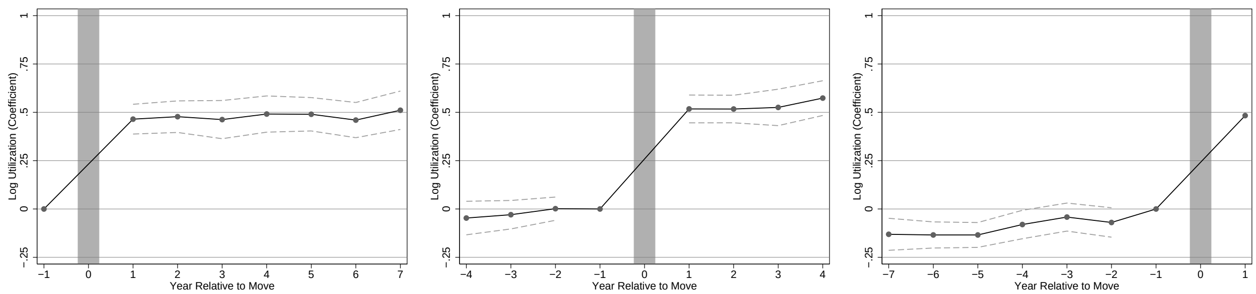
Figure 21: Event-study Analysis of Log Utilization, Adjusted for measurement error in timing of moves



The solid line in the figure connects the estimated $\hat{\theta}_r^{corrected}$ calculated following equation (11) from coefficients on the $\tilde{\theta}_{r(i,t)}\hat{\delta}_i$ terms from the estimation of Equation (5) shown in Figure 4. The dashed lines show the upper and low 95% confidence interval of the estimated $\hat{\theta}_r^{corrected}$, constructed using the same bootstrap approach as in Figure 4 with the only difference being that at each repetition, we now compute $\hat{\theta}_r^{corrected}$ from coefficients on the $\tilde{\theta}_{r(i,t)}\hat{\delta}_i$ terms according to Equation (11). $N = 3,780,960$ patient-years (movers only).

Figure 22: Balanced-Panel Event-study Analyses of Log Utilization, Adjusted for measurement error in timing of moves.

(a) Early Moves. (b) Middle Moves. (c) Late Moves.



These event studies and their confidence intervals are estimated in the same manner as in Figure 21 above, but limiting our baseline sample of movers to various balanced panels of movers whom we observe in each of a given set of relevant years. In addition, the $\hat{\delta}_i$ in Equation (5) are estimated only for movers in the respective panel and all non-movers. Panel (a) restricts to movers whom we observe in every relative year in [-1,7] (this consists of $49,022$ patients per year, so $N = 441,198$ patient-years).Panel (b) restricts to movers whom we observe in every relative year [-4,4] (this consists of $55,309$ patients for year, so $N = 497,781$ patient years). Panel (c) restricts the sample whom we observe in every relative year in [-7,1] (this consists of $64,337$ patients per year, so $N = 579,033$ patient years). The dashed lines show the upper and low 95% confidence interval of the estimated $\hat{\theta}_r^{corrected}$, constructed using the same bootstrap approach as in Figure 21.