What Causes Geographic Variation in Drug Prescribing? Evidence from Physician Migration*

David Beheshti†

Seth Neller‡

July 2, 2023

Abstract

In this paper, we examine the importance of individual physicians in explaining the significant variation in prescription drug spending in Medicare Part D. By tracking prescribing behavior before and after physician relocations, we find that movers' prescribing converges toward the average of their new location. However, this convergence is far from complete, highlighting the importance of idiosyncratic physician-specific factors. Overall, these physician-specific factors explain about 60 to 70 percent of the cross-sectional variation in prescription drug spending, suggesting that physicians are one of the most important supply-side determinants of this variation. We investigate several potential mechanisms behind this partial convergence.

*Thanks to Marika Cabral, Mike Geruso, Kevin Kuruc, and Melissa LoPalo for helpful conversations.
†University of Texas at San Antonio. Email: david.beheshti@utsa.edu.
‡University of Tennessee at Knoxville. Email: seth.neller@gmail.com.

1 Introduction

Spending on healthcare in the United States varies drastically by geographic location. This widespread variation in spending is prevalent among the privately insured as well as those enrolled in Medicare. Although price differences can explain a large fraction of the spending variation among the privately insured, the variation in Medicare spending is primarily the result of variation in intensity of treatment (Cooper et al., 2019). Despite minimal price differences, spending in Medicare still varies significantly. For example, in 2017 the highest spending hospital referral region (HRR) in Medicare spent over $13,000 per patient, approximately 74 percent more than the lowest spending HRR.[1] In Medicare Part D — the federal program which provides prescription drug benefits to Medicare beneficiaries — the disparity is even more striking, with the highest spending HRR spending over 300 percent more per enrollee than the lowest spending HRR (appendix Figure A1).

The proximate causes of this variation in treatment intensity are not entirely understood. The variation cannot be eliminated by controlling for observable differences in patient characteristics such as health status and income (Skinner, 2011), although this does not rule out other demand-side explanations such as unobservable variation in patient demand. Alternatively, the variation could be driven in large part by supply-side characteristics such as differential insurance reimbursement, pharmaceutical marketing, or hospital policies. Another potentially important supply-side factor is the behavior of individual physicians.[2] Physicians act as the point of contact between patients and the healthcare system and have prescriptive authority to determine the set of treatment options available to their patients. Due to large information asymmetries between physicians and patients, physicians play an important role in influencing patients' treatment decisions. However, physicians vary substantially in their beliefs regarding optimal treat-

[1] Price, age, sex, and race adjusted spending per Part A and B enrollee. Source: Dartmouth Atlas Project (https://www.dartmouthatlas.org/interactive-apps/medicare-reimbursements/#hrr).

[2] Throughout this paper, we use the term 'physician' to refer to anyone with prescriptive authority, even if they did not attend medical school (e.g., nurse practitioners).

ment options for similar patients (Cutler et al., 2019). In principle, these differences in physician practice styles could be a key driver behind the observed geographic variation in healthcare spending.

In this paper, we estimate the importance of individual physicians in explaining the significant cross-sectional variation in Medicare Part D prescription drug spending. Put another way, we ask what fraction of the observed variation in drug spending can be attributed to physician-specific factors as opposed to environmental factors?[3] For example, consider a thought experiment in which all physicians are randomly assigned to different geographic locations. On one extreme, if individual physicians are perfectly interchangeable and only environmental factors determine prescribing behavior, then this re-assignment would lead to no change in the observed cross-sectional variation. On the other extreme, if individual physicians are completely unresponsive to environmental factors, then this re-assignment would eliminate any geographic variation in prescribing. We estimate where physicians' responsiveness lies between these two extremes. Our estimate of physician responsiveness informs us regarding the extent to which physicians' prescribing habits are shaped by their environments, as opposed to being immutable characteristics of the physician.

There are a number of challenges in estimating the extent to which different factors contribute to variation in spending. Most notably, observed spending is jointly determined by the decisions of numerous agents including physicians, patients, hospitals, and insurers, making it difficult to isolate the importance of any given agent. However, understanding the sources of variation in spending is key for public policy. For example, areas with higher spending do not achieve better health outcomes relative to lower spending areas. If this higher spending is primarily due to overly aggressive prescribing practices, then policies which penalize high spending could reduce healthcare costs without harming patients. However, if higher spending is due to unobservably worse patient health or

[3]We use the term "environment" to refer to demand and supply-side determinants of prescription patterns unrelated to physician characteristics.

different patient preferences, then this type of intervention would be counterproductive.

In order to separately identify the relative contribution of physician-specific factors from other determinants of spending, we track spending associated with specific physicians who move from low-spending to high-spending HRRs (and vice versa) before and after they move. We compare these physicians' spending to other migrating physicians within the same origin HRR, but who moved to HRRs with different spending patterns. The idea behind this research design is that a physician's environment (patients, hospital, etc.) changes discretely upon moving, whereas factors unique to the physician (training, beliefs, etc.) trend smoothly across the move. By observing how the physician's behavior changes following the move, we can estimate the importance of physician-specific factors in determining healthcare spending. As an extreme example, suppose that idiosyncratic physician characteristics are the only relevant determinants of spending; i.e., physicians' treatment decisions are made independently of patient preferences or other supply-side factors. In this case we would expect to see no changes in the physician's spending after the move. However, in the opposite extreme in which physicians are completely malleable, we would expect average spending of physician movers to converge to the mean of their new environment. Observing where physician spending lies between these two extremes provides an estimate of the importance of individual physicians in explaining the observed variation in health spending.

We present the main findings in a series of event study figures. These figures reveal that, among physicians who ultimately move HRRs, prescribing behavior remains relatively constant in the years prior to moving. However, immediately upon moving to an HRR with higher (lower) average spending, physicians increase (decrease) their own spending in response. In the first year after the move, physicians who moved to HRRs with 100 percent higher spending than their origin HRR increase their own prescription spending by about 22 percent. This increase in spending continues to grow through the second year post-move, at which point it stabilizes at about 42 percent. This estimate suggests that physician-specific factors can explain approximately 58 percent of the cross-sectional variation in prescription drug spending in Medicare Part D, with the remaining 42 percent attributable to non-physician factors.[4]

Mechanically, this change in spending could be driven by two primary channels: (1) prescribing drugs to more patients—which we term the extensive margin, and/or (2) increasing spending per beneficiary—which we term the intensive margin. When we restrict our analysis to only consider differences in spending along the intensive margin, we find that physicians who move to HRRs with 100 percent higher drug spending *per beneficiary* ultimately increase their own per beneficiary spending by 29 percent. In a series of decomposition exercises, we find that changes in *overall* spending are primarily driven by the changes along the extensive margin, while differences in per-beneficiary spending are driven by both differences in the number of drug claims per beneficiary, as well as differences in the average cost of these claims.

While this analysis confirms that physicians' behavior responds to their environments, it is unclear whether this responsiveness is welfare enhancing. In principle, our results could be driven by physicians adopting positive or negative prescribing practices. As a proxy for appropriate prescribing behavior, we consider how physicians' decisions to prescribe brand name versus generic drugs change in response to a move. We find that physicians who move to areas with a higher fraction of spending on generic drugs shift their own prescriptions toward generics. However, we find asymmetrical results in how physicians adapt to local brand-vs-generic spending patterns. Physicians who move to areas where brand-name drugs are more prevalent are more adaptive to their new environment than physicians who move to areas were generic drugs are more prevalent. This finding indicates that physicians are susceptible to adopting inappropriate as well as appropriate prescribing practices in response to local norms.

We then turn our attention towards understanding the mechanisms driving the changes

[4]See the discussion in Section 4.1 for more details on the interpretation of these results.

in physicians' spending habits. We consider how a variety of physician and non-physician characteristics affect physicians' responsiveness to their environments. Among the observable physician-level characteristics, we find that experience is the most important factor in predicting how strongly physicians respond to environmental factors, with less experienced physicians exhibiting much more malleable behavior. We also find some evidence of differential responsiveness by specialty.[5]

We also consider how physicians' responsiveness is affected by non-physician "place" characteristics. We find that differences in observable patient characteristics, pharmaceutical marketing intensity, and market competitiveness have little impact. We find some suggestive evidence that differences in patient demand across areas may be an important factor in driving physician responsiveness.

Finally, we estimate econometric models to capture the causal effect of various HRRs on physician prescribing behavior. Following Finkelstein, Gentzkow and Williams (2016) and other papers from the place-effects literature, we correlate these causal effects with numerous factors to gather suggestive evidence regarding the mechanisms behind place-based drivers of spending. Our findings imply that, while observable characteristics explain only a very modest amount of variation in causal place effects, there may be a small but meaningful role for Part D insurers in controlling overall prescription drug costs. Given the importance of physician-specific factors implied by our analysis, we also correlate our causal physician effects with a vector of physician demographics, experience, and medical school information. We find that, aside from experience, observables characteristics explain a negligible amount of variation in physician effects, suggesting that idiosyncratic practice style plays a large role in the prescription drug spending variation.

To perform our analyses, we leverage physician-level data from the Centers for Medicare and Medicaid Services (CMS) that details prescription drug utilization for the near universe of physicians treating patients covered by Medicare Part D. This gives a sample

[5] Interestingly, we find no evidence that gender, credentials (e.g., MD vs DO), or medical school ranking affect the extent to which physicians alter their prescribing behavior.

of almost 1.2 million physicians (125,538 of which moved between HRRs at some point between 2014 and 2019).[6] Additionally, we link prescription data with detailed data on the physician's demographic characteristics, including medical training, enabling an examination of heterogeneous responses along several different margins.

This paper makes contributions to several related strands of literature. First, this paper complements several existing studies which examine the determinants of geographic variation in healthcare costs and utilization. By tracking individual patient's healthcare utilization before and after moves, Finkelstein, Gentzkow and Williams (2016) estimate that approximately 50 percent of the geographic variation in healthcare utilization among Medicare beneficiaries is driven by differences in demand, with the remaining 50 percent determined by supply-side factors.[7] Our work contributes to this paper by (a) identifying the importance of physician-specific factors, one of potentially many important components of their estimate, and (b) focusing on prescription drug spending, which is not covered in their paper. We find that physician-specific factors can account for 60 to 70 percent of the variation in prescription drug spending in Medicare Part D.

Second, our work contributes to the literature on the determinants of physician practice styles. Methodologically, this paper is most closely related to Molitor (2018), who tracks cardiologists across moves to estimate the effect of their environment on their choice of heart attack treatment. Relative to Molitor (2018), this work considers a broader range of outcomes over which physicians arguably have more scope for discretion. This research question relates most closely to Cutler et al. (2019), who examine the extent to which variation in physician beliefs elicited through vignettes can explain variation in end-of-life medical spending. The advantage of our setting is that we are able to observe actual changes in individual physician behavior as their environments change. This paper

[6]Our data begin in 2013, which we use to establish an origin HRR.

[7]Related papers track migrating patients in other countries and other healthcare settings (see appendix Table A1 for details). Callison, Kaestner and Ward (2018) also find evidence supporting a supply-side explanation of regional variation in health care utilization by following uninsured individuals who age into Medicare.

also contributes to other work which examines importance of physician practice styles in determining patient outcomes (Zhang, 2018; Tu, 2017; Fadlon and Van Parys, 2020; Staiger, Baker and Hernandez-Boussard, 2022). In contrast to these papers, we estimate how physician practice styles respond to new environments as opposed to estimating how those styles impact patients.[8]

Finally, our paper contributes to a relatively small literature on physician learning. Phelps and Mooney (1993) posit a model in which physicians' practice styles are influenced by their education and previous experience, and are then updated via a Bayesian learning process based on the norms of their new work environment. This model has two testable implications, which our findings broadly support. First, the Bayesian learning process suggests that changes in behavior should occur slowly over time, rather than immediately. Second, physicians with less experience should respond more strongly to their new environments. We observe that migrating physicians adjust their practice style to reflect their new location over the course of 2 to 3 years after their move, consistent with the slow learning process predicted by the model. We note that physicians never completely converge toward the mean of their new location, indicating that time-invariant physician-specific characteristics play an important role in determining practice styles. Our heterogeneity analysis suggests that more experienced physicians respond less to changes in environment, a finding that also supports a model of Bayesian learning, since this framework indicates that veteran physicians will place more weight on prior beliefs. Our findings contrast those of Molitor (2018), who found that cardiologists' behavior converges abruptly to the norms of their new location and did not vary by experience at move. These differences suggest that institutional norms for certain procedures, such as the cardiac catheterization studied by Molitor (2018), may result in more rapid convergence compared to drug prescriptions, which may allow for greater provider discretion.

[8]Recent work by Doyle Jr and Staiger (2022) exploit physician migration between groups within the same hospitals to estimate how physician practice styles are affected by peers, as well as how these styles affect patient health.

2 Research Design

The primary analysis consists of a series of event study regressions which track changes in physician-level prescribing behavior in the years leading up to and following a move. Specifically, we compare physicians from within the same origin HRR who move to areas with either higher or lower average prescribing behavior. Restricting our comparisons within origin HRRs allows for flexible trends in utilization across origin HRRs. Our regressions take the form

$$y_{its} = \alpha_i + \gamma_{to} + \sum_{t \neq -1} \beta_\tau \cdot 1(t = \tau) \cdot \delta_{idos} + \epsilon_{its}, \quad (1)$$

where y_{its} denotes the outcome for physician i in specialty s in year relative to move t. Subscripts d and o index the destination and origin location, respectively. Individual fixed effects are included in the α_i term. We also include relative year-by-origin fixed effects in the γ_{to} term. These fixed effects restrict our comparisons to physicians from within the same origin HRR who move to areas with differing levels of prescribing. Likewise, these fixed effects control for any common trends affecting prescribers surrounding a move. For example, if the logistics of a move lead to a reduction in prescribing, this would be absorbed by these fixed effects.[9] The independent variable of interest, δ_{idos}, is constructed at the individual physician level and represents the difference in mean prescribing between all physicians in specialty s in mover i's destination and origin. That is, $\delta_{idos} = \overline{y_{ids}} - \overline{y_{ios}}$, where $\overline{y_{ihs}}$ is the average value of y across all physicians in HRR h and specialty s.[10] This specification uses the full sample of prescribers to determine the values of δ_{idos}, but only includes migrating physicians in the regression. Standard errors

[9] In results not presented here, we find that our results are largely unchanged by including various other combinations of fixed effects. For example, fully interacting origin, year, relative year, and specialty group fixed effects makes little practical difference in our results.

[10] The $\overline{y_{ihs}}$ terms are constructed by first calculating averages of y in each HRR-calendar year-specialty cell. We then average this measure across all years in our sample period to arrive at the final pooled measure of place-specialty utilization. We explore alternate constructions of δ_{idos} in Section A.2.

throughout are clustered at the physician-level.[11]

The primary coefficients of interest are the β_τs, the coefficients on the interaction between time relative to move and δ_{idos}, where β_{-1} is normalized to zero. Each β_τ can be interpreted as the difference in y relative to the period immediately preceding the move, scaled by the difference in average y_{is} across the destination and origin locations. This scaling provides a convenient interpretation of β_τ as the relative importance of environmental factors relative to physician-specific idiosyncratic factors.

For example, consider the extreme case in which physicians make prescribing decisions completely independently of their environment. That is, a physician makes his or her prescribing decisions based entirely on factors unique to that particular physician. Then, we would expect to see values of β_τ close to zero in both the pre- and post-move periods, as prescribing is set independently of environment. However, if at the opposite extreme prescribing behavior was determined entirely by environmental factors, then we would expect to see values of β_τ close to zero in the pre-move period and close to one in the post-move period. This is because if only environment matters, then we should expect $y_{its} = \overline{y_{ihs}}$, where h is location. In other words, physicians simply prescribe at the average rate of their location. Values of β_τ in the post-period that are closer to zero or one allow us to infer the relative importance of environment versus idiosyncratic physician-specific factors in determining prescribing behavior.

The identifying assumption in this model is that, absent the move, trends in y_{its} would not have varied systematically with δ_{idos}, conditional on the other controls. For example, if physicians who ultimately move to higher prescribing areas were disproportionately trending up in their own prescriptions prior to moving, the identifying assumption would be violated. Fortunately, the event study specification allows us to assess the extent to which the results are driven by pre-existing trends in prescribing behavior. We discuss

[11] We note that, because δ_{idos} is a simulated regressor, we should instead bootstrap the standard errors. However, doing so is computationally expensive. In results not presented here, we find that bootstrapping the standard errors makes very little difference for our two main outcome variables.

this further for each of the main results in Section 4.

A related concern is endogenous migration. It is possible that abnormally high (or low) prescribers intentionally move to areas which conform more closely to their prescribing preferences. In order to investigate this possibility, we examine the relationship between a mover's rank in their HRR's pre-move prescribing distribution and the size of their ultimate move. Specifically, the bottom x-axis in each panel of Figure 1 displays the rank in the pre-move prescribing distribution, scaled between 0 and 100. The y-axis shows δ, the size of the move. For each percentile in the pre-move prescribing distribution, we compute the average move size, which is indicated by the blue circles. Since each outcome variable has a different δ, we also display a histogram for each associated δ in gray. In each panel, the plots reveal that there is essentially no relationship between a physician's rank in the pre-move distribution and the size of the move.[12] This indicates that prescribers are not moving in response to a perceived mismatch between their desired prescribing behavior and that of their peers.

That there is no relationship between a migrant's δ and pre-move spending habits does not mean that there is no sorting. For example, if physicians from high-spending HRRs tend to migrate to high-spending HRRs and vice versa, the value of δ for these migrants would be close to zero. In appendix Figure A2, we plot the average outcome of the origin HRR against the average outcome of the destination HRR, both centered around zero. This figure shows that migrants from low spending areas do tend to move to low spending areas, although still to significantly higher spending areas than where they originate. We reiterate that our regressions only compare migrants from within the same origin HRR, so our results are not biased by average differences in destination spending that are correlated with origin spending.

In addition to the event study regressions discussed in equation 1, we also present coefficients from more succinct regressions of the form:

[12] In results not presented here, we find nearly identical results if we focus on physicians' ranks in the national prescribing distribution as opposed to their origin-HRR distribution.

$$y_{its} = \alpha_i + \gamma_{to} + \beta_{transition} \cdot \mathbb{1}(0 \leq t \leq 1) \cdot \delta_{idos} + \beta_{post} \cdot \mathbb{1}(t > 1) \cdot \delta_{idos} + \epsilon_{its}, \qquad (2)$$

where all terms are defined as in equation 1, although instead of interacting time-to-move fixed effects with δ_{idos} we instead include interactions of δ_{idos} with both a "transition period" indicator $\mathbb{1}(0 \leq t \leq 1)$ and a post-move indicator $\mathbb{1}(t > 1)$. We chose this particular specification because, as we discuss in Section 4, our event study coefficients tend to grow between the year of the move and the second full year post-move, after which they tend to stabilize. We therefore report β_{post} as a summary measure of the long-run treatment effect.

3 Data

Medicare is a federal health insurance program administered by the Centers for Medicare and Medicaid Services (CMS) which primarily insures individuals over the age of 65. Medicare enrolled nearly 60 million beneficiaries in 2017, with total spending of over $700 billion, about 15 percent of the federal budget.[13] Traditional fee-for-service Medicare (Parts A and B) provides hospital and medical insurance, but does not include prescription drug coverage.

Prescription drug benefits are covered under Medicare Part D. Part D refers to both stand-alone private Medicare prescription drug plans (PDP) or Medicare Advantage plans with drug coverage (MAPD). Approximately three quarters of Medicare beneficiaries are enrolled in Part D, among whom a slight majority are enrolled in PDPs.[14] Spending on Part D accounted for $93.9 billion in 2017, approximately 13 percent of total Medicare

[13] Source: https://www.cms.gov/Research-Statistics-Data-and-Systems/Statistics-Trends-and-Reports/NationalHealthExpendData/NHE-Fact-Sheet.

[14] Source: https://www.kff.org/medicare/fact-sheet/an-overview-of-the-medicare-part-d-prescription-drug-benefit/.

spending.[15]

The primary sources of data are the 2013-2019 Part D Prescriber Public Use Files (PDPPUFs). These data are constructed from the CMS Chronic Conditions Data Warehouse, which contains information on all prescription drug events (PDEs) for Part D beneficiaries. Specifically, these data record all "final-action" PDEs—meaning that the patient filled the prescription, which was then paid for by their Part D plan. The PDPPUFs aggregate these PDEs to the physician-by-drug level and include information about the total cost, quantity, and number of beneficiaries who receive the drug each year. Any physician who prescribes a drug to a single Part D enrollee at any point in the sample is included in the dataset. However, if a physician prescribes to 10 or fewer beneficiaries in a particular year, then the exact number of beneficiaries is censored. We note that even if the number of beneficiaries is censored, we still observe the total cost and quantity of drugs dispensed.[16]

The PDPPUFs also include information about physician specialty, which we aggregate into the following groups: primary care, medical specialties (e.g., cardiology, nephrology), surgical specialties (e.g., general surgery, orthopedic surgery), nurse practitioners, and dental specialties. Finally, the PDPPUFs also include prescriber demographics—including complete address— which allows us to track physicians as they move across locations. To our knowledge, this is the largest publicly available panel of physician-level prescribing information. We supplement these data with additional information on medical school attendance from Physician Compare.

Outcome Variable Descriptions These data contain several different measures of prescribing behavior. We focus most of our analysis on two outcomes, log(spending) and

[15]Source: https://www.medpac.gov/wp-content/uploads/import_data/scrape_files/docs/default-source/fact-sheets/mar19_factsheet_sec.pdf.

[16]In Section A.2, we show that our results are not sensitive to the inclusion or exclusion of these low-volume prescribers.

log(spending per beneficiary).[17] The first, log(spending), provides a summary measure of a physician's aggregate prescribing behavior incorporating both the intensive and extensive margins. In contrast, the latter measure controls for the number of Part D enrollees receiving at least one prescription, thus measuring prescribing behavior only along the intensive margin.[18] The intensive margin is a function of the both the quantity of drugs prescribed to each beneficiary, as well as the cost of these drugs. While these measures are highly correlated ($\rho = 0.87$) at the individual physician level, the underlying drivers of variation in each measure are quite distinct, underscoring the need to consider each outcome separately.[19]

We explore the robustness of our results to using other prescribing measures, specifically the log of claims or "days supplied," both overall and per beneficiary. In principle, these outcomes could yield different results if, for example, differences in spending were driven in large part by differences in the typical number of days of medication included in a prescription claim. In practice, we find that these measures are very highly correlated and yield almost identical results as the baseline spending measures.

Sample Construction For the core sample used in our regression analyses, we focus only on movers. Non-movers are utilized only to create the geographic averages used to construct our δ measures. We include movers in our sample if they meet two criteria. First, we require that movers be observed over a set of contiguous years. For example, if a provider was in our sample from 2016-2019, they would be retained, as would a provider present from 2013-2017. This allows us to retain information from new providers entering the workforce, as well as retiring providers. However, a provider who appears in our sample in 2014-2016 and 2018 would be dropped, as they are not continuously

[17] We use the log transformation because the spending distribution has a very long right tail. This is also true of other measures such as claims and days supplied.

[18] In the remainder of the paper, we use the term "beneficiary" to refer to Part D enrollees with at least one drug claim. When we use the phrase "per beneficiary," we are dividing by the number of Part D enrollees to whom a particular physician prescribed at least one drug.

[19] We discuss this in more detail in Section 4.3.

present.[20] Our second criteria is that providers make exactly one move during our sample period; individuals making two or more moves are dropped. Imposing these two sample restrictions leaves us with a sample of 125,538 movers.

Descriptive Statistics Summary statistics from these data are shown in Table 1. The table shows several measures of prescribing intensity, followed by physician and then patient characteristics. The unit of observation in these data is a physician-year, with each cell reporting the average across physicians. Column (1) displays these means for the entire sample while column (2) includes only physicians who do not move over the sample period. Columns (3) and (4) show the summary statistics for movers both before and after they move. We have a total of just under 1.2 million physicians in the dataset, 10.5 percent of whom move across HRRs sometime during the sample period.

Comparing columns (2) and (3), we see that prescribing behavior is qualitatively similar for movers and non-movers. However, there are a few notable demographic differences among movers and non-movers. First, movers tend to be significantly less experienced, which we measure as the number of years since the physician finished medical school. The average non-mover in the sample has about 21 years of experience, double that of movers. Second, movers are significantly more likely to be female than non-movers. Third, movers are more likely to specialize in primary care. Finally, movers are slightly less likely to have attended a medical school ranked in the top 100.[21]

Movers also tend to treat a slightly different pool of patients than non-movers. In particular, they treat a higher fraction of black patients, dual Medicare-Medicaid eligible patients, patients with slightly higher risk scores, and patients under 65 years old. On average, neither movers' prescribing behavior nor the characteristics of their patients change substantially following their move. However, in Section 4 we show that this masks sub-

[20] In Section A.2, we show that we obtain nearly identical results if we impose balance in calendar years. Likewise, we show results separately for each cohort of movers and find that the pattern of results is similar for each cohort.

[21] Medical school rankings are taken from Schnell and Currie (2018).

stantial variation among those who move to higher or lower prescribing areas.

There are two important limitations to these data. First, the data do not cover all drugs that a particular physician prescribes, only those paid for by Part D. Therefore, we are unable to comment on physician prescribing patterns among the general population. Second, we only know the number of patients who received a prescription, not the total number of patients that the physician interacted with. While this does not cause any issues for the analysis along the intensive margin as we are normalizing by the number of patients receiving at least one prescription, it could cause problems for looking at total spending. For example, if physicians in certain areas see a higher fraction of Part D beneficiaries, then simply looking at the number of prescriptions or total spending would conflate the change in the physician's prescription behavior with the change in their patient population. However, this only causes bias if the number of Part D beneficiaries with whom physicians interact varies systematically with respect to δ_{idos}.[22] To the extent that the number of Part D beneficiaries per physician is relatively constant across HRRs or uncorrelated with the gap between destination and origin prescribing behavior, the estimates will remain unbiased.

To probe the robustness of the overall spending results, we compute the ratio of the total number of Part D enrollees to physicians in each HRR. We then interact this ratio with a set of time-to-move dummy variables in our main regressions. The idea here is to explicitly control for changes in the number of available patients. Results from this exercise are discussed in Section A.1, and are qualitatively similar to the main regression results.

Identifying Variation As discussed in Section 1, there is substantial geographic variation in physicians' prescribing behavior. Figure 2 demonstrates this variation for our two

[22] Of course, if we view the composition of patients as endogenous then this is not a bias.

main prescribing measures, averaged across all physicians in each HRR.[23] More specifically, panel (A) is constructed by first averaging total drug spending across all years for each physician. Then, we average across all physicians in each HRR. Panel (B) is constructed similarly, but averaging total drug spending per beneficiary with at least one claim, as opposed to total spending.[24] These maps demonstrate that there is significant variation in prescribing behavior across physicians in different HRRs.[25]

Our regression framework takes advantage of this variation by tracking physicians who originate in the same HRR, but move to different HRRs before and after the move. Specifically, we construct the measure of "move size", δ_{idos}, by computing the difference in the specialty-specific average value of the outcome in the destination and origin HRR for each mover using the entire sample of physicians.[26] The distribution of δ_{idos} for the two main outcome variables are shown in appendix Figure A3. The distributions are both approximately normally distributed and centered around zero, with larger (positive) values indicating more intensive prescribing in the destination HRR relative to the origin. Identification of the coefficients of interest comes from comparing how prescribing behavior changes for physicians who start in the same HRR, but move to HRRs with differing values of prescribing intensity, before and after the move.

[23] In our regressions, we take the log of these variables. Here we show the levels to facilitate easier comparisons.

[24] Note that the denominator is the number of beneficiaries who had at least one prescription drug claim, which is not necessarily the number of Part D enrollees that a particular physician interacted with.

[25] A map constructed at the HRR-level, rather than the physician level, is shown in appendix Figure A1. This map displays the total spending per Part D enrollee in each HRR.

[26] There are an average of 3,900 providers in each HRR. The smallest HRR in our sample has 408 providers.

4 Results

4.1 Main Event Studies

Log(spending) In this section, we present the results from regression equation 1 for our two main outcome variables, log(spending) and log(spending per beneficiary). We begin by considering log(spending). These results are shown in panel (A) of Figure 3. Immediately upon moving, the coefficients become positive and statistically significant. The change in spending during the year of the move is relatively small, but continues through the second year post-move, at which point it begins to stabilize. This indicates that physicians' prescribing behavior moves toward the average physician in their new destination, implying that environmental factors play a key role determining prescribing behavior. It is interesting to note that the growth in spending stops after the first few years post-move, indicating no further changes in behavior after an approximately three year adjustment period. In the top-right corner of this figure, we display the β_{post} coefficient from equation 2, which is equal to 0.42. This coefficient suggests that a physician moving to an HRR with 100 percent higher drug spending increases his or her own spending by 42 percent.

As discussed in Finkelstein, Gentzkow and Williams (2016), the exact magnitude of these regression coefficients provides an estimate of the relative importance of mover-specific characteristics versus environmental factors in explaining the gap between the origin and destination locations.[27] A value of 0 would indicate that all variation is due to physician-specific factors, while a value of 1 would indicate that all variation is due to environmental factors such as patients, insurers, and hospitals. Our β_{post} coefficient of 0.42 indicates that 42 percent of the geographic variation in drug spending is driven by environmental factors, with the remaining 58 percent $(1 - 0.42)$ of the of the geographic variation in drug spending attributable to idiosyncratic physician-specific characteristics.

[27] In their context, it is patients who are moving, as opposed to physicians.

Relative to the β_τ coefficients after the move, the coefficients in the pre-period are modest in magnitude. In fact, four of the five coefficients are statistically indistinguishable from zero at conventional levels. We view these coefficients as broadly consistent with the identifying assumption that movers were not systematically changing their prescribing behavior prior to moving. The -5 coefficient, however, just meets the threshold for statistical significance. We also note that, while statistically insignificant, the coefficients are slightly increasing during the periods prior to the move. While we cannot reject that this pattern is simply noise, one may view it as evidence of pre-existing trends. Specifically, this would indicate that migrants who move to higher spending HRRs were slightly (differentially) increasing in their spending prior to moving. This would suggest that some of the growth observed in the post-move period is not the result of the new environment, but the continuation of existing trends. We note, however, that even under this interpretation the growth in the pre-period coefficients is dwarfed by the growth observed in the years after moving.

Log(spending per beneficiary) Next, we present the results from equation 1 with log(spending per beneficiary) as the outcome variable. This variable is created by taking total drug spending for a specific physician, and dividing by the number of beneficiaries to whom the physician prescribed at least one drug. We note that δ is defined in terms of the outcome variable, so the identifying variation underlying this regression is different from the regression in panel (A). Because this variable accounts for the number of beneficiaries that a provider prescribes to, we can conceptualize it as a measure of spending along the intensive margin. These coefficients therefore indicate how responsive migrating physicians are to local prescribing norms, conditional on any change in the number of beneficiaries.

The coefficients from this regression are displayed in panel (B) of Figure 3. The pattern of coefficients is qualitatively quite similar to those in panel (A). This provides assurance that the results in panel (A) are driven by actual changes in prescribing behavior as opposed to changes in the composition of Part D versus non-Part D patients after moving. However, despite the similar pattern of the coefficients, the magnitude is somewhat smaller when examining spending per beneficiary as opposed to the total spending. The regression coefficient from equation 2 is shown in the top right-hand corner of panel (B). The value of 0.29 indicates that physicians' intensive margin spending behavior is stickier than their total spending behavior.

4.2 Describing the Response Dynamics

The regression coefficients in both panels of Figure 3 exhibit broadly similar dynamics. The coefficients begin to grow immediately upon moving, although the growth is initially quite modest. However, the coefficients continue to grow over time. For log(spending), this growth continues until the second full year after moving, at which point spending stabilizes. For log(spending per beneficiary), the growth continues throughout the five years that we observe after moving, although it slows notably after the second year. This pattern can be rationalized by a model of physician learning. Phelps and Mooney (1993) suggest that physicians' initial practice styles are formed during medical school and residency. These styles continue to evolve following a Bayesian learning process as physicians are exposed to new information.

Our findings align with the predictions of this model. Since practice styles are a function of both initial training and accumulated knowledge, this model predicts that changes in practice styles would occur slowly over time, which is precisely what we observe.[28] A second testable implication of this model is that physicians with more experience should be less responsive to changes in their environments. We examine this hypothesis in detail in Section 5.1, and again find that our results are consistent with the predictions of this model. These finding stand in stark contrast to Molitor (2018), who finds that cardiologists' treatment decisions for heart attack patients change discretely upon move with

[28]We consider alternative explanations in Section A.1.

no further growth, and do not vary by experience. However, Molitor (2018) finds that physicians are more responsive to their new environments in cases where the medical benefits of different treatment options are less certain.[29] Greater uncertainty surrounding appropriate drug prescribing relative to appropriate heart attack treatment could therefore partially reconcile the differences in our papers.

4.3 Spending Response Decomposition

Log(spending) In this subsection, we rigorously decompose the total spending response into component pieces. The goal of this exercise is to shed light on which specific behavioral changes drive our results. Mechanically, there are several different changes in prescribing behavior that could generate the spending response that we observe. To see this, note that we can re-write log(spending) as

$$log(spending) = log(\frac{spending}{claim}) + log(\frac{claims}{beneficiary}) + \log(beneficiaries)$$

At this point, we can substitute in regression equation 2 and derive that

$$\hat{\beta}^{spending}_{post} = \hat{\beta}^{spending/claim}_{post} + \hat{\beta}^{claim/beneficiary}_{post} + \hat{\beta}^{beneficiaries}_{post}$$

This shows that the observed spending response is itself a function of three separate responses. First, total spending is influenced by the amount of spending on each claim. This could be affected by either price differences in the same drugs across HRRs, or differences in the types of drugs prescribed. Second, total spending is influenced by the total number of claims for each beneficiary. At a high level, we can think about this as a proxy for the intensity of treatment. Finally, spending is influenced by changes in the number of beneficiaries receiving drug prescriptions. We show the results from this decomposition

[29] Clinical guidelines are clearer about appropriate treatment of STEMI heart attacks relative to NSTEMI heart attacks. For the latter, physicians are more likely to adopt the behavior of their new environment.

in panel (A) of Figure 4.[30]

This figure shows that the majority of the total spending response (79.2%) is driven by migrating physicians prescribing drugs to more beneficiaries. We also observe a smaller, but still highly significant increase in the number of claims per beneficiary (18.9%). Interestingly, there is no significant change in the cost per claim. Overall, this breakdown highlights that the most important mechanism driving our effects is an increase in extensive margin prescribing, that is, prescribing drugs to more patients. The remainder of the response is driven by increases along the intensive margin. In principle, the intensive margin response could be driven by increases in the number of drug claims per beneficiary, or the cost per claim. We find that the number of claims per beneficiary plays an important role in explaining our findings. This suggests that physicians not only increase the number of patients to whom they prescribe drugs, but also prescribe more drugs to each of them. In contrast, we find virtually no response in the cost per claim.[31]

Log(spending per beneficiary) In panel (B), we show the results from a similar breakdown for our log(spending per beneficiary) regression. Here, we break down the outcome variable into log(claims per beneficiary) and log(cost per claim). Interestingly, the decomposition is quite different from what we would have expected given the results in panel (A). In particular, the total change in spending per beneficiary is split somewhat equally in terms of claims per beneficiary and cost per claim. Considering just the intensive margin response in panel (A), the split is closer to 90-10. The reason for this apparent puzzle is that the underlying spatial variation in log(spending) is quite different from the variation in log(spending per beneficiary) (see Figure 2). In particular, the variation in

[30]We drop all prescribers for whom the number of beneficiaries is censored in these regressions so that the sample remains constant. We show in appendix Figure A4 that this does not meaningfully change our log(spending) results.

[31]In results not shown here, we find nearly identical results when substituting days supplied for claims in this breakdown. In a separate exercise, we also consider whether the length of claims (i.e., days supplied per claim) plays any role our regressions focusing on days supplied as our outcome variable, and find that virtually all of the variation in days supplied is driven by the number of claims as opposed to the length of the claims.

log(spending) is driven primarily by the number of beneficiaries receiving drug claims, with a smaller role for the number of claims per beneficiary and very little role for the cost per claim. In contrast, the cost per claim plays a much more important role in explaining the geographic variation in log(spending per beneficiary). Essentially, HRRs with high spending per beneficiary tend to be places where expensive drugs are being prescribed, whereas HRRs with high overall spending tend to be places with many beneficiaries receiving claims.

We demonstrate this in appendix Figure A5. In panel (A), we display a scatter plot of log(spending) against log(cost per claim) at the HRR level. In panel (B), we re-create this plot with log(spending per beneficiary) along the y-axis. Interestingly, higher costs per claim has very little association with total costs, highlighting the importance of the extensive margin in explaining log(spending). In contrast, more expensive claims are strongly associated with higher log(spending per beneficiary). Our findings are consistent with a story in which physicians adopt the behaviors that are prevalent in their new environments, but there are very different underlying behaviors in HRRs with high log(spending) versus high log(spending per beneficiary).

4.4 Brand Versus Generic Spending

To this point, we have established that migrating physicians partially adopt the prescribing behavior of their new locations. Whether this adoption is positive or negative, however, is unclear. If physicians are adopting practice styles more in-line with established best practices, then this convergence can be thought of as welfare enhancing. In contrast, if they are adopting inappropriate practice styles then this convergence could be welfare decreasing. In order to specifically test for the adoption of "better" prescribing behavior, we consider whether physicians become more likely to prescribe generic medications when moving to an area with a higher share of generic spending.[32] Specifically, we estimate equation 1 with y_{its} (and the corresponding δ) defined as the share of total spending spent on generic drugs.

We present the results from this exercise in Figure 5. The pattern of the coefficients broadly mimics what we observed in Figure 3. The coefficients grow slowly over time, before stabilizing approximately three years after the move. The summary coefficient from equation 2 is 0.33. This indicates that, upon moving from an area with no spending on generics to an area where all spending is on generics, migrating physicians increase their share of generic spending by 33 percentage points. Mechanically, this estimate pools the responses of physicians moving to areas with relatively more generic spending, as well as physicians moving to relatively less generic spending. When we estimate these responses separately, we find that physicians who move to areas with less generic spending than their origin respond more heavily than than vice versa. Specifically, we estimate $\hat{\beta}_{post} = 0.54$ when we restrict the sample to migrants with $\delta < 0$, as opposed to $\hat{\beta}_{post} = 0.23$ when we restrict the sample to migrants with $\delta > 0$. Interestingly, this implies that physicians are more likely to begin prescribing more brand name drugs in areas where prescribing brand drugs is common, as opposed to prescribing more generics in areas where generic prescribing is more prevalent. While this is not a complete examination of changes in appropriate versus inappropriate prescribing behaviors, it is suggestive that physicians are susceptible to picking up bad habits from their peers, as opposed to just good habits.[33]

A natural question is whether these responses represent actual changes in physicians' practice styles, as opposed to changes in pharmacy substitution regulations in different states. For example, some states have "mandatory substitution" laws that require phar-

[32]Generic drugs are pharmacologically identical to brand name drugs, but are less expensive. The American College of Physicians recommends the use of generic equivalents over brand name drugs whenever possible (Choudhry et al., 2016).

[33]Ideally, we would also examine other deviations from commonly accepted clinical guidelines. For example, in principle we could examine co-prescribing of contraindicated medications. However, without information about individual patients this is not possible.

macists to dispense the generic version of a drug if possible. Likewise, states vary based on whether patients need to explicitly consent to a substitution being made, or whether the pharmacist can assume that the patient consents unless explicitly stated otherwise ("presumed consent").[34] Song and Barthold (2018) finds that mandatory substitution laws have no effects, but presumed consent laws reduce consumers' probability of purchasing brand name drugs by 3.2 percentage points.

We conduct several analyses in order to examine whether these laws could drive our results. First, we re-estimate equation 1 while including interactions between relative year indicators and indicators for whether a state has a mandatory substitution law and whether the state has a presumed consent law.[35] We show these results in appendix Figure A6. The baseline coefficients are shown in blue, while the coefficients from the specification controlling for the regulatory environment of the state are shown in red. We note that the results are nearly identical, suggesting that differences in the regulatory environment are not driving our effects. Next, we estimate equation 2, restricting our sample to only include migrants for whom these laws did not differ between their origin and destination. If our results were in fact driven by differences in pharmacy regulations across states, then we would expect our results to be significantly attenuated by this restriction. When we keep only migrants whose presumed consent status did not change, we obtain an estimate of 0.33, nearly identical to our baseline. When we restrict the sample to include those whose mandatory substitution status did not change, our point estimate is 0.39.[36] In both cases, we observe broadly similar effects, suggesting that our findings are in fact the result of changes in the prescribing behavior of the physicians.[37]

[34] A state's regulatory environment is characterized by the combination of mandatory substitution and presumed consent laws. A state may have any combination. In all cases, the patient can refuse the substitution.

[35] We include the interactions because it may take time for migrants to adapt to the new regulatory environment.

[36] We do this in two separate regressions (as opposed to keeping only those with the exact same combination of laws) in order to maintain power.

[37] This does not necessarily imply that these laws do not affect brand versus generic spending. Rather, these results imply that these laws do not have any additional impact on physicians' behavior after conditioning on the change in the prescribing behavior of their peers. In results not presented here, we consider

Brand-Generic Costs and Total Spending Because brand name drugs tend to be more expensive than generic equivalents, it is possible that the decision to prescribe brands or generics could account for some of the observed spending response. We investigate this by estimating a version of equation 1 in which the δ term is defined in terms of log(spending) (or log(spending per beneficiary)), but the outcome is the share of total spending spent on generic drugs. These regressions measure how the fraction of spending on generics changes as physicians migrate to areas with higher or lower total spending. The results are shown in appendix Figure A7. The summary coefficient from panel (A) indicates that a 100 percent increase in destination spending is associated with a 1.2 percentage point decrease in the migrant's generic spending rate. This implies that, as providers migrate to areas with higher spending, they begin to prescribe relatively fewer generic drugs (and relatively more brand-name drugs). Recall that the coefficient from panel (A) of Figure 3 is 0.42, which indicates that a 100 percent increase in destination spending is associated with a 42 percent increase in the migrant's spending. At first glance, it is not obvious how to determine how much of the 42 percent increase in spending can be accounted for by a 1.2 percent decrease in the share of generic spending. As a simple back-of-the-envelope, we estimate the relationship between log(spending) and the generic spending share to help estimate how a decrease in the generic share translates into a change in spending.[38] We find that a one percentage point decrease in the generic share is associated with a 2.7 percent increase in spending and a 2.3 percent increase in spending

regressions similar to equation 2 where we replace the δ regressor with variables reflecting the change in the migrants legal environment. Specifically, we estimate

$$y_{it} = \alpha_i + \gamma_{to} + \beta_{transition} \cdot \mathbb{1}(0 \leq t \leq 1)\text{law switch}_{it} + \beta_{post} \cdot \mathbb{1}(t > 1) \cdot \text{law switch}_{it} + \epsilon_{it},$$

where law switch$_{it}$ is equal to 0 if the legal environment stays the same, 1 if moving from a permissive to mandatory substitution state, and -1 if vice versa. We estimate a coefficient of 0.011, which suggests that moving to a mandatory substitution state increases the generic spending share by about 1.1 percentage point. We find no statistically significant effects for presumed consent laws.

[38] Specifically, we estimate the following regression:

$$log(spending)_{it} = \beta_0 + \beta_1 generic\_share_{it} + \alpha_i + \gamma_t + \epsilon_{it},$$

where α_i and γ_t are individual and year fixed effects, respectively. We estimate this regression on the sample of movers, keeping only observations prior to the move year to avoid contamination.

per beneficiary. Using these estimates to scale our regression coefficients from appendix Figure A7, we estimate that the change in brand versus generic prescribing after moving can account for about 7.8 ((0.012*2.7%)/0.42%) percent of our primary log(spending) results and about 10.3 ((0.013*2.3%)/0.29%) percent of our log(spending per beneficiary) results.

5 Mechanisms

In the remainder of this paper, we explore two primary questions. First, what physician and non-physician characteristics are associated with larger spending responses? Second, we ask what physician and place characteristics are associated with higher *levels* of spending? The answers to these questions aid in our understanding of both why physicians change their behavior, as well as what types of behavior they are prone to adopting. Furthermore, they shed light on why spending is higher in certain areas. Understanding the drivers of both physician responsiveness to local spending patterns as well as the underlying correlates of spending can aid in identifying ways to reduce spending.

5.1 Exploration of Physician Characteristics

Experience We begin by exploring how physician-specific characteristics affect physicians' responsiveness to their new environments. The first characteristic that we consider is experience.[39] The learning model of Phelps and Mooney (1993) predicts that providers with less experience should respond more strongly to new information, as their priors will be given less weight. We test this hypothesis by interacting our transition and post dummies from equation 2 with indicators for different categories of experience. Specifically, we create indicators for bins of experience (0-4, 5-9, 10-14, 15-24, and 25+ years). We

[39] We define experience as the number of years since graduation from medical school. The following regressions do not include non-physician prescribers such as nurse practitioners, as they did not attend medical school and we have no information about when they completed their training.

present the results of this exercise in Figure 6. Consistent with the learning model, the coefficients monotonically decrease with experience. This is true for both log(spending) as well as log(spending per beneficiary), although the former displays a noticeably steeper gradient. In terms of magnitudes, it is useful to compare the coefficients for the least versus most experienced physicians. Upon moving to an HRR with 100 percent higher spending, a physician with fewer than five years of experience increases his or her own spending by slightly more than 60 percent, on average. In contrast, we would expect a provider with 25 or more years of experience to increase his or her own spending by less than 20 percent. Focusing on the intensive margin, this gap is smaller, but the difference is still a statistically significant 20 percentage point gap. These findings suggest that prior experience is a key determinant of the malleability of physicians' prescribing behavior.[40]

Specialty Next, we consider heterogeneity by specialty. We break prescribers into three broad categories: (1) primary care providers, (2) medical specialists, and (3) surgical specialists. For each category, we estimate equation 1 and present the results in Figure 7. In panel (A), we show the results when the outcome variable is log(spending). We display the coefficients for primary care providers in blue, medical specialists in red, and surgical specialists in gray. For each category, we also present the summary coefficients from equation 2 in the bottom-right corner of the panel. We find the strongest responsiveness among surgical specialties, with a summary coefficient of 0.61. Among primary care providers, the coefficient is 0.43, almost identical to our baseline results pooling all providers. Medical specialties appear to be the least responsive, with a summary coefficient of 0.33. The results for log(spending per beneficiary) are shown in panel (B). We again see that medical specialties respond less than surgical specialties, and among both sets of providers the response is smaller than in panel (A). We observe the largest summary coefficient among primary care providers, although in examining the event study coefficients it appears as

[40] In results not presented here, we examine whether more experienced physicians differ from less experienced physicians in terms of responsiveness along the intensive versus extensive margins following the discussion in Section 4.3. We find that there are no statistically significant differences.

though this is largely the result of the continuation of pre-existing trends. We therefore interpret these results with some caution.

In results not presented, we examine heterogeneity along a number of other characteristics. Interestingly, we find no evidence of heterogeneity by gender, credentials (e.g., M.D. versus D.O. versus D.D.S.), or medical school rank.[41] Overall, these results suggest somewhat limited scope for heterogeneous responses by demographic characteristics, although there does exist heterogeneity along some dimensions.

5.2 Exploration of Non-Physician Characteristics

In this section, we explore the extent to which non-physician-specific factors can explain the spending responses that we observe. The ultimate goal behind these exercises is to shed light on what drives the aggregate spending responses. Put differently, what is it about higher spending areas that causes physicians who move there to increase their own spending? We consider four hypotheses: (1) differences in pharmaceutical marketing intensity, (2) differences in patient characteristics, (3) differences in patient demand, and (4) differences in competitiveness.

Pharmaceutical Marketing First, we examine whether differences in pharmaceutical marketing can help to explain our findings. Prior research has found that increased spending by pharmaceutical companies affects physician prescribing behavior (e.g., Carey, Lieber and Miller, 2021). One possible explanation of our findings is that differences in spending are largely the result of differences in pharmaceutical marketing intensity, and that when providers migrate to areas where they are more intensely exposed to this marketing, they change their behavior accordingly. To test this hypothesis, we use data from Open Payments—a dataset which tracks payments from pharmaceutical companies to

[41] All results available upon request.

providers—to create measures of HRR-level marketing intensity.[42] We then construct δ terms for marketing intensity, analogous to the spending δ terms in equations 1 and 2. That is to say, the marketing intensity δ captures the change in pharmaceutical marketing from one location to another, such that a marketing $\delta > 0$ indicates an increase in pharmaceutical payments from a provider's previous location, while $\delta < 0$ represents a decrease. In appendix Figure A8, we present the results from equation 1 where we control for pharmaceutical payments in several different ways. Panel (A) shows the baseline results, equivalent to panel (A) of Figure 3. In panel (B) we include interactions between relative year and the marketing intensity δ term. Panels (C) and (D) restrict the sample to moves to higher and lower marketing HRRs, respectively. Including these measures of pharmaceutical marketing intensity has almost no impact on our results.[43] This suggests that differences in pharmaceutical marketing are not an important factor in determining the extent to which physicians change their prescribing behavior when moving across HRRs. This is broadly consistent with the findings of Carey, Lieber and Miller (2021), who find that pharmaceutical payments affect prescribing behavior in the very short term, but these effects fade away within a year.

Patient Characteristics Second, we consider the extent to which our results could be driven by differences in observable patient characteristics. While our Part D data has somewhat limited patient-side information, we are able to consider four different characteristics: (1) sex, (2) dual Medicare-Medicaid eligibility, (3) patient risk-score, and (4) age.[44] For each of these characteristics, we separately estimate equation 2 for migrants who move to HRRs with higher or lower levels of the associated characteristics. We show these results in Figure 8. Panel (A) displays the results for log(spending), while panel (B) displays the results for log(spending per beneficiary). In the first row, we examine

[42]Ideally, we would measure payments at the specific drug level. However, the vast majority of payments in these data do not specify a specific drug; they only list the pharmaceutical company.

[43]We find similar null impacts examining log(spending per beneficiary).

[44]The information on patient race is too heavily censored to use here.

migrants who move to HRRs with a higher or lower fraction of female patients (in blue and red, respectively). Conditional on similar changes in spending, migrants are more responsive to their new environments when moving to an area with a higher fraction of male patients.

In the next row, we consider migrants who move to HRRs with a higher or lower fraction of dual Medicare-Medicaid eligibles. These are primarily poor and disabled individuals. Interestingly, migrants are more responsive when moving to HRRs with a smaller fraction of dual eligibles. We see similar results when considering risk scores in the next row; physicians are more responsive when moving to areas with less sick patients. Ex ante, it is theoretically ambiguous whether we'd expect physicians to be more or less responsive when moving to areas with healthier patients. On the one hand, the care that these patients receive on the margin is likely less impactful, so providers may be more likely to default to local norms. This would lead one to expect more responsiveness to environmental factors when patients are relatively healthy. On the other hand, one might expect that sicker patients are more complicated, and in the absence of clear clinical guidelines providers might rely more heavily on the norms of the area, both as a guide for appropriate treatment and to minimize liability. This would suggest that physicians would be more responsive to environmental factors when patients are sicker. The fact that we find more responsiveness when patients are healthier is suggestive that, when decisions are more impactful, physicians are more likely to rely on their own personal practice style.

Lastly, we consider differences in responsiveness by patient age. We find that physicians are more responsive when moving to areas with older patients. At first glance, this may seem at odds with the previous results examining heterogeneity by patient risk score. Intuitively, one might expect that age is an important factor in determining risk scores. However, computing the correlation between average patient age and average risk score reveals that these two are largely uncorrelated.[45]

Patient Demand Next, we consider whether our results can be explained by differences in patient demand. We do this by examining heterogeneous responses by drug class. The idea here is that some prescriptions are likely driven by patient request, while many others are for drugs that patients do not have any pre-existing knowledge of. For example, a patient with chronic pain (or a painkiller addiction) is likely to request opioids (Finkelstein et al., 2022). In contrast, a patient is unlikely to have strong preferences for statins. A secondary benefit of this analysis is that it allows us to examine physician responsiveness as a function of the clarity of clinical guidelines. For example, in cases where there are no clear clinical guidelines describing appropriate prescribing practices, one may expect physicians to default to the behavior of their peers. In contrast, if clinical guidelines are clear then we may expect physicians to be less responsive to their environments. Molitor (2018) finds some evidence in support of this hypothesis, noting that physicians are more responsive to environmental factors when treating heart attacks where the clinical guidelines are weaker.

We classify drugs into one of fourteen different categories, and re-estimate equation 2 separately for each class.[46] We present the results of this exercise in Figure 9. The top row presents the baseline coefficient from equation 2. Each subsequent row presents the coefficient from a separate regression, with the outcome variable (and δ) defined in terms of spending on drugs in a particular class.[47] The bottom x-axis indicates the magnitude

[45] This is consistent with selection into old age. For example, very sick individuals do not tend to survive into old age, which negates the general trend of deterioration of health over time that is observed within individuals.

[46] The Medicare Part D detailed files include strings for the brand and generic drug names, but do not include NDC codes or information about therapeutic classes. In order to group these drugs into categories, we perform a fuzzy string merge with drug names from the Medical Expenditure Panel Survey (MEPS), which has information about both drug names and therapeutic classes. We then hand-checked each drug match and manually generated matches based on online searches for drugs which were un-merged by the fuzzy string merge algorithm.

[47] We do not present results for log(spending per beneficiary), as the number of beneficiaries receiving specific drugs is censored for over 60 percent of our prescriber-drug-year observations.

of the coefficient, while the top x-axis measures the total fraction of spending on drugs in a particular category.

The overall results are somewhat mixed. The coefficients for cardiovascular drugs (e.g., beta blockers), antineoplastics (e.g., chemotherapy), antihyperlipidemics (e.g., statins), and benzodiazepines (e.g., Xanax) are all statistically less than the baseline coefficient. The first three of these drug categories have relatively straightforward clinical indicators, which is supportive of the hypothesis that environmental factors matter less when guidelines are clear. These drugs also provide no recreational benefits, so we expect that patients would not have strong underlying demand for these drugs.

However, benzodiazepines are often prescribed based on somewhat subjective self-reported criteria and are commonly used recreationally, both of which would lead us to predict that the coefficient would be greater than the baseline. Interestingly, Ding (2022) finds that 60 percent of the geographic variation in mental health care claims are attributable to place-specific factors, although patients with prior mental health claims converge completely toward their destination average. This suggests that patient demand may not be a large driver of variation of mental health claims, which could explain the lack of convergence among migrating physicians.

The only class of drugs which is statistically greater than the baseline is opioids. This is consistent with both a story of higher patient demand as well as unclear prescribing guidelines leading prescribers to default toward the HRR average. Finkelstein et al. (2022) find that among opioid naive SSDI beneficiaries, the vast majority of variation in opioid abuse is driven by patient-specific factors. That patients are an important driver of variation in opioid abuse helps rationalize our finding of outsized importance of environmental factors in driving variation in opioid prescribing. In contrast, Finkelstein et al. (2022) find that prior opioid users see immediate increases in abuse upon moving to higher prescribing areas. The likelihood of abuse continues to rise in the following years, suggesting that place-specific factors play a critical role in opioid abuse among prior users.

Similarly, Eichmeyer and Zhang (2022, 2023); Staiger, Baker and Hernandez-Boussard (2022) and Laird and Nielsen (2016) find that providers play an important role in driving variation in opioid prescriptions. Our findings are consistent with this narrative, although our results suggest that the role of providers in opioid use may be smaller than the role of providers in other drug use.

In Section 5.1, we found that less experienced physicians are more responsive to their environment than more experienced physicians. We now examine whether this differential responsiveness is driven by any particular class of drugs. In appendix Figure A9, we plot the difference in responsiveness between the most and least experienced categories of physicians for each class of drugs, as well as the difference in responsiveness to generic spending. Interestingly, this figure shows that less experienced providers are more responsive uniformly across almost all drug classes. The one clear outlier is for erectile dysfunction medications (e.g., Viagra), for which more experienced providers appear relatively more responsive. In the last row, we also consider the choice to prescribe brand versus generic drugs, and again find that less experienced providers are more responsive to their environments.

HRR Competitiveness Lastly, we test for differential responsiveness by competitiveness. Recent work by Currie, Li and Schnell (2023) has found that general practice physicians increase their prescribing of opioids and anti-anxiety medications in response to increased competition. We examine this possibility in our context by constructing δ terms measuring the change in provider level HHI.[48] We then estimate equation 2 separately

[48]Since we don't possess information on total revenue (traditional Medicare + MA + Non-Medicare) among physicians and practice groups, we calculate a modified version of physician HHI using the total number of providers attributed to a given practice group, where practice group information is obtained from the Physician Compare database.

Specifically, we calculate (1) the number of providers in a group (g)-by-specialty (s)-by-HRR (j) cell and divide by (2) the *total* providers in a specialty-by-HRR cell to obtain our $share_{gsj}$ measure for the HHI calculation. Then we calculate provider HHI at the specialty-HRR level as follows:

$$PHHI_s j = \sum_{g \in G} (share_{gsj})^2$$

for physicians migrating to more and less competitive HRRs. Interestingly, we find no evidence of differential responsiveness across these moves. The summary coefficient for log spending for those moving to more competitive HRRs is 0.43, compared to 0.40 for those moving to less competitive HRRs. When considering log spending per beneficiary, these numbers are 0.28 and 0.29, respectively.[49] These results suggest that, conditional on the change in spending occurring upon a move, changes in the competitiveness of the environment are not an important determinant of changes in migrant spending.

5.3 Correlates of Place and Physician Effects

The analysis in the prior subsections concerns drivers of the *responsiveness* of physicians to their environments. However, in Section 4 we found that non-physician environmental factors–which we now refer to as "place"–account for approximately 30 to 40 percent of the observed variation in spending. Likewise, time-invariant physician-specific factors account for approximately 60 to 70 percent of the variation in spending. In this subsection, we analyze the characteristics of these physician and place factors.

To do so, we estimate the following regression:

$$y_{ijst} = \gamma_{jst} + \alpha_i + \varepsilon_{ijst}, \tag{3}$$

where y_{ijst} represents either provider i's log spending or log spending per beneficiary, γ_{jst} represents an HRR-by-relative year-by-specialty fixed effect, and α_i represents a physician fixed effect.

From this regression, we recover the estimated place-by-year-by-specialty effects ($\hat{\gamma_{jst}}$) as well as physician effects ($\hat{\alpha_i}$) for further analysis. We collapse our place effects down to the place-specialty level to obtain the average causal effect over our sample period ($\hat{\gamma_{js}}$). Following Finkelstein, Gentzkow and Williams (2016), we then explore the drivers

[49] Event studies are available upon request.

of these physician and place effects.[50]

Place Effects To better understand *associations* between our recovered place effects ($\hat{\gamma_{js}}$) and different environmental characteristics, we estimate bivariate regressions as follows:

$$\hat{\gamma_{js}} = \Gamma \cdot X_{js} + \xi_{js}, \tag{4}$$

where X_{js} is a single, selected HRR-level characteristic.

The results of this exercise are displayed in panels (A) and (B) of Figure 10 for log(spending) and log(spending per beneficiary), respectively. Each coefficient estimate and the associated confidence intervals are displayed in blue. Corresponding regressions displaying the relationship between raw averages ($\bar{y_{js}}$) and characteristics are displayed with gray crosses. Each point represents the association between a 1 standard-deviation change of the underlying variable and the outcome of interest.

Several notable relationships emerge from this analysis. First, the associations between place characteristics are much stronger for overall physician spending than for spending per beneficiary, consistent with our earlier analysis demonstrating that overall spending responds more strongly to place effects. The strongest correlate for overall spending is, perhaps unsurprisingly, the number of enrollees per provider in a given area, indicating that a larger base of Part D patients results in higher Part D spending per physician.

Interestingly, both overall spending (and spending per beneficiary) are negatively associated with a modified measure of Part D Insurer HHI, providing suggestive—but non-causal—evidence that more market power by Part D insurers may provide a meaningful way of controlling spending.[51] In contrast, the *provider* market power (as measured by

[50]We emphasize that our place effects are not directly comparable to those in Finkelstein, Gentzkow and Williams (2016) and other related work that identify place effects based on patient migration. In these papers, place effects measure the effect of all non-patient factors, inclusive of physicians. In contrast, our measured place effects include patient effects, but by construction exclude physician effects.

[51]Information on insurer enrollment comes from the CMS Part D enrollment files. This measure is a lower

Provider HHI) is not meaningfully associated with either spending measure. This suggests that physicians are not meaningfully changing their prescribing behavior (in terms of spending) in response to higher competition. Likewise, there do not appear to be clear or persuasive associations between place effects and aggregated patient characteristics.

It is worth noting that several associations in this exercise yield results that are counterintuitive. Notably both the average price of drugs in the HRR and the average payments to physicians from drug companies are *negatively* correlated with spending measures. This is opposite to the expected sign, given that higher drug prices should mechanically increase spending (all else held constant) and that higher drug spending has been shown to have modest and short-lived increases in physician prescribing behavior (Carey, Lieber and Miller, 2021). However, these counterintuitive correlations could likely be due to other spurious factors that are associated with higher prices and drug spending (such as a wealthier and healthier overall population), a known weakness of this type of analysis.

To further investigate the relationship between insurers and spending, we perform additional correlates analyses using the market shares (in terms of enrollees) of the seven largest Part D insurers during our sample period. This analysis utilizes the following regression specification:

$$\hat{\gamma}_{js} = S_j + \pi_j + \zeta_{js}, \tag{5}$$

where S_j is the market share of a given insurance company, and π_j are flexible controls for deciles of Insurer HHI.

The results of this exercise are displayed in panels (C) and (D) of Figure 10, below. Each point represents the association between a 10 percentage-point change in an in-

bound on true HHI because it only includes the top 7 insurers in terms of overall market share. However, these 7 combined make up 95% of all enrollees, so it is likely very close to the true measure. Specifically, for these 7 Insurers, HHI is calculated as:

$$IHHI_j = \sum_{i=1...7} (share_{ij})^2$$

where *share* is insurer i's share of enrollees in a given HRR.

surer's market share (with HHI held constant) and the outcome of interest.

As displayed in the graph, there are meaningful differences in the association between different insurers, where higher shares of United Health Care and Blue Cross Blue Shield are associated with lower spending overall. These estimates, while non-causal, do suggest that insurer's ability to control costs (through rationing and other means) varies meaningfully, even when controlling for overall market power.

Next, we ask how much of the *variation* in place effects can be explained by these observable factors. To do so, we estimate three different regressions: one controlling for the entire vector of patient, price, and competition characteristics (X_{js}, from panels (A) and (B) of Figure 10), one controlling for the vector insurer market share characteristics (S_j, from panels (C) and (D) of Figure 10), and one controlling for all the aforementioned covariates. The R^2 from these regressions is displayed in panels (A) and (B) of Figure 11.

For both overall and per-beneficiary spending, observable characteristics only explain a small portion of the variation (10.1% and 2.6%, respectively). Of this variation, 33% and 43% are attributable to insurer market share for overall and per-beneficiary spending, respectively.[52]

Physician Effects We extend our analysis to understand how much variation in *physician effects* can be explained by observable characteristics. Specifically, we estimate three different regressions with the "causal" physician effect ($\hat{\alpha}_i$) as the dependent variable. The first equation controls for a vector of covariates relating to the physician's gender, the rank of medical school attended, and their credentials (such as M.D. or D.O.). The second equation only controls for the physician's experience, while the third equation controls for all characteristics.

The R^2 from these regressions, which are broken out by physician specialty, is displayed in panels (C) and (D) of Figure 11. The results yield three primary findings. First,

[52]This method for calculating is based on Israeli (2007). The calculations are simple in the two-group case: 0.042/(0.087 + 0.042) and 0.015/(0.020 + 0.015), respectively.

explanatory characteristics explain relatively little regarding physician practice style: a range of 11.7% to 21.4% for overall spending and 2.8% to 15.3% for per-beneficiary spending. Additionally, there is substantial heterogeneity in the degree to which observables provide insight: they are much more relevant for medical specialties (e.g., cardiology or nephrology) than for primary care and surgical disciplines. Finally, of all the covariates, the most important is overwhelmingly physician experience. Using Israeli (2007) to determine the marginal contribution, we find that experience accounts for 78% to 85% of the explained variation for overall spending and 60% to 78% for per-beneficiary spending.

Combined with the results in Figure 6, this suggests that experience and early-career exposures may be a much larger determinant of practice style than formal education.

6 Discussion

Drug spending in Medicare Part D varies substantially across geographic locations. In this paper, we investigate the relative importance of environment versus idiosyncratic physician-specific factors in explaining this variation. We do so by tracking the spending behavior of individual physicians in the years leading up to and following a relocation. We find that physicians who move to higher spending areas tend to increase their own spending. This is consistent with a story of peer effects, but could also be driven by a number of other factors. We investigate a variety of different environmental characteristics, but find that these observable characteristics explain relatively little of the aggregate variation in spending.

Although physicians do adopt the practice style of their new peers to some extent, this convergence is far from complete. This implies that physicians have individual practice styles which they do not fully relinquish upon moving to an area with different practice styles. These physician-specific practice styles are quantitatively important. To illustrate this, we perform a simple back-of-the-envelope calculation in which we assign to each

HRR the spending associated with the average physician in the HRR in the p^{th} percentile of per enrollee spending. We then plot the associated change in aggregate, per enrollee spending for each value of p. We display the results from this exercise in appendix Figure A10. This figure shows that, if each physician behaved like the average physician in the least expensive HRR, spending per enrollee would fall by almost 20 percent. In contrast, if each physician behaved like the average physician in the most expensive HRR, spending per enrollee would increase by over 40 percent.

We also examine which physician-specific characteristics predict higher levels of malleability. Interestingly, most of the variation in responsiveness to environmental factors is not explained by observable characteristics of the physician. However, among the characteristics that we observe, experience is by far the most important. In fact, experience alone explains more of the variation in physician responsiveness than medical school rank, credentials, and gender combined. This has important implications for any policy efforts to reduce spending. Specifically, our results imply that physicians are most responsive to outside influences early in their careers, and that later interventions are likely to have muted effects.

We find that spending is malleable along both the extensive and intensive margins. In terms of overall spending, environmental factors account for about 42 percent of the geographic variation in spending. Along the intensive margin, this falls to 29 percent.[53] However, these estimates are identified off of migrating physicians, who tend to be younger and less experienced than the general population of physicians (see Table 1). If we weight our regressions to be representative of the full sample of physicians (including non-migrants), we find a smaller role for environmental factors and a larger role for idiosyncratic physician-specific practice styles.[54] Specifically, we find that environmental factors account for 30 and 24 percent of the variation in spending and spending per beneficiary, respectively.

Taken together, our results imply that efforts to target geographic variation in drug

[53]These numbers are taken from regression equation 2 as reported in Figure 3.
[54]Results are shown in appendix Figure A11.

spending—or even overall levels of drug spending—will need to heavily consider factors in the formation of idiosyncratic physician practice style. Our results demonstrating that young physicians are more malleable indicate that targeted interventions during the residency- and fellowship-training phases could potentially have meaningful long-term effects. Finally, while suggestive, our analysis of the correlates of causal place effects suggests that Part D insurer market power and idiosyncratic cost-management differences could be a useful tool in addressing overall spending.

References

Callaway, Brantly, and Pedro HC Sant'Anna. 2021. "Difference-in-differences with multiple time periods." *Journal of Econometrics*, 225(2): 200–230.

Callison, Kevin, Robert Kaestner, and Jason Ward. 2018. "A Test of Supply-side Explanations of Geographic Variation in Health Care Use." National Bureau of Economic Research.

Carey, Colleen, Ethan MJ Lieber, and Sarah Miller. 2021. "Drug firms' payments and physicians' prescribing behavior in Medicare Part D." *Journal of Public Economics*, 197: 104402.

Choudhry, Niteesh K, Thomas D Denberg, Amir Qaseem, and Clinical Guidelines Committee of the American College of Physicians. 2016. "Improving adherence to therapy and clinical outcomes while containing costs: opportunities from the greater use of generic medications: best practice advice from the Clinical Guidelines Committee of the American College of Physicians." *Annals of internal medicine*, 164(1): 41–49.

Cooper, Zack, Stuart V Craig, Martin Gaynor, and John Van Reenen. 2019. "The price ain't right? Hospital prices and health spending on the privately insured." *The Quarterly Journal of Economics*, 134(1): 51–107.

Currie, Janet, Anran Li, and Molly Schnell. 2023. "The Effects of Competition on Physician Prescribing." National Bureau of Economic Research.

Cutler, David, Jonathan S Skinner, Ariel Dora Stern, and David Wennberg. 2019. "Physician beliefs and patient preferences: a new look at regional variation in health care spending." *American Economic Journal: Economic Policy*, 11(1): 192–221.

De Chaisemartin, Clément, and Xavier d'Haultfoeuille. 2020. "Two-way fixed effects estimators with heterogeneous treatment effects." *American Economic Review*, 110(9): 2964–96.

Ding, Hui. 2022. "Geographic Variation in Mental Health Care: Evidence from Migration."

Doyle Jr, Joseph J, and Becky Staiger. 2022. "Physician Group Influences on Treatment Intensity and Health: Evidence from Physician Switchers." National Bureau of Economic Research.

Eichmeyer, Sarah, and Jonathan Zhang. 2022. "Pathways into Opioid Dependence: Evidence from Practice Variation in Emergency Departments." *American Economic Journal: Applied Economics*, 14(4): 271–300.

Eichmeyer, Sarah, and Jonathan Zhang. 2023. "Primary care providers' influence on opioid use and its adverse consequences." *Journal of Public Economics*, 217: 104784.

Fadlon, Itzik, and Jessica Van Parys. 2020. "Primary care physician practice styles and patient care: Evidence from physician exits in medicare." *Journal of Health Economics*, 71: 102304.

Finkelstein, Amy, Matthew Gentzkow, and Heidi Williams. 2016. "Sources of geographic variation in health care: Evidence from patient migration." *The quarterly journal of economics*, 131(4): 1681–1726.

Finkelstein, Amy, Matthew Gentzkow, Dean Li, and Heidi L Williams. 2022. "What Drives Risky Prescription Opioid Use? Evidence from Migration." National Bureau of Economic Research.

Godøy, Anna, and Ingrid Huitfeldt. 2020. "Regional variation in health care utilization and mortality." *Journal of Health Economics*, 71: 102254.

Israeli, Osnat. 2007. "A Shapley-based decomposition of the R-square of a linear regression." *The Journal of Economic Inequality*, 5: 199–212.

Laird, Jessica, and Torben Nielsen. 2016. "The effects of physician prescribing behaviors on prescription drug use and labor supply: evidence from movers in Denmark." *Working Paper*.

Molitor, David. 2018. "The evolution of physician practice styles: evidence from cardiologist migration." *American Economic Journal: Economic Policy*, 10(1): 326–56.

Moura, Ana, Martin Salm, Rudy Douven, and Minke Remmerswaal. 2019. "Causes of regional variation in Dutch healthcare expenditures: Evidence from movers." *Health economics*, 28(9): 1088–1098.

Phelps, Charles, and Cathleen Mooney. 1993. "Variations in medical practice use: causes and consequences."

Schnell, Molly, and Janet Currie. 2018. "Addressing the opioid epidemic: is there a role for physician education?" *American journal of health economics*, 4(3): 383–410.

Skinner, Jonathan. 2011. "Causes and consequences of regional variations in health care." In *Handbook of health economics*. Vol. 2, 45–93. Elsevier.

Song, Yan, and Douglas Barthold. 2018. "The effects of state-level pharmacist regulations on generic substitution of prescription drugs." *Health economics*, 27(11): 1717–1737.

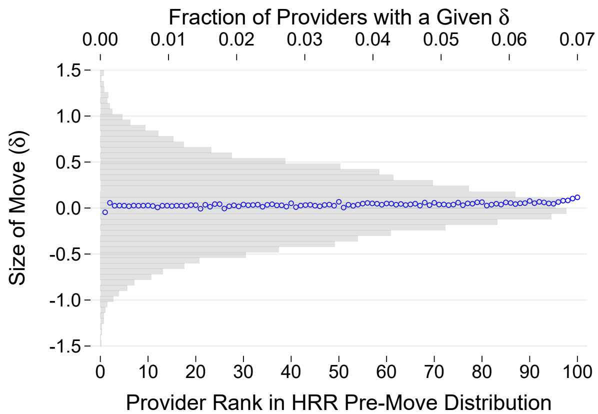
Staiger, Becky, Laurence Baker, and Tina Hernandez-Boussard. 2022. "Provider Opioid Prescribing Behaviors and Opioid Use in Medicaid."

Sun, Liyang, and Sarah Abraham. 2021. "Estimating dynamic treatment effects in event studies with heterogeneous treatment effects." *Journal of Econometrics*, 225(2): 175–199.

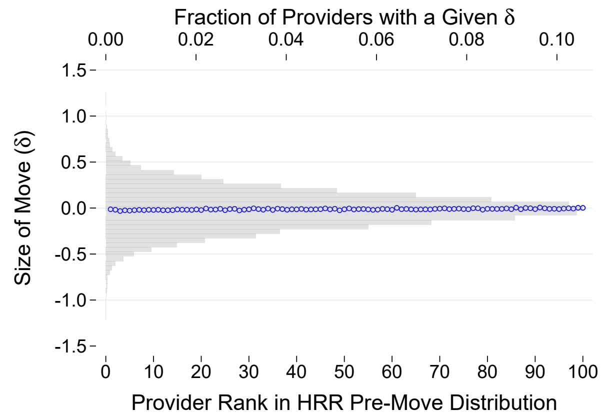
Tu, Peter. 2017. "Quasi-experimental Evidence of Physician Effects." *Harvard University Job Market Paper*.

Zhang, Xuan. 2018. "Disruption in Primary Care and Patient Outcomes: Evidence from Physician Retirement." Working Paper.

Figure 1: Physician Rank in Within HRR Pre-move Distribution and Move Size



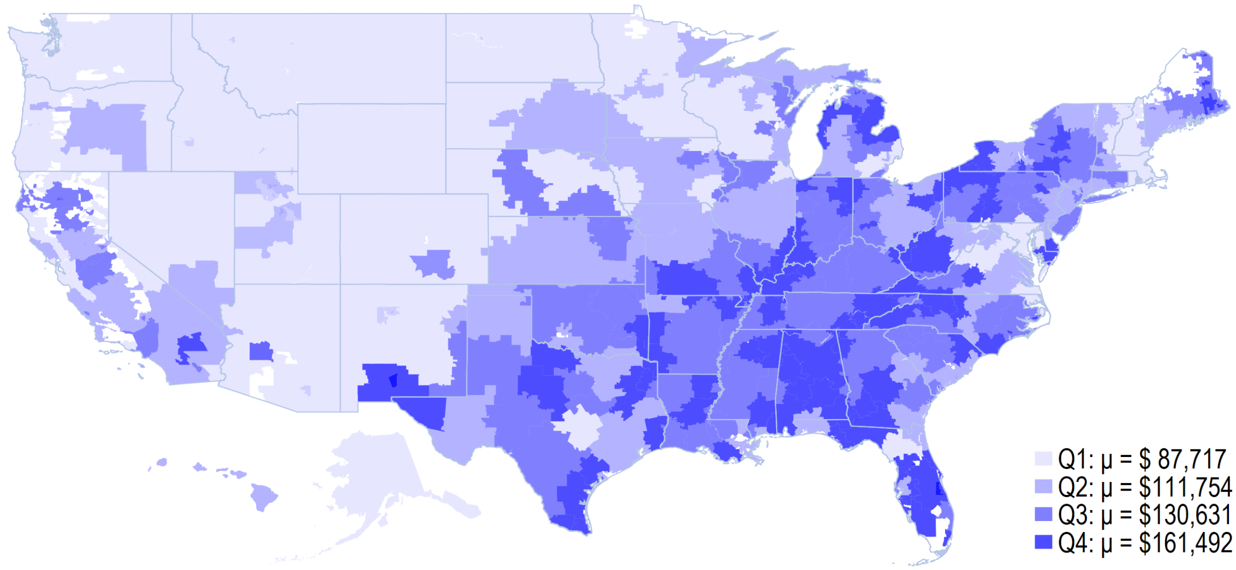
(A) Log(spending)



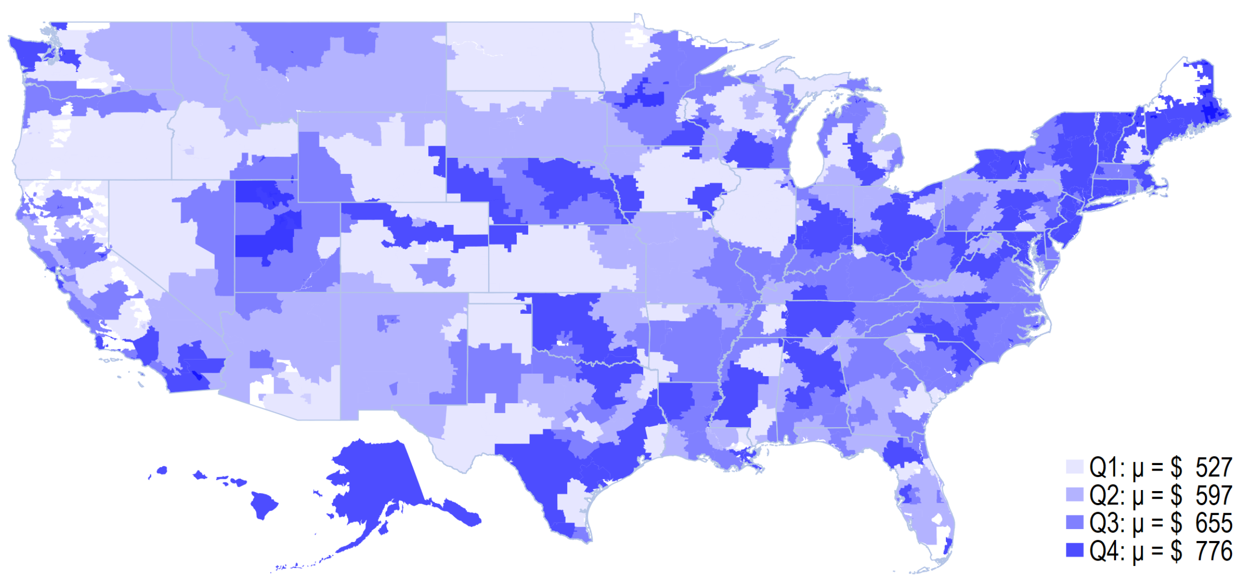
(B) Log(spending per beneficiary)

Note: These figures examine the extent to which physicians' pre-move prescribing behavior predicts the characteristics of their post-move HRR. The blue circles plot the average δ for each percentile of the physicians rank in the HRR (specialty-specific) pre-move prescribing distribution, scaled from 0 to 100. For a sense of scale, we also show the histograms of δ.

Figure 2: Geographic Variation in Provider-Level Prescribing Behavior



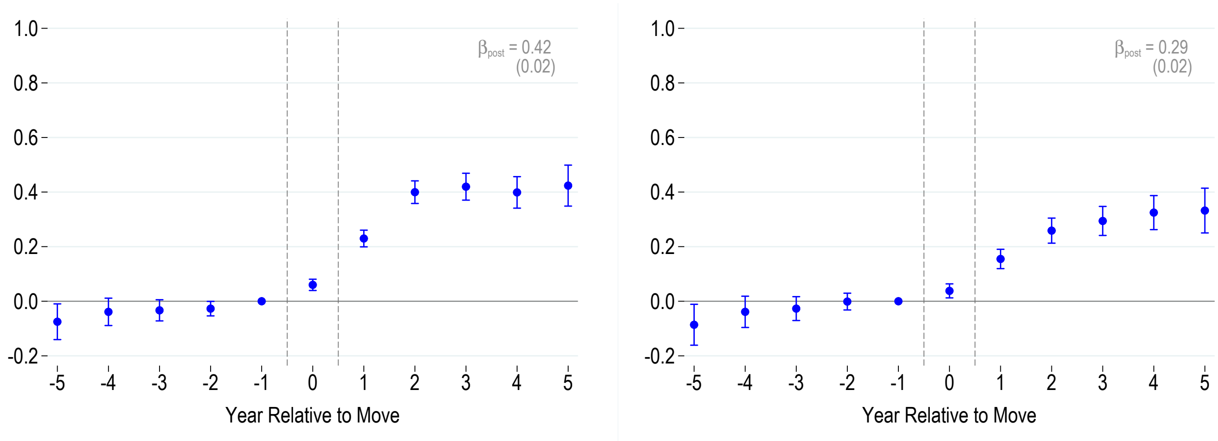
(A) Drug Spending



(B) Drug Spending per Beneficiary

Note: This figure displays two different measures of physician-level prescribing behavior. For each measure, the corresponding map shows the average across all physicians in each hospital referral region (HRR). Specifically, these maps are constructed by first averaging the corresponding measure across all years for each physician. Then, we average across all physicians in each HRR. In each panel, the map is colored according to quartile, with darker shades of blue indicating higher spending HRRs. The legend includes information on the average (μ) within each quartile.

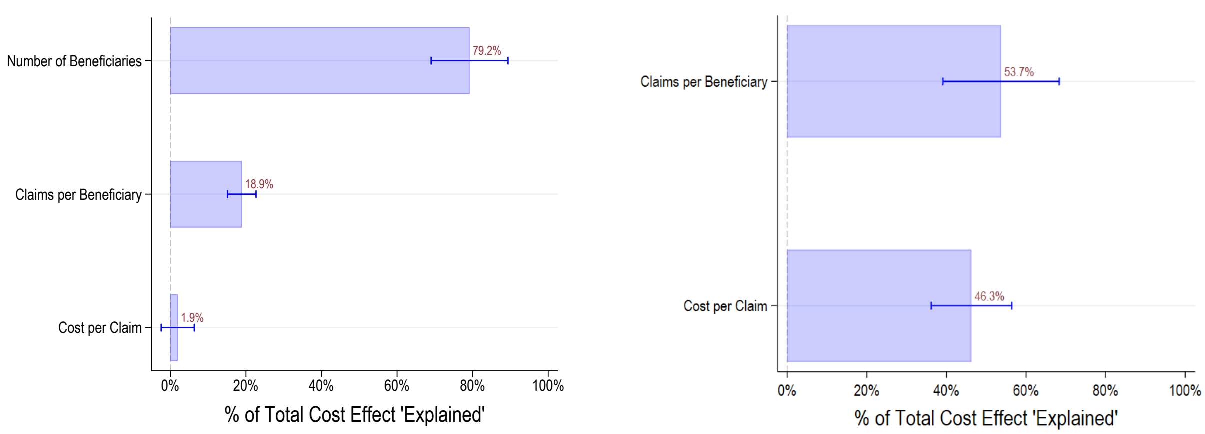
Figure 3: Changes in Prescribing Upon Move



(A) Log(spending) **(B)** Log(spending per beneficiary)

Note: These figures display the β_τ coefficients from equation 1. The outcome variables are log(spending) and log(spending per beneficiary) in panels (A) and (B), respectively. Standard errors are clustered at the physician-level. The top-right corner of each sub-figure displays the β_{post} coefficient from equation 2, with the associated standard error in parentheses.

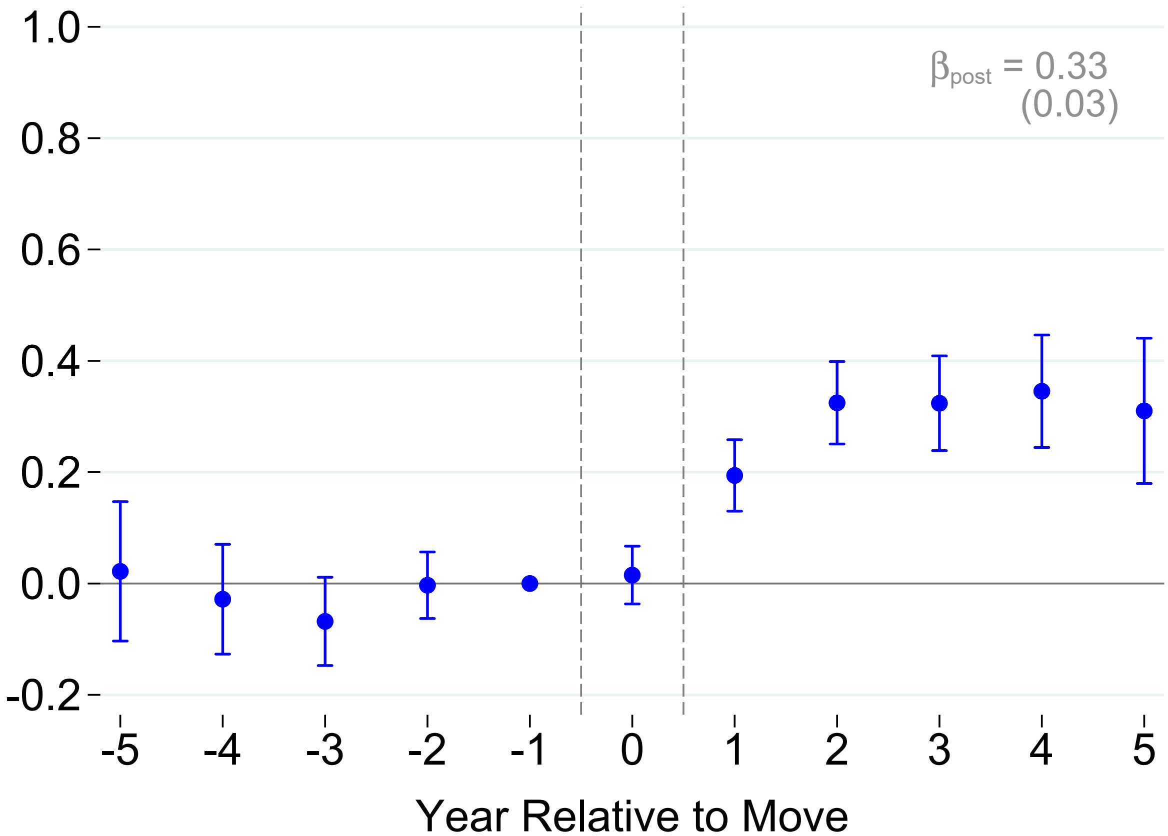
Figure 4: Decomposition of Total Spending Response



(A) Log(spending) **(B)** Log(spending per beneficiary)

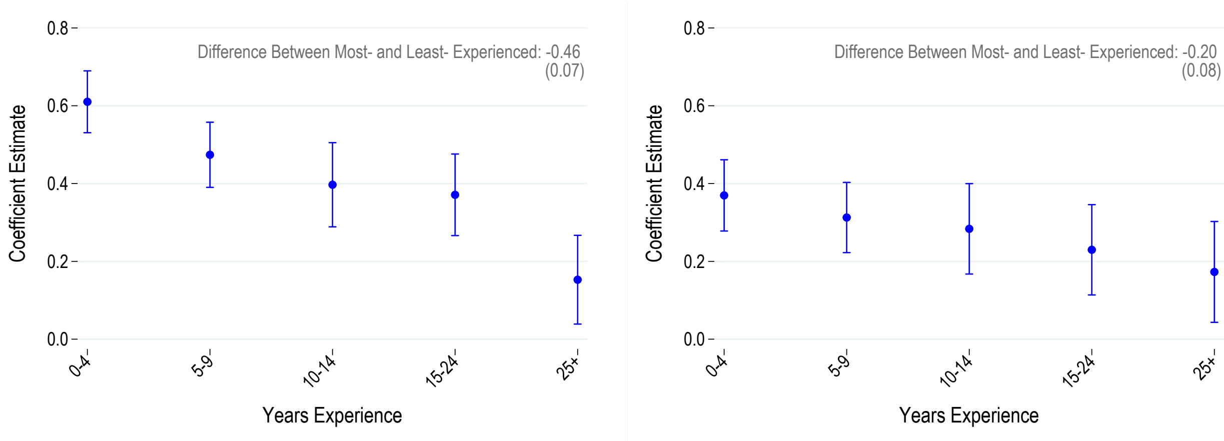
Note: This figure shows the $\hat{\beta}_{post}$ coefficient from equation 2, broken down into component pieces. Panel (A) shows the log(spending) regression broken down into log(spending per claim), log(claims per beneficiary), and log(number of beneficiaries). Panel (B) shows the log(spending per beneficiary) regression broken down into log(spending per claim) and log(claims per beneficiary).

Figure 5: Event Study: Fraction of Spending on Generic Drugs Before and After Move



Note: This figure displays the β_τ coefficients from equation 1. The outcome variables is the share of spending on generic drugs. Standard errors are clustered at the physician-level. The top-right corner displays the β_{post} coefficient from equation 2, with the associated standard error in parentheses.

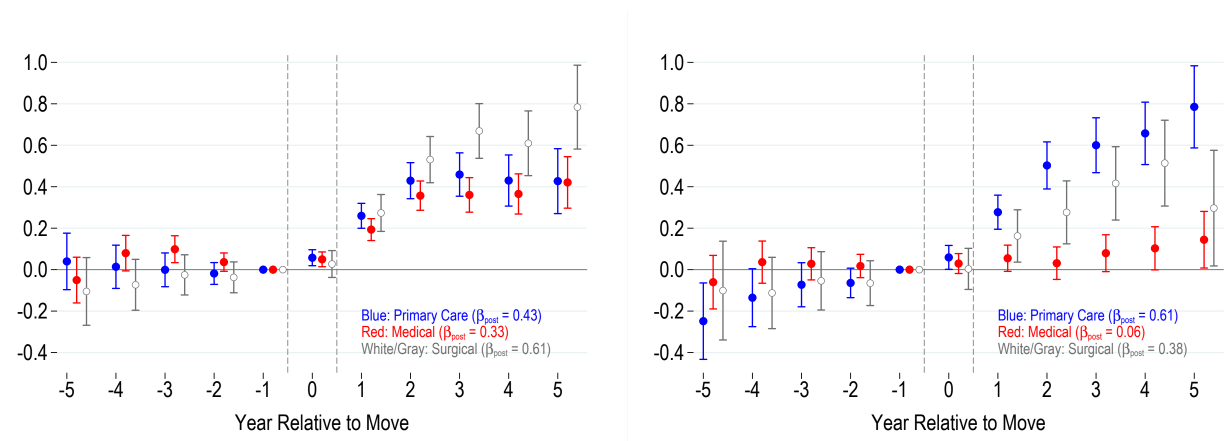
Figure 6: Changes in Prescribing Upon Move: Interacted with Experience



(A) Log(spending) **(B)** Log(spending per beneficiary)

Note: These figures display the coefficients on the interactions of δ_{idos}, a post indicator, and experience indicators from equation 2. The outcome variables are log(spending) and log(spending per beneficiary) in panels (A) and (B), respectively. Standard errors are clustered at the physician-level.

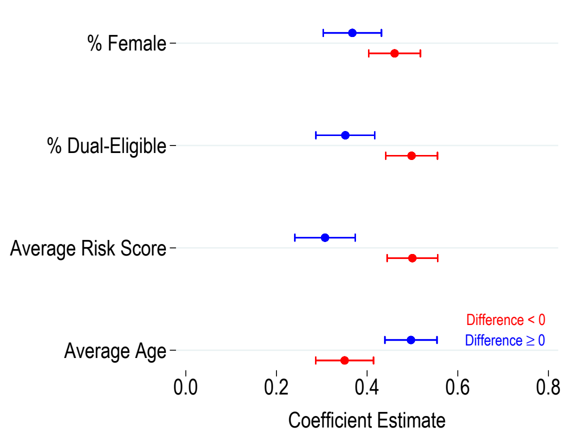
Figure 7: Changes in Prescribing Upon Move: Heterogeneity by Physician Type



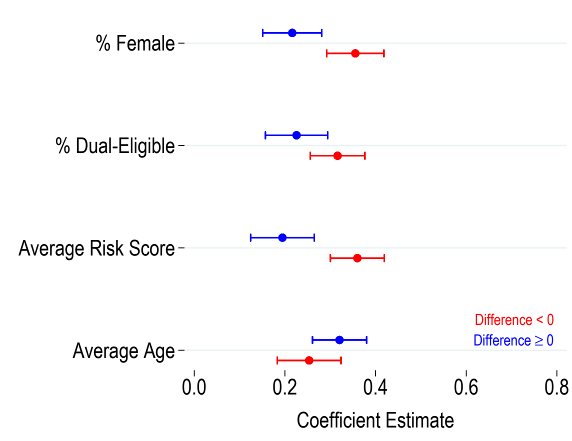
(A) Log(spending) **(B)** Log(spending per beneficiary)

Note: These figures display the β_τ coefficients from equation 1, estimated separately by type of physician. The outcome variables are log(spending) and log(spending per beneficiary) in panels (A) and (B), respectively. Standard errors are clustered at the physician-level.

Figure 8: Changes in Prescribing Upon Move: Heterogeneity by Patient Characteristics



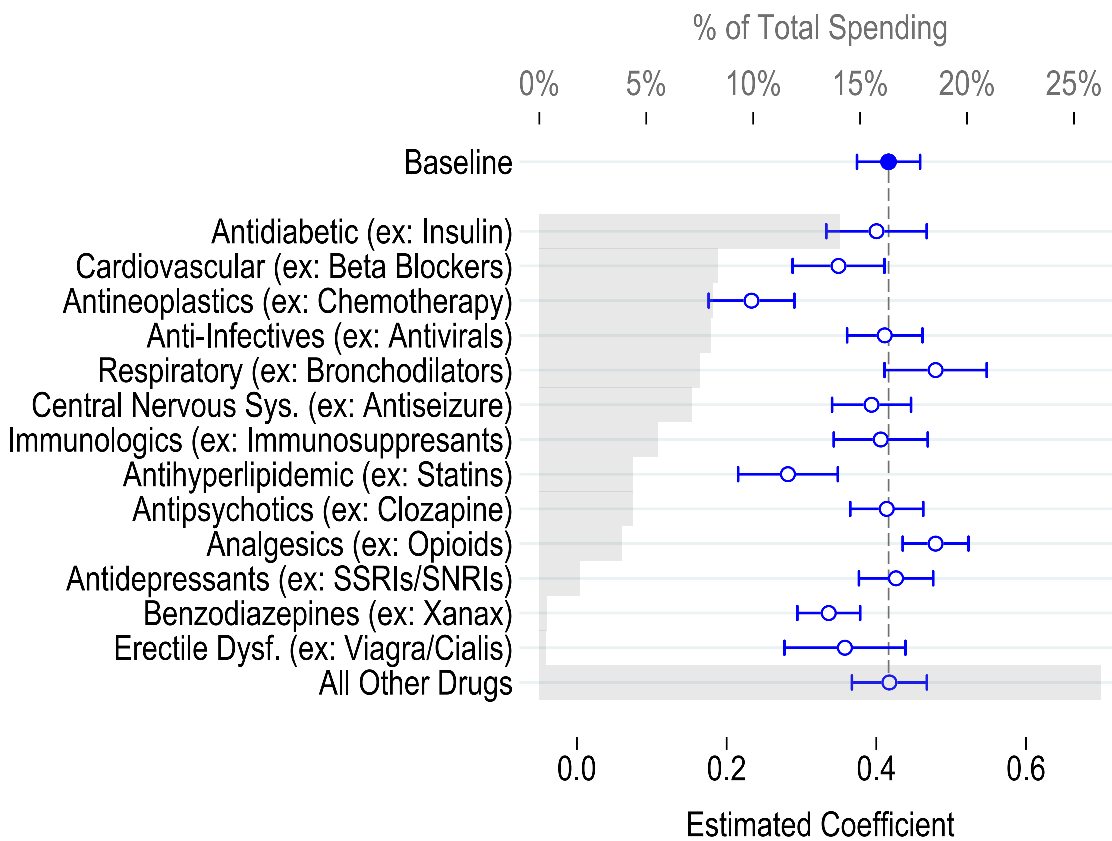
(A) Log(spending)



(B) Log(spending per beneficiary)

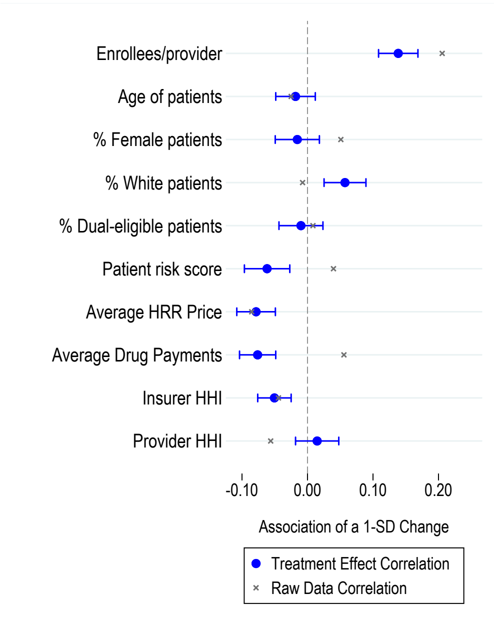
Note: These figures display the β_{post} coefficients from equation 2, splitting the sample by migrants who move to HRRs with higher or lower levels of the associated characteristic. Red and blue coefficients are estimated based on physicians who migrate to HRRs with lower or higher levels of the associated characteristic, respectively. The outcome variables are log(spending) and log(spending per beneficiary) in panels (A) and (B), respectively. Standard errors are clustered at the physician-level.

Figure 9: Heterogeneous Responses by Drug Class

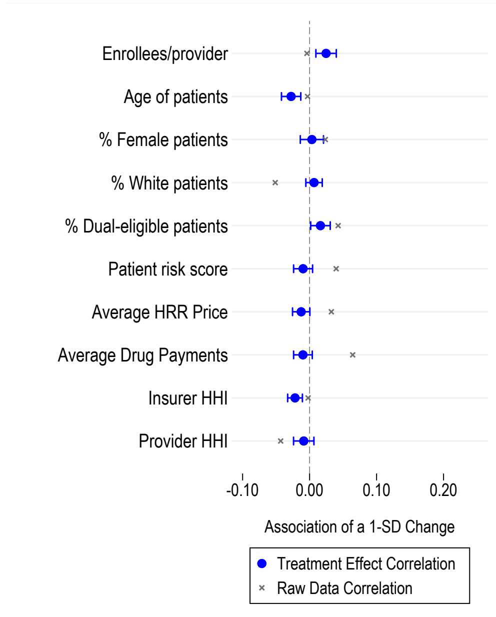


Note: This figure displays the $\hat{\beta}_{post}$ coefficient from equation 2 from the baseline regression in the top row, then estimated separately for each drug category in the following rows. The outcome variable is the log of spending on that particular drug category. The fraction of spending on each drug category is indicated by gray bars.

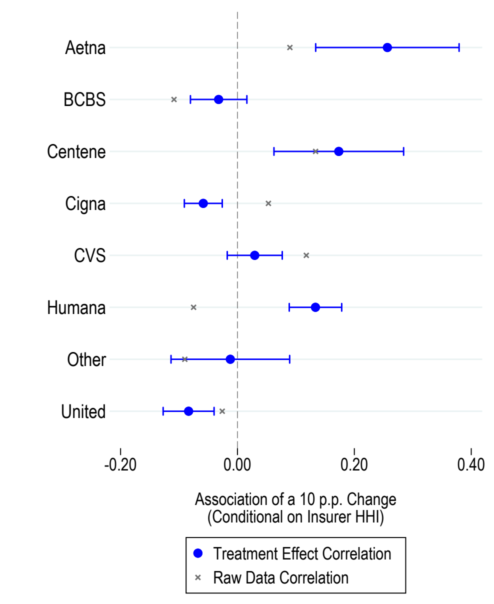
Figure 10: Place Effect Correlates and Insurer Market Share Characteristics



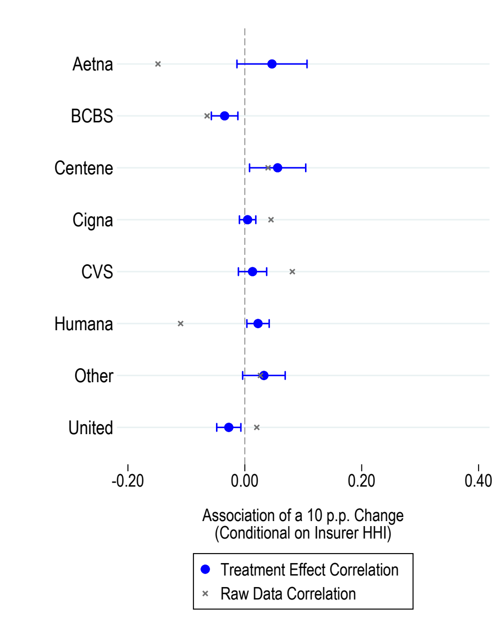
(A) Log(spending)



(B) Log(spending per beneficiary)



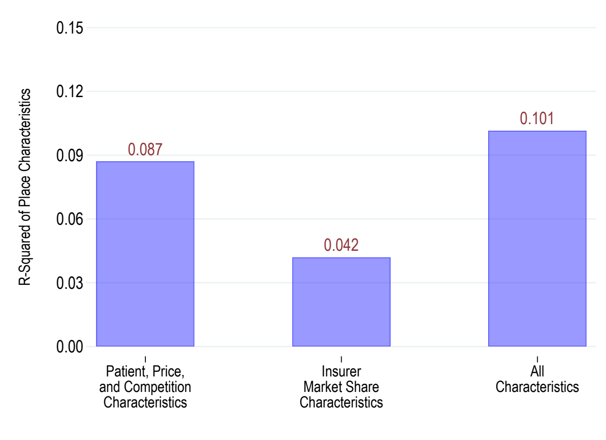
(C) Log(spending)



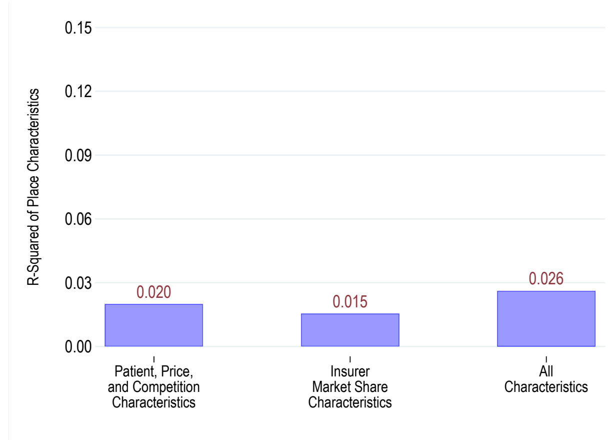
(D) Log(spending per beneficiary)

Note: Each row of these sub-figures displays the Γ coefficients from equation 4 for a particular regressor in blue. Regressions showing the relationship between raw averages of the outcome and characteristic are displayed as gray crosses. Panels (A) and (B) estimate $\hat{\gamma}_{js}$ based on equation 3 with log(spending) and log(spending per beneficiary) as the outcome variable, respectively. Panels (C) and (D) are constructed similarly, except the characteristics of interest are insurance market shares.

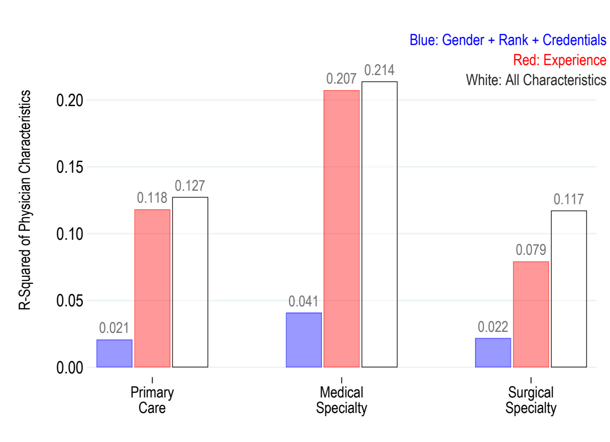
Figure 11: Variation in Place and Physician Effects Explained by Observables



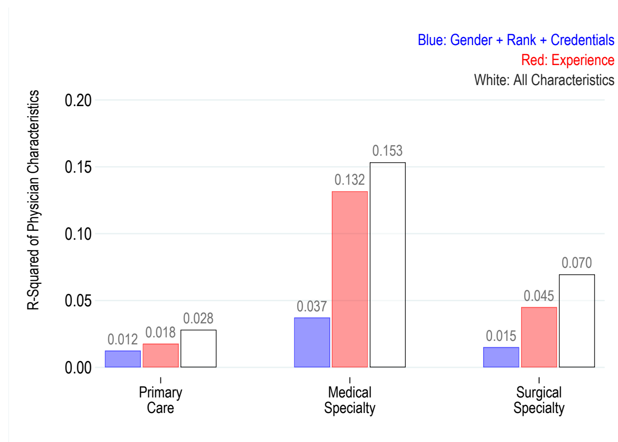
(A) Log(spending)



(B) Log(spending per beneficiary)



(C) Log(spending)



(D) Log(spending per beneficiary)

Note: These figures show the fraction of variation in place (top row) and physician (bottom row) effects explained by observables. Panels (A) and (B) show the R^2 from three regressions: the first regresses the place effect on the entire vector of patient, price, and competition characteristics. The second only controls for the vector of insurer market shares, and the third includes all variables. Panels (C) and (D) regress the physician effect demographic characteristics, experience, and then all characteristics. The left panels are based on log(spending), while the right panels are based on log(spending per beneficiary).

Table 1: Summary Statistics

	(1) All	(2) Stayers	(3) Movers (pre)	(4) Movers (post)
Prescribing				
Log total drug cost	9.56	9.57	9.21	9.73
Log cost per Part D benef.	5.32	5.32	5.29	5.42
Generic cost as % of total cost	39.58	40.10	35.34	35.89
Provider Characteristics				
% Male	60.41	61.25	54.36	53.78
Experience	19.66	21.00	10.12	12.60
% Top 100 Med. School	37.38	38.34	31.31	31.29
Primary Care	27.62	26.62	35.50	34.89
Medical Specialty	35.94	35.56	38.26	39.27
Surgical Specialty	8.65	8.53	9.97	9.30
Other Specialty	27.78	29.29	16.27	16.53
Patient Characteristics				
% Male	39.88	39.86	40.03	40.00
% White	74.28	74.70	69.26	72.13
% Black	14.94	14.30	21.45	19.67
% Hispanic	10.51	10.40	9.68	12.76
% Medicaid	39.24	38.52	42.03	46.04
% Under 65	25.16	24.53	31.98	27.80
Average Risk Score	1.48	1.45	1.66	1.67
# of Providers	1,193,504	1,067,966	125,538	125,538
Observations	6,404,294	5,652,152	346,826	405,316

Note: This table presents summary statistics for prescribing behavior, provider characteristics, and patient characteristics. Column (1) displays the summary statistics for the entire sample. Column (2) reports the same statistics only for the sample of non-movers. Columns (3) and (4) report statistics for movers, split by pre- and post-move.

A APPENDIX

A.1 Alternative Explanations

In this section, we consider several two alternative explanations of our findings, both in terms of overall magnitudes and the dynamics of physicians' spending responses.

Changes in Patient Composition One of the limitations of our data is that we do not know the number of Part D patients that each physician interacts with, only the number that receive at least one drug prescription. For example, a patient who visits a provider but does not fill a drug prescription over the course of the year will not be included in our measure of beneficiaries. This could, in theory, pose some challenges in interpreting the event study coefficients if the number of Part D enrollees that each physician interacts with varies by area. In particular, this could lead us to conflate changes in physician behavior with changes in the number of available Part D enrollees. Ideally, we would explicitly control for the number of Part D enrollees that each physician interacts with in order to account for this possibility. Unfortunately, we do not have this information. We do, however, have data on the total number of Part D enrollees in each HRR, regardless of whether they filled a drug prescription. We therefore proxy for the number of Part D patients that each provider interacts with by taking the total number of Part D enrollees in each HRR and dividing it by the total number of physicians in each HRR. While obviously not a perfect proxy, it does allow us to control to some extent for the number of potential Part D patients in a given geography.

We then reestimate equation 1, adding interactions between this measure of the number of enrollees per provider with time-to-move fixed effects. This allows for a separate effect of changes in the number of available patients for each physician. The results from this exercise are included in appendix Figure A12. These regressions reveal a nearly identical pattern qualitatively, although the magnitudes of the coefficients for log(spending) are somewhat smaller. This suggests a smaller role for environmental factors to affect prescribing behavior after taking into account the number of potential Part D patients. The results for log(spending per beneficiary) are nearly identical, which is unsurprising given that this latter regression only examines changes in prescribing along the intensive margin.

Logistics of Migration We argue that the dynamics that we observe in our event studies are evidence of physicians slowly learning over time. An alternative explanation for these dynamics is that they are simply the result of the time that it takes to build up a new practice. For example, one could argue that physicians cannot immediately match the spending of their new peers because it takes time to accumulate a sufficiently large patient pool to do so. We discuss three reasons why this explanation is unlikely to explain the pattern of our results.

First, we reiterate that equation 1 includes relative year-by-origin fixed effects. This means that any common trends in spending among migrating physicians are accounted for in our model. To the extent that all migrating physicians undergo a "rebuilding" process, this would not affect the pattern of our event study coefficients.[55] Second, we note that we observe similar dynamics for both log(spending) and log(spending per beneficiary). While log(spending) may be affected by the actual act of migration, log(spending per beneficiary) measures only intensive margin spending. This outcome should therefore be immune to any mechanical effect of setting up a new practice. Finally, we estimate equation 1 separately for migrant moving to higher or lower spending areas. The idea

[55]In appendix Figure A13, we show the raw trends in log(spending) pooling across all move-year cohorts, broken down by quartile of δ and rescaled to equal zero in the period before the move. This figure shows that spending is increasing over time among all cohorts prior to moving. After moving, spending among the higher quartile migrants rapidly diverges from the lower quartile migrants. This figure largely alleviates any concerns that the results may be driven by issues related to setting up a new practice. The figure also illustrates the importance of including the origin-by-relative year fixed effects, as higher quartile migrants had slightly higher spending growth prior to moving.

here is that, while it could potentially take several years to *increase* spending, there is no clear reason while it would take time to *decrease* it. We present the results of this exercise in appendix Figure A14. For both outcome variables, we observe broadly similar dynamics over time, regardless of whether the destination is higher or lower spending than the origin.[56]

Another potential concern is the extent to which migrants continue to prescribe to patients in their origin after moving. Ideally, we would identify migrants whose total spending outside of their current HRR exceeded some threshold and drop them from our sample. However, in our data we are only able to observe the location of the prescriber; we do not observe where the prescription is filled. If migrants are slow to completely switch their practice to their new HRR, this could potentially account for the learning pattern in our event studies. While we are unable to test for this explicitly, we are able to perform a limited analysis using the publicly available Physician Shared Patient Patterns ("PSPP") dataset. This dataset, which covers the years 2013 through 2015, details events when a provider delivers health care services to a patient within 30 days after another provider provided health care services to the same patient. To provide a specific example, if a patient had a visit with Dr. Smith (a primary care provider) and then saw Dr. Jackson (a cardiologist) 21 days later, this "dyad" would be captured as a single interaction in the PSPP data.

We utilize this PSPP data to see what percentage of a provider's "shared" patients are seeing another provider in the same HRR before and after the move. If the provider is still treating patients in their origin HRR, we should expect the provider's share of ''same-HRR" patients to be lower after the move (as a meaningful fraction of their patients reside in their origin HRR, where they no longer live). However, if a provider is meaningfully shifting their practice away from their old patients, we should expect to see the percentage

[56]We do note that the magnitudes are somewhat smaller for migrants who move to lower spending HRRs, which suggests that physicians who move to lower spending HRRs adjust their behavior less than physicians who move to higher spending HRRs.

of patients "shared" with other providers in their current HRR (the destination HRR) rebound immediately to pre-move levels.

To determine the effect of the move on patient composition, we utilize all providers that moved within our sample time frame (capturing moves from 2014 through 2019). We then run a regression using the following equation:

$$\text{shared patients}_{it} = \sum_{k \neq -1} \rho_k \cdot \mathbb{1}(\text{relative year}_{it} = k) + \lambda_t + \alpha_i + \epsilon_{it} \tag{6}$$

where shared patients$_{it}$ is the percentage of patients shared by providers in the same *current* HRR as provider i, and relative year$_{it}$ is the year relative to the provider's move. Calendar year and individual fixed effects are represented by λ_t and α_i, respectively. The coefficients of interest, ρ_k, display the change in the percentage of shared patients relative to the year immediately preceding the move.

The results of this analysis are displayed in appendix Figure A15. Note that, within the figure, there is only one post-move period because this data set only extends through 2015, and our earliest movers changed location in 2014. However, there are many pre-periods, because the data set begins in 2013, and there are several cohorts of movers who changed HRRs after that year. As shown in the figure, there is very little evidence of a change in the percentage of patients treated by same-HRR providers prior to move (i.e., a flat pre-trend), but there is a sharp decline during the year of the move, as the moving provider will share some patients with providers in their old (origin) HRR during the year, and some in their new (destination) HRR. This quickly recovers in the post-move year, where the estimate is statistically indistinguishable from the pre-move period, providing suggestive evidence that the provider has shifted their practice to fully match the pre-move practice patterns.

These results suggest that physicians gradually switching away from patients in their old HRR toward patients in their new HRR is not driving the learning pattern that we observe in our main results. However, we acknowledge that we also cannot completely rule this out. As a final thought, we also note that how exactly such a gradual switching pattern would affect our estimates is not clear ex ante. For example, suppose that physicians gradually transition from their origin HRR to their destination HRR. Furthermore, suppose that no learning takes place; any changes in behavior are immediate. It is possible that this would generate the learning pattern that we observe in our regressions, but only if the physician's behavioral changes only apply to new patients. If instead, the physician changes his or her prescribing practices with the origin patients as well (reflecting the new environment), then we wouldn't expect to see growing coefficients over time.

A.2 Robustness

In this section, we investigate the robustness of our main results to a variety of specification checks and sample restrictions.

Treatment Effect Heterogeneity Numerous recent papers have highlighted issues arising from treatment effect heterogeneity in two-way fixed effects models (Sun and Abraham, 2021; Callaway and Sant'Anna, 2021; De Chaisemartin and d'Haultfoeuille, 2020). We investigate whether our results are potentially affected by this issue in appendix Figure A4. Specifically, we re-estimate equation 2 separately for each cohort of movers, with cohorts defined by the year in which they move. The idea behind this exercise is to test whether there are differential responses by cohort, which could lead to issues in the interpretation of our main findings. The results from this exercise are under the heading "Regressions by Move Year." Panel (A) presents the results for log(spending), while panel (B) shows the results for log(spending per beneficiary). Each dot represents the $\hat{\beta}_{post}$ coefficient from equation 2.[57] In all cases, the coefficient is similar to what we find in the baseline model. The coefficients using only movers from 2017 are a bit smaller, but again not statistically different from the baseline. We interpret this as evidence that our results

[57] We cannot estimate this regression for 2018 movers, as they do not have two full years post-move. We show the corresponding event studies for each move-year cohort in appendix Figure A16.

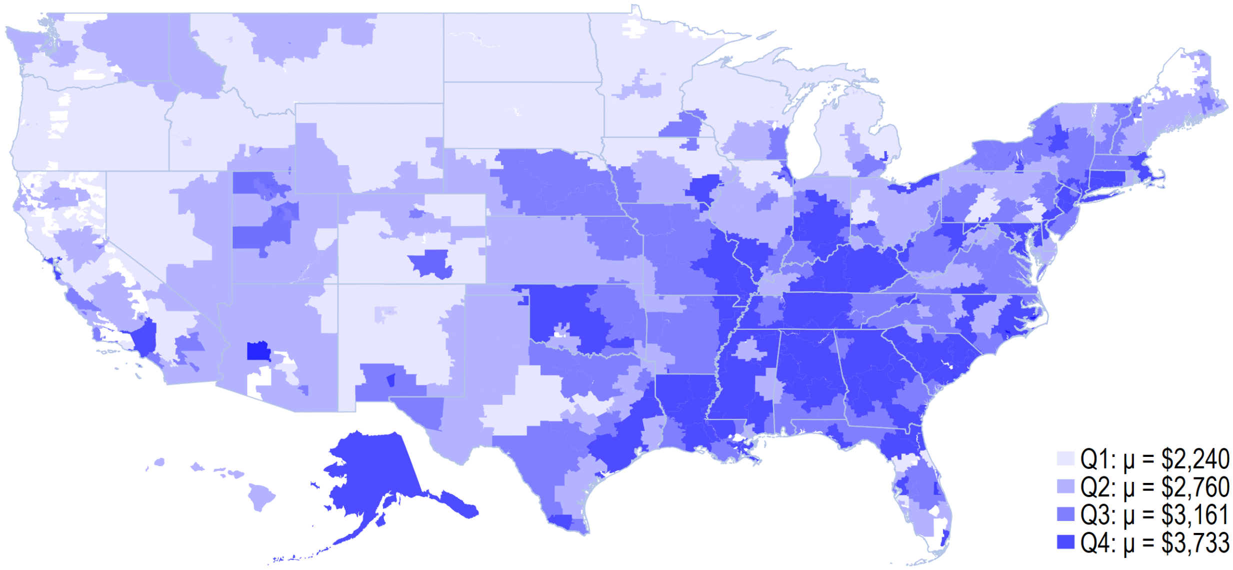
are not the driven by the presence of treatment effect heterogeneity.

Alternate δ Definitions and Samples Next, we consider the sensitivity of our results to various sample restrictions and alternative ways of defining our treatment variable δ. The results are shown in appendix Figure A4 under the heading "Different δ Definitions." First, we construct the δ terms for each mover using data only from their move year. In principle, prescribing practices could vary across time, so pooling together all years in the construction of our δs could mask important differences in the changes that physicians experience upon moving. In practice, redefining our δs in this way has little effect on our findings. Second, we reconstruct our δs for each mover using a leave-one-out mean.[58] This has almost no effect on our results, because each individual mover is a small fraction of overall spending in a given HRR-specialty.[59] Next, we omit all physicians who prescribe to fewer than 11 beneficiaries in a calendar year from the δ computation. This allows us to see whether the presence of a large number of very low prescribers affects our results by distorting the relative difference in environment that migrants experience upon moving. Once again, the results are similar to the baseline. Finally, we restrict our sample to include only physicians who we observe in every year of the sample. The results are shown under the heading "Different Samples," and again make very little practical difference.

[58] We note that this is the correct way to compute these terms. However, doing so is computationally expensive so we show it only as a robustness test.

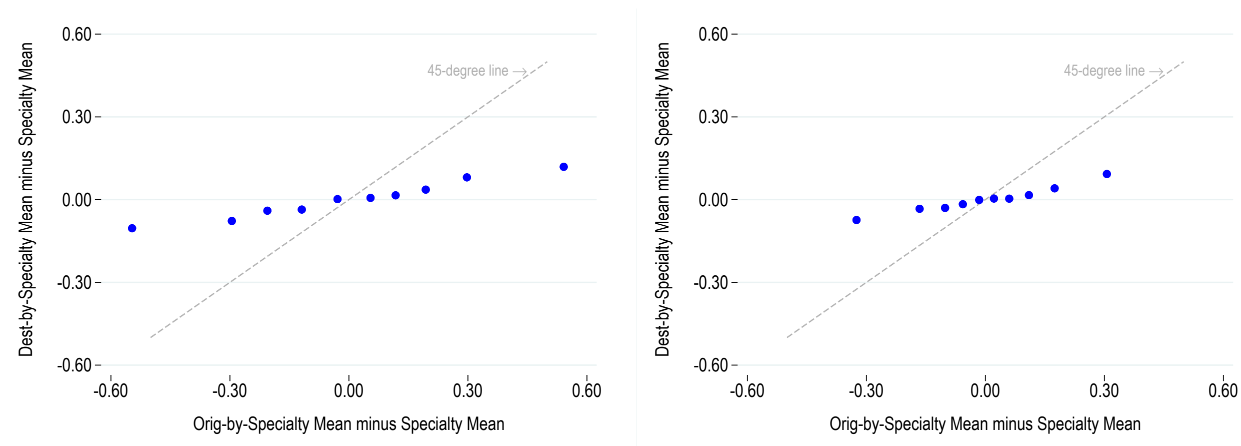
[59] If we compute the ratio of each physician's spending to total spending in his or her HRR-specialty, the median is 0.004 percent, with a mean of 0.05 percent.

Figure A1: Geographic Variation in HRR-Level Prescribing Behavior



Note: This figure displays total spending in each HRR, divided by the number of Part D enrollees in that HRR. This differs from Figure 2 panel (B) in two key ways. First, the average is computed at the HRR-level, rather than the physician level. Second, the denominator is the total number of Part D enrollees, as opposed to the number of beneficiaries who received a drug prescription. The map is colored according to quartile, with darker shades of blue indicating higher spending HRRs. The legend includes information on the average (μ) within each quartile.

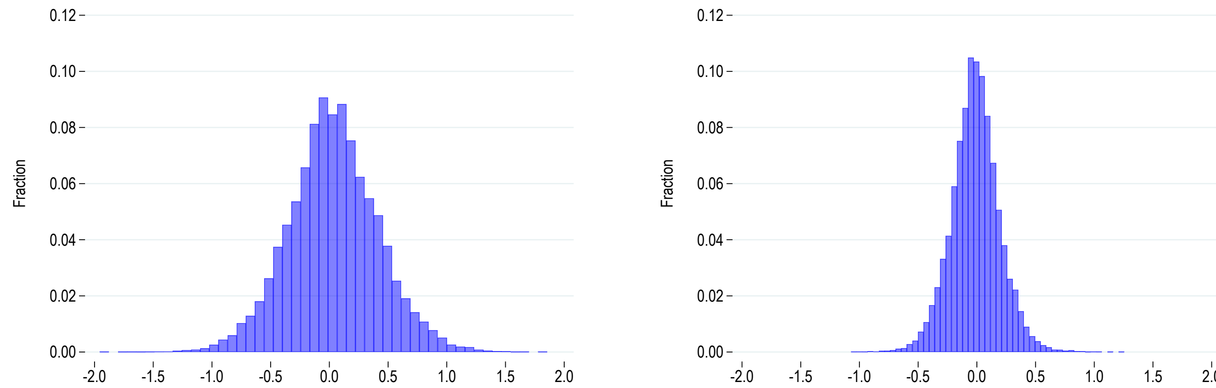
Figure A2: HRR-Level Sorting



(A) Log(spending) **(B)** Log(spending per beneficiary)

Note: These figures display binned scatterplots of origin-by-specialty mean spending against destination-by-specialty mean spending. Each axis is re-centered around zero by subtracting the national specialty mean spending. Panel (A) shows the results for log(spending), while panel (B) shows the results for log(spending per beneficiary). In each panel, the 45 degree line is shown as a dashed gray line.

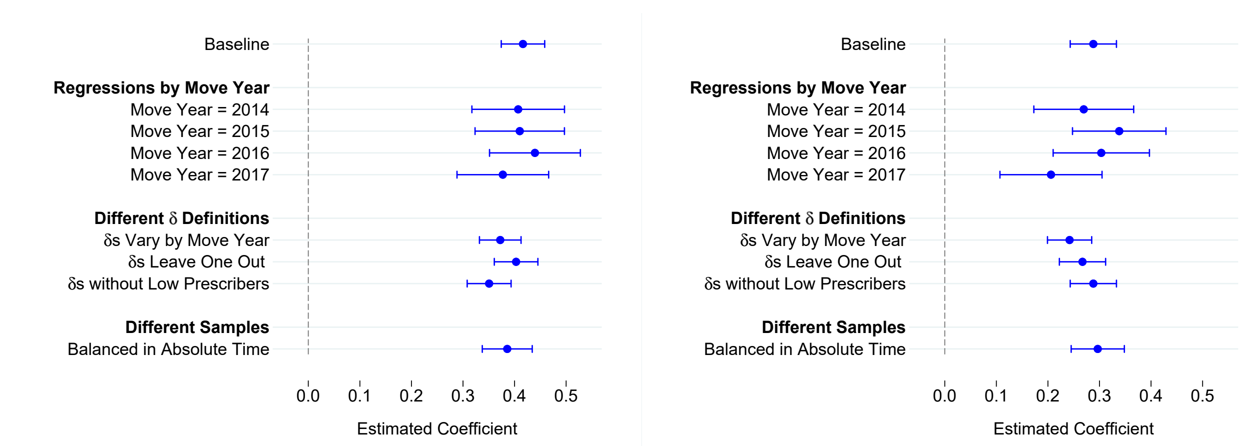
Figure A3: Histograms of δ's



(A) Log(spending) (δ) **(B)** Log(spending per beneficiary) (δ)

Note: This figure displays the treatment variable δ_{idos} for our two primary outcome variables, where $\delta_{idos} = \overline{y_{ids}} - \overline{y_{ios}}$. Panel (A) shows the histogram of the δ_{idos} term for log(spending), while panel (B) shows the histogram of the δ_{idos} term for log(spending per beneficiary).

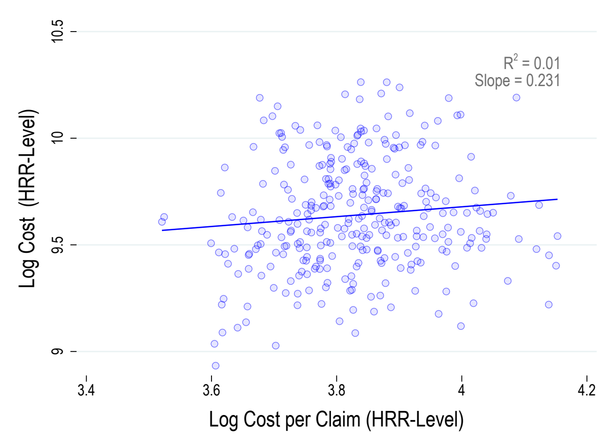
Figure A4: Regression Estimates with Alternate Samples and δ Definitions



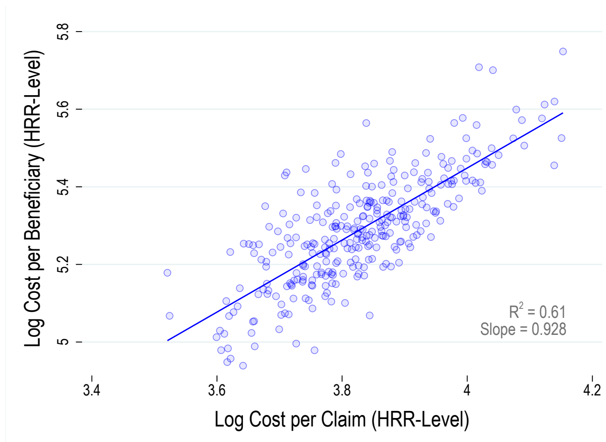
(A) Log(spending) **(B)** Log(spending per beneficiary)

Note: These figures display the value of $\hat{\beta}_{post}$ from equation 2 and the associated 95 percent confidence intervals for several different samples and alternate definitions of δ. Panels (A) and (B) display the results for log(spending) and log(spending per beneficiary), respectively. The first row shows the baseline regression coefficients, while the next several rows show the results restricted to physicians who move in a particular year. Next, we present results where we vary the definition of δ. Specifically, we compute move year-specific δs, δs that leave out the mover, and δs that omit all low prescribers (i.e., prescribers who prescribe drugs to fewer than 11 patients in a year). Finally, we re-estimate our regression including only physicians who we observe in every year of our sample.

Figure A5: Log(Cost per Claim) Versus Log(Spending) and Log(Spending per Beneficiary)



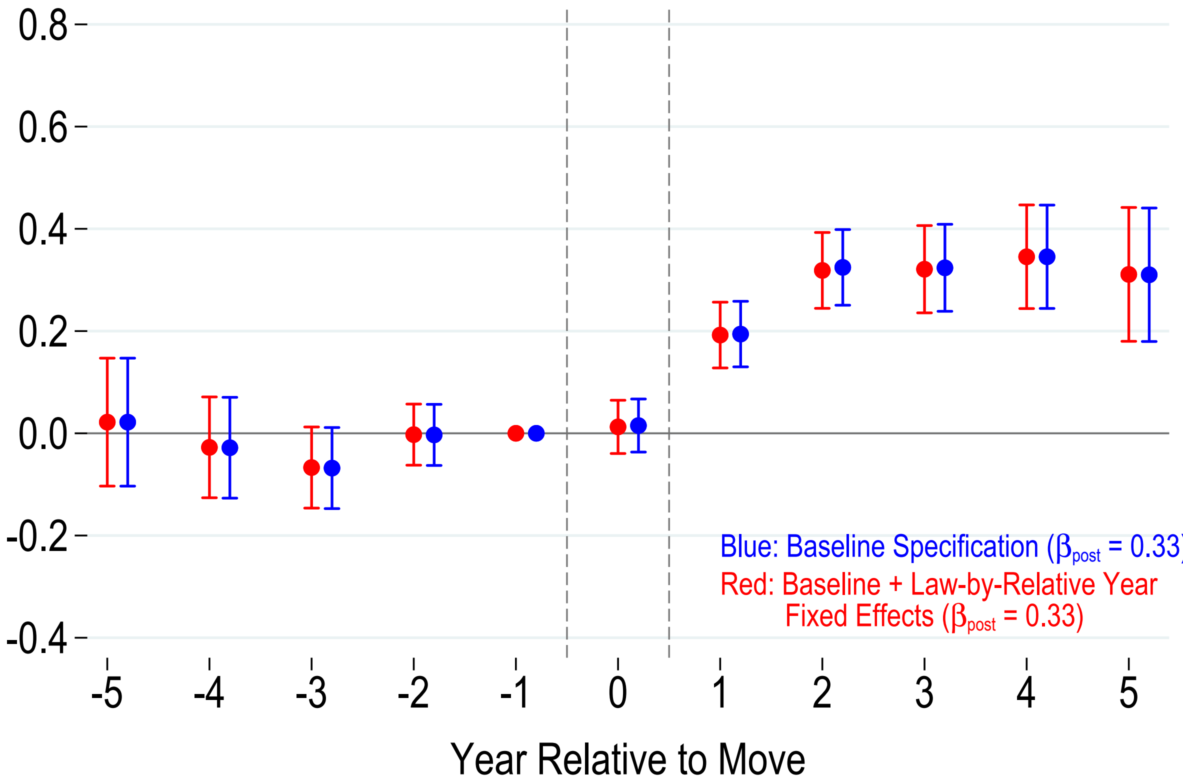
(A) Log(spending)



(B) Log(spending per beneficiary)

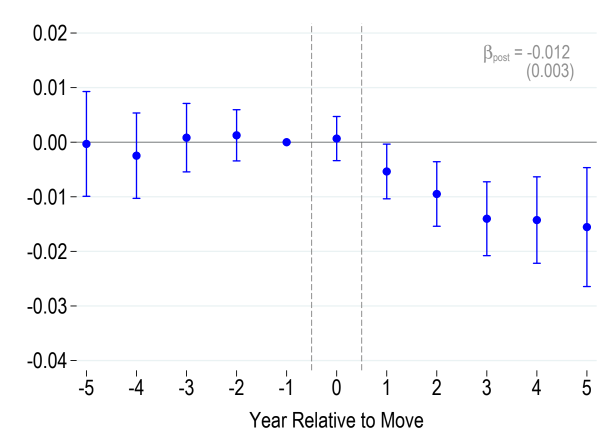
Note: This figure shows scatter plots of log(spending) and log(spending per beneficiary) against log(cost per claim) in panels (A) and (B), respectively. The regression line is shown in blue. Each circle represents a single HRR. Each panel displays the R^2 and slope from a simple linear regression.

Figure A6: Changes in Generic Share Upon Move: Controlling for Presumed Consent and Mandatory Substitution Laws

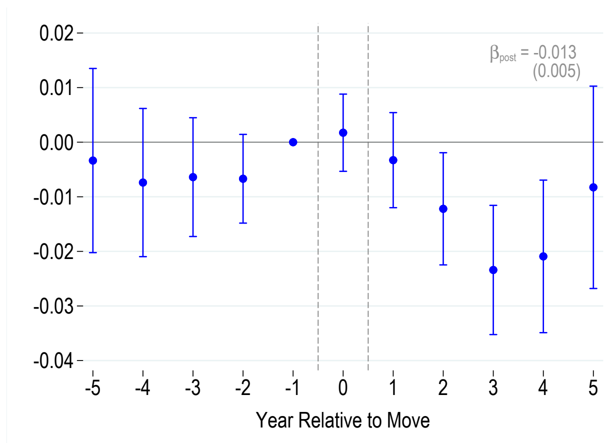


Note: This figure shows coefficients from equation 1 in blue. The outcome variable is generic drug spending as a fraction of total spending. The coefficients in red are estimated from a version of equation 1 that also includes interactions between relative year and indicators for presumed consent laws and mandatory substitution laws.

Figure A7: Change in Generic Spending Share Across Moves



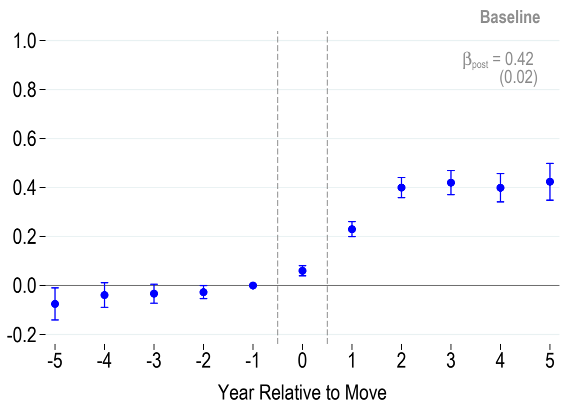
(A) δ based on log(spending)



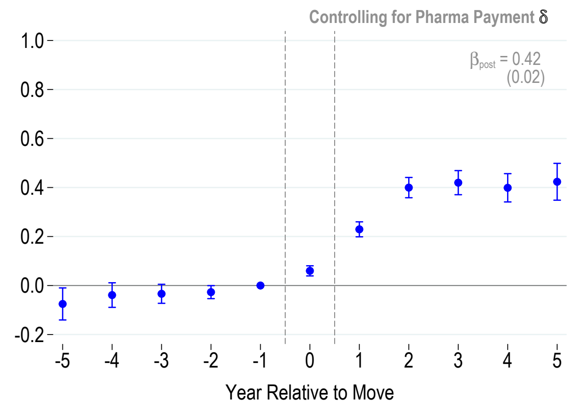
(B) δ based on log(spending per beneficiary)

Note: These figures display the β_τ coefficients from a version of equation 1. The outcome variable in each panel is the share of spending on generic drugs, but the δ terms are based on log(spending) and log(spending per beneficiary) in panels (A) and (B), respectively. Standard errors are clustered at the physician-level. The top-right corner of each sub-figure displays the β_{post} coefficient from equation 2, with the associated standard error in parentheses.

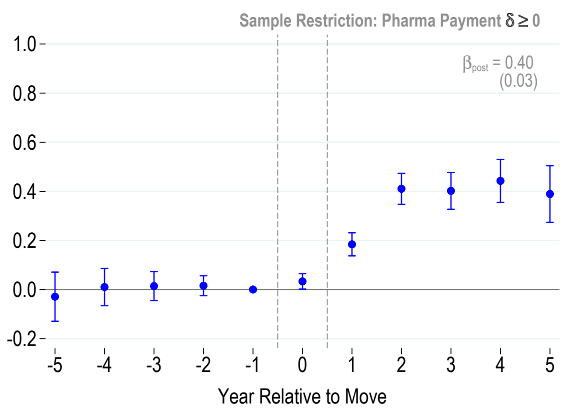
Figure A8: Changes in Prescribing Upon Move: Controlling for Pharma Payments



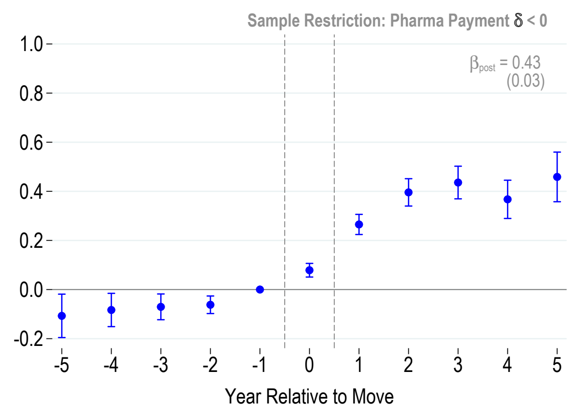
(A) Baseline



(B) Baseline + Controlling for Payments



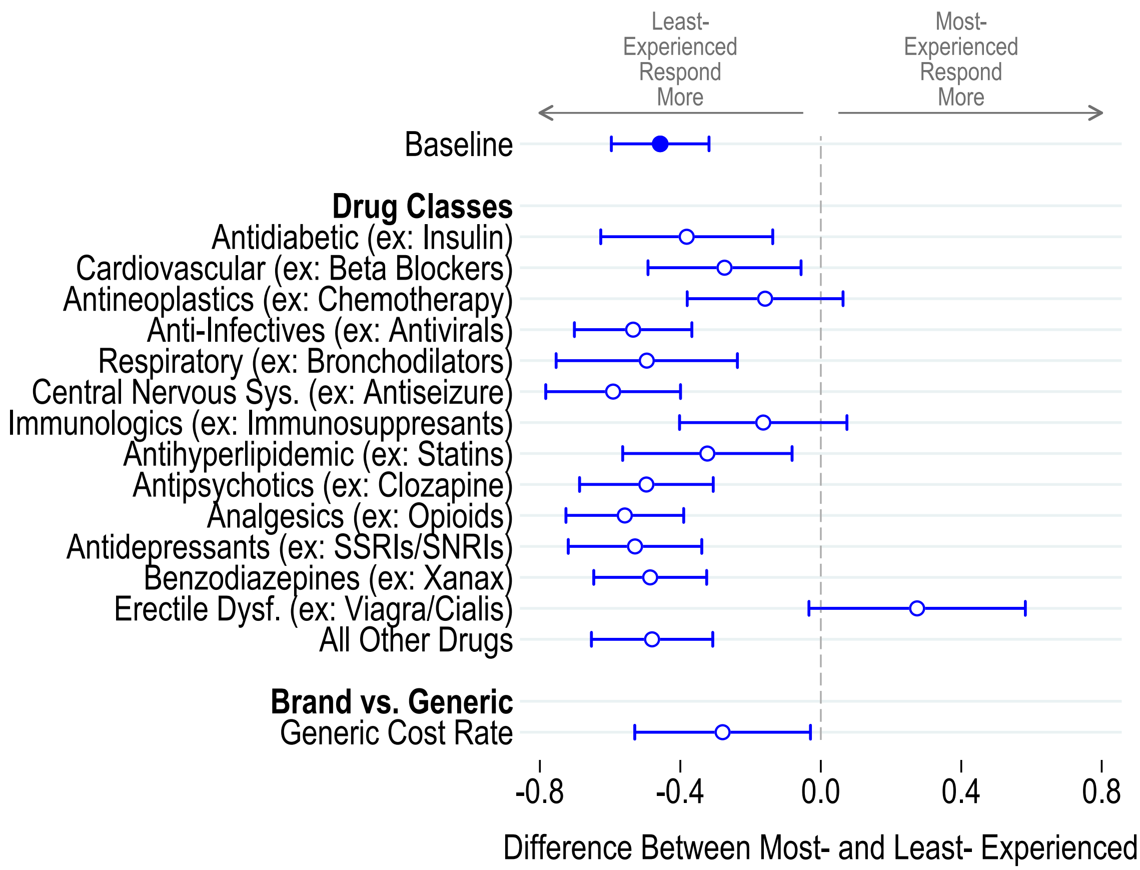
(C) Migration to Higher Payment Areas



(D) Migration to Lower Payment Areas

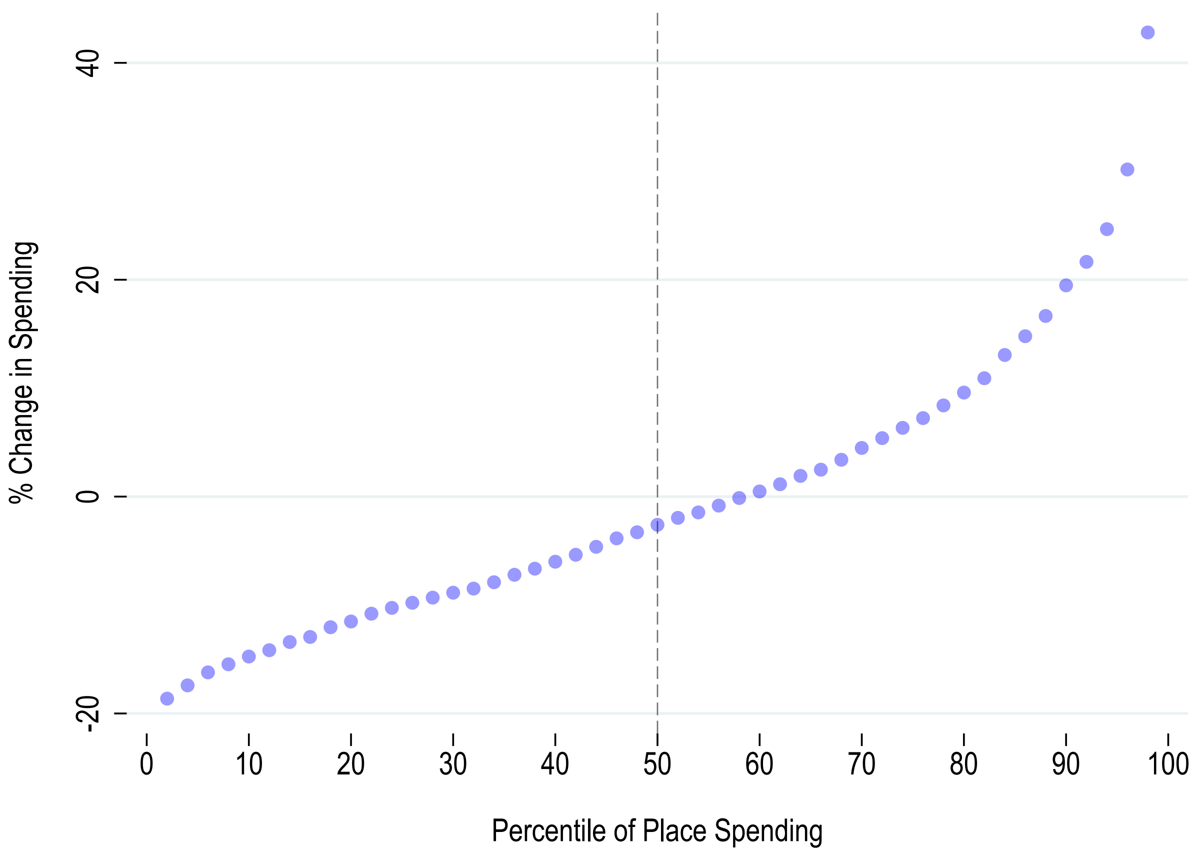
Note: This figure displays the β_τ coefficients from equation 1 for log(spending). Panel (A) is our baseline measure. In panel (B), we add interaction terms between relative year and the pharmaceutical payment δ. Panels (C) and (D) restrict the sample to include only migrants who move to higher and lower payment HRRs, respectively.

Figure A9: Heterogeneous Responses by Drug Class and Experience



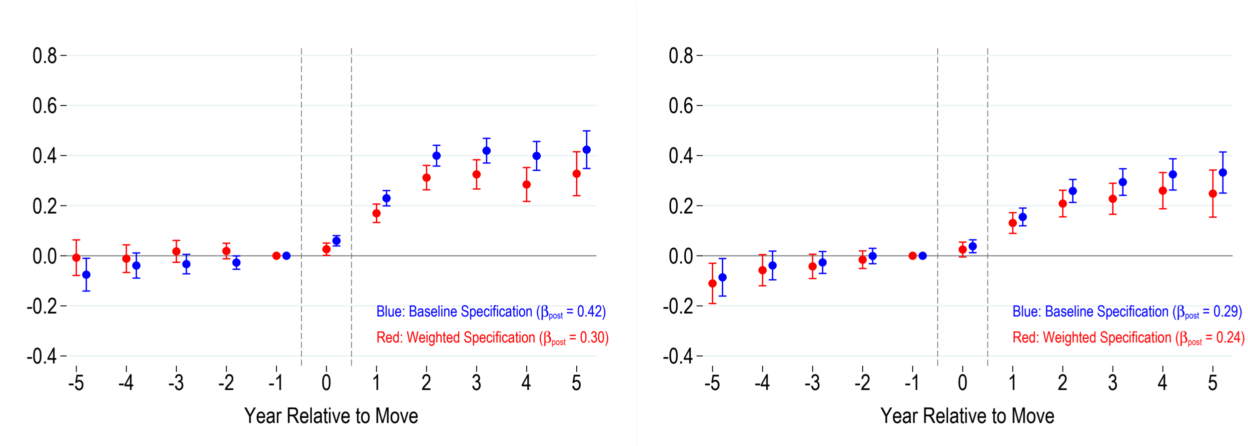
Note: This figure displays the difference in the $\hat{\beta}_{post}$ coefficients from equation 2 for the most versus the least experienced providers, for each drug category.

Figure A10: Estimated Savings under Counterfactual Spending Patterns



Note: This figure displays estimated savings under various counterfactual spending patterns. Each dot represents the savings that would occur if we were to assign to each HRR the spending associated with the average physician in the HRR in the p^{th} percentile of per enrollee spending. Values of p are plotted along the x-axis, and savings are plotted along the y-axis.

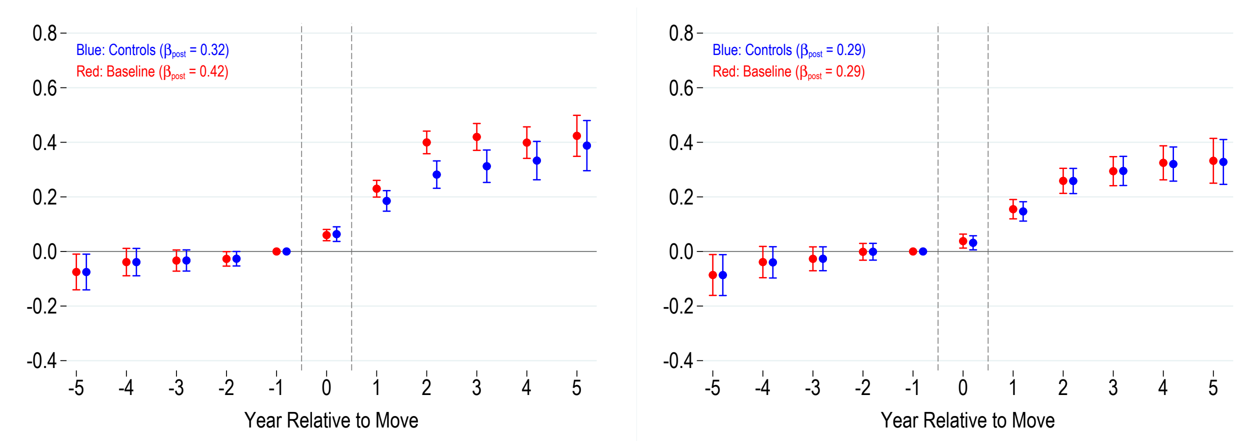
Figure A11: Regression Estimates Weighting by Experience



(A) Log(spending) **(B)** Log(spending per beneficiary)

Note: These figures display the value of $\hat{\beta}_\tau$ from equation 1 and the associated 95 percent confidence intervals from our baseline regressions (in blue) and our weighted specification (in red). Panels (A) and (B) display the results for log(spending) and log(spending per beneficiary), respectively. Weights are chosen to mimic the experience distribution of the entire sample or physicians, including non-migrants.

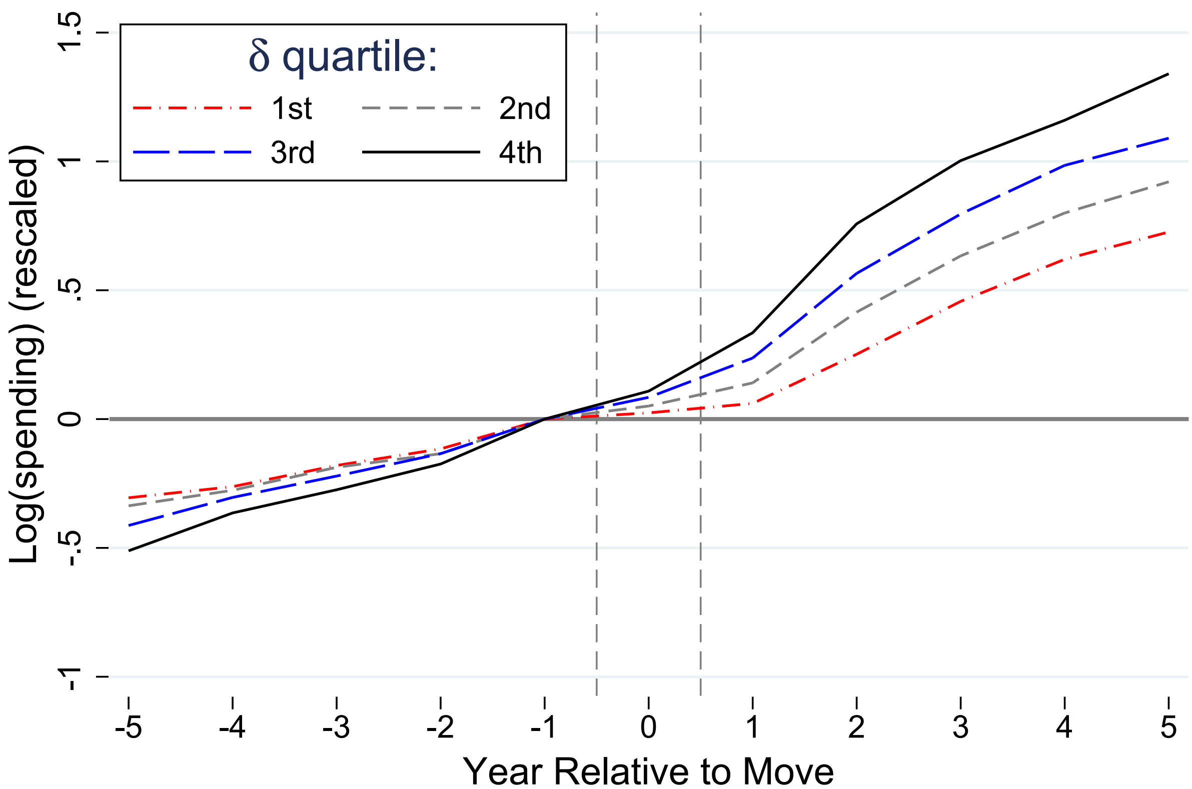
Figure A12: Changes in Prescribing Upon Move: Controlling for Enrollees per Provider



(A) Log(spending) **(B)** Log(spending per beneficiary)

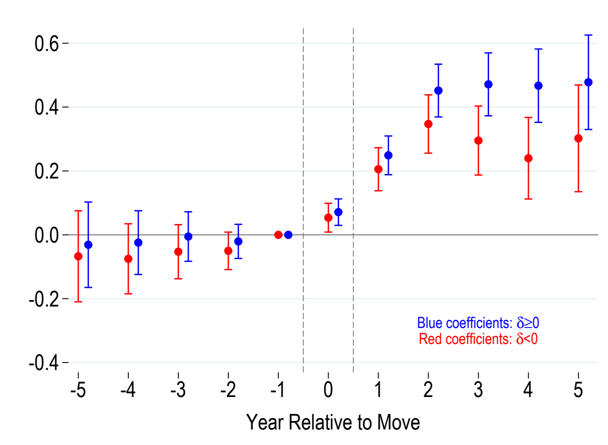
Note: These figures display the β_τ coefficients from equation 1. The outcome variables are log(spending) and log(spending per beneficiary) in panels (A) and (B), respectively. Standard errors are clustered at the physician-level. The red coefficients are identical to those in Figure 3, while the blue coefficients come from a version of equation 1 that also include interactions between the number of enrollees per provider and time-to-move fixed effects.

Figure A13: Trends in Log(spending)

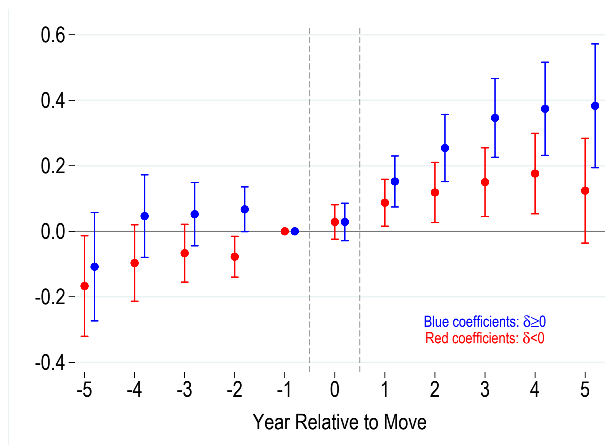


Note: This figure displays trends in log(spending) broken down by quartile of δ, pooling across all move-year cohorts. Each quartile is re-centered around zero in the year prior to the move.

Figure A14: Separate Regression Estimates for Opposite Signs of δ



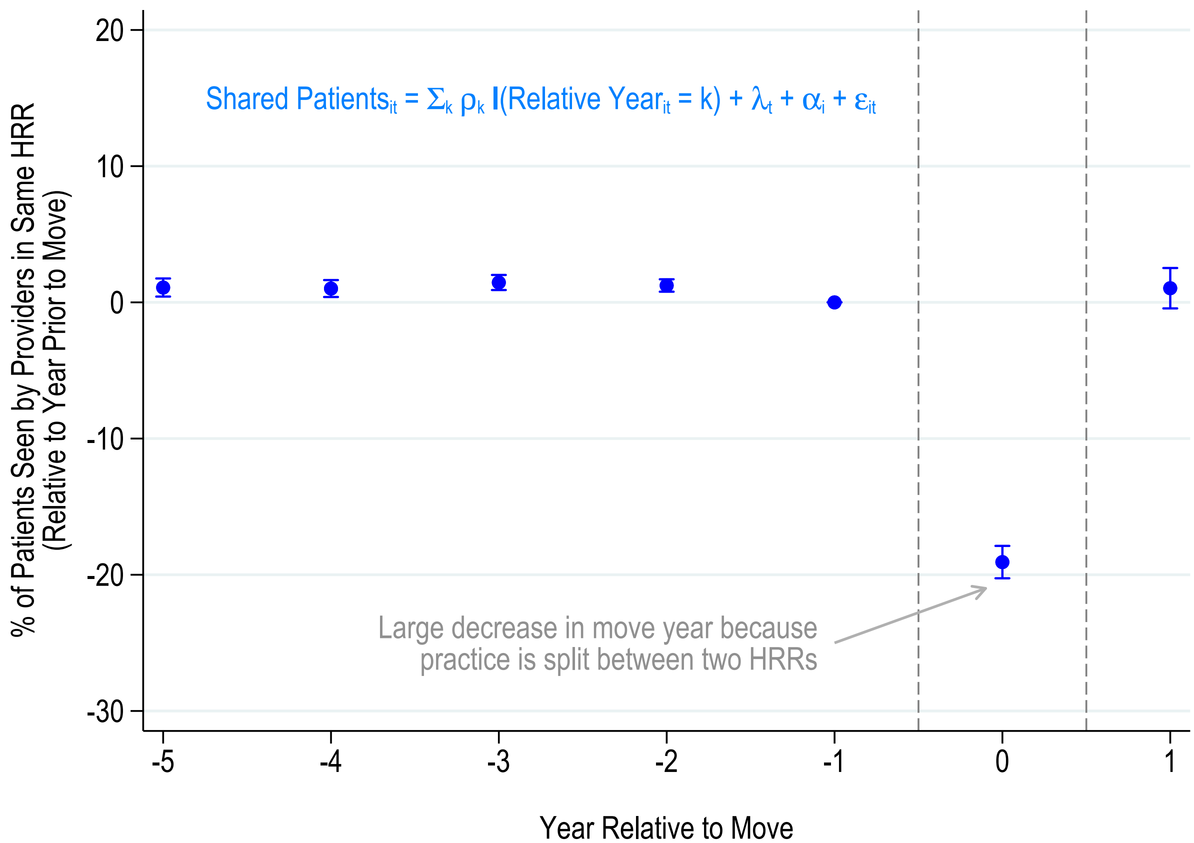
(A) Log(spending)



(B) Log(spending per beneficiary)

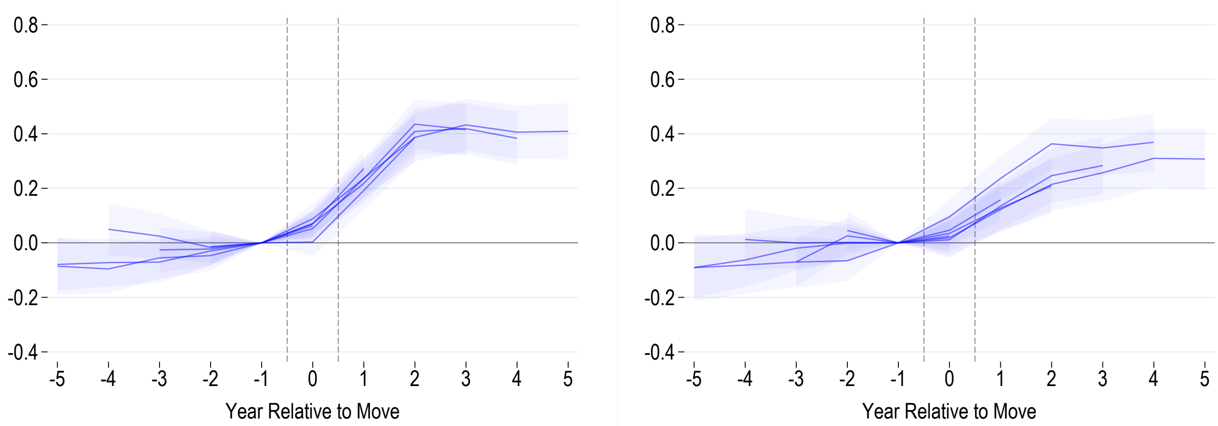
Note: These figures display the β_τ coefficients from equation 1 and the associated 95 percent confidence intervals. The blue coefficients correspond to regressions in which we only include movers with values of $\delta \geq 0$, while the red coefficients correspond to regressions only including migrants for whom $\delta < 0$. Panels (A) and (B) display the results for log(spending) and log(spending per beneficiary), respectively.

Figure A15: Event Study: Fraction of Shared Patients in Same HRR Pre and Post Move



Note: This figure displays the ρ_k coefficients from equation 6. These coefficients represent the change in the fraction of shared patients, relative to one year before the move.

Figure A16: Event Study Estimates by Move-Year Cohort



(A) Log(spending) **(B)** Log(spending per beneficiary)

Note: These figures display the β_τ coefficients from equation 1 and the associated 95 percent confidence intervals estimated separately for each move-year cohort from 2014 to 2017.

Table A1: Comparison of Effect Sizes with Related Literature

Paper	Identification	Sample/Outcome(s)	Magnitude
Current paper	Physician Migration	Medicare Part D/Drug Spending	10% increase in HRR drug spending $\implies$ 4.2% increase in migrant spending
Finkelstein, Gentzkow and Williams (2016)	Patient Migration	Medicare A & B/Total Utilization	10% increase in HRR utilization $\implies$ 4.7% increase in migrant utilization
Molitor (2018)	Cardiologist Migration	Medicare A & B/Catheterization	10% increase in HRR catheterization rate $\implies$ 6-8% increase in migrant catheterization rate
Godøy and Huitfeldt (2020)	Patient Migration	Danish Population/Hospital Utilization	10% increase in local utilization $\implies$ 4 % increase in migrant utilization
Moura et al. (2019)	Patient Migration	Dutch Population/Healthcare Spending	10% increase in local spending $\implies$ 3% increase in migrant spending
Ding (2022)	Patient Migration	Medicare A & B/Mental Health Care Utilization	10 p.p. increase in HRR mental health utilization $\implies$ 6 p.p. increase in migrant mental health utilization
Finkelstein et al. (2022)	Patient Migration	SSDI Beneficiaries/Opioid Abuse	10 p.p. increase in state opioid abuse $\implies$ 2.7 p.p. increase in migrant opioid abuse upon move, increasing to 5.96 p.p. increase after 5 years
Laird and Nielsen (2016)	Patient Migration	Danish Population/Opioid Use and Labor Force Participation	10% increase in prescriber prescription rate $\implies$ 4.5 p.p. increase in probability of migrant opioid use and 1.5 p.p. decrease in labor force participation

Note: This table summarizes the main results from several different "movers" papers. The first column lists the authors of the paper. The second provides a brief description of the identification strategy in each paper. The third column summarizes the sample and outcome variables studied. The fourth column describes the magnitudes of the main results. Where possible, we rescale the authors findings for easy comparison with this paper.