Providers, Places, and Children's Mental Health Care

Gabriel Swagel*

May 12, 2025

Abstract

Children's mental health is the defining public health crisis of our time. Using insurance claims for a national sample of 8 million privately insured children, I provide the first systematic quantification of the drivers of variation in children's mental health prescribing in the United States. I separate variation in pediatric ADHD medication and antidepressant prescribing due to differences in: 1) primary care provider (PCP) prescribing intensities, 2) regional practice environments, and 3) child health and demand. I find that eliminating differences in PCP prescribing intensities would reduce the variance of provider prescribing rates by 50 percent for ADHD medication and 65 percent for antidepressants. Geographic variation analyses understate the extent of treatment variation and the role of providers in driving overall treatment variation. I also find suggestive evidence that higher-quality PCPs tend to have higher ADHD prescribing intensities but lower antidepressant prescribing intensities.
Keywords: Children's mental health, healthcare variation, providers
JEL Codes: I1, I11, J13

*Princeton University. Email: gswagel@princeton.edu. I am grateful to Janet Currie for her guidance and support. I also thank Nick Buchholz, Jacob Dorn, Judith Hellerstein, Kate Ho, Kaan Cankat, Henrik Kleven, Linda Ouyang, Owen Zidar, Yuci Zhou, and seminar participants at Princeton University for invaluable feedback. Additional thanks to Emily Cuddy for sharing data on mental health NDC and ICD code classifications, Linda Ouyang for sharing the medical school rankings data, and to Emily Cuddy, Boriana Miloucheva, and Kaycee Mendoza for assistance navigating the BCBS data.

1 Introduction

According to the U.S. Surgeon General, children's mental health is the "defining public health crisis of our time" (Richtel 2023). Poor childhood mental health has been linked to delinquency, reduced test scores, lower income, and increased take-up of welfare in adulthood (Currie 2020). Two of the most common childhood mental health diagnoses are ADHD and depression. One in five children experience an episode of major depressive disorder in adolescence (Dwyer and Bloch 2019), while nearly 10 percent are diagnosed with ADHD (Danielson 2023). There is no precise diagnostic technology for ADHD or depression, leaving providers to exercise a relatively high level of discretion in diagnosis and treatment decisions.[1]

There are dual concerns over the substantial variation in pediatric mental health care that has been documented across regions and within regions across providers (Cuddy and Currie 2020). On one hand, many children with mental health issues remain undiagnosed or are under-prescribed (Jorm et al. 2017; Whitney and Peterson 2019). On the other, many children appear to be incorrectly diagnosed and over-prescribed (Merten et al. 2017; Piper et al. 2018), with clear evidence of inappropriate and harmful "red-flag" prescribing for children with mental health issues (Currie and Cuddy Forthcoming). Devising policies to address these concerns and reduce treatment variation requires understanding the factors that drive different prescribing rates of mental health drugs across regions and providers.

In this paper, I provide the first systematic quantification of the drivers of variation in children's mental health prescribing in the United States. I use the largest repository of health insurance claims data for privately insured children across the United States to estimate a model of ADHD medication and antidepressant prescribing that decomposes the drivers of regional and provider-level prescribing rate variation into three classes of factors: 1) differences in primary care provider (PCP) prescribing intensities; 2) regional practice environments; and, 3) underlying child health and demand. A PCP's prescribing intensity for a given condition measures their relative propensity to prescribe for the average child. Practice environments encompass factors that affect prescribing that are common across children and PCPs in a given region, such as the supply of mental health specialists in an area, state special education financing laws that incentivize mental health testing (Morrill 2018), or other local cultural and environmental factors such as climate or stigma surrounding mental health treatment.

[1] See Persson, Qiu and Rossin-Slater (Forthcoming) and Currie and MacLeod (2020) for discussions of provider discretion. The medical literature also finds evidence that ADHD and depression may be genetically linked (Riglin et al. 2021; Garcia-Argibay et al. 2024).

Using the claims, I construct a matched PCP-child panel of ADHD and antidepressant prescribing. Following the literature, the geographic unit of analysis is a hospital referral region (HRR).[2] I document that prescribing rates exhibit substantial variation across HRRs; the difference in prescribing rates between the top 10 percent of HRRs and the bottom 10 percent of HRRs is 6.25 percentage points for ADHD medication and 4.76 percentage points for antidepressants (36 and 18 percent greater than the respective average rates). However, I present new evidence that the vast majority of variation in prescribing occurs within regions across providers: the variance of PCP prescribing rates is 15 to 36 times the variance of prescribing rates across HRRs.

For the main results, I estimate AKM-style (Abowd, Kramarz and Margolis 1999) models of prescribing that include three-way additive child, PCP, and HRR fixed effects. I find that provider prescribing intensities drive a substantial share of prescribing variation. Equalizing prescribing intensities across PCPs would reduce variation in ADHD medication and antidepressant prescribing rates across PCPs by 50 and 65 percent, respectively. Consistent with the descriptive evidence, the variation in provider prescribing intensities within regions is an order of magnitude larger than the variation across regions. When decomposing variation across HRRs, I find that differences in average PCP prescribing intensities explain little of the regional variation, while differences in practice environments (captured by HRR fixed effects) account for 40 to 50 percent of the geographic variation. These results highlight the importance of going beyond geography in studying variation in medical care. Provider prescribing intensities account for a large share of mental health treatment heterogeneity across otherwise similar children, a role that is masked when focusing only on regional averages.

The models exploit child and provider *migration* across regions and within-region child *switching* between PCPs to identify PCP prescribing intensities and HRR practice environment effects, building on the approach from Badinski et al. (2023). I validate the models through three sets of analyses. First, to build support for the identifying assumptions and provide intuition for the variation underlying the main results, I estimate event studies of prescribing around child and provider geographic moves and children switching between providers. Second, I conduct analyses that show that switching providers is common, implying that the results are not subject to bias from small cell sizes and limited mobility that plagues matched firm-worker studies (Bonhomme, Lamadon and Manresa 2019). Finally, I implement a series of tests for the validity of the AKM model's additive

[2] See, for example, Finkelstein, Gentzkow and Williams (2016); Ding (2023); Badinski et al. (2023). The Dartmouth Atlas of Health Care divides the United States into 306 Hospital Referral Regions (HRRs).

structure, following Card, Heining and Kline (2013).

I use the model estimates to show several additional results. First, I find that a quarter of providers have different relative prescribing intensities across conditions, contrary to the standard uni-dimensional "cowboy" or "comforter" characterization of provider practice styles (Cutler et al. 2019; Clemens et al. 2024). I then explore the provider and regional characteristics associated with provider prescribing intensities and HRR practice environment characteristics. The estimated provider prescribing intensities show that there is substantial heterogeneity in intensity across providers, but the model does not identify the "optimal" provider prescribing intensity. Nonetheless, I find systematic patterns by specialty, experience, and training that provide suggestive evidence that higher-quality providers tend to have lower antidepressant prescribing intensity and higher ADHD prescribing intensity.

The paper proceeds as follows. In Section 2, I summarize the related literature and discuss the key contributions of this paper. In Section 3, I specify the AKM-style model of prescribing, discuss the identifying assumptions, and derive the variance decompositions. I then introduce the data and present descriptive statistics in Section 4. In Section 5, I estimate the illustrative moving and switching event studies. Section 6 presents the main estimates. In Section 7, I explore the characteristics associated with practice environments and provider prescribing intensities. Section 8 concludes.

2 Background and contribution

In this paper, I provide the first systematic quantification of the drivers of variation in children's mental health prescribing in the United States. I apply a version of the migration empirical strategy previously used to study geographic variation in healthcare for the elderly in a new setting - children's mental health care. Building on the latest methodology that adds providers as a third source of heterogeneity (in addition to places and providers), I extend the analytical framework to decompose provider-level variation in addition to geographic differences. I show that the conclusions on the drivers of variation can depend on the level of aggregation of the analysis. Finally, I contribute to the literature on physician practice styles by providing evidence that a substantial share of providers have opposing relative practice intensities for different conditions. In this section, I describe the related literature and this paper's contributions in more detail.

2.1 Variation in health care

It is well-documented that observationally similar patients often receive different care (Chandra, Cutler and Song 2011). A large literature has focused on the substantial regional differences in health care utilization across the United States (Skinner and Fisher 1997; Fisher et al. 2003*b*,*a*; Baicker and Chandra 2004; Sirovich et al. 2006; Song et al. 2010).[3] To separate supply and demand drivers of healthcare variation, the empirical literature has exploited quasi-experimental migration. Previous work has quantified place effects in Medicare spending (Finkelstein, Gentzkow and Williams 2016), cardiologist practice intensity (Molitor 2018), mortality (Finkelstein, Gentzkow and Williams 2021), opioid usage (Finkelstein et al. 2022), and mental health care for the elderly (Ding 2023).[4] These papers have mostly studied care provided to the elderly, a focus that reflects the availability of high quality data and the importance of Medicare in government spending. However, the conclusions from these studies cannot be extrapolated to other settings, especially in light of the weak regional correlation between Medicare spending and health care spending for the privately insured (Cooper et al. 2019). Applying these strategies in a new health care context of mental health care among privately insured children, I find that the drivers of geographic variation are different across settings (when studying pediatric mental health prescribing instead of overall utilization for the elderly).

There are concerns that the literature has over-emphasized geographic variation in health care. In their National Academies report, Newhouse et al. (2013) call for research and policy that targets providers, as most health care decisions are made at the individual provider or organization (e.g. hospital or physician group) level. As a step in this direction, Cutler et al. (2019) use vignettes to show that differences in physician practice styles – particularly physician beliefs about appropriate care – across regions explain (in a statistical sense) a large share of geographic variation in Medicare utilization. Most recently, Badinski et al. (2023) develop an extension of the model from Finkelstein, Gentzkow and Williams (2016) to separate place effects on Medicare utilization into a physician component and other non-physician supply-side factors. While both papers attempt to unpack the "black-box" of place effects, they remain primarily focused on outcomes at the regional level.

Methodologically, I build on the contribution of Badinski et al. (2023) to the standard

[3] The classic example compares high-utilization McAllen to low-utilization El Paso, two cities in Texas that are similar along a number of socioeconomic characteristics, noting that the observed quality of care and health outcomes are no better in McAllen than El Paso (Gawande 2009).

[4] Medicare is the U.S. public health insurance program for those aged 65 and older.

movers design. They add providers as a third dimension of heterogeneity in treatment intensity (in addition to patients and places) and show how patient and provider migration and within-HRR patient-provider switching can separately identify the three components of variation in Medicare utilization. My analysis extends their framework in three ways. First, I decompose provider-level variation in addition to geographic variation. As I show, the conclusions one can draw when decomposing variation depend on the level of aggregation being studied. Given that variation in pediatric mental health prescribing rates across providers dwarfs across-region variation, understanding the drivers of provider-level variation is crucial for understanding overall treatment heterogeneity. Second, I use provider switching event studies to evaluate the validity of the conditionally idiosyncratic switching assumption. Finally, I innovate by estimating the model separately for two outcomes (ADHD medication and antidepressant prescribing) and comparing the joint distribution of estimated effects. This approach could be useful in settings beyond healthcare in which there are multiple outcomes of interest, such as applications in education value-added estimation where math and reading scores are measured separately or there is also interest in non-test score outcomes (Chetty, Friedman and Rockoff 2014; Jackson 2018).

2.2 Mental health care and provider practice styles

This paper adds to the growing economics literature studying mental health care. Previous work highlights regional shortages of mental health providers (McBain et al. 2019), concentrated prescribing by physician (Berndt et al. 2015), differences in provider skill (Currie and MacLeod 2020), and the proclivity for "red-flag" prescribing (Currie and Cuddy Forthcoming). Marquardt (2021) documents substantial variation in physician practice intensity for ADHD among providers in the same healthcare system. Another strand of closely related work uses physician switching in Denmark (Laird and Nielsen 2016) to study physician effects and Medicare patient moving in the United States (Ding 2023) to study place effects in mental health care. I nest these empirical strategies into a more general framework and apply the model to high-quality data to quantify the role of providers, places, and patients in driving variation in mental health prescribing for a broad swath of children in the United States.

I also contribute to the broader literature on provider practice styles. A number of papers provide evidence of heterogeneity in primary care physicians (PCP) practice styles using a variety of methods, including: vignettes (Sirovich et al. 2008), controlling for patient characteristics (Grytten and Sørensen 2003), random assignment of patients to physicians (Currie and Zhang 2023), and patient switching PCPs associated with PCP exits or patient geographic moves (Kwok 2019; Ahammer and Schober 2020; Fadlon and Van Parys 2020). Other studies document similar results for physician propensity to perform a Cesarean-section (Epstein and Nicholson 2009; Currie and MacLeod 2017), utilize invasive procedures for heart attack patients (Currie, MacLeod and Van Parys 2016), or diagnose pneumonia in the emergency room (Chan, Gentzkow and Yu 2022). With the exception of Gowrisankaran, Joiner and Léger (2023)'s case study of emergency department physicians in Montreal, previous research has considered physician practice style in terms of the effect on overall healthcare expenditure or in the context of a particular condition. I show that a substantial share of pediatric PCPs have opposing relative practice intensities for different conditions. These results indicate that a richer, multidimensional understanding of practice styles is important to consider when designing provider-targeted interventions.

Additionally, I contribute further evidence that a provider's training and experience is associated with their practice style. Doyle, Ewer and Wagner (2010) and Schnell and Currie (2018) show in two different settings that a provider's medical school ranking is associated with care utilization. Focusing on specialty training, Doyle (2020) finds that EDs have better outcomes for heart failure patients when a cardiologist is on staff, controlling for a variety of confounding factors. Similarly, Chan and Chen (2022) show that being treated by an nurse practitioner in emergency departments increases length of stay and costs relative to being treated by a physician, although there is little effect on health outcomes. Evidence on the relationship between experience and outcomes is more mixed (Parys 2016; Facchini 2022). My results further suggest that interventions targeting provider training can make a difference in healthcare practice.

3 Empirical Model

3.1 Setup

I specify an AKM-style (Abowd, Kramarz and Margolis 1999) model of ADHD stimulant and antidepressant prescribing. This approach has been applied in a variety of settings, including wage inequality (Card, Heining and Kline 2013), brand loyalty (Bronnenberg, Dubé and Gentzkow 2012), and health care utilization (Finkelstein, Gentzkow and Williams 2016). While most previous studies have decomposed variation between two sources (in health applications, patients and places), the model below builds on the recent contribution of Badinski et al. (2023), who include physicians as a third dimension of variation.

I consider a child i who visits their PCP j in year t. Define two indicator variables, a_{ijt} and d_{ijt}, for whether the child is prescribed ADHD medication or antidepressants during year t (regardless of whether the medications were prescribed by the PCP or a specialist). I model $y \in \{a, d\}$ as an additive linear function of region ($r(ijt)$), primary care provider, and child-specific components:[5]

$$y_{ijt} = \alpha_i + \delta_j + \gamma_{r(ijt)} + \tau_t + \beta \mathbb{X}_{it} + \epsilon_{ijt} \tag{1}$$

The child-specific effects, α_i, capture the child's underlying health and their (or perhaps more accurately, their parents') preferences for pharmaceutical treatment with regards to condition y.[6] The provider-specific effects, δ_j, represent provider prescribing intensity for condition y, capturing the relative propensity of the provider to prescribe for the average child. Differences across providers in prescribing intensity may reflect differences in diagnostic skill (Currie and MacLeod 2020) or diagnostic thresholds (Marquardt 2021), variation in beliefs about the effectiveness and side effects of pharmaceutical treatment (e.g. the "cowboy" or "comforter" models from Cutler et al. (2019)), or differences in specialist referral networks (Zeltzer 2020). The regional effects, $\gamma_{r(ijt)}$, encompass factors driving prescribing that are common to all child-provider pairs in region r. These regional prescribing effects may reflect the local supply of mental health specialists or the supply of psychiatrists relative to psychologists (Cuddy and Currie 2020), special education financing laws that incentivize mental health testing (Morrill 2018), or other non-provider-specific healthcare and cultural factors. I also control for common time trends (τ_t^y) and the child's age in bins of 5-9, 10-14, and 15-21 ($\mathbb{X}_{it}$).

Although the outcome variables measure prescriptions regardless of the prescribing provider, the model attributes any prescribing to the child's PCP. I focus on the PCP because they are often the main provider of mental health care (Jetty et al. 2021), and they are broadly responsible for managing their patients' overall care and referring to specialists. The PCP prescribing intensities I estimate thus capture both the PCP's inclination

[5]I estimate all models separately for each condition $y \in \{a, d\}$. For clarity, I often suppress y specific notation on right-hand side variables (e.g. fixed effects).

[6]The child fixed effects may also capture variation in moral hazard from potential out-of-pocket costs. Given that, by construction, all children in the data have similar (private) health insurance and standard ADHD medication and antidepressants are widely available as generics with minimal out-of-pocket costs, the impact of such variation is likely minimal.

to prescribe in isolation and their *relative* proclivity to refer children in their care to high prescribing intensity specialists. Differences in the overall supply of specialists and the *average* prescribing intensity of those specialists across regions are captured by the region fixed effects.[7] Finally, differences in children's likelihood of visiting a specialist regardless of a referral from their PCP are captured by the child fixed effects.

3.2 Identification and inference

The model exploits variation in prescribing from child and provider *migration* across regions, and variation in prescribing from children *switching* between providers. Child and provider effects (α_i and δ_j) are separately identified off of prescribing changes as children switch providers and prescribing variation within a provider across patients. Intuitively, consistent increases in prescribing as children switch from provider A to provider B and decreases as children switch from B to A identify the relative intensities for these two providers. A child seeing provider A who has consistently lower prescribing relative to A's average prescribing then identifies the child's fixed effect. In Section 4, I show that switching between providers is common and that being prescribed is not an absorbing state (i.e children switch off their medications after being prescribed), providing sufficient variation to empirically exploit these sources of identification. Place effects are identified from changes in prescribing as patients and providers move across regions. These moves also allow the estimated child and provider effects to be comparable across regions, rather than being identified only as within-region relative prescribing intensities.

The first identifying assumption is a "parallel trends" assumption for movers across different origin-destination pairs. Unobservable shocks in health or pharmaceutical preferences (for child movers) and shocks in practice intensity (for provider movers) must be uncorrelated with the difference in prescribing intensity between the origin and destination. This parallel trends assumption is present in previous health applications of movers designs using Medicare data for the elderly (Finkelstein, Gentzkow and Williams 2016, 2021; Finkelstein et al. 2022; Ding 2023). The assumption is *a priori* more plausible in this context of children's mental health care relative to the Medicare data. In the Medicare context, health concerns and the desire to be closer to family aid are oft-cited reasons for moving. Here, it would be surprising if privately insured, non-elderly parents system-

[7] To the extent that PCPs have overlapping referral networks, the estimated share of prescribing variation driven by PCP intensities is a conservative bound on the variation driven by differences across all providers because some of the differences across specialists are averaged out across PCPs. Nonetheless, I find empirically that PCP intensities drive a substantial share of prescribing variation across providers.

atically move across regions because of changes in their children's mental health status. In Section 5, I show that there are no systematic trends in prescribing of either drug class before or after child or provider moves.

The second assumption is conditionally idiosyncratic switching of children between providers. The model explicitly allows for sorting between higher proclivity patients and higher intensity providers. However, I assume that sudden changes in child health or demand do not drive provider switching. This assumption would be violated if, for example, parents respond to a child exhibiting ADHD symptoms by switching to a PCP who is more likely to prescribe ADHD medication. Below, I test this assumption with event studies of patient switching. I show that prescribing jumps in the direction of the old-new provider prescribing rate difference even when switching is clearly driven by exogenous forces (e.g. the old provider moves). Relatedly, the additive structure of the model also rules out sorting on match-effects between children and providers. Following Card, Heining and Kline (2013), I find that upward effects when a child switches from a low- to a high-intensity provider are symmetric to the downward effects when a child switches in the opposite direction. Moreover, prescribing is stable before and after the switches.

A growing literature has highlighted potential pitfalls for inference in high-dimensional fixed effects models, focusing particularly on matched establishment-worker wage decompositions. A substantial share of small cell sizes and limited mobility of workers across firms introduce sampling errors in the estimated worker and establishment effects that can upwardly bias the variance of these components and negatively bias the covariance between components (Andrews et al. 2008, 2012). Bonhomme, Lamadon and Manresa (2019) and Kline, Saggio and Sølvsten (2020) introduce procedures to compute bias-corrected estimates of the variance components when these issues arise. In Section 4, I present descriptive evidence that small cell size and limited mobility concerns are less relevant in my context. Bonhomme et al. (2023) show that there is little bias in the share of wage variance due to establishment effects when restricting to establishments with more than 30 matched workers. I therefore restrict my estimation sample to providers who are matched to at least 30 children and observed in the data for at least three years. In contrast to the wage data, where this restriction drops a substantial share of workers, I retain over 80 percent of the child-year observations in the data. There is a high degree of switching providers in the data. Over 60 percent of children in the estimation sample see multiple providers and 90 percent of providers see at least 30 switching children (see Figure A1 and Figure A2).

3.3 Variance decompositions

Using the model, I can decompose geographic and provider-level variation in prescribing into the shares attributable to providers, regional practice environments, and child health and demand.

Following Finkelstein, Gentzkow and Williams (2016) for the geographic decomposition, define $\bar{y}_r$ as the prescribing rate for medication $y \in \{a, d\}$ in a region r. Define the child-year weighted average of provider effects in region r as $\bar{\delta}_r$ and the average of child effects as $\bar{\alpha}_r$. When aggregating to groups of regions, we define can define the analogous measures. The share of the difference in prescribing between region r and r' attributable to regional practice environments (non-provider place effects) is then (suppressing some of the condition y notation for clarity):

$$S_{place}(r, r') = \frac{\gamma_r - \gamma_{r'}}{\bar{y}_r - \bar{y}_{r'}} \tag{2}$$

The share attributable to differences in average provider prescribing intensity is:

$$S_{docs}(r, r') = \frac{\bar{\delta}_r - \bar{\delta}_{r'}}{\bar{y}_r - \bar{y}_{r'}} \tag{3}$$

And the share attributable to differences in average child health and demand is:

$$S_{child}(r, r') = \frac{\bar{\alpha}_r - \bar{\alpha}_{r'}}{\bar{y}_r - \bar{y}_{r'}} \tag{4}$$

Any remaining share is attributable to differences in average time trends (regions have different relative sample sizes across years), average age effects (regions have different age compositions), and average residuals (these are empirically negligible).

For the provider-level decomposition, define a provider j's prescribing rate as $\bar{y}_j$. From Equation 1:

$$\bar{y}_j = \delta_j + \bar{\alpha}_j + \bar{\gamma}_j + \bar{\tau}_j + \beta \bar{\mathbb{X}}_j + \bar{\epsilon}_j \tag{5}$$

Where, for example, $\bar{\alpha}_j$ denotes the average child fixed effect across provider j's matched child-years. The variance of provider-level prescribing rates can therefore be written as:

$$V(\bar{y}_j) = V(\delta_j) + V(\bar{\alpha}_j) + V(\bar{\gamma}_j) + V(\bar{\tau}_j + \beta\bar{\mathbb{X}}_j + \bar{\epsilon}_j) + \\ 2cov(\delta_j, \bar{\alpha}_j) + 2cov(\delta_j, \bar{\gamma}_j) + 2cov(\bar{\alpha}_j, \bar{\gamma}_j) + \aleph_{cov} \quad (6)$$

Where $\aleph_{cov}$ represents the additional covariance terms. In addition to decomposing provider-level variation separately by condition, the estimated δ_j's also recover the joint distribution of provider prescribing intensities across conditions, allowing me to provide new evidence on the extent to which practice intensities vary between conditions within a provider.

4 Data

4.1 Sample and variable definitions

The primary data are insurance claims and membership files from the Blue Cross Blue Shield (BCBS) Alliance for Health Research, a collaborative effort involving most of the regional BCBS plans, from 2012 through 2022.[8] These data offer several advantages. The large sample size and broad national coverage of BCBS plans – BCBS plans cover nearly one-third of all Americans (BCBS 2024) – allow me to follow children over time and across regions. The data contain detailed information on all claims for inpatient, outpatient, and pharmaceutical care.[9] Children with BCBS health insurance tend to live in younger, less racially diverse, and higher socioeconomic status areas than the average American (Currie and Cuddy Forthcoming); the results below may not reflect prescribing patterns for children covered by Medicaid. Nonetheless, children covered by BCBS constitute a large population of interest.

Although BCBS plans are not fully uniform, coverage for ADHD medication and antidepressant prescribing is standard across plans and the most commonly prescribed drugs were available as generics with low out-of-pocket costs throughout the sample

[8]This limited data set is made available through a secure data portal and is drawn from Blue Cross Blue Shield (BCBS) Axis®, the largest source of commercial insurance claims data in the U.S. Accessing insurance claims data often requires extended negotiations with individual insurance carriers, or with government entities. Further information about the BCBS Health of American Initiative, including information about their Axis® data base and contact information is available at: https://www.bcbs.com/the-health-of-america/about.

[9]BCBS plans typically offer some coverage for visits to out-of-network providers, such that most primary care encounters are likely to be captured in the data.

period. Any residual variation driven by differences across the regional BCBS plans will be captured by the regional practice environment effects. Empirically, I find that within-region prescribing variation is an order of magnitude greater than across-region variation – evidence that differences in coverage across BCBS plans plays (at most) a minor role in driving prescribing variation.

I focus on children aged 5 through 21, restricting to child-years in which the child is covered by BCBS (including prescription drug coverage) for at least six months. This base BCBS child sample includes 11,495,405 children observed over 47,427,135 child-years. The child-year panel is unbalanced; children enter the sample as they turn 5 or obtain BCBS coverage, and exit the sample as they turn 21 or transition off of BCBS coverage.[10]

Following the standard practice in the literature (Finkelstein, Gentzkow and Williams 2016, 2021; Ding 2023; Badinski et al. 2023), the geographic unit of analysis is a hospital referral region (HRR), as defined by the 1998 Dartmouth Atlas. I use the child's ZIP code of residence associated with each claim to assign the HRR of residence for each child in each calendar year. See Appendix A.1 for more details. I define child *movers* as those whose HRR of residence changes exactly once. In the analysis sample (defined below), I observe 243,940 child-movers over 1,326,094 child-years. Given the data source, I only observe movers who maintain or reacquire coverage from a BCBS plan after their move. The ubiquity and national coverage of BCBS plans ensure I observe a sufficient number of movers between HRRs. As I report in Table A1, the distribution of origin and destination Census Regions and the share of moves across state boundaries are roughly similar to those in Finkelstein, Gentzkow and Williams (2016).

As children may exit the sample and then reappear some years later, I can only determine the year of the move for a child mover who is observed in her destination immediately following her final observed year in her origin. I use this subset of 170,188 child movers to estimate the child mover event study in Section 5, but include all movers when estimating the main model. I define relative-year 0 as the first calendar year the mover is assigned to their destination HRR. As children move throughout the year, relative-years -1 and 0 represent partially treated years. In the child mover event study, I account for this by normalizing the coefficient for relative-year -2 to 0.

I match children to their PCP in each year and construct indicators for whether the

[10] For entry or exit to bias estimates, it would require parents to be obtaining or dropping BCBS coverage – likely switching jobs – in response to their children's mental health shocks. I argue that this degree of selection is implausible, particularly given that BCBS coverage of pediatric mental health care is comparable to other private insurers.

child was prescribed an ADHD stimulant or antidepressant in that year. For each claim, I observe the date of service, the diagnosis, procedure, and BETOS codes, and the National Provider Identifier (NPI) of the rendering or prescribing provider. I link these to provider characteristics, including speciality, from the National Plan and Provider Enumeration System (NPPES) and data provided by IQVIA. I first consider a set of typical PCPs: physicians with specialties in General Practice, Family Medicine, Internal Medicine, and Pediatrics, as well as Nurse Practitioners (NPs) and Physician Assistants (PAs). I assign the child's PCP as the provider from these specialties with whom they have the greatest number of office visits in that calendar year (if they see a provider with one of these specialties but have no recorded office visits, I assign the PCP based on the number of overall claims). If a child does not see any of these specialties, I assign their PCP as the provider with whom they have the most office visits, excluding a subset of specialties such as dentists or physical therapists. See Appendix A.2 for additional details. 28 percent of child-years are not matched to a PCP, roughly in line with survey evidence for the share of children who have not had a well-visit checkup in the past 12 months (NHIS 2016).[11]

I define a provider's location based on the location of the patients she treats. I construct a provider-patient-year dataset for the providers in my sample using all BCBS members (including adults), and then follow the procedure developed by Badinski et al. (2023) to define a provider as a "mover" only if the primary location of residence for their patients clearly shifts exactly once during the sample.[12] There are 15,842 provider movers in the sample I use for analysis.[13] The provider mover identification procedure results in the calendar year associated with year 0 relative to the move as the only partially treated year. In Figure A4, I show that, on average, over 80 percent of a mover provider's patients are in their assigned origin HRR prior to relative-year 0, 40 percent are in the origin in relative-year 0, and over 80 percent are in the provider's destination HRR in the following years. I therefore normalize the coefficient on relative-year -1 to 0 in the provider mover event study below.

Some non-mover providers are observed treating patients from multiple HRRs in a given year, likely reflecting that providers may practice at multiple locations and that patients may travel across HRR boundaries for care. To avoid identifying place effects

[11] I include these observations in the model estimation by constructing HRR-specific "provider" IDs for unmatched observations.

[12] See Appendix A of Badinski et al. (2023) for the details of the algorithm.

[13] Table A2 shows the distribution of origin and destination Census Regions for provider movers. The number of provider movers makes it infeasible to estimate the model using alternative, smaller geographic units, such as Primary Care Service Areas, given that there are 6,102 Primary Care Service Areas in the United States.

from within-provider variation in location that is not clearly associated with migration, when estimating the model I include a separate fixed effect for each non-mover provider-HRR combination. For provider movers, I estimate a single fixed effect for child-years in the origin HRR during the provider's pre-move years, in the origin and destination HRRs in the move year, and in the destination HRR after the move year. Other child-years are associated with provider-HRR specific fixed effects. A provider's estimated prescribing intensity is then the average of their estimated fixed effects, weighted by the number of corresponding child-years.

I impose additional restrictions to arrive at my analysis sample from the base BCBS sample. To minimize potential bias from sampling error in the AKM variance decomposition and ensure sufficient power to identify provider prescribing intensities, I drop providers who are matched to fewer than 30 children or observed for fewer than three years in the panel (Bonhomme et al. 2023).[14] I also drop the small number of children who move more than once. Finally, I restrict to the largest connected set of children, places, and providers. These restrictions reduce the number of providers in the sample from 717,724 to 130,616, but retain 82.6 percent of the child-year observations. In Table 1, I show how the age distribution, prescribing rates, and healthcare utilization measures compare between the base BCBS sample and the sample I use for analysis. The child-years excluded from the primary analysis due to these restrictions have higher healthcare utilization along a number of dimensions. In Table A3 and Table A4, I show that the main decomposition results are robust to including child-years matched to smaller providers.

4.2 Descriptive statistics

Figure 1 maps the HRR prescribing rates for ADHD medication and antidepressants, illustrating the regional variation in prescribing across the United States. The average HRR has an ADHD prescribing rate of 4.21 percent and antidepressant prescribing rate of 3.75 percent, with standard deviations of 1.80 and 1.37 percentage points, respectively. ADHD prescribing rates are highest in the Southeast (HRRs in Louisiana, in particular, have consistently high ADHD prescribing rates), while antidepressant prescribing rates are highest in the Northeast and Plains states. The correlation coefficient of HRR prescribing rates across drug classes is 0.64.

Although there is a high degree of variation in pediatric mental health prescribing

[14] Similar small cell restrictions are common practice (Grytten and Sørensen 2003; Epstein and Nicholson 2009; Sorkin 2018).

rates across HRRs, averaging by region masks the full extent of treatment variation. Figure 2 shows the distributions of provider-level prescribing rates. The provider-level variation is 15 to 36 times as large as the geographic level variation. The (child-year weighted) average provider prescribing rate for ADHD medication is 5.81 percent, with a standard deviation across providers 6.74. Similarly, for antidepressants, the mean provider prescribing rate is 6.87 percent, while the standard deviation is 7.55.

The empirical model relies on within-child variation in prescribing as children switch between providers (and move across regions). One potential concern is that prescribing may exhibit habit formation; once a child is prescribed, they stay prescribed. Such state-dependence would limit the identifying variation and bias the provider fixed effects in finite samples. In Figure A3, I investigate how prescribing evolves in the years following a child's first prescription. For both ADHD medication and antidepressants, five years after the initial prescription, only half of children are still prescribed the medication, suggesting that the role of habit formation is limited.[15]

As discussed above, an additional concern is that limited switching across providers would provide insufficient power to identify the provider prescribing intensities. However, the provider graph is highly connected, allowing for unbiased inference using standard methods. In Figure A2, I show the distribution across providers of the number of switchers they see in the sample. Almost 90 percent of providers are matched to at least 30 switchers in the sample.[16] For comparison, in the Swedish linked establishment-worker data used in Bonhomme, Lamadon and Manresa (2019), only 0.8 percent of establishments have at least 50 employee switchers. Additionally, Bonhomme et al. (2023) show that in the canonical AKM earnings framework with firm and worker effects, if all firms have more than 30 workers (including switchers and non-switchers), the standard two-way fixed effects estimation recovers unbiased estimates of the variance of firm effects.

[15]That the patterns of prescribing are similar for ADHD medication and antidepressants raises additional questions about mental health treatment adherence that are beyond the scope of this paper but worthy of additional investigation. For most children, antidepressants are meant be taken as a course until a few months after symptoms abate. However, children with ADHD are often recommended to continue taking medication for longer periods (or for life).

[16]In Figure A1 I present this result from the persepective of the children switching; the figure shows the distribution of children by the number of different PCPs they are matched to in the sample. Over 60 percent of children switch providers at least once during the sample, with a large share switching multiple times.

5 Illustrative event studies

Before turning to the main results from estimating Equation 1, I present a series of event studies to illustrate how child and provider moves can identify place effects and child switching can identify provider effects. These event studies also serve to test the identifying assumptions of parallel trends across origin-destination pairs and conditionally idiosyncratic switching between providers.

5.1 Mover event studies

The child mover event study plots how prescribing changes in proportion to the origin-destination difference in prescribing rates. Consider a child i who moves from origin $orig(i)$ to destination $dest(i)$. Define $\bar{y}_{-i,orig(i)}$ and $\bar{y}_{-i,dest(i)}$ as the leave-one-out average prescribing rates in the origin and destination, respectively. Define $\hat{\delta}_i = \bar{y}_{-i,dest(i)} - \bar{y}_{-i,orig(i)}$ as that child's origin-destination difference in prescribing rates. The child event study specification is:

$$y_{it} = \alpha_i + \theta_{w(it)}\hat{\delta}_i + \phi_{w(it)} + \tau_t + \beta\mathbb{X}_{it} + \epsilon_{it} \tag{7}$$

which is estimated separately for each condition on the sample of child movers. The regression includes child fixed effects (α_i) to control for individual heterogeneity, year fixed effects (τ_t), age bin controls (5-9, 10-14, 15+) and relative-year (to move) effects ($\phi_{r(i,t)}$). The coefficients of interest, $\theta_{w(it)}$, are the relative-year specific effects of the origin-destination difference in prescribing. As noted above, relative-years -1 and 0 represent partially treated years. I therefore normalize the coefficient for $w = -2$ to 0. If the parallel trends assumption holds, the event study plot should be flat prior to the move, (potentially) exhibit a jump at the move, and be flat again after the move. The jump in the event study coefficient at the move represents the joint contribution of local practice environments effects and average provider prescribing intensity to the prescription rate difference across origin-destination pairs for movers in the sample (Finkelstein, Gentzkow and Williams 2016).

Figure 3 plots the $\theta_{w(it)}$ estimated from Equation 7 for ADHD medication and antidepressants.[17] Prior to the move, average prescribing is roughly flat relative to the origin-

[17] In the regressions, observations before relative-year -5 are binned together with the relative year -5 indicator, and observations after relative-year 5 are binned together with the relative-year 5 indicator. The relative-year coefficients for -5 and 5 show the estimated θ_{wit} for these relative-year bins.

destination difference in prescribing; the pre-move coefficients are not statistically significant. After the move, average prescribing changes in the direction of the difference in HRR prescribing rates, and remains shifted in that direction. Although the estimates are somewhat imprecise, the post-move estimates are nearly all statistically distinguishable from 0. The point estimates show a 0.2 point jump after the move for both ADHD medication and antidepressants, consistent with the supply-side factors (average provider intensity and local practice environments) jointly explaining 20 percent of the variation in prescribing across origin-destination pairs for movers in the sample. This share is lower than the 60 percent supply share for variation in overall utilization among Medicare beneficiaries estimated by Finkelstein, Gentzkow and Williams (2016) but roughly in line with the 30 percent supply share for mental health prescribing among Medicare beneficiaries found by Ding (2023).

The provider mover event study has a similar setup. Consider provider mover j who moves from origin $o(j)$ to destination $d(j)$. Define $\bar{y}_{-j,orig(j)}$ and $\bar{y}_{-j,dest(j)}$ as the average prescribing rates in the origin and destination, respectively, excluding child-years with provider j as the PCP. Define $\hat{\delta}_j = \bar{y}_{-j,dest(j)} - \bar{y}_{-j,orig(j)}$ as that provider's origin-destination difference in prescribing rates. I estimate:

$$\bar{y}_{jt} = \zeta_j + \theta_{w(j,t)}\hat{\delta}_j + \phi_{w(j,t)} + \tau_t + \epsilon_{jt} \tag{8}$$

where $\bar{y}_{jt}$ is the provider's prescribing rate for medication y in year t. I include provider fixed effects (ζ_j), year fixed effects (τ_t), and relative-year effects ($\phi_{w(j,t)}$). The relative-year specific coefficients of interest, $\theta_{w(j,t)}$, capture how a provider's average prescribing changes before and after a move, as a share of the origin-destination prescribing rate difference.

Figure 4 plots the provider mover event study estimates for ADHD medication and antidepressants.[18] Provider prescribing rates are stable prior to and after the move, and jump sharply at the move in the direction of the origin-destination prescribing rate difference. When providers move, they not only change local practice environments, but also face a different distribution of child health and demand. The 0.7 to 0.8 jump in the provider mover event study implies that local practice environments and average child health and demand jointly explain roughly 70 to 80 percent of the variation in origin-destination prescribing rates for provider movers. This result is thus the mirror image

[18] In the regressions, observations before relative-year -5 are binned together with the relative year -5 indicator, and observations after relative-year 5 are binned together with the relative-year 5 indicator. The relative-year coefficients for -5 and 5 show the estimated θ_{wit} on these relative-year bins.

of the large child health and demand share estimated in the child event study. It is reassuring that the child and provider mover event studies provide similar estimates of the supply and demand shares of geographic variation, as they are estimated on distinct subsamples of the data at different levels of aggregation.

5.2 Switching event studies

The switching event studies analyze how prescribing changes upon a switch between providers relative to the difference in prescribing rates between a child's old and new providers. For child i who switches from provider $o(i)$ to $n(i)$, define κ_i as the difference in the leave-one-out prescribing rates between $o(i)$ and $n(i)$. I estimate:

$$y_{it} = \alpha_i + \theta_{w(it)}\hat{\kappa}_i + \phi_{w(it)} + \tau_t + \beta\mathbb{X}_{it} + \epsilon_{it} \tag{9}$$

The main coefficients of interest, $\theta_{w(it)}$, capture how average prescribing changes in the years preceding and following a child switching providers as a share of the difference in prescribing rates between the old and new providers. The regression also includes relative-year ($\phi_{w(it)}$) and age bin controls ($\mathbb{X}_{it}$). As in the child mover event study, switches occur between relative-years -1 and 0; I thus normalize the coefficient for relative-year -2 to 0. If the conditionally idiosyncratic switching assumption holds, the jump at the switch is an estimate of the contribution of provider prescribing intensities (as opposed to differences in unobservable child health) to the prescribing rate differences between old and new providers for consistent switchers.

I first present results from a set of switching event studies estimated on all switchers for whom I can identify the year of the switch.[19] As children may switch PCPs multiple times, I focus on each switcher's first switch and drop observations following a potential second switch to clarify the comparison between the pre-period with the old provider and the post-period with the new provider. The resulting "first switch" sample is comprised of 3,875,841 children observed over 14,870,961 child-years.

Figure 5 plots the $\theta_{w(it)}$ from estimating Equation 9 on the first switch sample separately for ADHD medication and antidepressant prescribing. Following a switch, the probability of being prescribed changes sharply in the direction of the difference in provider prescribing rates. The magnitude of the estimated jump is 0.3 for both ADHD medica-

[19] As was the case for movers, I can only identify the year of a switch for children who are observed seeing a new PCP in the next calendar year because children may drop out of the sample for a period and then reappear with a new provider (because they lose BCBS coverage for an intervening period).

tion and antidepressants, implying that provider prescribing intensities explain 30 of the variation in prescribing rates across old-new provider pairs.[20]

While the ADHD estimates are relatively stable before and after the switch, the antidepressant estimates exhibit a small but statistically significant pretrend. The large sample size provides a high degree of statistical precision such the confidence intervals will not contain zero even for small point estimates (indeed, the pre-switch ADHD estimates, which display no visual pretrends, are statistically significant). Nonetheless, the antidepressant pretrend raises concerns that some switching is endogenously related to the old-new provider prescribing difference. To further evaluate the switching estimates, I identify and estimate event studies for switches that occur after children move HRRs or after their provider moves HRRs – two situations where switches are clearly unrelated to systematic shocks to child health. If the pretrends in the first switch antidepressant event study indicate that estimated jump is driven by endogenous switching, there should be no jump in the event study coefficients estimated from exogenous switches. Reassuringly, as Figure 6 shows, the exogenous switching event studies exhibit a similar jump as estimated in Figure 5.

5.3 Summary

Together, the moving and switching event studies suggest that differences in local practice environments and provider prescribing intensities play a substantial role in driving geographic and provider level variation in pediatric mental health prescribing. As children and providers move across regions and when children switch providers, their prescribing clearly shifts towards the average prescribing in their destination or of their new provider. However, the event studies are illustrative; they do not quantify the share of variation explained by each component. For example, the child mover event study captures the joint contributions of differences in local practice environments and average provider intensities to geographic variation, whereas the provider mover event study captures the joint contributions of differences in local practice environments and average child health. Additionally, the mover event studies estimate mover-weighted effects, and the main switching event study focused only on each switchers' first switch. Applying the AKM-style model to the full sample provides a framework to combine these sources of variation and estimate component shares that reflect the overall distribution of children,

[20] In Figure A5, I present results from estimating the switching event studies separately by the timing of treatment (i.e. switch year), in the spirit of Callaway and Sant'Anna (2021); Sun and Abraham (2021). I find no evidence of meaningfully heterogeneous treatment effects by year.

regions, and providers.

6 Results

6.1 Geographic variation estimates

Figure 7 maps the HRR practice environment effects and HRR average provider prescribing intensities from estimating Equation 1. ADHD practice environment effects tend to be higher in the Southeast and lowest in the West. Antidepressant practice environments tend to be higher in the Mid-Atlantic, Midwest, and Mountain West, and lower in the South. Average provider intensities are more geographically dispersed, but there are pockets of high intensity ADHD providers in the Northeast and on the West Coast, and high intensity antidepressant providers in Appalachia and along the Gulf Coast. Notably, the high ADHD prescribing rates in HRRs in Louisiana are clearly driven more by practice environment than (on-average) high-intensity providers, whereas the high antidepressant prescribing rates in the Plains states are driven by a mix of practice environment effects and providers. These comparisons provide some intuition for the results of the quantitative geographic decomposition.

I use the model estimates to decompose the difference in pediatric mental health prescribing between high and low prescribing regions. For ADHD medication and antidepressants separately, I compare groups of HRRs above and below the median prescribing rate, in the top and bottom quartile, and the top and bottom decile. For each comparison, I report the observed difference in prescribing rates, and the estimated share of the difference attributable to average provider prescribing intensity (Equation 3), regional practice environment effects (Equation 2), and average child health and demand (Equation 4).

Table 2 presents the geographic decompositions.[21] For ADHD, the average provider share is negative across all three comparisons, ranging from -4.78 percent for above/below median HRRs to -9.36 percent for the top/bottom deciles. This implies that higher ADHD prescribing regions tend to have slightly lower ADHD prescribing intensity providers, although the differences in average prescribing intensity across regions are small. In contrast, differences in local practice environments explain approximately 50 percent of the

[21] Bootstrapped standard errors clustered at the child level are reported in parentheses. I use the Bayesian bootstrap (Rubin 1981; Badinski et al. 2023). Specifically, for each child in the sample, I draw 50 weights from a Dirichlet distribution. I repeat the AKM model estimation 50 times, weighting each child-year observation by the child's associated Dirichlet weight. The reported standard errors are the standard deviations of the resulting bootstrap estimates.

geographic variation in ADHD prescribing. Together, I find that supply-side factors explain roughly 45 percent of the geographic variation in ADHD prescribing rates across HRRs. Demand-side factors, including estimated average child health and demand and observable characteristics (the age bins and time trends) explain the remaining 55 percent of the variation. For antidepressants, higher prescribing rate areas do tend to have higher prescribing intensity providers. Nonetheless, the average provider share is only half the practice environment share (20 percent vs. 40 percent). The supply-side factors explain a total of 60 percent of the differences in antidepressant prescribing between high and low rate regions. The impact of observables (the $X\beta$ share) on geographic variation is larger for antidepressants than ADHD medication, reflecting the age gradient in antidepressant prescribing and the secular rise in antidepressant prescribing over the sample period captured by the calendar year fixed effects.

The supply-side factors together explain between 40 to 60 percent of the geographic variation in pediatric mental health prescribing, comparable to the overall supply-side share of geographic variation in Medicare utilization estimated in Finkelstein, Gentzkow and Williams (2016). However, the supply side share in pediatric mental health prescribing is comprised primarily of differences in local practice environments. Differences in average PCP prescribing intensities play only a small role (for ADHD, they actually mitigate geographic variation). This result contrasts sharply with the findings from Badinski et al. (2023), who estimate that average physician practice intensities are roughly three times as important as other supply side factors in explaining geographic variation in Medicare utilization. The results highlight that the drivers of geographic variation are context dependent and emphasize the importance of quantifying the components across various settings. Indeed, even in the same context of pediatric mental health prescribing for the same sample of children and providers, the drivers of geographic variation are different across conditions.

Although differences in average PCP prescribing intensity explain, at most, a small share of geographic variation in pediatric mental health prescribing, the interpretation is not that providers do not matter. As described in Section 3, the practice environment share not only captures differences in cultural and environmental factors that affect mental health and demand for pharmaceutical treatment, but also geographic heterogeneity in mental health specialist supply and average specialist practice intensity. Furthermore, the vast majority of variation in prescribing occurs within regions across providers and is not addressed by the geographic decomposition. In the following section, I present provider-level decompositions that emphasize the important role PCPs play in driving pediatric mental health treatment variation.

6.2 Provider-level results

The variance of prescribing rates across providers is over 15 times the magnitude of the variance across HRRs for ADHD prescribing and over 36 times the magnitude for antidepressant prescribing. Provider prescribing rates vary due to differences in the health and demand of the children they see, the provider's prescribing intensity, and their practice environment. Using the model estimates, I decompose the variation in provider prescribing rates into these components, as derived in Equation 6.

Table 3 presents the provider decomposition results; the first row reports the variance of provider prescribing rates, while the remainder report the share of the variance explained by each component.[22] For ADHD, variation in provider prescribing intensities explains over 30 percent of provider-level variation in prescribing rates, while variation in child health and demand accounts for almost 45 percent. The shares are reversed for antidepressants; provider intensities explain 45 percent of provider-level antidepressant rate variation, while child effects account for approximately 30 percent. HRR practice environment effects explain little of the variation (2 to 3.5 percent), consistent with the importance of within region variation relative to across region differences. Differences in child observables ($X\beta$) across providers also explain little of the variation.

For both ADHD medication and antidepressants, there is assortative matching between providers and children. 20 percent of the variation in provider prescribing rates is driven by higher intensity providers seeing children with more mental health needs and/or higher demand for pharmaceutical treatment relative to lower intensity providers.[23] All together, the results imply that if providers all had the same prescribing intensity, variation in prescribing rates of ADHD medication and antidepressants across providers would be reduced by 50 percent and 65 percent, respectively.

These findings only provide evidence on relative prescribing intensities, they do not identify the "optimal" provider prescribing intensity. In Section 7, I present patterns of associations between provider prescribing intensities and provider characteristics that suggest that higher-quality PCPs may have relatively lower antidepressant prescribing intensities and relatively higher ADHD stimulant prescribing intensities.

[22] As in Table 2, bootstrapped standard errors clustered at the child level are reported in parentheses.

[23] Recall from Section 3 that the model explicitly allows and accounts for assortative matching at baseline, but assumes that switches are not driven by resorting on shocks.

Comparing the provider prescribing intensities across conditions, I find that provider practice styles exhibit variation within providers. Figure 8 presents the joint distribution of provider prescribing intensities for ADHD medication and antidepressants. I divide providers into 25 cells by quintiles of their estimated prescribing intensities, and report the share of providers within each cell. As providers are divided by quintile in each dimension, summing across a given row or column contains 20 percent of providers. Cells along the diagonal have the highest share of providers, indicating that providers' prescribing intensities are generally correlated across conditions. However, there is a nontrivial share of providers with opposite relative prescribing intensities across conditions. 24 percent of providers (over 31,000 providers) are in the top two quintiles of prescribing intensity for one condition but in the bottom two quintiles for the other. For comparison, 20 percent of providers are in the top two quintiles or bottom two quintiles of prescribing intensity for both. That a given provider's prescribing intensities display meaningful variation across conditions just within the context of pediatric mental health emphasizes the importance of studying practice style more broadly as a multi-dimensional object.

6.3 Model validation

I conduct a series of additional tests to further validate the model estimates. To test the additive structure of the model, I implement an event-study of switching providers that classifies providers based on the quartile of their estimated prescribing intensity, in the spirit of the analysis in Card, Heining and Kline (2013). Using the first switch sample, in Figure 9 I show the average prescribing rate by the year relative to the switch for children leaving quartile 1 providers (with the lowest prescribing intensity) and quartile 4 providers (with the highest prescribing intensity). Prescribing is stable before and after the switch, and the switching effects are remarkably symmetric when a child switches up or switches down. To test the assumption of separability between children and providers, I estimate a saturated model that includes a separate dummy for each child-provider-HRR combination. If match effects are important, the saturated model should fit the data much better than the baseline model. I find that the fit does improve, but the reduction in root mean squared error is small, from roughly 0.12 to 0.08. Finally, I re-estimate the model on the full BCBS sample, including child-years matched to small providers by adding an HRR-specific small provider fixed effect. The geographic and provider variation decomposition results remain qualitatively the same, as shown in Table A3 and Table A4.

7 Explaining differences in practice environments and prescribing intensities

Given that regional practice environments and provider prescribing intensities play such a key role in driving variation in pediatric mental health prescribing, it is natural to ask what causes places and providers to have different effects on pediatric mental health prescribing. As a first pass, I explore how HRR and provider characteristics are associated with place and provider effects.

7.1 Correlates of HRR practice environment

Figure 10 shows the correlation between the estimated HRR practice environment effects and HRR characteristics. Each bolded row represents the coefficients (or sets of coefficients, in the case of the *Region* indicators) and 95 percent cofidence intervals from separately regressing the ADHD and antidepressant HRR fixed effects on that characteristic, weighting HRRs by their associated number of child-years.[24] The regional differences mirror the maps in Figure 7; The Midwest has the highest practice environment effects for antidepressant prescribing (and relatively high effects for ADHD), while the South has the highest practice environment effects (but relatively low antidepressant practice environments effects). The West has lower intensity practice environments across the board.

The first HRR characteristic I consider is an indicator for whether the HRR's state has a special education financing system that allocates funding based on the number of students receiving special education services. This funding mechanism incentivizes schools to classify additional students as requiring special education services, often through ADHD diagnoses (Morrill 2018). HRRs in states with special education finance incentives have higher ADHD practice environments, increasing average ADHD stimulant prescribing rates in those HRRs by 0.27 percentage points, or 5.9 percent of the overall prescribing rate.[25] However, the difference in antidepressant practice environment effects is weaker and statistically indistinguishable from 0. These results support the conclusion from previous work that state special education finance incentives affect whether children are treated for ADHD, and provide evidence that the estimated practice environments reflect

[24] The 95 percent confidence intervals are constructed from two-step bootstrapped standard errors clustered at the child-level.

[25] The standard deviation of HRR ADHD fixed effects is 1.39. 1.39 multiplied by 0.2 is 0.27. The overall ADHD prescribing rate of 4.57 percent.

underlying causes.

The next three regressions analyze how mental health specialist supply and PCP competition are associated practice environments. I use the BCBS claims to construct measures of mental health specialists per child, the share of those specialists who are psychiatrists, and PCPs per child.[26] An advantage of constructing these measures from the claims is that these are the providers observed treating children with BCBS insurance, so the measures better reflect the true availability for a child in the sample. I standardize these characteristics in the regressions, so coefficients report the association with a one standard deviation change in the measure. While the density of available mental specialists per child is uncorrelated with the practice environment effects, HRRs with a higher share of psychiatrists have lower practice environment intensities for both ADHD medication and antidepressants. Although this finding is perhaps unintuitive, it is consistent with shortages of psychiatrists leading PCPs to take a larger role in mental health prescribing with resulting looser prescribing. Finally, there is little evidence of competition between PCPs driving prescribing effects at the HRR level.

The final two characteristics relate the pediatric mental health HRR practice environment intensities to the most commonly studied geographic measure of healthcare - Medicare spending per beneficiary (which are similarly standardized across HRRs).[27] Total Medicare spending is positively associated with ADHD stimulant practice environment effects but negatively associated with antidepressant practice environment intensity; the magnitudes are small but the associations are statistically significant. Perhaps more relevant is Medicare physician spending, as interventions targeting high-spending Medicare physicians are the most likely to spill over to pediatric mental health prescribing. Both conditions' practice environment intensities are negatively correlated with Medicare physician spending per beneficiary, indicating that any spillovers would reduce pediatric mental health prescribing in already low-intensity areas.

[26] For each HRR in each year, I compute the ratio of the number of mental health specialists who see a BCBS-covered child residing in that HRR to the number of BCBS-covered children in that HRR, and use the average number of specialists per child across the years in the sample. I also divide these specialists between psychiatrists and other mental health specialists (e.g. psychologists, therapists, and mental health social workers), and compute the average psychiatrist share. Finally I count the average number of PCPs in the analysis sample who see a child residing in that HRR.

[27] Medicare spending data is drawn from the Dartmouth Atlas (Dartmouth Atlas of Health Care 2024). I use price-adjusted measures for 2011 through 2019.

7.2 Correlates of provider prescribing intensity

Figure 11 presents parallel correlations for provider prescribing intensities. I show results for associations with provider gender and specialty. I also show results for experience and medical school training for physicians in the sample – using data on physician training provided by IQVIA.[28] Female providers tend to be higher intensity for antidepressant prescribing relative to male providers, but lower intensity for ADHD. On average, family medicine physicians, internal medicine physicians, and NPs and PAs have similar practice intensities (within a tenth of a standard deviation). However, pediatricians, who see just over 50 percent of the child-years in the sample, tend to have higher ADHD prescribing intensity and over 0.3 standard deviations lower antidepressant prescribing intensity. For antidepressants, this difference in intensities translates to a 1.5 percentage point lower antidepressant prescribing rate for pediatricians relative to family medicine physicians (holding fixed patient and practice environment characteristics), over 20 percent of the average antidepressant prescribing rate across providers. That pediatricians – who specialize in treating children – have such different prescribing intensities suggests that PCPs trained in other specialties may be prescribing sub-optimally. In this case, the results suggest that non-pediatricians over-prescribe antidepressants but under-prescribe ADHD medication.

The patterns of prescribing intensities by physician experience and medical school rank are consistent with the theory that relatively lower antidepressant prescribing intensity and higher ADHD prescribing intensity represent optimal provider behavior. More experienced physicians and physicians who attended higher ranked medical schools tend to have lower antidepressant prescribing intensity and higher ADHD prescribing intensity. These measures are only available for the subset of providers matched from the sample to the IQVIA data (see Table A5). Nonetheless, to the extent that experience and medical school rank proxy for physician quality, these results suggest higher-quality physicians are less likely to prescribe antidepressants and more likely to prescribe ADHD medication.

Finally, I compare physicians trained at medical schools in the United States to those trained at medical schools outside the United States. US-trained physicians have higher prescribing intensities for both ADHD medication and antidepressants, perhaps reflect-

[28] Provider gender and specialty are measured in the NPPES. Provider experience and medical school information are constructed using data provided by IQVIA. Provider experience is defined as years since medical school graduation in 2023. Medical school ranks are based on the 2023 US News & World Report ranking. Table A5 shows the distribution of providers across the various characteristics.

ing differences in cultural stigma surrounding mental health treatment.

8 Conclusion

I use health insurance claim data for a large, national sample of children to quantify the drivers of variation in pediatric mental health prescribing. Childhood mental health care is a pressing area of focus for healthcare practitioners and policymakers due to the effects of mental health issues in childhood on long-term health and labor market outcomes. There is substantial heterogeneity in rates of mental health diagnoses and treatment across regions and providers that has raised concerns of some children going untreated while others are improperly or unnecessarily prescribed mental health drugs. However, determining whether differences in treatment rates reflect differences in the supply of mental health care (and if so, what components of supply?) or differences in children's underlying mental health needs and demand for care is difficult because the latter is not easily observable.

I leverage the migration of children and providers across regions and the switching of children across primary care providers to separately identify the roles of PCP prescribing intensities, regional practice environments (non-provider-specific supply side factors), and child health and demand in driving variation in pediatric ADHD medication and antidepressant prescribing. I find that differences in average PCP prescribing intensities across regions explain little of the geographic variation in prescribing, while differences in practice environments explain 40-50 percent. However, I show that the geographic decomposition understates the role of providers in driving heterogeneity in mental health prescribing. Provider-level variation in prescribing rates dwarfs geographic variation, implying that the geographic decomposition does not address the bulk of the overall heterogeneity in treatment. Decomposing variation at the provider-level, I estimate that 35 percent of the variation in provider prescribing rates of ADHD medication and 45 percent of the antidepressant prescribing variation is due to differences in the intensity with which providers treat otherwise similar patients. Another 20 percent of the variation for both drug classes is driven by the assortative matching of higher intensity providers and higher demand (and/or treatment need) children.

The importance of differences in practice environments relative to differences in average provider practice styles in explaining geographic variation in pediatric mental health prescribing contrasts with the results from Badinski et al. (2023), who show that differences in average provider practice styles are three times as important as practice environments in explaining geographic variation in Medicare utilization. The diverging results are not contradictory, but rather emphasize that the drivers of variation in healthcare are context-specific. My results also highlight the importance of going beyond geography in studying variation in medical care. Provider prescribing intensities drive a large share of mental health treatment heterogeneity across otherwise similar children, but the role of differences across providers is not evident when studying regional average outcomes. Additionally, I show that a substantial share of providers have non-uniform relative prescribing intensities across conditions, in contrast to the standard uni-dimensional characterization of providers as either high-intensity "cowboys" or low-intensity "comforters."

The variance decompositions alone do not provide strong conclusions about optimal mental health prescribing. The model estimates *relative* provider prescribing intensities, but does not identify the "correct" provider treatment practices. Nonetheless, the patterns of prescribing intensities by provider specialty, experience, and training suggest some normative implications. Pediatricians, more experienced physicians, and physicians trained at top-ranked medical schools tend to have higher ADHD prescribing intensities and lower antidepressant prescribing intensities than other providers. To the extent that these measures proxy for provider quality, interventions that shift provider practice styles in these directions could be beneficial for children's mental health treatment. However, I emphasize that these findings are only suggestive. A crucial direction for future research is to evaluate the impact of provider pediatric mental health practice styles on children's outcomes, and provide definitive evidence for optimal mental health treatment practices.

References

Abowd, John M., Francis Kramarz, and David N. Margolis. 1999. "High Wage Workers and High Wage Firms." *Econometrica*, 67(2): 251–333. Publisher: John Wiley & Sons, Ltd.

Ahammer, Alexander, and Thomas Schober. 2020. "Exploring variations in health-care expenditures—What is the role of practice styles?" *Health Economics*, 29(6): 683–699. Publisher: John Wiley & Sons, Ltd.

Andrews, M. J., L. Gill, T. Schank, and R. Upward. 2008. "High Wage Workers and Low Wage Firms: Negative Assortative Matching or Limited Mobility Bias?" *Journal of the Royal Statistical Society. Series A (Statistics in Society)*, 171(3): 673–697.

Andrews, M.J., L. Gill, T. Schank, and R. Upward. 2012. "High wage workers match with high wage firms: Clear evidence of the effects of limited mobility bias." *Economics Letters*, 117(3): 824–827.

Badinski, Ivan, Amy Finkelstein, Matthew Gentzkow, and Peter Hull. 2023. "Geographic Variation in Healthcare Utilization:The Role of Physicians."

Baicker, Katherine, and Amitabh Chandra. 2004. "Medicare Spending, The Physician Workforce, And Beneficiaries' Quality Of Care." *Health Affairs*, 23(Suppl1): W4–184.

Berndt, Ernst R., Robert S. Gibbons, Anton Kolotilin, and Anna Levine Taub. 2015. "The heterogeneity of concentrated prescribing behavior: Theory and evidence from antipsychotics." *Journal of Health Economics*, 40: 26–39.

Blue Cross Blue Shield. 2024. "The Blue Cross Blue Shield System." https://www.bcbs.com/about-us/the-blue-cross-blue-shield-system.

Bonhomme, Stéphane, Kerstin Holzheu, Thibaut Lamadon, Elena Manresa, Magne Mogstad, and Bradley Setzler. 2023. "How Much Should We Trust Estimates of Firm Effects and Worker Sorting?" *Journal of Labor Economics*, 41(2): 291–322. Publisher: The University of Chicago Press.

Bonhomme, Stéphane, Thibaut Lamadon, and Elena Manresa. 2019. "A Distributional Framework for Matched Employer Employee Data." *Econometrica*, 87(3): 699–739. Publisher: John Wiley & Sons, Ltd.

Bronnenberg, Bart J., Jean-Pierre H. Dubé, and Matthew Gentzkow. 2012. "The Evolution of Brand Preferences: Evidence from Consumer Migration." *American Economic Review*, 102(6): 2472–2508.

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

Card, David, Jörg Heining, and Patrick Kline. 2013. "Workplace Heterogeneity and the Rise of West German Wage Inequality*." *The Quarterly Journal of Economics*, 128(3): 967–1015.

Chan, David C, Matthew Gentzkow, and Chuan Yu. 2022. "Selection with Variation in Diagnostic Skill: Evidence from Radiologists*." *The Quarterly Journal of Economics*, 137(2): 729–783.

Chandra, Amitabh, David Cutler, and Zirui Song. 2011. "Chapter Six - Who Ordered That? The Economics of Treatment Choices in Medical Care." In *Handbook of Health Economics*. Vol. 2, , ed. Mark V. Pauly, Thomas G. Mcguire and Pedro P. Barros, 397–432. Elsevier.

Chan, Jr, David C, and Yiqun Chen. 2022. "The Productivity of Professions: Evidence from the Emergency Department." *National Bureau of Economic Research Working Paper Series*, No. 30608.

Chetty, Raj, John N. Friedman, and Jonah E. Rockoff. 2014. "Measuring the Impacts of Teachers I: Evaluating Bias in Teacher Value-Added Estimates." *American Economic Review*, 104(9): 2593–2632.

Clemens, Jeffrey, Pierre-Thomas Léger, Yashna Nandan, and Robert Town. 2024. "Physician Practice Preferences and Healthcare Expenditures: Evidence from Commercial Payers." *National Bureau of Economic Research Working Paper Series*, No. 33090.

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.

Cuddy, Emily, and Janet Currie. 2020. "Treatment of mental illness in American adolescents varies widely within and across areas." *Proceedings of the National Academy of Sciences*, 117(39): 24039–24046. Publisher: Proceedings of the National Academy of Sciences.

Currie, Janet. 2020. "Child health as human capital." *Health Economics*, 29(4): 452–463. Publisher: John Wiley & Sons, Ltd.

Currie, Janet, and Emily Cuddy. Forthcoming. "Rules vs. Discretion: Treatment of Mental Illness in U.S. Adolescents." *Journal of Political Economy*.

Currie, Janet, and Jonathan Zhang. 2023. "Doing More with Less: Predicting Primary Care Provider Effectiveness." *The Review of Economics and Statistics*, 1–45.

Currie, Janet, and W. Bentley MacLeod. 2017. "Diagnosing Expertise: Human Capital, Decision Making, and Performance among Physicians." *Journal of Labor Economics*, 35(1): 1–43. Publisher: The University of Chicago Press.

Currie, Janet M., and W. Bentley MacLeod. 2020. "Understanding Doctor Decision Making: The Case of Depression Treatment." *Econometrica*, 88(3): 847–878. Publisher: John Wiley & Sons, Ltd.

Currie, Janet, W. Bentley MacLeod, and Jessica Van Parys. 2016. "Practice style and patient health outcomes: The case of heart attacks." *Journal of health economics*, 47: 64–80.

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.

Danielson, Melissa L. 2023. "Trends in Stimulant Prescription Fills Among Commercially Insured Children and Adults — United States, 2016–2021." *MMWR. Morbidity and Mortality Weekly Report*, 72.

Dartmouth Atlas of Health Care. 2024. "Medicare Spending Data." *https://www.dartmouthatlas.org*, Available at: https://www.dartmouthatlas.org/data/topic/spending/.

Ding, Hui. 2023. "Geographic Variation in Mental Health Treatment Utilization: Evidence from Migration."

Doyle, Joseph J., Steven M. Ewer, and Todd H. Wagner. 2010. "Returns to physician human capital: evidence from patients randomized to physician teams." *Journal of Health Economics*, 29(6): 866–882.

Doyle, Jr., Joseph J. 2020. "Physician Characteristics and Patient Survival: Evidence from Physician Availability." *National Bureau of Economic Research Working Paper Series*, No. 27458.

Dwyer, Jennifer B., and Michael H. Bloch. 2019. "Antidepressants for Pediatric Patients." *Current psychiatry*, 18(9): 26–42F.

Epstein, Andrew J., and Sean Nicholson. 2009. "The formation and evolution of physician treatment styles: An application to cesarean sections." *Journal of Health Economics*, 28(6): 1126–1140.

Facchini, Gabriel. 2022. "Forgetting-by-not-doing: The case of surgeons and cesarean sections." *Health Economics*, 31(3): 481–495. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/hec.4460.

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, and Heidi Williams. 2021. "Place-Based Drivers of Mortality: Evidence from Migration." *American Economic Review*, 111(8): 2697–2735.

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 Working Paper Series*, No. 30471.

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

Fisher, Elliott S., David E. Wennberg, Thérèse A. Stukel, Daniel J. Gottlieb, F. L. Lucas, and Etoile L. Pinder. 2003*b*. "The implications of regional variations in Medicare spending. Part 2: health outcomes and satisfaction with care." *Annals of Internal Medicine*, 138(4): 288–298.

Garcia-Argibay, Miguel, Isabell Brikell, Anita Thapar, Paul Lichtenstein, Sebastian Lundström, Ditte Demontis, and Henrik Larsson. 2024. "Attention-Deficit/Hyperactivity Disorder and Major Depressive Disorder: Evidence From Multiple Genetically Informed Designs." *Biological Psychiatry*, 95(5): 444–452.

Gawande, Atul. 2009. "The Cost Conundrum." *The New Yorker*. Section: annals of medicine.

Gowrisankaran, Gautam, Keith Joiner, and Pierre Thomas Léger. 2023. "Physician Practice Style and Healthcare Costs: Evidence from Emergency Departments." *Management Science*, 69(6): 3202–3219. Publisher: INFORMS.

Grytten, Jostein, and Rune Sørensen. 2003. "Practice variation and physician-specific effects." *Journal of Health Economics*, 22(3): 403–418.

Jackson, C. Kirabo. 2018. "What Do Test Scores Miss? The Importance of Teacher Effects on Non–Test Score Outcomes." *Journal of Political Economy*, 126(5): 2072–2107. Publisher: The University of Chicago Press.

Jetty, Anuradha, Stephen Petterson, John M. Westfall, and Yalda Jabbarpour. 2021. "Assessing Primary Care Contributions to Behavioral Health: A Cross-sectional Study Using Medical Expenditure Panel Survey." *Journal of Primary Care & Community Health*, 12.

Jorm, Anthony F., Scott B. Patten, Traolach S. Brugha, and Ramin Mojtabai. 2017. "Has increased provision of treatment reduced the prevalence of common mental disorders? Review of the evidence from four countries." *World Psychiatry*, 16(1): 90–99.

Kline, Patrick, Raffaele Saggio, and Mikkel Sølvsten. 2020. "Leave-Out Estimation of Variance Components." *Econometrica*, 88(5): 1859–1898. Publisher: John Wiley & Sons, Ltd.

Kwok, Jennifer. 2019. "How Do Primary Care Physicians Influence Healthcare? Evidence on Practice Styles and Switching Costs from Medicare."

Laird, Jessica, and Torben Nielsen. 2016. "The Effects of Physician Prescribing Behaviors on Prescription Drug Use and Labor Supply: Evidence from Movers in Denmark."

Marquardt, Kelli. 2021. "Physician Practice Style for Mental Health Conditions: The Case of ADHD."

McBain, Ryan K., Aaron Kofner, Bradley D. Stein, Jonathan H. Cantor, William B. Vogt, and Hao Yu. 2019. "Growth and Distribution of Child Psychiatrists in the United States: 2007–2016." *Pediatrics*, 144(6): e20191576.

Merten, Eva Charlotte, Jan Christopher Cwik, Jürgen Margraf, and Silvia Schneider. 2017. "Overdiagnosis of mental disorders in children and adolescents (in developed countries)." *Child and Adolescent Psychiatry and Mental Health*, 11: 5.

Molitor, David. 2018. "The Evolution of Physician Practice Styles: Evidence from Cardiologist Migration." *American Economic Journal: Economic Policy*, 10(1): 326–56.

Morrill, Melinda Sandler. 2018. "Special Education Financing and ADHD Medications: A Bitter Pill to Swallow." *Journal of Policy Analysis and Management*, 37(2): 384–402. Publisher: John Wiley & Sons, Ltd.

National Center for Health Statistics. 2016. "National Health Interview Survey." *https://www.cdc.gov/nchs/nhis/index.htm*, 2015 Release. Hyattsville, Maryland: U.S. Department of Health and Human Services, Centers for Disease Control and Prevention.

Newhouse, Joseph P., Alan M. Garber, Robin P. Graham, Margaret A. McCoy, Michelle Mancher, and Ashna Kibria, ed. 2013. *Variation in Health Care Spending: Target Decision Making, Not Geography*. Washington (DC):National Academies Press (US).

Parys, Jessica Van. 2016. "Variation in Physician Practice Styles within and across Emergency Departments." *PLOS ONE*, 11(8): e0159882. Publisher: Public Library of Science.

Persson, Petra, Xinyao Qiu, and Maya Rossin-Slater. Forthcoming. "Family Spillover Effects of Marginal Diagnoses: The Case of ADHD." *American Economic Journal: Applied Economics*.

Piper, Brian J., Christy L. Ogden, Olapeju M. Simoyan, Daniel Y. Chung, James F. Caggiano, Stephanie D. Nichols, and Kenneth L. McCall. 2018. "Trends in use of prescription stimulants in the United States and Territories, 2006 to 2016." *PloS One*, 13(11): e0206100.

Richtel, Matt. 2023. "The Surgeon General's New Mission: Adolescent Mental Health." *The New York Times*.

Riglin, Lucy, Beate Leppert, Christina Dardani, Ajay K. Thapar, Frances Rice, Michael C. O'Donovan, George Davey Smith, Evie Stergiakouli, Kate Tilling, and Anita Thapar. 2021. "ADHD and depression: investigating a causal explanation." *Psychological Medicine*, 51(11): 1890–1897.

Rubin, Donald B. 1981. "The Bayesian Bootstrap." *The Annals of Statistics*, 9(1): 130–134. Publisher: Institute of Mathematical Statistics.

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. Publisher: The University of Chicago Press.

Sirovich, Brenda E., Daniel J. Gottlieb, H. Gilbert Welch, and Elliott S. Fisher. 2006. "Regional variations in health care intensity and physician perceptions of quality of care." *Annals of Internal Medicine*, 144(9): 641–649.

Sirovich, Brenda, Patricia M. Gallagher, David E. Wennberg, and Elliott S. Fisher. 2008. "Discretionary Decision Making By Primary Care Physicians And The Cost Of U.S. Health Care." *Health Affairs*, 27(3): 813–823. Publisher: Health Affairs.

Skinner, Jonathan, and Elliott S. Fisher. 1997. "Regional Disparities in Medicare Expenditures: An opportunity for reform." *National Tax Journal*, 50(3): 413–425. Publisher: The University of Chicago Press.

Song, Yunjie, Jonathan Skinner, Julie Bynum, Jason Sutherland, John E. Wennberg, and Elliott S. Fisher. 2010. "Regional Variations in Diagnostic Practices." *New England Journal of Medicine*, 363(1): 45–53.

Sorkin, Isaac. 2018. "Ranking Firms Using Revealed Preference." *The Quarterly Journal of Economics*, 133(3): 1331–1393.

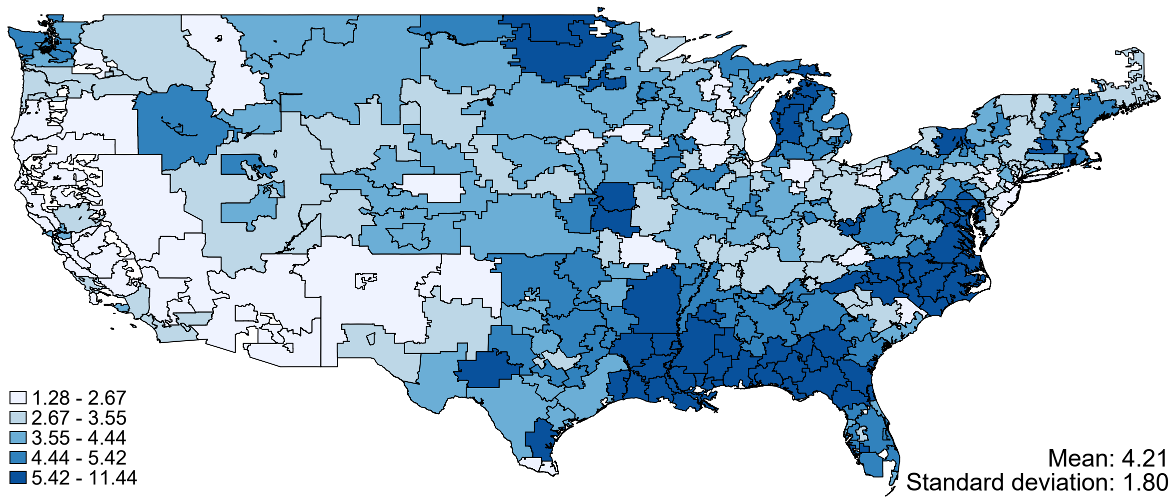
Sun, Liyang, and Sarah Abraham. 2021. "Estimating dynamic treatment effects in event studies with heterogeneous treatment effects." *Journal of Econometrics*, 225(2): 175–199.

Whitney, Daniel G., and Mark D. Peterson. 2019. "US National and State-Level Prevalence of Mental Health Disorders and Disparities of Mental Health Care Use in Children." *JAMA Pediatrics*, 173(4): 389–391.

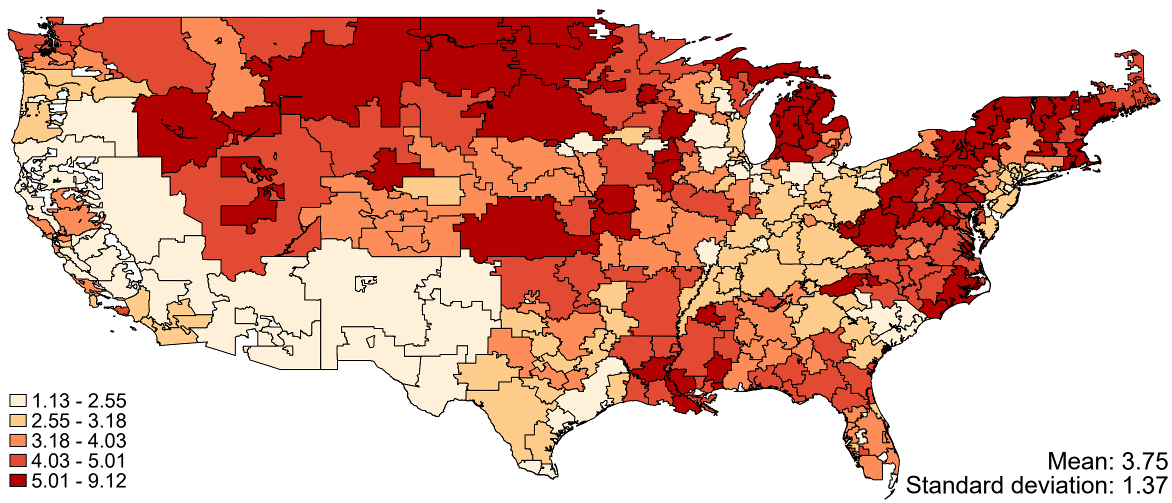
Zeltzer, Dan. 2020. "Gender Homophily in Referral Networks: Consequences for the Medicare Physician Earnings Gap." *American Economic Journal: Applied Economics*, 12(2): 169–97.

Figure 1: Prescribing rates across HRRs

(a) ADHD medication



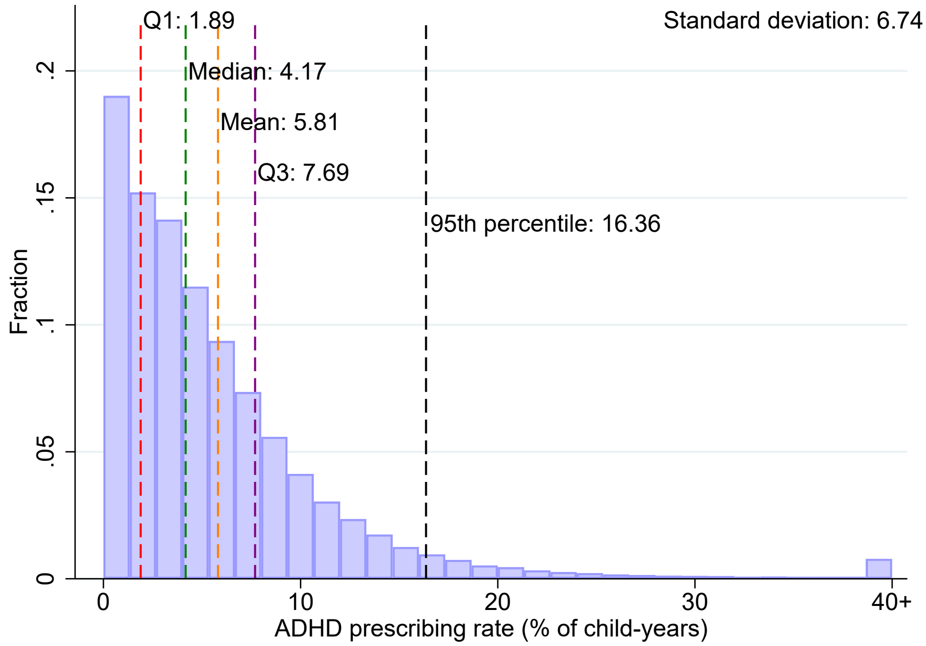
(b) Antidepressants



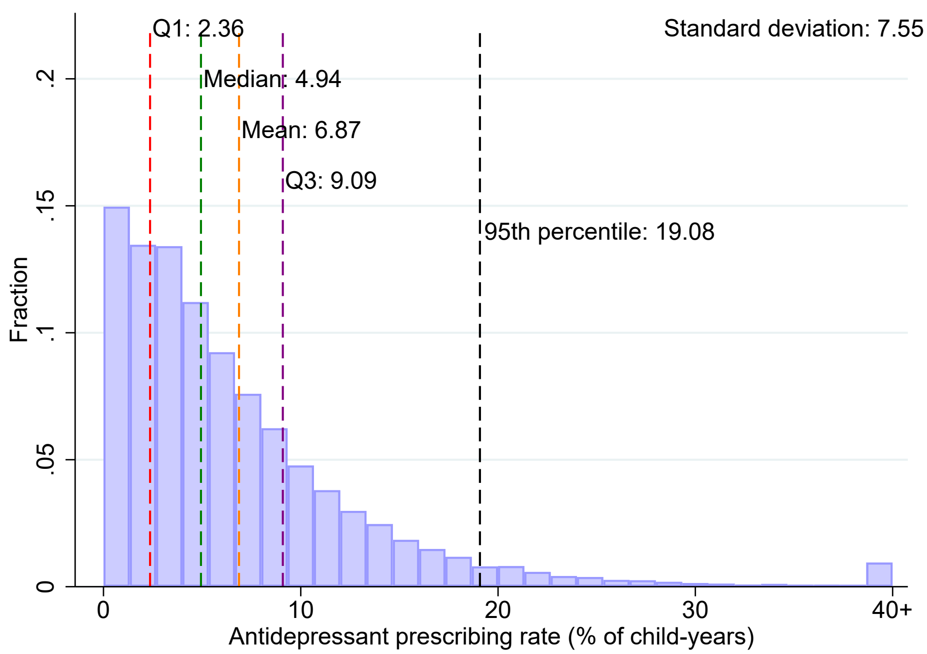
Notes: The figures map the distribution of HRR prescribing rates for ADHD medication (Panel A) and antidepressants (Panel B), shaded by quintile as defined in the legend. Means and standard deviations are calculated across HRRs (unweighted). The sample is the analysis sample, comprising 8,066,999 children observed over 39,188,740 child-years. The Alaska and Hawaii HRRs are not mapped for clarity, but are included in the sample.

Figure 2: Distribution of provider prescribing rates

(a) ADHD medication

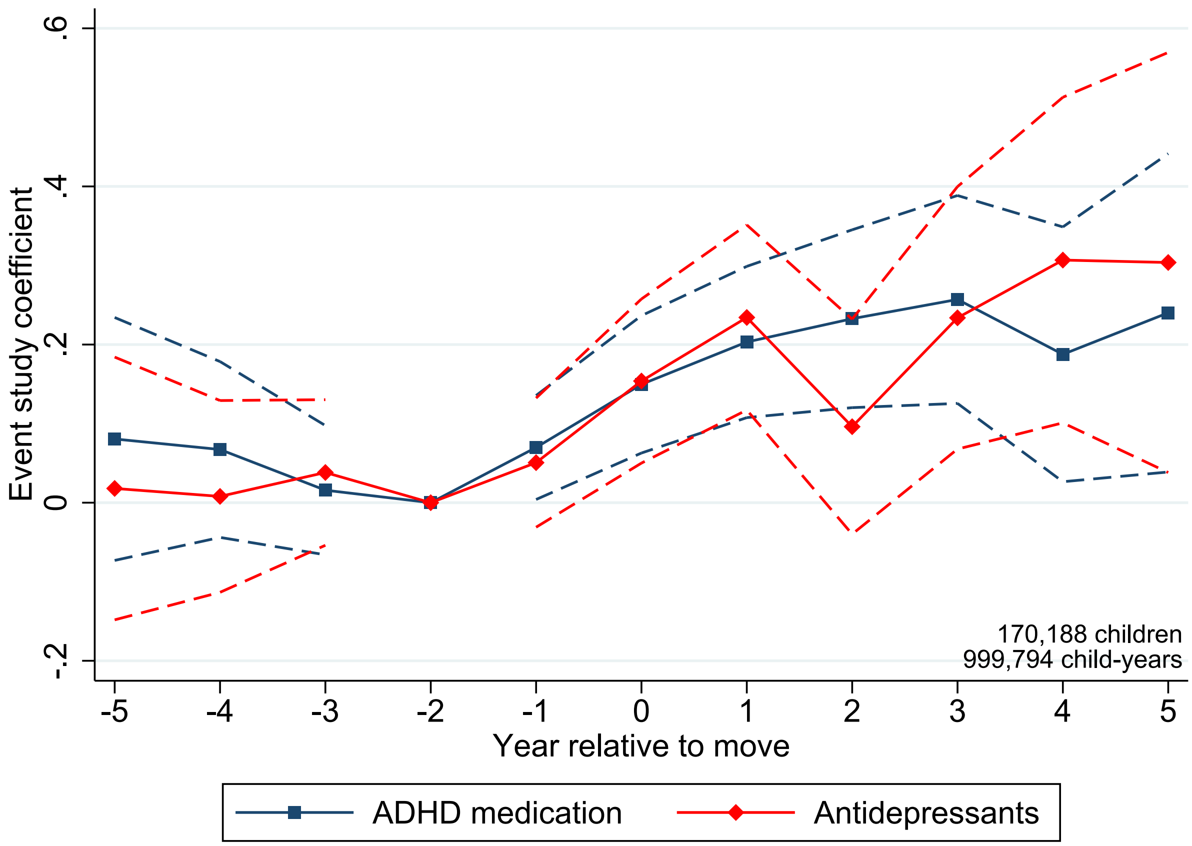


(b) Antidepressants



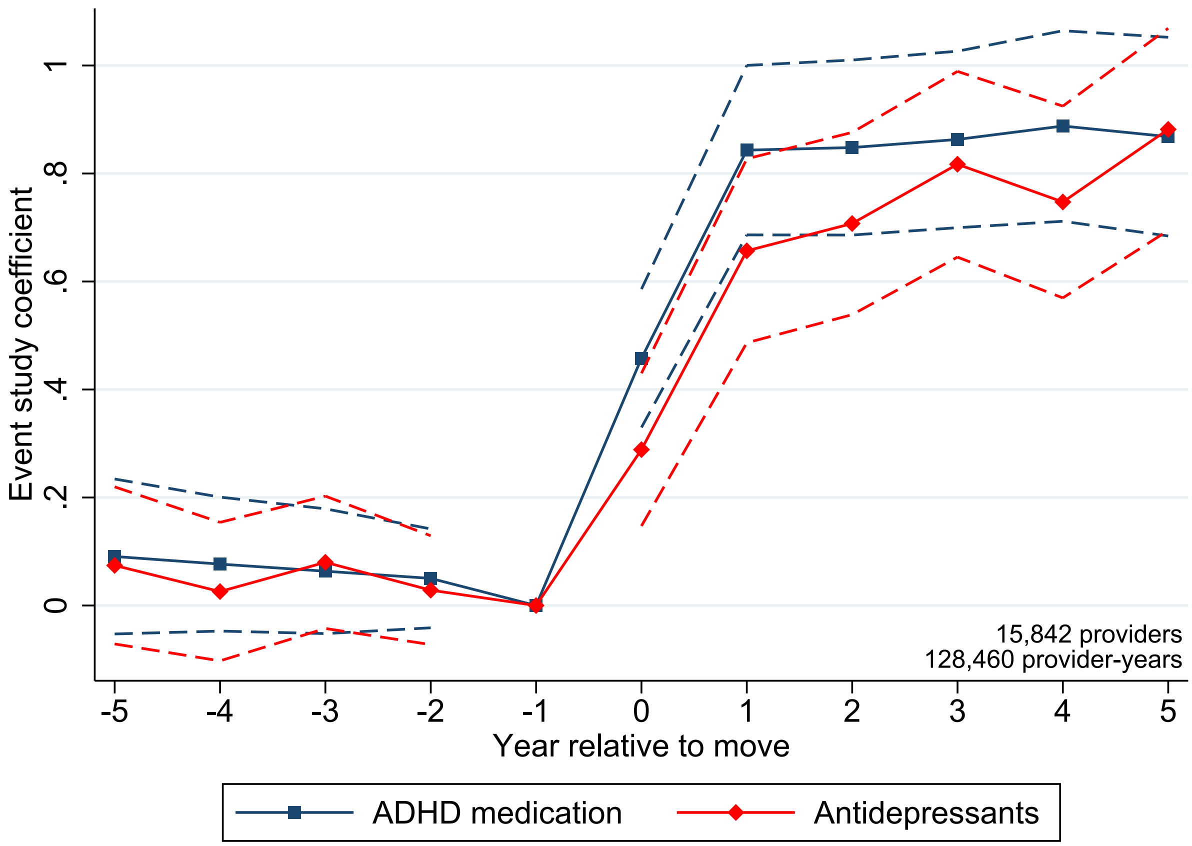
Notes: The histograms show the distribution of provider prescribing rates for ADHD medication (Panel A) and antidepressants (Panel B). The sample is providers in the analysis sample (N = 130,616). Statistics shown on the plots are calculated over the full distribution; providers with prescribing rates of 40 percent and above are grouped together in the 40+ bin for clarity.

Figure 3: Child mover event study



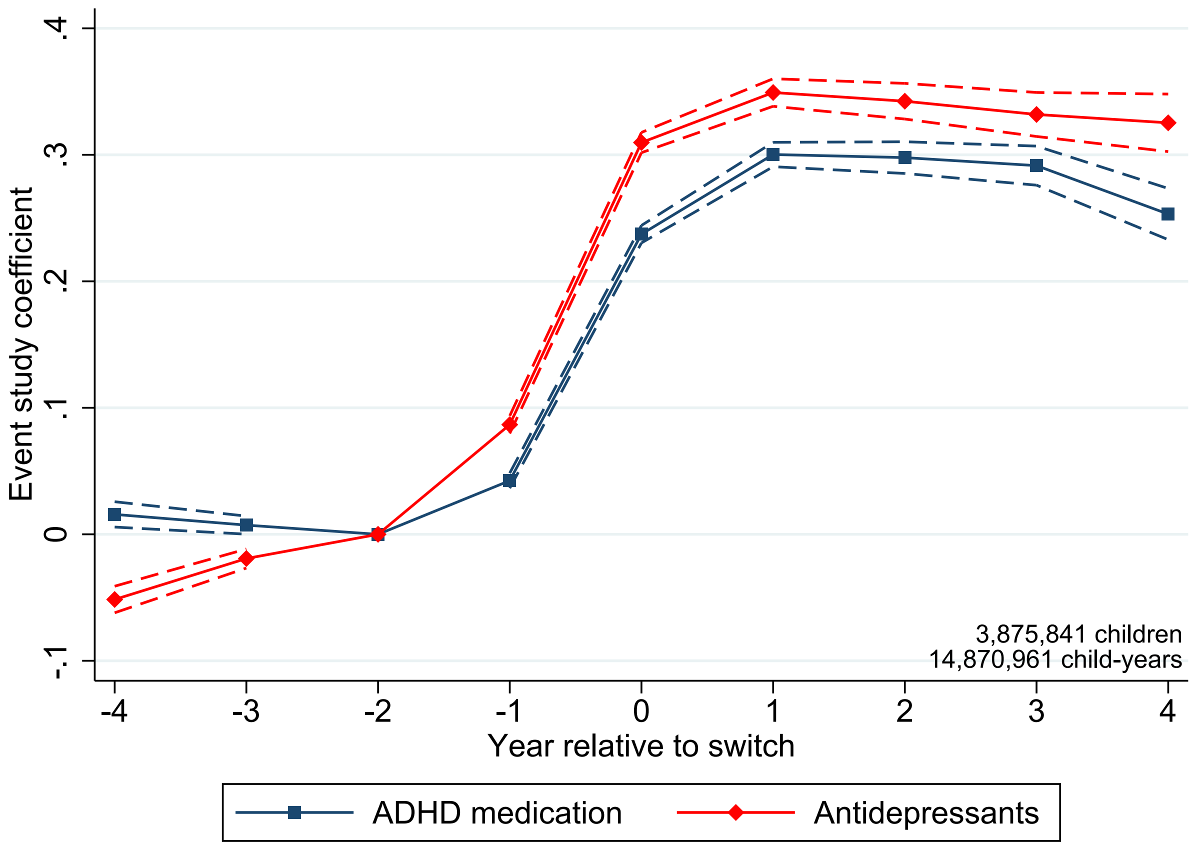
Notes: Figure plots the estimated θ_w coefficients from the child mover event study of Equation 7. The blue squares show estimates with ADHD stimulant prescribing as the dependent variable. Red diamonds show estimates for antidepressant prescribing. Relative year zero is defined as the year a child mover's HRR residence switches from their origin HRR in the previous year to their destination HRR. The coefficients for relative year -2 are normalized to 0 as the exact timing of moves occur in year -1 and year 0; years -1 and 0 are partially treated. Observations before relative-year -5 are binned together with the relative year -5 indicator, and observations after relative-year 5 are binned together with the relative-year 5 indicator. The relative-year coefficients for -5 and 5 show the estimated θ_{wit} on these binned indicators. The regression includes controls for age-bins (5-9, 10-14, 15-21). Dashed lines indicate upper and lower bounds of the 95 percent confidence intervals, calculated using standard errors clustered at the child level. The sample includes child movers from the analysis sample for whom it is possible to determine the year of the move (170,188 children observed over 999,794 child-years).

Figure 4: Provider mover event study



Notes: Figure plots the estimated θ_w coefficients from the provider mover event study of Equation 8. The blue squares show estimates with the provider's ADHD stimulant prescribing rate as the dependent variable. Red diamonds show estimates for the antidepressant prescribing rate. The coefficient for relative year -1 is normalized to 0. observations before relative-year -5 are binned together with the relative year -5 indicator, and observations after relative-year 5 are binned together with the relative-year 5 indicator. The relative-year coefficients for -5 and 5 show the estimated θ_{wit} on these binned indicators. Dashed lines indicate upper and lower bounds of the 95 percent confidence intervals, calculated using standard errors clustered at the provider level. The sample includes provider movers from the analysis sample (15,842 providers observed over 128,460 provider-years).

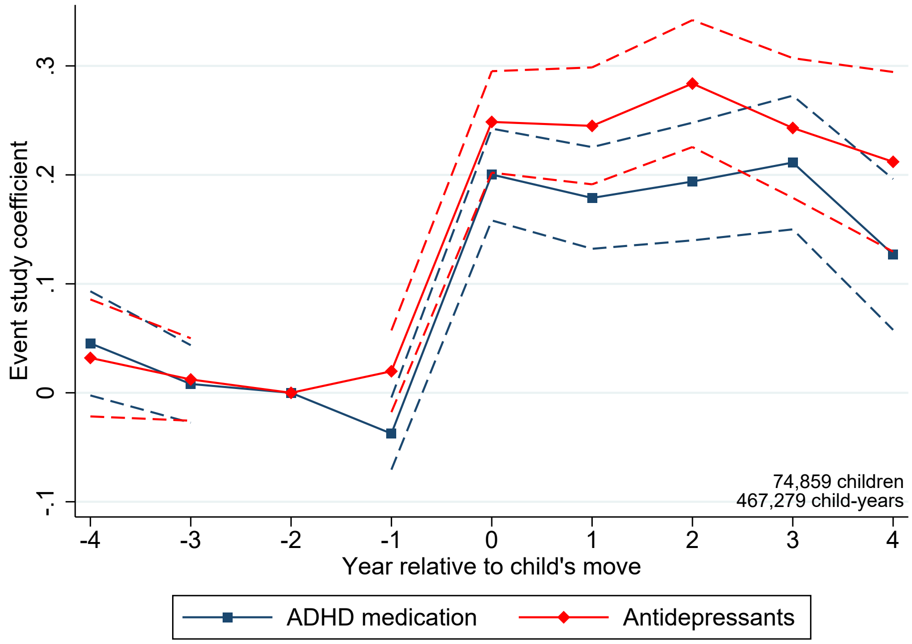
Figure 5: Switching event study - first switch sample



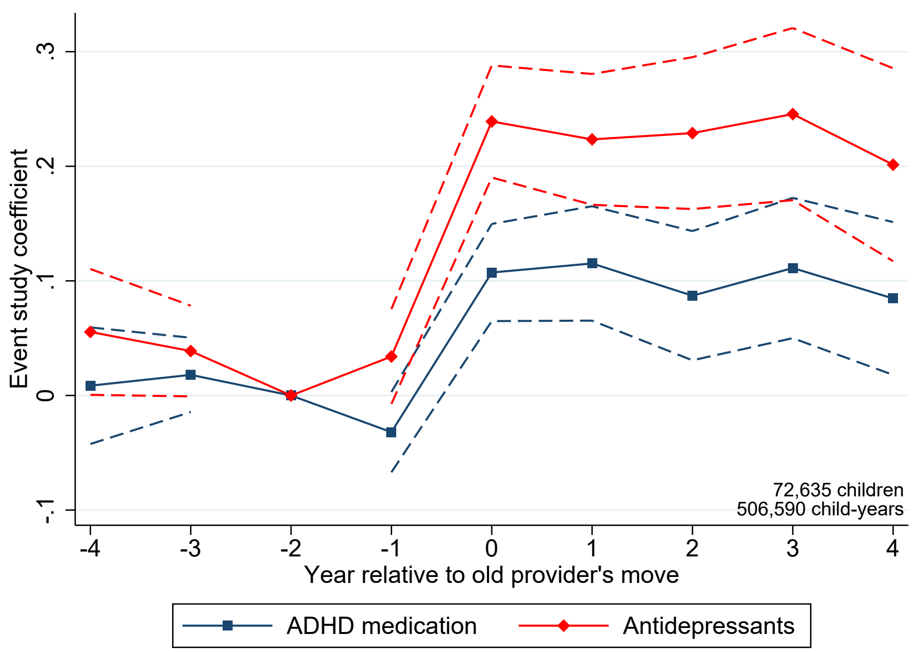
Notes: Figure plots the estimated θ_w coefficients from the switching event study of Equation 9. The blue squares show estimates with ADHD stimulant prescribing as the dependent variable. Red diamonds show estimates for antidepressant prescribing. As switches occur over relative years -1 and 0, the coefficient for relative year -2 is normalized to 0. Regression includes controls for age bins (5-9, 10-14, 15-21). Dashed lines indicate upper and lower bounds of the 95 percent confidence intervals, calculated using standard errors clustered at the child level. A switcher is defined as a child who is matched to multiple primary providers in the estimation sample. The figure sample includes switchers for whom it is possible to determine the year of their first switch, and restricts to the observations prior to any subsequent switches. This sample includes 3,875,841 children observed over 14,870,961 child-years.

Figure 6: Exogenous switching event studies

(a) Child move driven switches

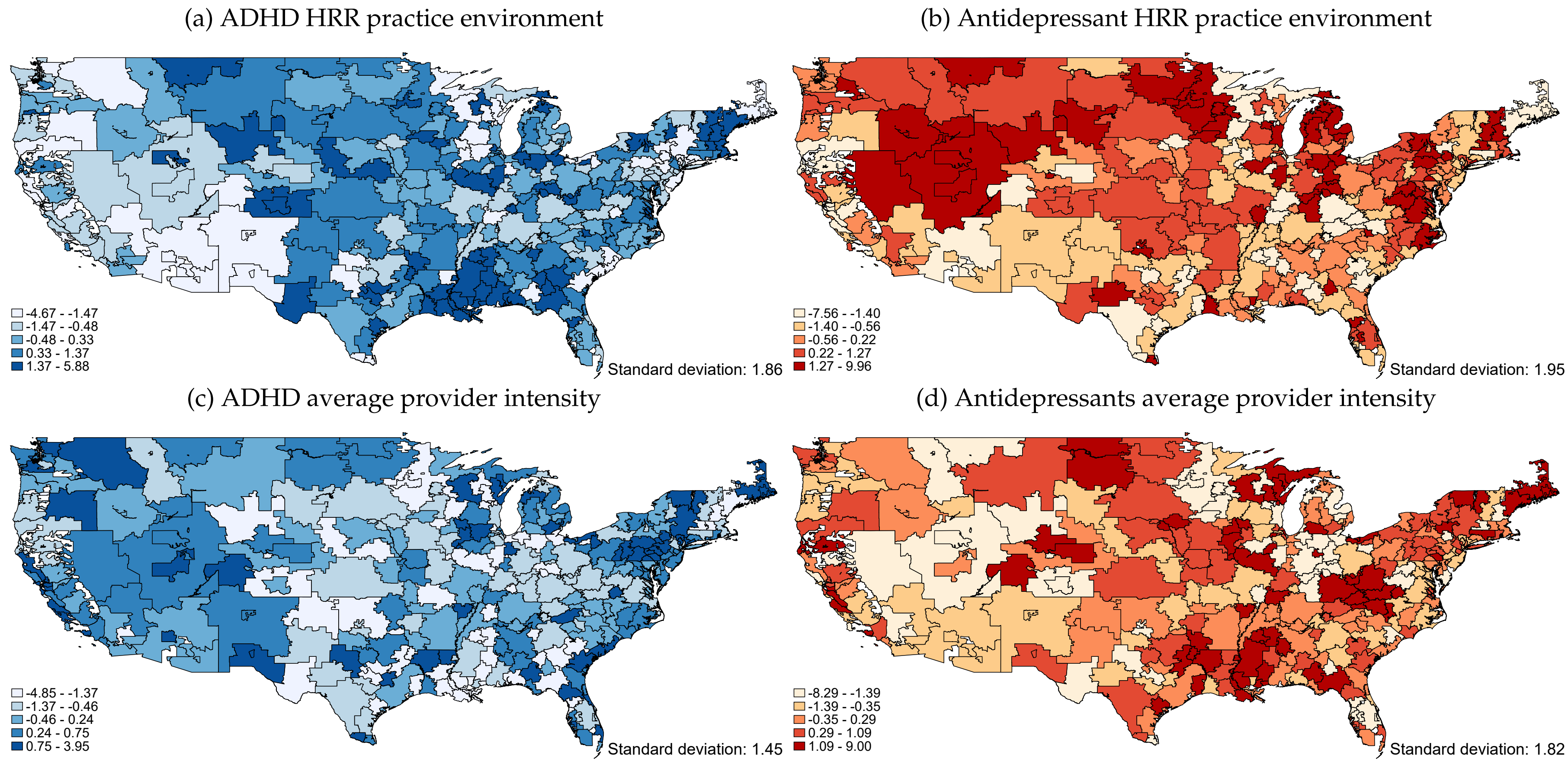


(b) Provider move driven switches



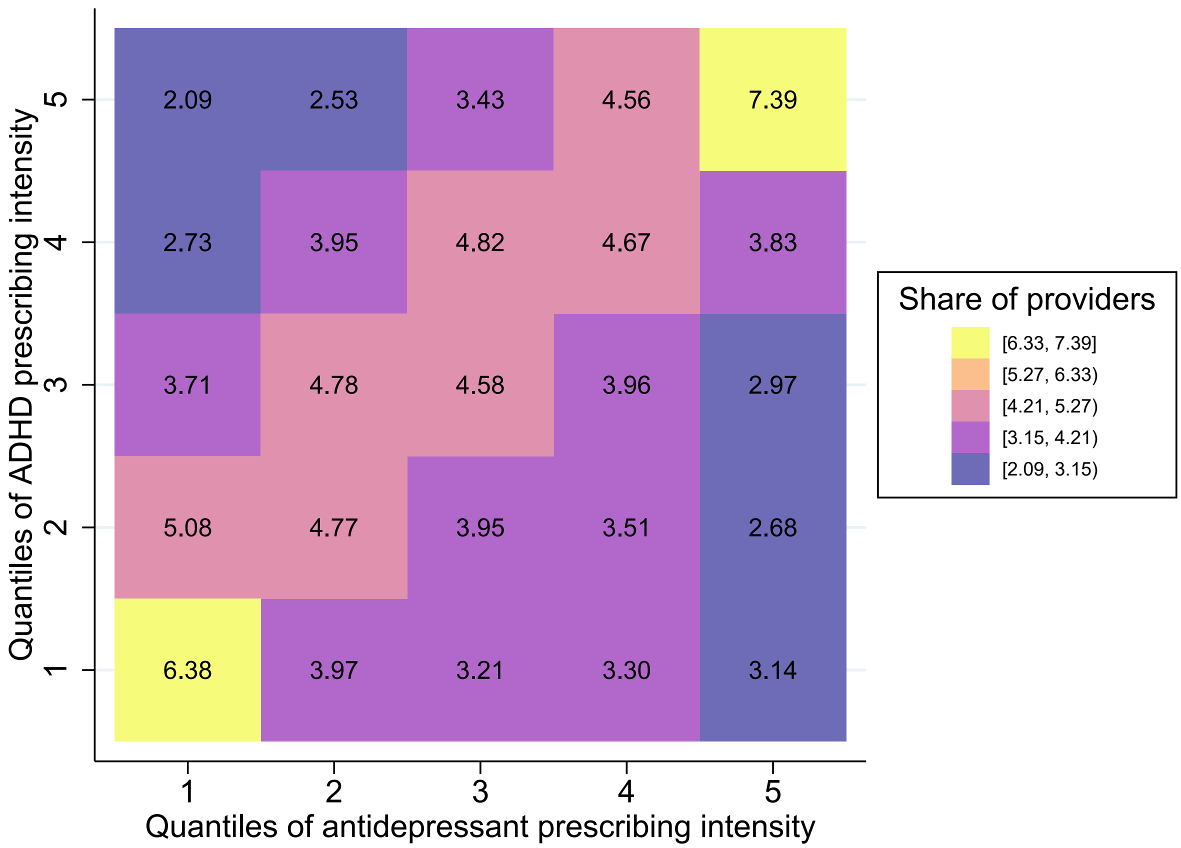
Notes: Figure plots the estimated θ_w coefficients from versions of the switching event study in Equation 9, estimated on subsamples of switchers. Panel A shows results for switches that occurred as a child moved HRRs. A child's "old" provider is defined as the last provider she saw while residing in their origin HRR, while her new provider is defined as the first provider she saw in her destination HRR. Panel B shows results for switches that occurred when a child's provider moved HRRs. I identify children who saw provider movers in the year prior to the provider's move year. For these children, their "old" provider is defined as the mover provider, and their new provider is defined as the next provider they saw. Regression includes controls for age bins (5-9, 10-14, 15-21). For both sets of regressions, observations before relative-year -4 are binned together with the relative year -4 indicator, and observations after relative-year 4 are binned together with the relative-year 4 indicator. The coefficients for year -2 are normalized to 0. Dashed lines indicate upper and lower bounds of the 95 percent confidence intervals, calculated using standard errors clustered at the child level.

Figure 7: Distribution of practice environment and average provider intensities across HRRs



Notes: Panels A and B map the estimated practice environment effect (γ_r, as defined in Equation 1) for ADHD medication and antidepressants, respectively. Panels C and D map the average provider prescribing intensity ($\bar{\delta}_r$) for each drug class. HRRs are shaded by quintiles of the distribution within each panel. The sample is the analysis sample, comprising 8,066,999 children observed over 39,188,740 child-years. The Alaska and Hawaii HRRs are not mapped for clarity, but are included in the sample.

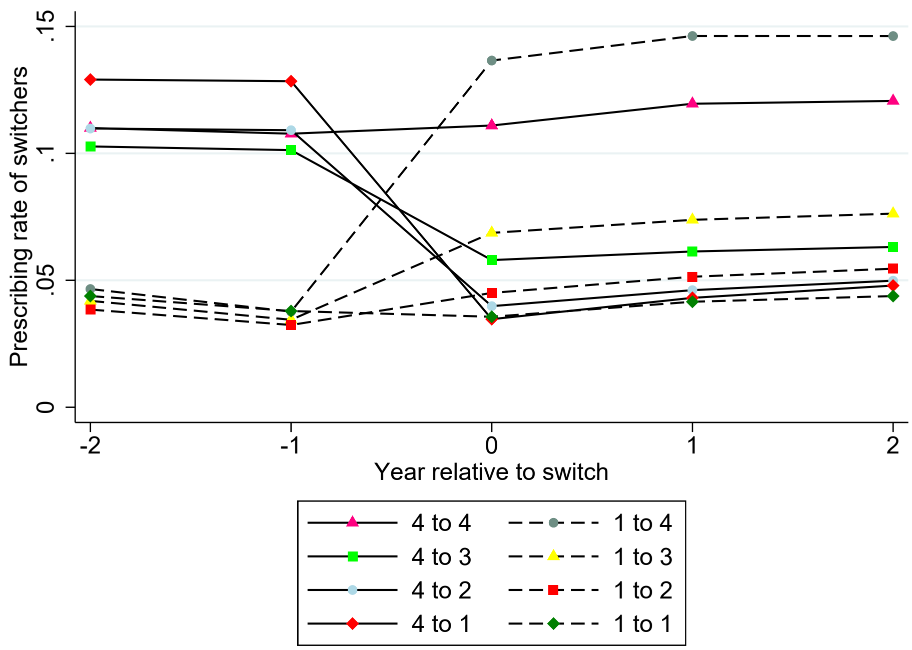
Figure 8: Provider prescribing intensities across conditions



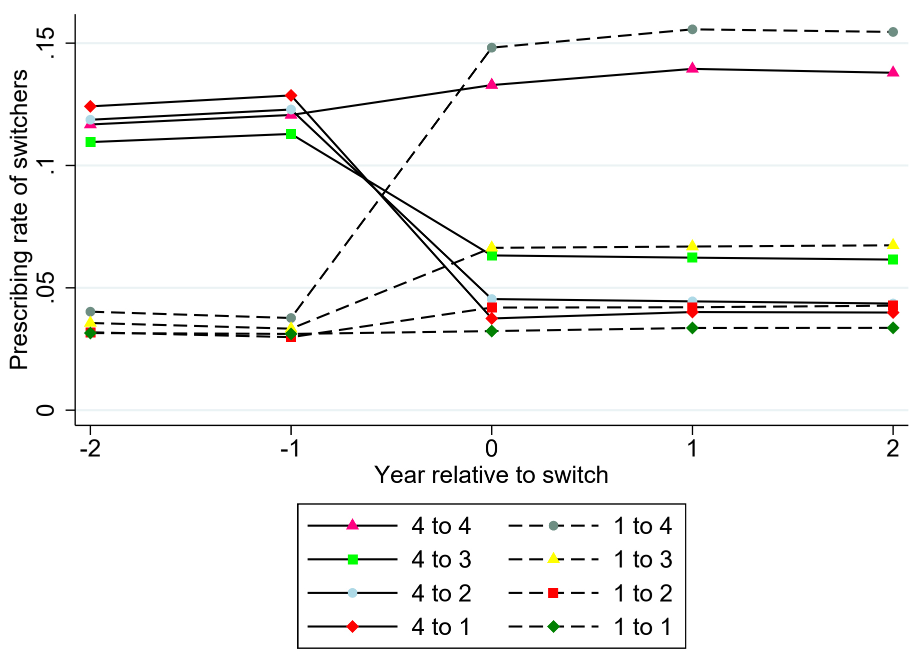
Notes: Figure shows the cross-condition distribution of provider prescribing intensities by quintile. Providers are grouped by quintiles of their ADHD stimulant prescribing intensity on the Y-axis and quintiles of their antidepressant prescribing intensity on the X-axis. Each cell reports the share of providers in the given ADHD and antidepressant quintile; rows and columns each sum up to 20 percent. Cells are shaded per the legend in the figure. The sample is the analysis sample (N providers = 130,616).

Figure 9: Average prescribing of switchers, classified by quartile of prescribing intensity for old and new providers

(a) ADHD medication

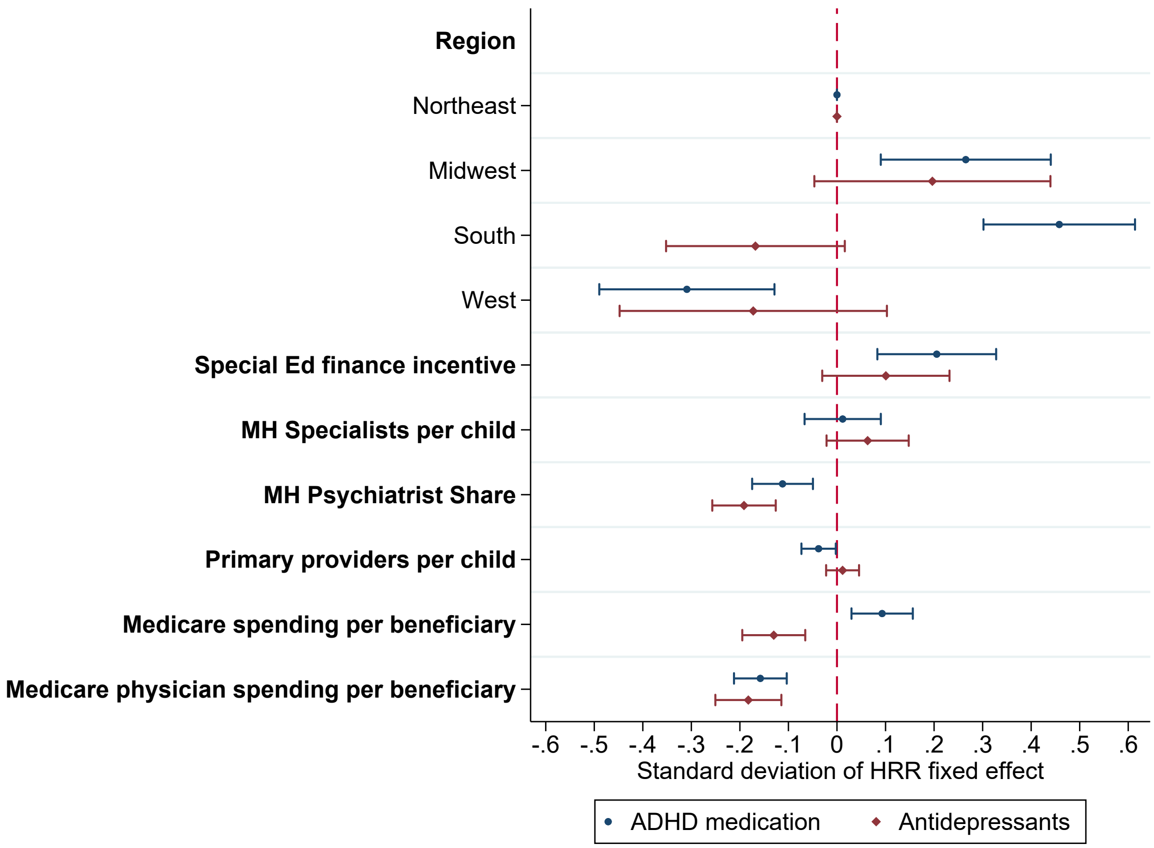


(b) Antidepressants



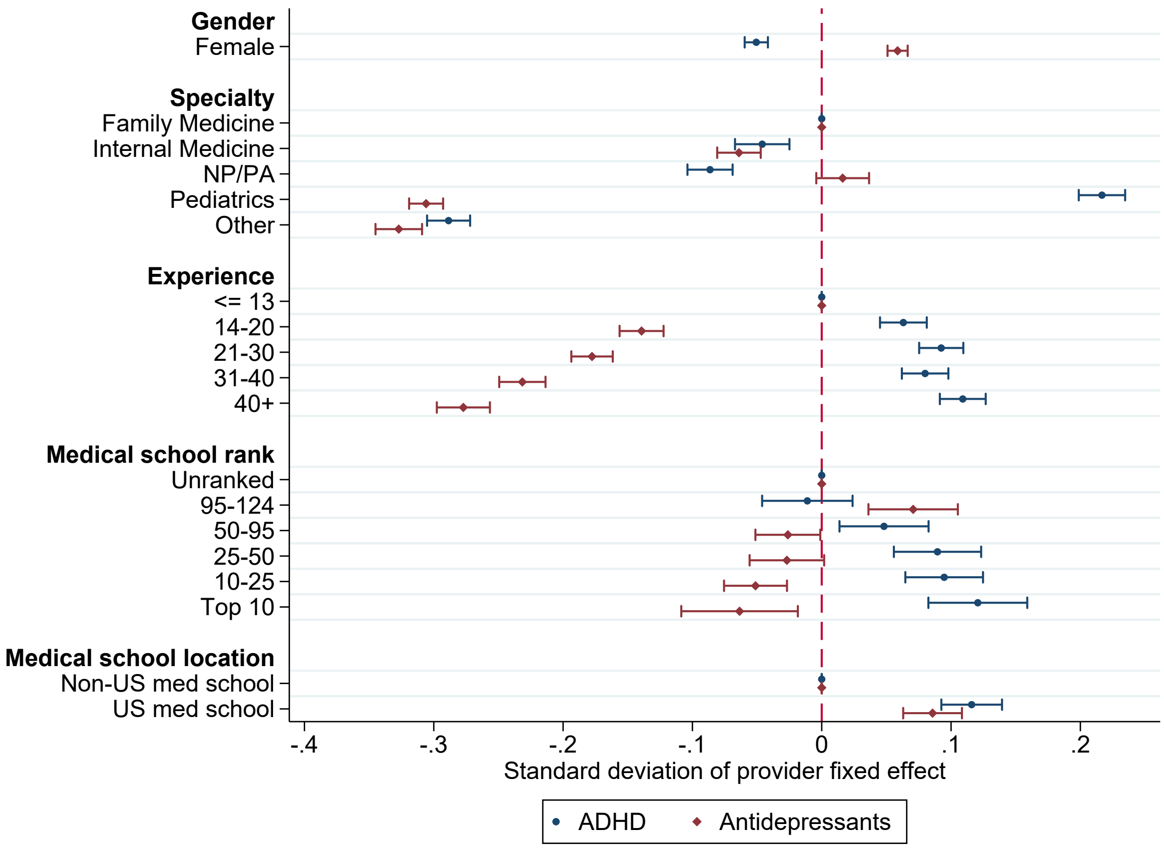
Notes: Figure shows average prescribing rate of ADHD medication (Panel A) and antidepressants (Panel B) by year relative to the switch for the first switch sample (N children = 3,875,841). Switchers are divided based on the quartiles of their old and new providers estimated prescribing intensities (separately for each drug class).

Figure 10: HRR-level correlates with practice environment effects



Notes: Figure plots HRR-level regression coefficients and 95 percent confidence intervals from regressions of the HRR practice environment fixed effects against various HRR-level characteristics. The fixed effects are standardized in the regression, all coefficients are reported in standard deviation units. Regressions are weighted by the number of child-years in each HRR. Special education finance incentive is an indicator for whether the HRR's state allocates additional funding as a function of the number of children receiving special education services (including children diagnosed with ADHD). States are coded based on a manual update of Morrill (2018)'s database of state statutes. Pennsylvania is the sole state that switched financing mechanism during the sample period, and thus HRRs in Pennsylvania are excluded from this regression. The remainder of the explanatory variables are also standardized across HRRs. Mental health specialists per child is the average across years of number of different mental health specialists who are the rendering or prescribing provider on a claim for a child residing in that HRR divided by the number of children. The psychiatrist share calculates the ratio of the number of specialists who are psychiatrists to the number who are non-prescribing psychologists and mental health social workers, following Cuddy and Currie (2020). Primary providers per child measures the number of large providers in the analysis sample matched to children residing in the HRR, divided by the number of children and averaged across years. Medicare spending and medicare physician spending measures are drawn from the Dartmouth Atlas. 95 percent confidence intervals are constructed using two-step Bayesian bootstrap standard errors clustered at the child level (Rubin 1981; Badinski et al. 2023). Specifically, for each child in the sample, I draw 50 weights from a Dirichlet distribution. I repeat the AKM model estimation 50 times, weighting each observation by its respective Dirichlet weight. I repeat the regressions above for each of the 50 sets of estimated fixed effects. The reported standard errors are the standard deviations of the resulting bootstrap estimates.

Figure 11: Correlates of provider prescribing intensity



Notes: Figure plots provider-level regression coefficients and 95 percent confidence intervals from regressions of the provider prescribing intensity fixed effects against various provider characteristics. The fixed effects are standardized in the regression, all coefficients are reported in standard deviation units. Regressions are weighted by the provider's number of associated child-years. Provider gender and specialty are drawn from the NPPES. Experience and medical school information are drawn from IQVIA data covering only physicians (i.e. not NP/PAs). 95 percent confidence intervals are constructed using two-step Bayesian bootstrap standard errors clustered at the child level (Rubin 1981; Badinski et al. 2023). Specifically, for each child in the sample, I draw 50 weights from a Dirichlet distribution. I repeat the AKM model estimation 50 times, weighting each observation by its respective Dirichlet weight. I repeat the regressions above for each of the 50 sets of estimated fixed effects. The reported standard errors are the standard deviations of the resulting bootstrap estimates.

Table 1: Sample construction

	Base BCBS child sample	Analysis sample
Sample size:		
N Children	11,495,405	8,066,999
N Child-Years	47,427,135	39,188,740
N Providers	717,724	130,616
Age composition:		
5-9	25.10	26.00
10-14	28.76	30.61
15+	46.14	43.39
Healthcare utilization:		
ADHD medication rate (%)	4.64	4.57
Antidepressant rate (%)	4.63	4.02
Average healthcare spending ($000s)	69.84	43.84
Average mental health care spending ($000s)	10.27	6.87
Hospitalized (%)	1.04	0.82
Visited ER (%)	9.75	8.79

Notes: The first column reports the sample size and summary statistics for the Base BCBS child sample of all child-years of children aged 5 through 21 covered for at least 6 months by a BCBS plan (including prescription coverage) from 2012 through 2022. The second column reports sample statistics for the sample I use for analysis. Relative to the Base BCBS child sample, I drop children who move HRRs multiple times in the sample period and child-years matched to providers who are matched to fewer than 30 children or observed for fewer than 3 years in the panel, and restrict to the largest connected set of children, places and providers. Age composition and healthcare utilization outcomes are measured at the child-year level. Spending measures are converted to 2019 dollars using the CPI.

Table 2: Geographic variation decomposition

	Above/below median		Top/bottom 25%		Top/bottom 10%	
	ADHD meds	Antidepressants	ADHD meds	Antidepressants	ADHD meds	Antidepressants
Observed difference (pp)	2.65	2.21	4.00	3.44	6.25	4.76
Avg. provider share (%)	-4.78	19.96	-5.39	20.51	-9.36	20.88
	(4.95)	(7.39)	(4.66)	(6.83)	(4.51)	(6.87)
Practice environment share (%)	50.05	40.96	49.64	38.94	53.76	36.75
	(5.56)	(8.48)	(5.30)	(8.08)	(5.60)	(8.30)
Avg. child share (%)	57.98	54.15	59.18	52.97	58.28	53.17
	(2.87)	(3.98)	(2.78)	(3.37)	(3.16)	(3.29)
Xβ share (%)	-3.24	-15.07	-3.43	-12.42	-2.69	-10.81
	(0.04)	(0.07)	(0.05)	(0.05)	(0.04)	(0.05)

Notes: Table is based on estimating Equation 1 separately for ADHD medication prescribing and antidepressant prescribing, and the geographic decomposition described in subsection 3.3. Each columns divides HRRs into two groups based on percentiles of their prescribing rate of the specified drug type. The first row reports the observed difference in prescribing rates between the two groups, measured in percentage points. The following rows report the share of this observed difference explained by each component, as derived in Equation 3, Equation 2, and Equation 4. The sample is the analysis sample, comprising 8,066,999 children observed over 39,188,740 child-years. Standard errors (in parentheses) are clustered at the child level and calculated using a Bayesian bootstrap (Rubin 1981; Badinski et al. 2023). Specifically, for each child in the sample, I draw 50 weights from a Dirichlet distribution and repeat the AKM model estimation and decomposition 50 times, weighting each observation by its respective Dirichlet weight. The reported standard errors are the standard deviations of the resulting bootstrap estimates.

Table 3: Provider variation decomposition

	ADHD meds	Antidepressants
Variance of prescribing rates (pp)	45.45	56.99
Share of variation explained by (%):		
Variance of provider effects	31.76	45.39
	(0.30)	(0.45)
Variance of avg. child effects	44.81	32.15
	(0.25)	(0.15)
Variance of avg. HRR effect	3.48	2.29
	(0.36)	(0.30)
Variance of avg. $X\beta$	0.29	2.74
	(0.01)	(0.02)
Variance of avg. residual	0.00	0.00
	(0.04)	(0.04)
2cov(avg. child, provider)	20.53	19.59
	(0.39)	(0.38)
2cov(avg. child, avg. HRR)	3.17	1.15
	(0.33)	(0.36)
2cov(provider, avg. HRR)	-3.21	-1.63
	(0.50)	(0.55)
$\aleph_{cov}$	-0.83	-1.68

Notes: Table is based on estimating Equation 1 separately for ADHD medication prescribing and antidepressant prescribing, and the provider-level decomposition described in subsection 3.3. The first row reports the observed variance of prescribing rates (in percentage points) across providers. The following rows report the share of the variance explained by each component, as derived in Equation 6. $\aleph_{cov}$ aggregates the share of variation explained by the remaining covariance terms. The sample is the analysis sample, comprising 8,066,999 children observed over 39,188,740 child-years. Standard errors (in parentheses) are clustered at the child level and calculated using a Bayesian bootstrap (Rubin 1981; Badinski et al. 2023). Specifically, for each child in the sample, I draw 50 weights from a Dirichlet distribution and repeat the AKM model estimation and decomposition 50 times, weighting each observation by its respective Dirichlet weight. The reported standard errors are the standard deviations of the resulting bootstrap estimates.

A Data Appendix

A.1 Defining HRR of residence

In this section, I describe how I define the HRR of residence for each child-year. The BCBS member files include monthly ZIP codes of residence, which I match to the corresponding HRRs using the Dartmouth Atlas crosswalk. However, there are a number of cases, particularly around potential moves, where children are observed to be in multiple HRRs for overlapping periods that can be non-negligible.

For each child, I aggregate all claims to the HRR-half-year level. Within each half-year, I keep the HRR associated with the most claims, and calculate the length of residence in that HRR. If there are multiple HRRs within a given half-year and all have no claims, I first attempt to assign the reference half-year's HRR as the most recent previous half-year with an HRR assignment. If there is no previous half-year with an HRR assignment, I assign the reference half-year's HRR as the closest future half-year with an HRR assignment. If there is also no future half-year with an HRR assignment, I code the HRR as missing for that half-year. For a given year, if the HRR of residence is the same across the two havles of the year, assign that as the HRR. If the HRR is different across the two half-years within a given year, I keep the HRR with the longer length of residence.

A.2 Assigning PCPs

In this section, I describe how I assign each child a PCP in each year. For each non-facility claim, I observe the date of service and the NPI of the rendering (for a non-facility, non-pharmaceutical claims) or prescribing (for pharmaceutical claims) provider. Additionally, I observe the Berenson-Eggers Type of Service (BETOS) codes associated with non-pharmaceutical claims (based on the HCPCS procedure codes, which I also observe). I link the provider NPIs to specialty data from the NPPES.

I collapse the data to count total number of claims, number of office visits (*BETOS: M1a/b*), number of well-visit claims (*ICD-10: Z00, Z13.31, Z13.39, Z71.1, Z71.2, Z71.3, Z71.8; ICD-9: V20, V70.0, V65.3, V65.41, V65.5, V65.8, V65.9; CPT: 99381, 99382, 99384, 99385, 99391, 99392, 99393, 99394, 99395)*, and number of prescriptions for each child-provider-year.

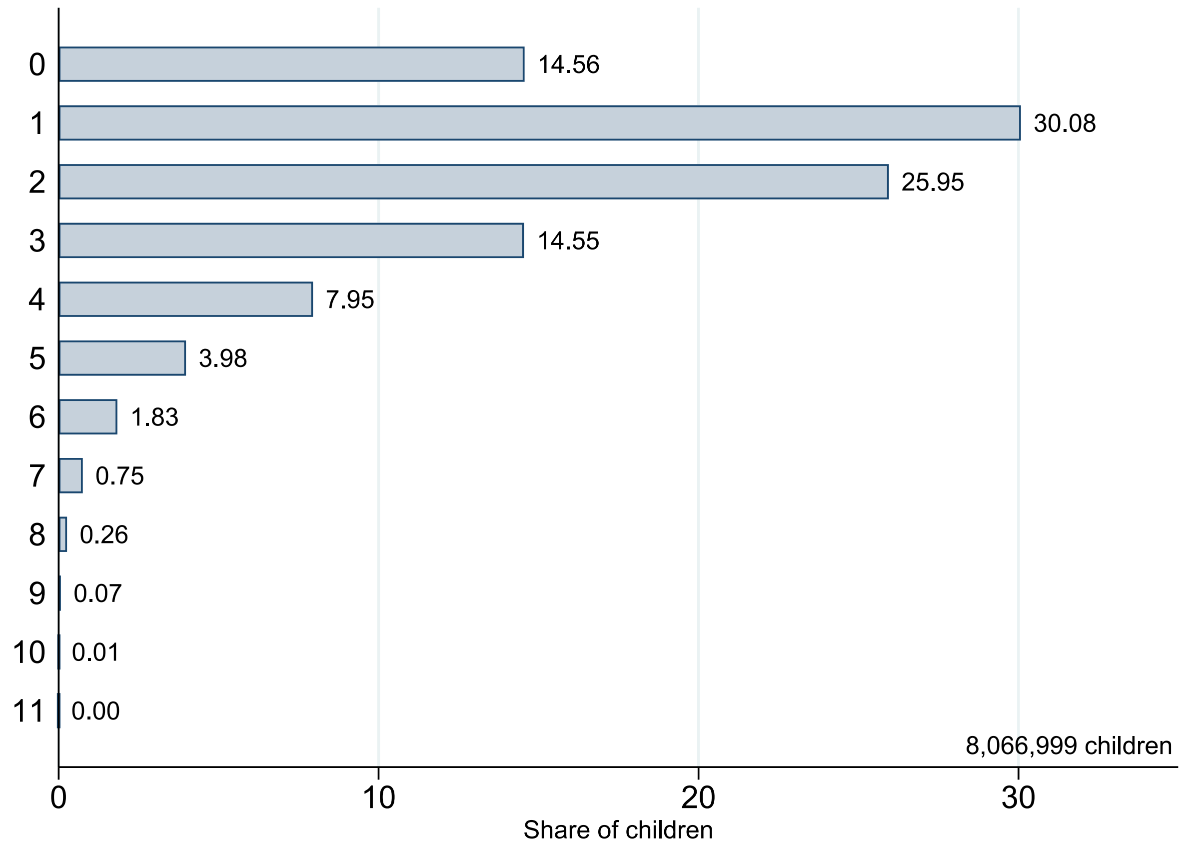
I then categorize providers into three groups based on their specialty codes, and keep only the first two groups as potential PCPs.

- **Group 1 - Typical PCP specialties:** General Practitioners (*208D00000X, 208M00000X*), Family Medicine, Internal Medicine, Pediatrics
- **Group 2 - All other providers, excluding the following specialties:**
  - Ambulance, Anesthesiology, Assisted Living Facility, Audiologist, Chiropractor, Contractor, Counselor, Dentist, Dermatology, Dietitian, Registered, Durable Medical Equipment and Medical Supplies, Medical Genetics, Nuclear Medicine, Nutritionist, Obstetrics and Gynecology, Occupational Therapist, Occupational Therapy Assistant, Ophthalmology, Optometrist, Orthopaedic Surgery, Pathology, Pharmacist, Pharmacy, Physical Medicine and Rehabilitation, Physical Therapist, Plastic Surgery, Podiatrist, Psychologist, Radiologic Technologist, Radiology, Rehabilitation Counselor, Rehabilitation Hospital, Respiratory Therapist, Certified, Respiratory Therapist, Registered, Respite Care, Skilled Nursing Facility, Social Worker, Specialist, Specialist/Technologist, Specialist/Technologist Cardiovascular, Specialist/Technologist Other, Specialist/Technologist Pathology, Substance Abuse Rehabilitation Facility, Surgery, Technician, Technician, Health Information, Technician, Other, Technician, Pathology, Technician/Technologist.
- **Group 3 - Specialties excluded in Group 2**

I then assign the PCP for each child-year as follows. Assign the PCP as the Group 1 provider with the most office visits, break ties with the number of well-visit claims. If the child had no office visits to Group 1 providers in that year, but saw a Group 1 provider, I assign the PCP as the Group 1 provider with the most claims, breaking ties with the number of prescriptions. If the child did not see a Group 1 provider in that year, I assign the PCP as the Group 2 provider with the most office visits, breaking ties with the number of prescriptions. If the child-year remains unmatched, I assign their PCP as missing.

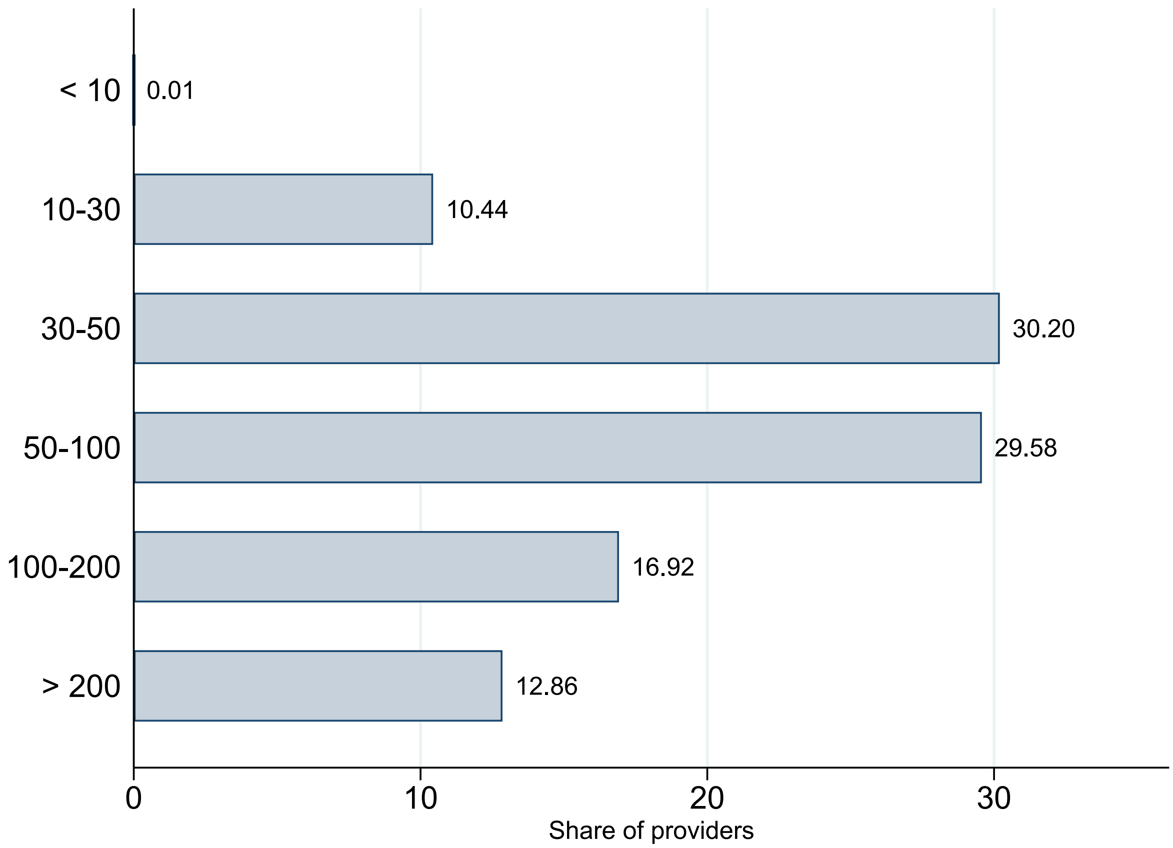
B Appendix Figures and Tables

Appendix Figure A1: Distribution of number of PCPs per child



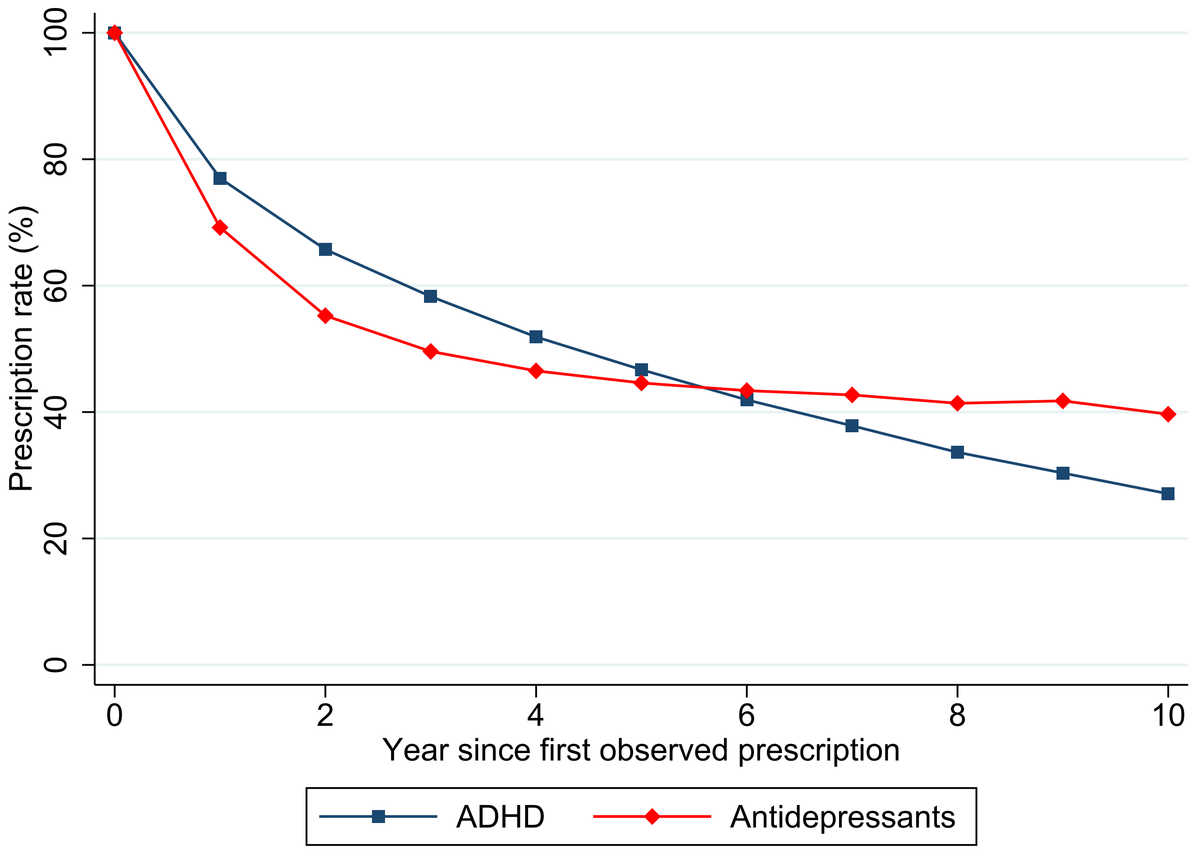
Notes: Figure shows the distribution across children of the number of different PCPs they are matched to in the sample. See Appendix A.2 for details of the PCP matching algorithm. The sample is the analysis sample (N = 8,066,999).

Appendix Figure A2: Distribution of number of switchers per provider



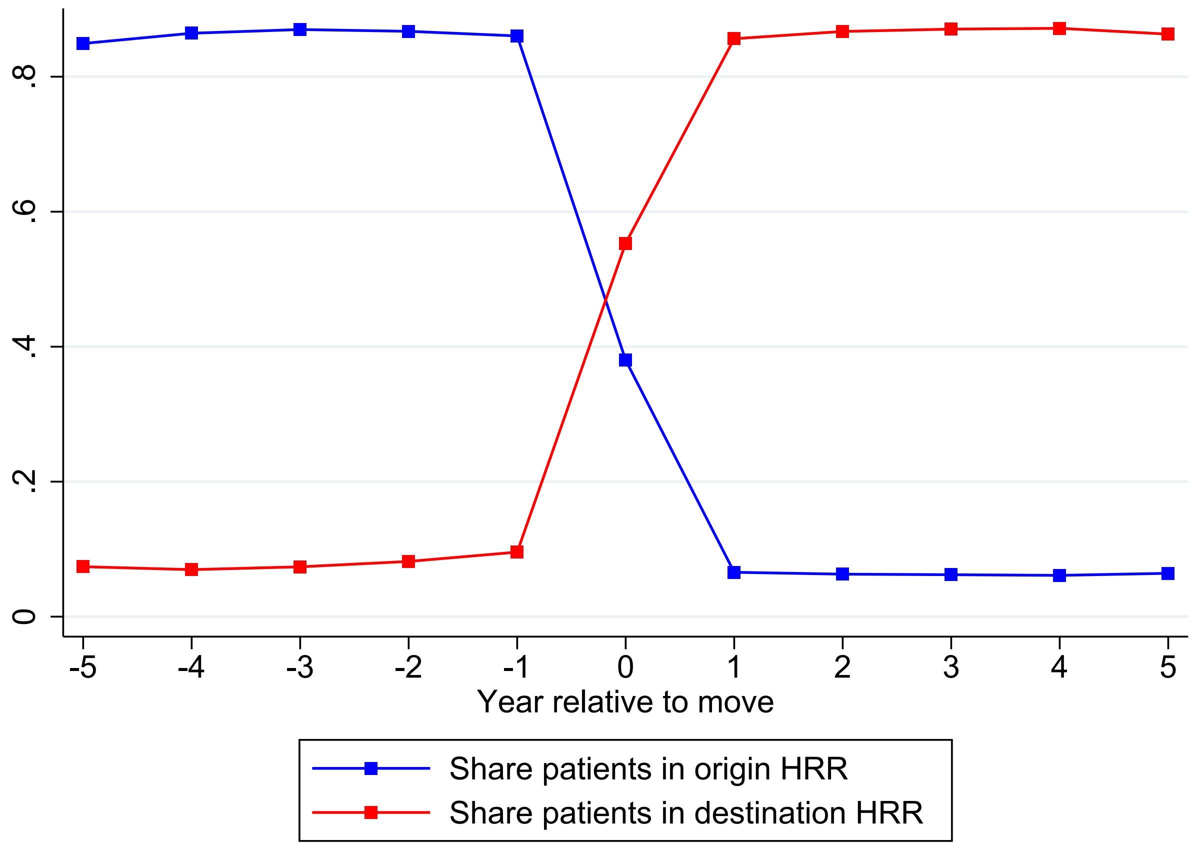
Notes: Figure shows the distribution of the number of switcher children providers are matched to in the sample. A child is defined as a switcher if they are matched to at least two different PCPs. Almost 90 percent of providers in the sample are matched to more than 30 switchers. The sample is providers in the analysis sample (N = 130,616). For comparison, in the Swedish linked establishment-worker data used by (Bonhomme, Lamadon and Manresa 2019), only 74 of 8,794 firms (0.8 percent) have at least 50 employees who switch firms.

Appendix Figure A3: Child prescription hazard rates



Notes: Figure shows the prescribing rates for ADHD medication and antidepressants in the years following a child's first observed prescription.

Appendix Figure A4: Provider mover origin-destination patient shares

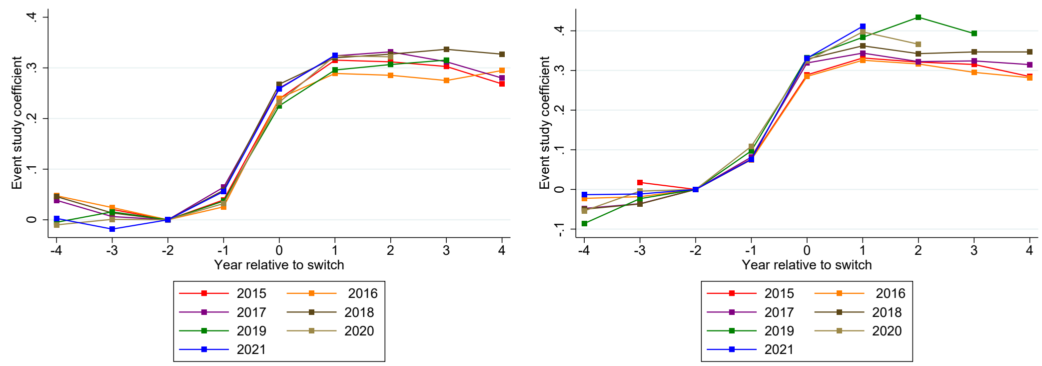


Notes: Figure displays the average provider-year share of patients residing in the origin HRR and destination HRR by relative year for provider movers. These measures include all BCBS-covered patients provider saw, not only children. Movers are determined by identifying providers with a clear shift in the residence of their patient population (all patient, not only children) from one HRR to another, following the algorithm from (Badinski et al. 2023). Shares are computed on a sample of 15,842 provider movers.

Appendix Figure A5: Switching event studies separately by year

(a) ADHD

(b) Antidepressants



Notes: Figure plots the estimated θ_w coefficients from the first switch sample event studies estimated separately by the year of the switch. Regression includes controls for age bins (5-9, 10-14, 15-21), but do not include relative-year fixed effects as these would be collinear with the calendar year fixed effects when estimating the event study separately by switch year. Panel A shows the results for ADHD medications, Panel B shows the results for antidepressants. The coefficients for year -2 are normalized to 0.

Appendix Table A1: Movement between Census Regions - child movers

		Destination				
		Midwest	Northeast	South	West	Total
Origin	Midwest	16.98	0.65	5.27	1.76	24.66
	Northeast	0.78	7.68	4.57	0.76	13.78
	South	3.21	2.09	37.06	2.28	44.63
	West	1.45	0.61	3.45	11.41	16.93
	Total	22.42	11.03	50.34	16.21	100.00

Notes: Table shows the percentage of child moves that take place between each of the 16 origin-destination pairs of census regions. 55 percent of child moves occur across state lines across non-neighboring HRRs. The sample size is all child movers (N = 243,940 children)

Appendix Table A2: Movement between Census Regions - provider movers

		Destination				
		Midwest	Northeast	South	West	Total
Origin	Midwest	13.45	1.22	4.90	3.41	22.98
	Northeast	0.93	10.35	4.89	1.77	17.94
	South	2.54	2.23	34.48	3.76	43.01
	West	1.11	0.61	2.14	12.21	16.08
	Total	18.03	14.41	46.40	21.15	100.00

Notes: Table shows the percentage of providers moves that take place between each of the 16 origin-destination pairs of census regions. The sample size is provider movers (N = 15,842 providers).

Appendix Table A3: Geographic variation decomposition with small providers

	Above/below median		Top/bottom 25%		Top/bottom 10%	
	ADHD meds	Antidepressants	ADHD meds	Antidepressants	ADHD meds	Antidepressants
Observed difference (pp)	2.65	2.21	4.00	3.44	6.25	4.76
Avg. provider share (%)	2.54	2.19	3.92	3.40	6.06	5.39
Practice environment share (%)	47.84	33.43	45.77	31.99	48.69	27.37
Avg. child share (%)	66.53	65.77	64.40	67.50	66.51	67.74
$X\beta$ share (%)	-2.62	-12.92	-2.45	-13.08	-2.61	-9.37

Notes: Table reports results from re-estimating the geographic decomposition from Table 2 using the Base BCBS child sample that includes child-years matched to small providers. Specifically, I group child-years matched to small providers by HRR into a set of "small provider" fixed effects and re-estimate Equation 1 when including these observations and report the resulting geographic decomposition. The sample is the Base BCBS child sample (N = 11,495,405 children, 47,427,135 child-years).

Appendix Table A4: Provider variation decomposition with small providers

	ADHD medication		Antidepressants	
	Var. Component	Share	Var. Component	Share
Variance of prescribing	45.45	100.00	56.99	100.00
Share of variation explained by:				
Variance of provider effects	13.41	29.51	23.81	41.78
Variance of avg. child effects	20.79	45.73	20.49	35.96
Variance of avg. HRR effect	1.37	3.02	1.16	2.04
Variance of avg. $X\beta$	0.13	0.28	2.10	3.69
Variance of avg. residual	0.00	0.00	0.00	0.00
2cov(avg. child, provider)	10.14	22.31	12.00	21.06
2cov(avg. child, avg. HRR)	1.37	3.01	0.60	1.05
2cov(provider, avg. HRR)	-1.39	-3.06	-1.01	-1.77
$\aleph_{cov}$		-0.80		-3.81

Notes: Table reports results from re-estimating the geographic decomposition from Table 3 using the Base BCBS child sample that includes child-years matched to small providers. Within each medication, the first column reports the variance (or covariance) of the specified component, and the second column reports that component's share of the overall variance of provider prescribing rates. I group child-years matched to small providers by HRR into a set of "small provider" fixed effects and re-estimate Equation 1 when including these observations and report the resulting provider decomposition. The sample is the Base BCBS child sample (N = 11,495,405 children, 47,427,135 child-years).

Appendix Table A5: Characteristics of providers in the sample

	N Providers	Share of providers	Share of child-years
Gender:			
Male	52,054	39.85	43.87
Female	69,492	53.20	47.33
Missing	9,070	6.94	8.80
Specialty:			
Family Medicine	37,424	28.65	22.54
Internal Medicine	8,610	6.59	4.94
NP/PA	24,005	18.38	10.15
Pediatrics	38,079	29.15	52.50
Other	18,967	14.52	7.86
Missing	3,531	2.70	2.00
Physician experience:			
≤ 13	3,442	2.64	1.87
14-20	10,776	8.25	8.46
21-30	21,351	16.35	22.33
31-40	17,264	13.22	18.28
40+	11,914	9.12	11.52
Missing	65,869	50.43	37.53
Physician medical school rank:			
Unranked	18,408	14.09	16.84
95-124	4,814	3.69	3.56
50-95	15,064	11.53	16.12
25-50	7,818	5.99	8.23
10-25	3,922	3.00	4.00
Top 10	1,708	1.31	1.68
Missing	78,882	60.39	49.56
Physician medical school country:			
Non-US med school	6,814	5.22	5.93
US med school	44,920	34.39	44.50
Missing	78,882	60.39	49.56

Notes: Table reports summary statistics of provider characteristics for providers in the analysis sample (N = 130,616). Provider gender and specialty are drawn from the NPPES. Experience and medical school information are drawn from IQVIA data covering only physicians (i.e. it does not cover Nurse Practitioners and Physician Assistants).