Providers, Places, and Children's Mental Health Care

Gabriel Swagel

Discussant: Emily Cuddy

November 7, 2025

2025 Annual Health Econometrics Workshop, Ann Arbor, MI

Children's Mental Health

- Children's mental health is the "defining public health crisis of our time" (US Surgeon General)

Children's Mental Health

- Children's mental health is the "defining public health crisis of our time" (US Surgeon General)
- Two of the most common childhood mental health diagnoses: ADHD and depression (Dwyer and Bloch 2019, Danielson 2023)

Children's Mental Health

- Children's mental health is the "defining public health crisis of our time" (US Surgeon General)
- Two of the most common childhood mental health diagnoses: ADHD and depression (Dwyer and Bloch 2019, Danielson 2023)
- Concerns over the substantial variation in pediatric mental health care across regions and within regions across providers (Cuddy and Currie 2020)

This paper

Using BCBS health insurance claims data from 2012 through 2022, this paper

- Provides systematic quantification of the drivers of variation in children's mental health prescribing in the United States

This paper

Using BCBS health insurance claims data from 2012 through 2022, this paper

- Provides systematic quantification of the drivers of variation in children's mental health prescribing in the United States
- Decomposes the drivers of regional and provider-level prescribing rate variation into
  - Regional practice environments
  - Differences in PCP prescribing intensities
  - Child health and demand

This paper

Using BCBS health insurance claims data from 2012 through 2022, this paper

- Provides systematic quantification of the drivers of variation in children's mental health prescribing in the United States
- Decomposes the drivers of regional and provider-level prescribing rate variation into
  - Regional practice environments
  - Differences in PCP prescribing intensities
  - Child health and demand
- Shows additional results on provider practice styles
  - Different practice intensities for different conditions
  - How is practice style of PCPs related to their training and experience?

Preview of Results

With child and provider *migration* across regions and within-region child *switching* between PCPs, an AKM-style model is estimated.

Results show that

- Provider prescribing intensities drive a substantial share of prescribing variation

Preview of Results

With child and provider *migration* across regions and within-region child *switching* between PCPs, an AKM-style model is estimated.

Results show 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%

Preview of Results

With child and provider *migration* across regions and within-region child *switching* between PCPs, an AKM-style model is estimated.

Results show 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%
- However, differences in average PCP prescribing intensities explain little of the regional variation, while differences in HRR fixed effects account for 40-50% of the geographic variation

Preview of Results

With child and provider *migration* across regions and within-region child *switching* between PCPs, an AKM-style model is estimated.

Results show 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%
- However, differences in average PCP prescribing intensities explain little of the regional variation, while differences in HRR fixed effects account for 40-50% of the geographic variation
- A quarter of PCPs have different relative prescribing intensities across conditions
  - Higher quality providers tend to have lower antidepressant prescribing intensity and higher ADHD prescribing intensity

Literature Review

Literature Review

1. **Variation in health care**
   - **Geographic variation:** *Skinner and Fisher 1997; Fisher et al. 2003a; Fisher et al. 2003b; Baicker and Chandra 2004; Sirovich et al. 2006; Song et al. 2010; Finkelstein, Gentzkow, and Williams 2016; Molitor 2018; Finkelstein, Gentzkow and Williams 2021; Ding 2023*
   - **Provider effects:** *Cutler et al. 2019; Badinski et al. 2024*
2. **Mental Health Care**
   - *Berndt et al. 2015; Laird and Nielsen 2016; McBain et al. 2019; Currie and MacLeod 2020; Marquardt 2021; Currie and Cuddy 2026*
3. **Provider practice styles**
   - *Grytten and Sorensen 2003; Epstein and Nicholson 2009; Currie, MacLeod, and Van Parys 2016; Currie and MacLeod 2017; Kwok 2019; Ahammer and Schober 2020; Fadlon and Van Parys 2020; Chan, Gentzkow, Yu 2022; Gowrisankaran, Joiner, and Leger 2023; Currie and Zhang 2023*

Model

Model

Consider a child i who visits their PCP j in year t,

$$y_{ijt} = \alpha_i + \delta_j + \gamma_{r(ijt)} + \tau_t + \beta X_{it} + \varepsilon_{ijt} \tag{1}$$

- $y \in \{a, d\}$: whether the child is prescribed ADHD medication or antidepressants during year t (by PCP *or* other doc*)
- $r(ijt)$: region of ijt (here, HRR)
- X_{it}: age bins (5-9, 10-14, and 15-21)

Fixed Effects

- Child-specific fixed effects α_i:
  - Capture child's underlying health and/or their parents' preferences for Rx treatment for condition y
  - Could also capture variation in moral hazard from potential out-of-pocket costs (assumed to be minimal)

Fixed Effects

- Child-specific fixed effects α_i:
  - Capture child's underlying health and/or their parents' preferences for Rx treatment for condition y
  - Could also capture variation in moral hazard from potential out-of-pocket costs (assumed to be minimal)
- Provider-specific effects δ_j:
  - Represent PCP prescribing intensity for condition y, capturing the relative propensity of the PCP to prescribe
  - Reflect differences in diagnostic skill (Currie and MacLeod 2020), diagnostic thresholds (Marquardt 2021), beliefs (Cutler et al. 2019), and specialist referral networks (Zeltzer 2020)

Fixed Effects

- Child-specific fixed effects α_i:
  - Capture child's underlying health and/or their parents' preferences for Rx treatment for condition y
  - Could also capture variation in moral hazard from potential out-of-pocket costs (assumed to be minimal)
- Provider-specific effects δ_j:
  - Represent PCP prescribing intensity for condition y, capturing the relative propensity of the PCP to prescribe
  - Reflect differences in diagnostic skill (Currie and MacLeod 2020), diagnostic thresholds (Marquardt 2021), beliefs (Cutler et al. 2019), and specialist referral networks (Zeltzer 2020)
- Regional effects $\gamma_{r(ijt)}$:
  - Encompass factors driving prescribing that are common to all child-provider pairs in r
  - Reflect differences in local supply (Cuddy and Currie 2020), special education financing laws (Morrill 2018), and non-provider-specific healthcare and cultural factors

Identification

The model exploits child and provider *migration* across regions and within-region child *switching* between PCPs

- α_i and δ_j are separately identified off of prescribing changes as children switch providers and prescribing variation within a provider across patients.

Identification

The model exploits child and provider *migration* across regions and within-region child *switching* between PCPs

- α_i and δ_j are separately identified off of prescribing changes as children switch providers and prescribing variation within a provider across patients.
- $\gamma_{r(ijt)}$ are identified from changes in prescribing as patients and PCPs move across regions

Identification

The model exploits child and provider *migration* across regions and within-region child *switching* between PCPs

- α_i and δ_j are separately identified off of prescribing changes as children switch providers and prescribing variation within a provider across patients.
- $\gamma_{r(ijt)}$ are identified from changes in prescribing as patients and PCPs move across regions
- Thought experiment
    - (1) child switch PCP in region $\rightarrow$ PCP FE + child FE
    - (2) child move across region $\rightarrow$ region FE (net out PCP FE)
    - (3) PCP move across region $\rightarrow$ better region FE + PCP FE

Assumptions and issues

- "Parallel trends": Unobservable shocks in health or Rx preferences (for child movers) and shocks in practice intensity (for PCP movers) must be uncorrelated with the difference in prescribing intensity between the origin and destination.
  - Why moving? Here, likely different than oft-used Medicare setting.
- Conditionally idiosyncratic switching of children between PCPs + no sorting on match-effects between children and providers
  - A priori possible $\rightarrow$ Hunting for desired treatment?
- Pitfalls of high-dimensional fixed effects models $\rightarrow$ small cell sizes and limited mobility of children and PCPs
  - Not so much here... lots of mobility. Too much?

Data

Data

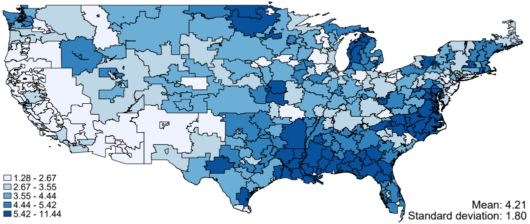
- Insurance claims and membership files from the Blue Cross Blue Shield (BCBS) Alliance for Health Research, from 2012 through 2022
    - Focus on children aged 5 thru 21 with coverage spells of at least 6 months (11.5M)
    - Match children to their "modal" PCP and HRR regions for each year (about 28% are unmatched)
- **Movers**
    - 243,940 children yielding 1.3M child-years
    - 15,842 PCP movers (adapted procedure from Badinski et al. 2023)
    - Almost 90 percent of PCPs are matched to at least 30 switchers in the sample ▸ Switchers
- **Additional restrictions**
    - Providers need 30+ child patients + 3 years; drop multi-mover children; limit to largest connected set (717,724 → 130,616 providers + 82.6% of child-years)

Results

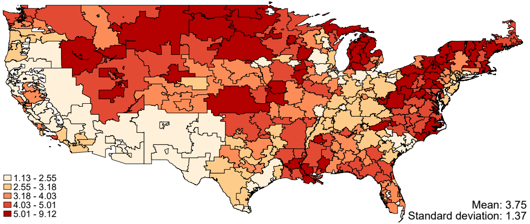
HRR prescribing rates (ADHD = 4.21%, AD = 3.75%)

Figure 1: Prescribing rates across HRRs

(a) ADHD medication



(b) Antidepressants

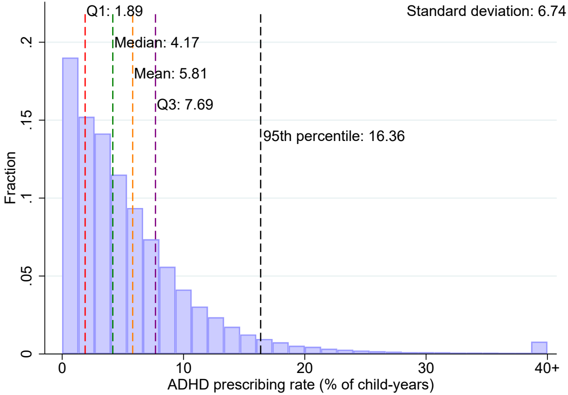


Notes: The figures map the distribution of HRR prescribing rates for ADHD medication (Panel A) and antidepressants (Panel B), shaded by quintile as defined in the legend. Means and standard deviations are calculated across HRRs (unweighted). The sample is the analysis sample, comprising 8,066,999 children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

Provider prescribing rates (ADHD = 5.81%, AD = 6.87%)

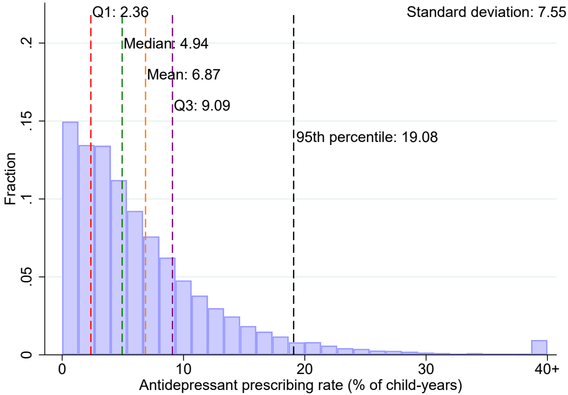
Figure 2: Distribution of provider prescribing rates

(a) ADHD medication



Provider prescribing rates (ADHD = 5.81%, AD = 6.87%)

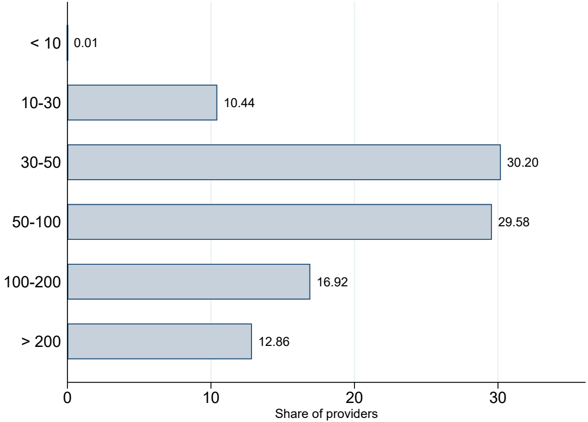
(b) Antidepressants



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.

Extent of "connectedness"

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.

Illustrative event studies

$$y_{ut} = \alpha_u + \tau_t + X_{ut}\beta + \sum_{k\neq 0} \phi_k \, \mathbf{1}\{t - T_u = k\} + \sum_{k\neq 0} \theta_k \Big(\mathbf{1}\{t - T_u = k\} \cdot \Delta_u \Big) + \varepsilon_{ut}$$

- α_u: unit FE (child FE for movers/switchers; provider FE for provider movers)
- τ_t: year FE; X_{ut}: age bin controls
- ϕ_k: raw event-time effects (tests for pre-trends / timing validity)
- θ_k: fraction of the origin→destination gap realized at event time k
- T_u: event year (move / first switch)

Event Study	Δ_u (leave-one-out difference)
Child moves HRR	$\overline{y}_{\text{dest},-i} - \overline{y}_{\text{orig},-i}$
Provider moves HRR	$\overline{y}_{\text{dest},-j} - \overline{y}_{\text{orig},-j}$
Child switches PCP	$\overline{y}_{\text{new},-i} - \overline{y}_{\text{old},-i}$

Intuition: What Each Event Study Identifies

Core idea: When a child or provider crosses environments, prescribing adjusts. θ_k measures **how much of the prescribing difference** between origin and destination is due to **supply-side factors**, rather than underlying child need.

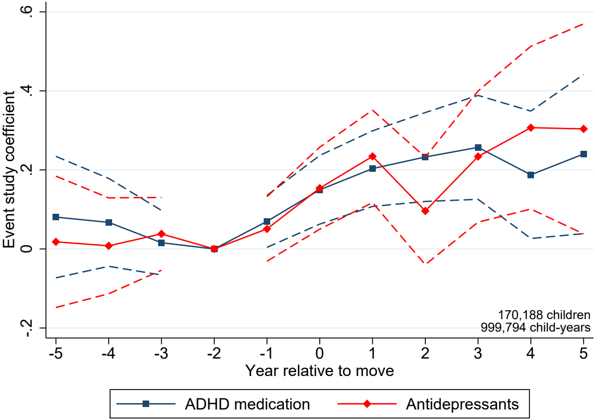
Interpretation of the jump in θ_k:

- Child moves HRR: Jump = **provider practice style + regional practice environment** contribution to the origin$\rightarrow$ destination prescribing difference.
- Provider moves HRR: Jump = **regional practice environment + destination patient mix** contribution.
- Child switches PCP: Jump = **provider prescribing intensity (practice style)** holding child and region fixed.

Flat pre-period $\Rightarrow$ timing not driven by changes in underlying child need or provider behavior.

Child mover event study

Figure 3: Child mover event study

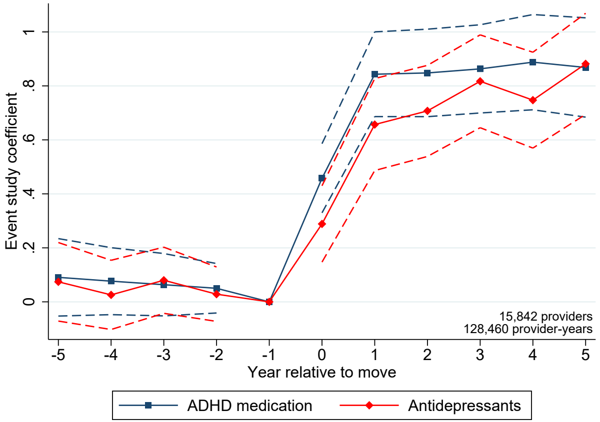


1

[1]Supply-side explains about 20% of variation in prescribing across origin-destination pairs in the sample. (versus 60% in Finkelstein, Gentzkow, and Willaism (2016)).

PCP mover event study

Figure 4: Provider mover event study

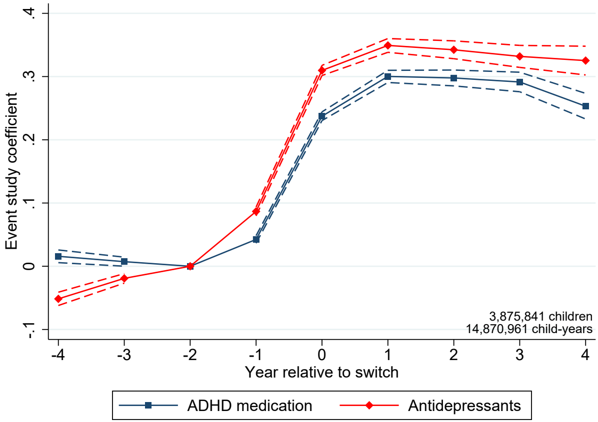


2

[2] Local practice environments and average child health/demand explains about 70-80% of variation in prescribing across origin-destination pairs in the sample.

PCP switching event study (v1 versus v2)

Figure 5: Switching event study - first switch sample



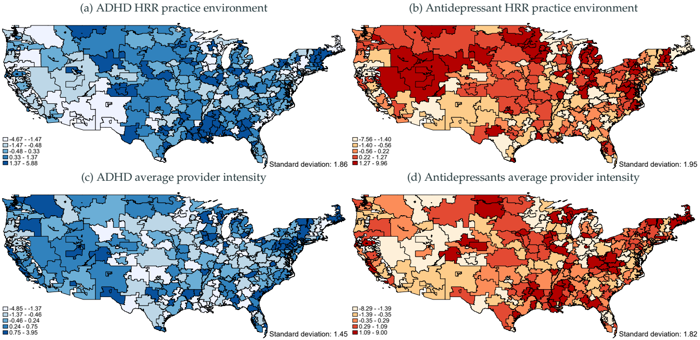
3

[3] Provider prescribing intensities explain about 30% of the variation in prescribing rates across old-new provider pairs.

What is driving the variation in prescribing intensity?

Main results

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.

Geographic Variation Decomposition

Table 1: Decomposing variation across HRRs

	Above/below median	
	ADHD meds	Antidepressants
Observed difference (pp)	2.65	2.21
Avg. provider share (%)	-4.78	19.96
	(4.95)	(7.39)
Practice environment share (%)	50.05	40.96
	(5.56)	(8.48)
Avg. child share (%)	57.98	54.15
	(2.87)	(3.98)
$X\beta$ share (%)	-3.24	-15.07
	(0.04)	(0.07)

Provider Variation Decomposition

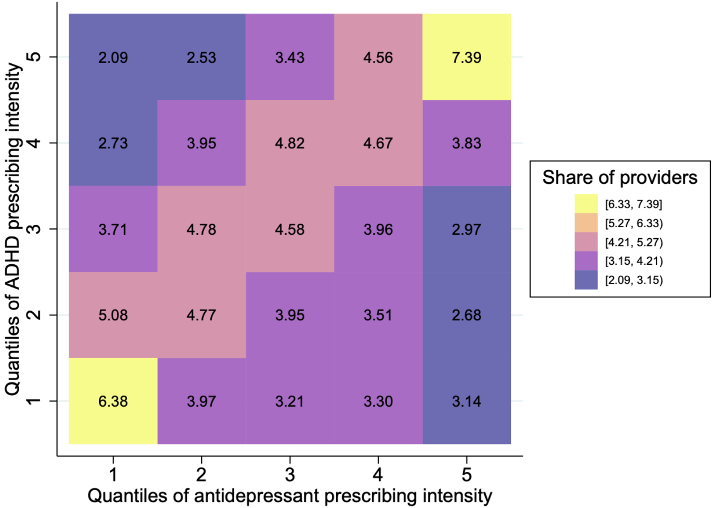
Note: Variance of prescribing rates across providers is 15 (36) times the magnitude of the variance across HRRs for ADHD (antidepressant) prescribing!

Table 2: 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
Variance of avg. child effects	44.81	32.15
Variance of avg. HRR effect	3.48	2.29
Variance of avg. $X\beta$	0.29	2.74
Variance of avg. residual	0.00	0.00
2cov(avg. child, provider)	20.53	19.59
2cov(avg. child, avg. HRR)	3.17	1.15
2cov(provider, avg. HRR)	-3.21	-1.63
$\mathbb{N}_{\text{cov}}$	-0.83	-1.68

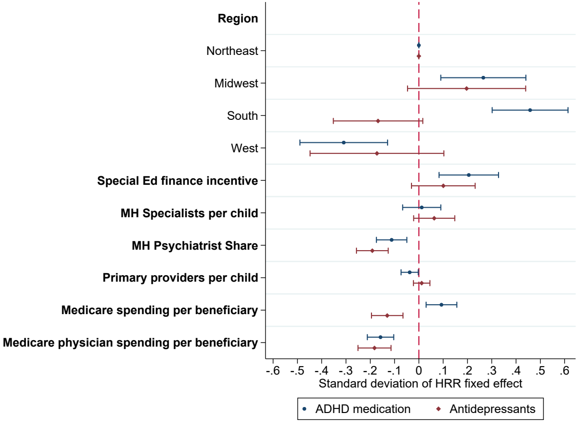
Do providers vary how they prescribe across conditions? Cowboys versus comforters?

Provider Prescribing Intensities across Conditions



What's driving places and providers to have different effects on pediatric mental health prescribing?

Correlations for HRR Prescribing Intensities



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

Correlations for Provider Prescribing Intensities

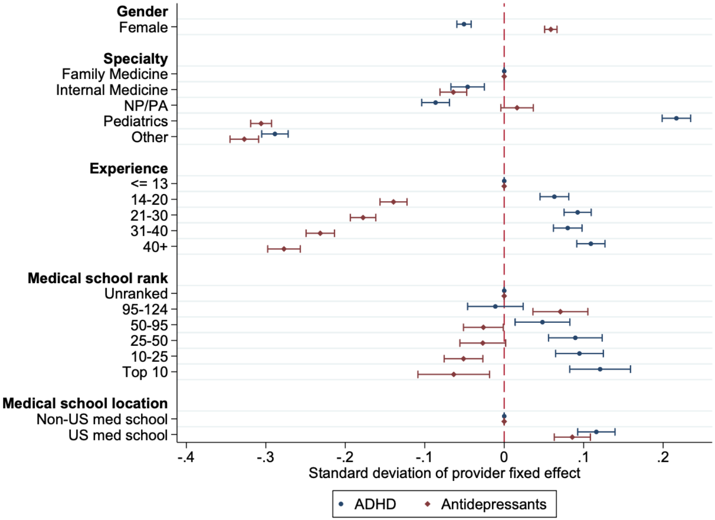


Figure 1: Correlates of Provider Prescribing Intensity

Comments

Discussion

- First off, very nice paper! Unconditional on being a graduate student.
- Polished and very comprehensive. All of the checks I would expect to find for this type of design are present (even if not mentioned by me today!).
- **Comment #1:** y_{ijt}: extensive margin only. Limitations?
- **Comment #2:** Treatment intensity versus treatment thresholds.
  - The paper blends differences in diagnostic thresholds with differences in treatment aggressiveness. The event studies identify a combined effect, not prescribing intensity per se.
- **Comment #3:** *One* margin of treatment here...
  - Guidelines would say that prescribing isn't the first-best form of treatment.

Discussion, continued

- **Comment #4:** δ_j: fixed and fully portable? Maybe not?
  - In the AKM setup, PCP and HRR effects are assumed to be additively separable and *stable* across environments. But prescribing behavior is likely *shaped* by referral networks, specialist availability, practice-group norms, and local reimbursement and coordination structures. So, "style" is actually an equilibrium adaptation to the environment.
- **Comment #5:** Peer effects among providers
  - The paper treats providers as independent units, but prescribing behavior may be networked. Provider effects could actually be group effects, meaning the AKM model is mis-specifying the object of identification.
- **Comment #6:** What's under the hood of rematching of patients and providers?
  - Under the hood of these movers is a patient–provider matching process — not random assignment. Paper demonstrates timing of the move/switch is not driven by changes in symptoms or Rx shocks, but what about the other end?

Conclusion

Conclusion

With a large, national sample of children, this paper

- Studies the driver of substantial heterogeneity in rates of mental health diagnoses and treatment
- Leverages the migration of children and providers across regions and the switching of children across PCPs

Conclusion

With a large, national sample of children, this paper

- Studies the driver of substantial heterogeneity in rates of mental health diagnoses and treatment
- Leverages the migration of children and providers across regions and the switching of children across PCPs

Findings:

- Provider prescribing intensities drive 50-65% of provider-level variation in children's mental health treatment
- The role of differences across providers is not evident when studying regional average outcomes
- A substantial share of providers have non-uniform relative prescribing intensities across conditions

Thank you!

Appendix: Number of Switchers per Provider

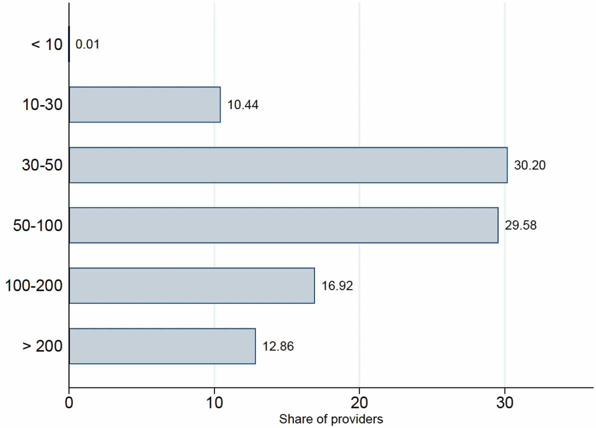


Figure 2: Distribution of Number of Switchers per Provider

Back