Physicians' Location Choice and Access to Healthcare: The Case of *Dobbs* 2022

Sheikh Atiya Islam*

This version: November 9, 2026

Job Market Paper

Abstract

I study how local legal environments affect physicians' geographic practice location choices. I use changes in state abortion law before and after the 2022 *Dobbs* decision to examine how restrictions on clinical practice affect where OB/GYNs choose to locate. I construct a novel state-year panel of abortion provisions and combine it with longitudinal physician practice-location data. In a first-difference triple-difference analysis, a 1-SD increase in permissiveness is associated with a 0.089 increase in the OB/GYN annual arrivals relative to a control specialty, General Surgery, per 100,000 residents, about 26% of mean pre-2022 OB/GYN arrivals as a scale benchmark. I then develop and estimate a discrete location-choice model, allowing for persistent differences in incumbent-location attachment. I find that a 1-SD increase in permissiveness is associated with 6.3% higher relative destination-choice odds, while incumbent persistence is greater in more restrictive environments. A counterfactual policy path removing measured net tightening since 2015 reallocates about 427 OB/GYN physician-equivalents across counties by 2024, a magnitude equivalent to roughly 29% of a typical year's physician mover volume. The counterfactual also suggests that restrictive legal environments modestly reduce physician availability in counties already facing low physician supply. The results show how legal environments shape physicians' geographic allocation through destination sorting even as physician stocks adjust gradually.

JEL classification: I11, I18, J44, J61, R23

Keywords: Physician location choice; abortion regulation; *Dobbs*; discrete choice; geographic access to healthcare.

*Department of Economics, Pennsylvania State University, University Park, PA 16802. Email: smi5203@psu.edu. I am grateful to my dissertation committee, Conor C. Ryan, Joris Pinkse, and Mark Roberts, for their guidance and feedback. This work received support from the Institute for Humane Studies. All errors are my own.

1 Introduction

The geographic distribution of physicians is an important determinant of access to medical care. How physicians choose where to practice in response to financial and market amenities like wages, shortage-area incentives, loan-forgiveness programs, local demand, and other market characteristics has been extensively studied (Hurley, 1991; Bolduc et al., 1996; Holmes, 2005; Chou and Lo Sasso, 2009). Recent evidence further shows that early-career physicians respond to these incentives in their location choices more strongly than the relocation decisions of established physicians (Khoury et al., 2025). At the same time, aggregate physician supply usually adjusts slowly because of lengthy residency or medical training. Entry into both the profession and individual specialties is also institutionally constrained (Gottlieb et al., 2025; Gottlieb and Nicholson, 2026). In such a labor market, legal provisions and regulations that can alter the scope of the services a physician can provide can add an additional channel through which the geographic allocation of physicians can be shaped. These regulations can alter the attractiveness of a location by changing which services physicians may provide, the circumstances under which they may provide them, and the legal risks associated with clinical practice. As a result, changes in the legal environment may affect access to healthcare through the geographic reallocation of existing practicing physicians.

I study this relationship in the market for obstetrician–gynecologists (OB/GYNs) during a period of unusually large changes in state abortion law. The Supreme Court's June 2022 decision in *Dobbs v. Jackson Women's Health Organization* removed the federal constitutional protection for abortion which returned the primary authority over abortion policy to the states (Supreme Court of the United States, 2022). It resulted in an abrupt cross-state difference in the legality of abortion. However, this divergence did not begin in June 2022. State legislatures had already enacted substantial waves of abortion restrictions before *Dobbs* along with several minor provisions that made the overall legal environment more difficult for providing certain services, while other states moved to protect or expand access (Nash et al., 2017, 2019, 2020; Nash and Naide, 2021). I collect all such major and minor provisions and construct a state-level legal permissiveness index where the mean cumulative change relative to 2015 is -2.25 by 2021 and declines further to -3.53 by 2024, where negative values denote net tightening. Seventy-two percent of states are more restrictive than their own 2015 baseline by the end of the sample. This continuous policy evolution provides a more comprehensive description of the variation physicians faced compared to a simple pre/post *Dobbs* analysis.

I combine this legal variation with annual practice locations for Medicare-enrolled

clinicians from the Centers for Medicare & Medicaid Services (CMS) Provider Data Catalog (Centers for Medicare & Medicaid Services, 2026). The analysis follows OB/GYNs from 2017 through 2024 at the county level. The final panel contains 44,209 OB/GYNs and 248,687 physician-year observations. The structural estimation sample contains 41,066 physicians and 225,108 annual location choices. A central feature of the data is that location is highly persistent. Physicians remain in the same county in 94.6 percent of observed consecutive-year choices. Among the 5.4 percent who move, roughly 38 percent cross a state line. Destination choices can therefore shift substantially while county physician stocks adjust slowly.

The empirical analysis first examines how county-level physician flows change with legal permissiveness. I compare OB/GYNs with general surgeons in a triple-difference design that accounts for local market shocks shared by the two specialties. My primary specification relates annual changes in the specialty-specific arrivals gap to annual changes in the permissiveness index. I also use an externally calibrated three-category policy classification as an exploratory comparison.

The reduced-form evidence is strongest for physician arrivals. In the preferred first-difference specification, a one-standard-deviation increase in permissiveness corresponds to a 0.089 increase in the annual OB/GYN-minus-general-surgery arrivals gap per 100,000 residents, a scale benchmark equivalent to about 26 percent of the 2017–2021 mean OB/GYN arrival rate. The arrival differential remains positive when counties with fewer than 10,000 or 20,000 residents are excluded, although those estimates become less precise. The corresponding departure differential is small, so the net-flow estimate largely reflects the arrivals margin. These flow responses are visible even while the estimated relationship with the stock of practicing physicians remains imprecise.

I then develop a discrete-choice model of annual county practice choices. A physician weighs a destination county's attractiveness against distance, the advantage of remaining within state, and the additional utility of staying in her previous-year practice county. I allow this stay premium to vary persistently across physicians, capturing differences in location attachment. The legal environment enters through two channels: the attractiveness of counties to prospective movers and the additional value of the incumbent location to physicians already practicing there. I estimate the model in two stages, first recovering county-year destination utilities and the behavioral parameters of geographic mobility, then relating recovered destination attractiveness to legal permissiveness.

The structural estimates identify both destination sorting and incumbent-location inertia. A one-standard-deviation increase in permissiveness raises relative destination-

choice odds by about 6.3 percent, holding other utility components fixed. Physicians also have substantial and heterogeneous attachment to their current practice locations. The negative incumbent-policy interaction indicates that the additional value of staying is greater in more restrictive environments. The two margins explain how differences in recruitment and destination choices can emerge before substantial changes in the geographic stock of incumbent physicians. I also examine heterogeneity by career stage and medical-school background; the subgroup estimates are imprecise and are discussed with the structural robustness results.

Finally, I use the estimated model to examine how alternative legal environments change the geographic distribution of practicing OB/GYNs. Removing all measured net tightening relative to the 2015 state baselines reallocates about 427 physician-equivalents across counties by 2024. The 427 physician-equivalents represent 1.57 percent of the active simulated workforce and are equivalent in magnitude to approximately 29 percent of one typical year's geographic mover volume, using the observed 5.41 percent annual move rate to benchmark 27,181 active simulated physicians. The latter compares a 2024 stock-distribution difference with an annual flow, rather than the fraction of individual movers whose decisions change. The counterfactual also produces a net shift of about 31 physician-equivalents toward counties with lower initial OB/GYN availability. A narrower experiment holding states at their 2021 policy levels reallocates about 130 physician-equivalents. These counterfactuals trace changes in where the existing workforce practices, rather than changes in national physician supply.

This paper contributes to three strands of literature. First, I extend research on abortion policy and the practicing physician workforce. Staiger et al. (2025) examine national OB/GYN practice locations around *Dobbs* and find limited changes in aggregate state-level OB/GYN counts alongside considerable persistence in physicians' locations. I construct a novel state-year panel of major and minor abortion-related legal provisions spanning the pre- and post-*Dobbs* periods. Linking that detailed policy history to longitudinal physician-location microdata, I study county-level arrivals and departures and identify choices that are not apparent from aggregate stock comparisons. The structural model then connects these geographic flows to the incentives governing destination choice and incumbent attachment.

Second, the paper relates to work on abortion policy and the physician training pipeline. Hammoud et al. (2024) document changes in OB/GYN residency applications following changes in abortion access, Mermin-Bunnell et al. (2025) study residency applications across specialties under abortion restrictions, and Ganguly et al. (2026) find larger post-*Dobbs*

declines in applications to programs in restrictive states, particularly among specialties more closely connected to abortion-related care. These studies establish that legal environments affect where future physicians seek training. I complement this evidence by studying realized practice-location decisions among physicians already in the workforce. The distinction matters because training choices can respond relatively quickly, whereas the geographic distribution of practicing physicians reflects both destination choices and substantial persistence among incumbents.

Third, the paper contributes to the broader literature on physician location and geographic variation in health-care supply. Existing work shows that physicians respond to local demand, compensation, shortage-area incentives, and other market characteristics when choosing where to practice (Hurley, 1991; Bolduc et al., 1996; Holmes, 2005; Chou and Lo Sasso, 2009), and recent evidence shows that such incentives operate particularly strongly through the location choices of early-career physicians rather than relocation by established physicians (Khoury et al., 2025). More broadly, location-specific supply factors account for an important share of geographic variation in health-care utilization (Finkelstein et al., 2016), while physician supply adjusts slowly because of lengthy training and entry constraints (Gottlieb et al., 2025; Gottlieb and Nicholson, 2026). I extend this literature by treating the local legal environment as an economically relevant attribute of physician location choice and by estimating a discrete-choice model that separates destination attractiveness from incumbent attachment. This framework connects observed changes in physician flows to counterfactual changes in geographic allocation and provides a mechanism through which regulation can affect access to care even when aggregate physician supply responds only gradually.

The rest of the paper proceeds as follows. Section 2 describes the institutional setting and permissiveness measure. Section 3 presents the physician-location data and descriptive facts. Section 4 discusses the motivating facts from the reduced-form analysis. Sections 5 and 6 develop and estimate the location-choice model. Section 7 presents the structural estimates and model fit. Section 8 reports the policy-path counterfactuals. Section 9 discusses robustness and the scope of the analysis, and Section 10 concludes.

2 Institutional Background and Policy Measurement

2.1 Dobbs and the evolution of state policy

On June 24, 2022, the Supreme Court decided *Dobbs v. Jackson Women's Health Organization*, overturning *Roe v. Wade* and *Planned Parenthood v. Casey* and ending the federal constitutional right to abortion (Supreme Court of the United States, 2022). States subsequently diverged sharply in the legality of abortion, especially through total bans and early gestational limits. This divergence built on an earlier accumulation of medication, procedural, provider, and other abortion-related restrictions, as well as protective legislation in other states. Guttmacher's state-policy reviews document this longer evolution during 2016–2021 (Nash et al., 2017, 2019, 2020; Nash and Naide, 2021). The combination of earlier regulatory changes and the more salient post-*Dobbs* bans motivates a continuous legal-policy measure.

Figure 1 displays the median and cross-state distribution of cumulative changes in permissiveness relative to 2015. The median is zero in 2017–2018, −1.5 in 2019–2021, and −3.0 in 2022–2024. State-level means show the same broad tightening, moving from −0.90 in 2017 to −2.25 in 2021 and −3.53 in 2024. Negative values indicate additional recorded restrictions relative to 2015. Forty percent of jurisdictions had tightened by 2017, 54 percent by 2021, and 72 percent by 2024. The figure also shows how the cross-state distribution widens after *Dobbs*. I use 2015 as the common pre-*Dobbs* reference year because it is the first year of the consistently assembled state-policy panel; annual changes in the first-difference specification are unaffected by this reference-year subtraction.

2.2 The permissiveness index

I construct an annual state-level panel of abortion-related legal provisions by combining a hand-compiled chronology of legal developments with a historical state-policy workbook. The chronology records the state, provision, direction of change, and recorded year of each event. Within each state and provision, its coded status carries forward until a later recorded event changes or reverses it. The workbook contributes five additional annually observed provision categories: Medicaid funding restrictions, telehealth restrictions, targeted regulation of abortion providers (TRAP), waiting periods, and parental involvement. The panel covers the 50 states and the District of Columbia from 2015 through 2025. It records 13 restrictive provision categories in total, with eight from the event chronology and five from the workbook. Eight protective categories are tracked separately. Appendix A.1

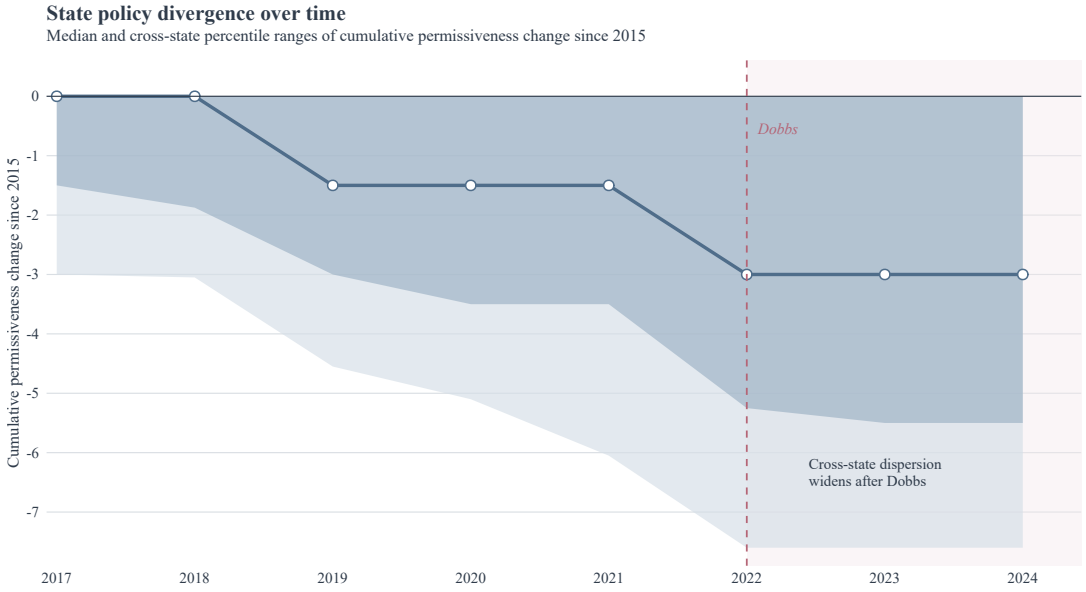


Figure 1: Evolution of the state abortion-policy restriction index, 2017–2024. The solid line reports the median cumulative change relative to each state's 2015 legal environment, while the shaded regions report the interquartile and 10th–90th percentile ranges across states. Higher values indicate fewer measured restrictions relative to the state baseline; negative values indicate net tightening. The shaded post-2022 period marks the post-*Dobbs* environment.

provides the construction rules and source details.

The continuous permissiveness index summarizes the accumulated severity of restrictive legal provisions. Its eight event-derived categories are total bans, criminal provisions, civil enforcement provisions, method bans, early gestational limits, medication restrictions, genetic restrictions, and ultrasound requirements. The five workbook categories are Medicaid, telehealth, TRAP, waiting periods, and parental involvement. Protective legislation is recorded separately because legal protections and restrictive requirements can coexist; enactment of a protection does not necessarily repeal a separately recorded restriction. This construction allows the primary index to measure cumulative restrictive burden consistently across states and years.

I assign higher weights to provisions with broader implications for the services an OB/GYN may provide and the legal risks of clinical practice, and lower weights to more specific procedural requirements. A total ban receives weight 3, criminal provisions 2.5, civil enforcement, method bans, and Medicaid restrictions 2, six intermediate provisions 1.5, and parental-involvement and ultrasound requirements 1. The weights are fixed features of the measurement design and represent an ordering of restrictive severity, rather than estimates of the separate effect of each provision. Table 1 lists all 13 categories, their sources, and weights. The permissiveness index is the negative weighted sum; higher values therefore indicate a less restrictive recorded environment.

Table 1: Restrictive legal provisions in the permissiveness index

Restrictive provision	Source	Weight
Total ban	Event chronology	3.0
Criminal provision	Event chronology	2.5
Civil enforcement	Event chronology	2.0
Method ban	Event chronology	2.0
Medicaid funding restriction	Annual workbook	2.0
Early gestational limit	Event chronology	1.5
Medication restriction	Event chronology	1.5
Genetic restriction	Event chronology	1.5
TRAP requirement	Annual workbook	1.5
Waiting period	Annual workbook	1.5
Telehealth restriction	Annual workbook	1.5
Parental-involvement requirement	Annual workbook	1.0
Ultrasound requirement	Event chronology	1.0

Notes: The primary index is the negative severity-weighted sum of these 13 restrictive indicators. Eight additional protective provision categories are recorded separately and do not enter this sum. The annual workbook supplies five categories; the event chronology supplies eight.

Let A_{st} denote the absolute restriction index for state s in year t. The primary treatment is the cumulative change relative to the state's own 2015 environment,

$$P_{st}^{\text{raw}} = A_{st} - A_{s,2015}. \tag{1}$$

The empirical specifications use

$$P_{st} = \frac{P_{st}^{\text{raw}} - \overline{P}}{\sigma_P}, \qquad \sigma_P = 2.8638, \tag{2}$$

where the normalization is held fixed throughout the analysis. Subtracting the centering constant $\overline{P}$ does not change the estimated slopes in specifications with fixed effects or annual differences. A higher P_{st} indicates that a state's measured restrictive burden has fallen relative to its own 2015 level. Most observed changes are additional tightening; the index rises when recorded restrictive requirements are relaxed or repealed. The estimated policy gradient therefore captures changes in this cumulative restrictive burden.

The normalization measures changes from each state's own 2015 legal environment. Accordingly, $P_{st}^{\text{raw}} = 0$ denotes no recorded net change from that state's baseline, while states with different starting legal regimes can have the same cumulative change. The separate baseline legal measure used in the structural model preserves information about these initial differences.

The event-derived indicators are assigned to the years recorded in the legal-event chronology, with subsequent events updating the annual provision status. This gives the policy measure a consistent state-year interpretation throughout the period. Because enactment, implementation, and enforcement can occur at different times, I also examine targeted alternative timing assignments while retaining the original policy series as the baseline. The annual raw-index change has a standard deviation of 1.2834 in the 2017–2024 estimation panel, equivalent to 0.4481 of the fixed 2.8638-unit policy normalization. Appendix A.1 describes the timing conventions, and Section 9 summarizes the robustness estimates.

Figure 2 illustrates the distinction between absolute policy levels and subsequent policy changes. The left panel reports the absolute restriction index in 2021, while the right panel reports the change from 2021 to 2024. The comparison shows that substantial cross-sectional differences were already present before *Dobbs* and that additional state-level policy changes occurred afterward.

I use the continuous cumulative-change measure as the primary treatment because it captures the timing and severity of both major and more specific restrictive provisions.

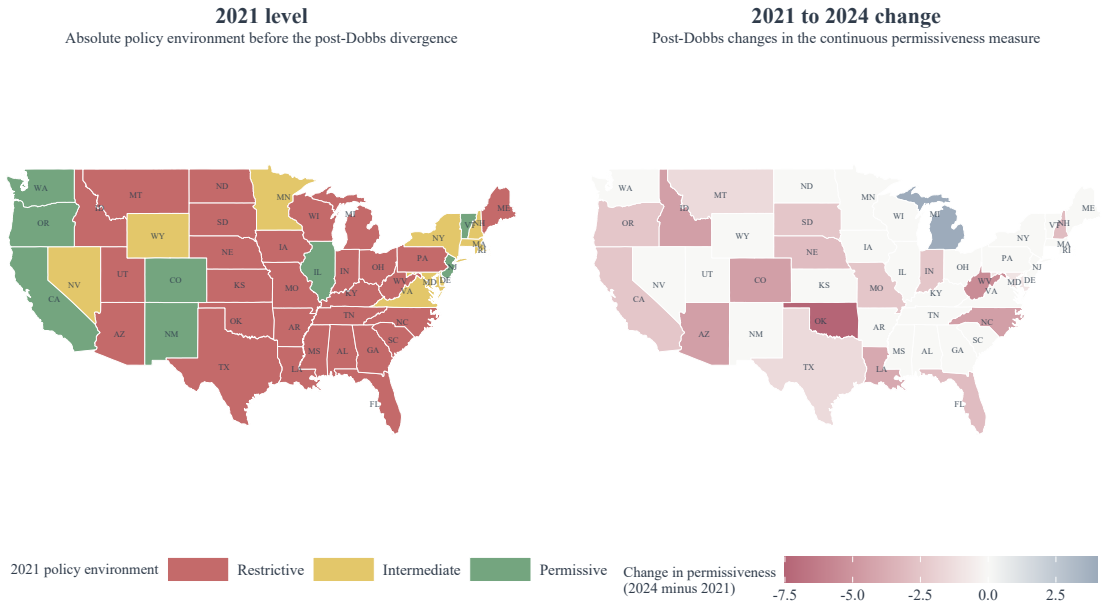


Figure 2: State abortion-policy environments before and after the post-*Dobbs* divergence. The left panel displays the absolute 2021 restriction index. The right panel reports the change in the continuous measure from 2021 to 2024; negative values indicate additional measured tightening.

For comparison, I classify each state's absolute index into restrictive, intermediate, and permissive categories using external calibration to Guttmacher's January 2020 classification. The calibrated cutoffs are $A_{st} \leq -4.75$ for restrictive, $A_{st} \geq -1.00$ for permissive, and intermediate otherwise. The index agrees with the external three-category benchmark in 42 of the 49 overlapping states (85.7 percent) in 2020. This agreement summarizes calibration to an independently developed policy classification, while the continuous treatment retains finer within-state variation for estimation. Appendix A.3 reports the calibration and category support.

2.3 Why the legal environment can affect physician location

The legal environment of a practice location can affect an OB/GYN practice through several channels. State policy directly determines which abortion services may legally be provided and under what conditions, while more restrictive environments have historically been associated with fewer facilities and more limited modes of abortion-service delivery (Jones et al., 2018). Following *Dobbs*, uncertainty over the practical scope of life and health exceptions has added another dimension of legal and administrative risk to clinical

decision-making (Felix et al., 2024). These features can affect the organization of care, the range of services available within a practice, and the expected costs and benefits of locating in a particular market.

If legal rules enter the location-specific payoff to practice, their first-order effect need not be an immediate reduction in the stock of incumbent physicians. Established physicians face substantial relocation costs, whereas physicians already changing locations can respond on the destination margin by choosing among markets with different legal environments. This distinction is consistent with evidence that physician location choices are sensitive to market incentives (Bolduc et al., 1996; Chou and Lo Sasso, 2009; Holmes, 2005), that responses to shortage-area designation are concentrated among earlier-career physicians (Khoury et al., 2025), and that abortion restrictions affect prospective training choices (Hammoud et al., 2024; Mermin-Bunnell et al., 2025; Ganguly et al., 2026). At the same time, practicing OB/GYNs exhibit substantial persistence in their state-policy environments (Staiger et al., 2025). Taken together, these patterns motivate an empirical design centered on geographic sorting: whether changes in the legal environment alter where practicing physicians choose to locate when they move, even before those responses generate large changes in the existing physician stock.

3 Data and Descriptive Facts

3.1 Physician locations

The physician data come from the CMS Doctors and Clinicians National Downloadable File in the Provider Data Catalog (Centers for Medicare & Medicaid Services, 2026). The file identifies Medicare-enrolled clinicians by National Provider Identifier (NPI), specialty, and practice location. I construct one county practice location per physician-year and classify specialties consistently across annual files. OB/GYNs are the primary workforce of interest. General surgeons provide a comparison specialty: both groups are specialized physicians who choose among the same local healthcare markets, but abortion-specific provisions bear more directly on OB/GYN clinical practice. General surgeons are not entirely insulated from abortion-policy changes: they can share hospital service lines with OB/GYNs and may participate in obstetric emergencies. Comparing county-level flows helps account for local hiring and healthcare-market conditions shared by both specialties.

The final OB/GYN stock panel spans 2017–2024 and contains 44,209 physicians and 248,687 physician-years on the common county geography. The structural decision sample

is narrower because a choice occasion requires observing a physician in consecutive years. It contains 41,066 physicians and 225,108 annual location choices. Physicians whose specialty coding overlaps across the final specialty panels are excluded consistently across the reduced-form and structural samples. Connecticut is excluded from the final structural geography due to substantial changes in its county classification. This removes 4,194 OB/GYN physician-years, or 1.66 percent of the raw stock panel. Appendix A.1 documents the exact geographic support and assignment rules used in the replication files.

3.2 County markets and auxiliary data

A market in the final analysis is a county, and both the structural and reduced-form specifications use the same universe of 3,135 counties. County-year population measures are drawn from StatsAmerica Population Estimates and Decennial Census data, which are based on U.S. Census Bureau population estimates, and I use the cleaned county-year series to construct per-capita outcomes (StatsAmerica, 2025; U.S. Census Bureau, 2025). I also use the Area Health Resources Files (AHRF) from the Health Resources and Services Administration as a source of county-level market characteristics used in structural robustness specifications (Health Resources and Services Administration, 2026).

3.3 Sample construction and descriptive facts

Table 2 summarizes the final analysis sample. The physician panel contains 44,209 OB/GYNs observed over 248,687 physician-years, while the structural decision sample contains 41,066 physicians and 225,108 annual choice occasions. Practice locations are highly persistent. The annual stay rate is 94.6 percent, compared with 3.34 percent for within-state moves and 2.07 percent for cross-state moves. Conditional on moving, the median move is 116.4 kilometers, although the distribution has a long right tail. Figure 3 shows that these mobility patterns are persistent across years. Within-state moves remain more common than cross-state moves, even though both account for a small share of annual choices. This is why the model allows for an incumbent stay premium, differences in attachment across physicians, and separate distance and same-state mobility costs.

The county distribution of OB/GYNs is sparse. Across 25,080 county-years, 51.9 percent contain no observed OB/GYN. The pooled number of OB/GYNs per 100,000 residents declines from 10.03 in 2017 to 8.94 in 2024. The structural estimation is conducted on 11,789 county-year destination cells with identified support. Although these cells account for 47.0 percent of county-years, they contain 91.5 percent of the population

and 99.87 percent of the observed OB/GYN stock. By construction, every observed chosen destination in the structural decision sample lies on identified support.

Table 2: Final analysis sample

Statistic	Value
OB/GYN physicians, stock panel	44,209
OB/GYN physician-years	248,687
Physicians, structural decision sample	41,066
Annual choice occasions	225,108
Stay rate	94.59%
Within-state move rate	3.34%
Cross-state move rate	2.07%
Counties	3,135
County-years, 2017–2024	25,080
County-years with zero OB/GYNs	51.93%
OB/GYNs per 100,000 residents, 2017	10.03
OB/GYNs per 100,000 residents, 2024	8.94
Structural county-year destination cells	11,789

Notes: Movement rates are based on physicians observed in consecutive years. Sample and geographic-support rules are described in Section 3.1 and Appendix A.1.

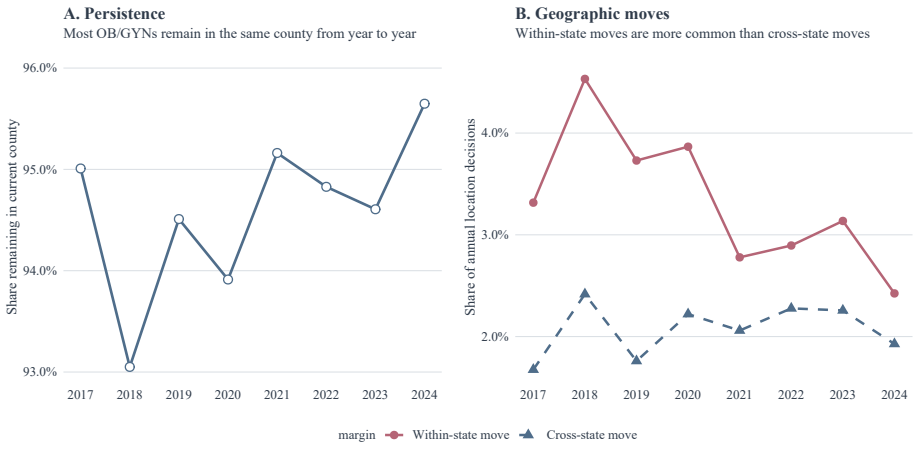


Figure 3: OB/GYN geographic persistence and annual move margins. The figure reports the annual shares of physicians who remain in the same county, move within state, or move across state lines over 2017–2024.

4 Reduced-Form Evidence

I first examine how state legal permissiveness relates to geographic changes in OB/GYN practice locations. The reduced-form analysis compares OB/GYNs with general surgeons within the same county-year, allowing local healthcare-market conditions shared by the two specialties to be absorbed by fixed effects. General surgery is an economically useful comparison because it faces similar physician labor-market conditions while being less directly exposed to abortion-specific rules. The identifying assumption is that, conditional on the fixed effects, other changes associated with the legal environment do not differentially shift OB/GYN outcomes relative to general surgery. The comparison allows shared local shocks to cancel, but an opposite-direction change specific to general surgery would increase the estimated specialty differential. The general-surgery panel contains 26,986 physicians and 150,502 consecutive-year transitions, compared with 41,066 OB/GYNs and 225,108 transitions; every stacked county-year has one observation for each specialty.

4.1 Empirical specification and identification

For the OB/GYN-only panel, I estimate

$$Y_{ct} = \beta P_{s(c)t} + \alpha_c + \lambda_t + \varepsilon_{ct}, \tag{3}$$

where Y_{ct} is a county-year outcome, $P_{s(c)t}$ is standardized cumulative permissiveness change, and α_c and λ_t are county and year fixed effects. State-clustered standard errors allow outcomes for counties within the same policy jurisdiction to be correlated (Cameron and Miller, 2015). As higher P denotes more permissiveness, the coefficient for an additional one-standard-deviation tightening is $-\beta$.

I examine geographic arrivals, departures, net flow, physician stocks per 100,000 residents, and the extensive margin of having any physician in the county. A move is recorded only when the physician is observed in consecutive years. Arrivals and departures are measured per 100,000 county residents; net geographic flow equals arrivals minus departures. Within-state and cross-state arrivals partition total arrivals. These outcomes distinguish the adjustment of existing practice locations from changes in the observed national workforce.

The cross-specialty triple-difference specification stacks OB/GYNs and general surgeons:

$$Y_{cgt} = \beta^{DDD}\left(P_{s(c)t} \times \mathbf{1}\{g = \text{OB}\}\right) + \zeta_{ct} + \psi_{cg} + \tau_{gt} + e_{cgt}, \tag{4}$$

where $g \in \{\text{OB}, \text{GS}\}$. County-by-year effects ζ_{ct} absorb all local shocks common to the two specialties, including the permissiveness main effect. County-by-specialty effects ψ_{cg} absorb permanent local specialty composition, and specialty-by-year effects τ_{gt} absorb national specialty-specific trends. With exactly two specialties, equation (4) is algebraically equivalent to regressing the county-year OB/GYN-minus-general-surgery outcome gap on $P_{s(c)t}$ with county and year fixed effects; the final code verifies this coefficient equivalence exactly. The DDD therefore does not create new state-policy variation. It changes the comparison by differencing out shocks shared by the two specialties.

For the preferred arrivals specification, I first difference this specialty gap over consecutive years. Let $G_{ct} = Y_{c,\text{OB},t} - Y_{c,\text{GS},t}$. I estimate

$$\Delta G_{ct} = \beta^{FD} \Delta P_{s(c)t} + \lambda_t + \Delta e_{ct}, \tag{5}$$

where $\Delta G_{ct} = G_{ct} - G_{c,t-1}$ and $\Delta P_{s(c)t} = P_{s(c)t} - P_{s(c),t-1}$. Year fixed effects λ_t absorb common annual changes, and standard errors are clustered by state. This is the first-difference transformation of the two-specialty gap specification: permanent county differences drop out, while the general-surgery comparison remains built into G_{ct}. I give this specification priority because it directly relates year-to-year changes in the specialty-specific outcome gap to year-to-year changes in the legal environment, rather than attributing changes in the gap to accumulated differences in the policy index. It is also much less sensitive than the levels estimate to excluding small counties.

The triple-difference identifies changes in the OB/GYN-versus-general-surgery outcome gap associated with the policy index, after accounting for county-year conditions shared across specialties and national changes specific to each specialty. I use one- and two-year future-policy values, a 2017–2019 pre-period comparison, and partial-residual plots to examine the timing and specification of this relationship. The corresponding estimates and inference are reported in Appendix B.6.

4.2 Reduced-form results

The first-difference specification is the primary analysis of geographic flows. It relates annual changes in the OB/GYN-minus-general-surgery arrivals gap to annual changes in permissiveness, net of common annual shocks. Table 3 reports the preferred estimates, and Table 4 reports the supporting levels results.

A one-standard-deviation increase in legal permissiveness is associated with 0.0892 additional annual OB/GYN arrivals per 100,000 residents *relative to general surgery*

Table 3: Preferred first-difference DDD: OB/GYN relative to general surgery

Outcome per 100,000 residents	Coefficient	State-cluster SE	Wild p
Geographic arrivals	0.0892**	(0.0288)	0.0140
Within-state arrivals	0.0690**	(0.0235)	0.0217
Cross-state arrivals	0.0202	(0.0130)	0.1127
Geographic departures	0.0098	(0.0257)	0.7063
Net geographic flow	0.0795*	(0.0385)	0.0666
County-year differences	21,945		
State-level jurisdictions / clusters	50 (49 states plus DC)		
Year fixed effects	Yes		

Notes: Outcome and treatment are annual first differences of the OB/GYN-minus-general-surgery county-year gap and standardized permissiveness index, respectively. Arrivals equal within-state plus cross-state arrivals; net flow equals arrivals minus departures. The component coefficients are accounting decompositions, not independent pieces of evidence. Standard errors are clustered by jurisdiction. Wild p-values use the same restricted state-level Rademacher wild-cluster-t algorithm, 9,999 draws, and a prespecified seed for each reported outcome and weighting specification. Stars follow the reported wild p-values: $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.10$.

(wild-cluster $p = 0.0140$). This is equivalent in magnitude to 0.89 arrivals per million residents, or about 26 percent of the unweighted 2017–2021 mean OB/GYN arrivals rate of 0.3369 per 100,000. The percentage is a scale comparison, since the coefficient is estimated from the specialty difference rather than from OB/GYN arrivals alone. I also estimate the first-difference specification separately for OB/GYNs and general surgeons. The OB/GYN-only coefficient is +0.0312 (wild-cluster $p = 0.2985$), while the general-surgery coefficient is −0.0581 ($p = 0.0243$). Taking the difference gives the preferred DDD estimate of +0.0892, up to rounding. Each single-specialty regression can also reflect changes in local healthcare markets that are correlated with policy. The DDD compares the two specialties within the same county-year, removing local shocks that affect both groups in the same way. The departure differential is 0.0098 ($p = 0.706$); consequently, the net-flow coefficient is 0.0795 ($p = 0.067$). The within-state arrivals differential is 0.0690 ($p = 0.0217$), about 77 percent of the 0.0892 point estimate; the cross-state component is 0.0202 ($p = 0.1127$). These are additive coefficient components, not shares of individual physician moves. These flow margins are consistent with policy-related changes in where physicians establish practice, with the strongest estimated differential on arrivals.

In supporting levels regressions, the arrivals differential is 0.0503 per 100,000 (wild-cluster $p = 0.136$). Since net flow is arrivals minus departures, the levels net-flow coefficient of 0.0463 reflects the arrivals estimate less the 0.0041 departure differential. The coefficient of the preferred first-difference design is a differential in the OB/GYN-minus-general-

Table 4: Levels specification: continuous permissiveness change and county-level physician outcomes

	OB/GYN only (1)	OB/GYN relative to general surgery (2)
Geographic arrivals per 100,000	0.0170	0.0503
	(0.0180)	(0.0261)
Net geographic flow per 100,000	0.0281*	0.0463*
	(0.0165)	(0.0213)
Geographic departures per 100,000	−0.0111	0.0041
	(0.0174)	(0.0217)
Physicians per 100,000	−0.0296	−0.0197
	(0.0349)	(0.0495)
County fixed effects	Yes	–
Year fixed effects	Yes	–
County × year fixed effects	–	Yes
County × specialty fixed effects	–	Yes
Specialty × year fixed effects	–	Yes
State-clustered standard errors	Yes	Yes
Observations	25,080	50,160
County-year pairs	25,080	25,080

Notes: Column (1) estimates equation (3). Column (2) estimates equation (4); each county-year contains one OB/GYN observation and one general-surgery observation. The restriction-index change is standardized using the raw-index standard deviation of 2.8638. Coefficients report the response to one standard deviation of the index. In the observed data, most variation reflects differential tightening, so the tightening effect is the negative of the displayed coefficient. Standard errors clustered by state are in parentheses. Stars in column (1) use state-clustered asymptotic p-values; stars in column (2) use the 9,999-draw restricted Rademacher wild-cluster bootstrap reported in Appendix Table B.9. Appendix Tables B.8–B.10 report policy-index and timing diagnostics. Additional extensive-margin, cross-state-flow, raw-count, and stock-accounting outcomes are reported in Appendix Table B.2. $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

surgery outcome gap, rather than an OB/GYN-only proportional effect.

The preferred regression weights county-year observations equally to describe geographic allocation across local markets. Weighting by 2017 county population gives an arrivals coefficient of 0.0420 (state-cluster SE 0.0161; wild-cluster $p = 0.0220$) on the identical first-difference sample. Thus the estimated per-capita response is smaller under resident weighting, although positive under either weighting rule. Winsorizing the first-differenced arrivals rate gives 0.0473 ($p = 0.0321$), reinforcing the importance of reporting magnitude sensitivity rather than treating the 26-percent benchmark as invariant. The corresponding population-weighted departures differential is 0.0349 ($p = 0.0673$), so weighted arrivals and departures largely offset, leaving a net-flow differential of 0.0071 ($p = 0.673$). The arrivals

differential remains positive and statistically significant under both weighting schemes, while the implied geographic net-flow differential is sensitive to county weighting. These estimates describe OB/GYN flows *relative to general surgery*, not OB/GYN-only turnover. Table B.5 reports the complete comparisons.

The positive arrivals differential is present under alternative severity-index constructions, population restrictions, and outcome winsorization. In the preferred first-difference timing diagnostics, the one- and two-year future-policy-change coefficients are small and not statistically distinguishable from zero (joint state wild-cluster $p = 0.8064$); the contemporaneous estimate remains stable. Appendix B.7 reports the specifications and same-sample comparisons. The joint policy-component estimates point toward a role for the accumulation of non-ban restrictions rather than total bans alone: the non-ban coefficient is positive, while the total-ban coefficient is imprecise. This is suggestive of the importance of the broader regulatory environment, not a separately identified causal effect of individual provisions or a statistically established difference between the two components. The exploratory categorical comparison provides a further contrast in the net-flow outcome: the restrictive-versus-permissive net-flow coefficient is approximately -0.2105 for OB/GYNs and $+0.0017$ for general surgeons, giving a specialty differential of -0.2121. This supports an OB/GYN-specific net-flow relationship in the categorical comparison; the preferred continuous first-difference arrivals result remains a differential between both specialties. The corresponding specialty-specific FD table, structural-support sample comparison, and alternative specifications are documented in Appendices B.3–B.6 and B.1.

These results show that physician flows respond along margins that aggregate stocks can conceal. To connect county-level inflows with the incentives behind physicians' location decisions and to evaluate alternative policy paths, I next estimate a discrete-choice model of annual practice locations.

5 Model

5.1 Decision problem

Each year, physician i chooses a county in which to practice. For a physician observed in consecutive years, the previous-year county $o_{it} \equiv j_{i,t-1}$ is the incumbent location, and the current county j is chosen from the feasible set $\mathcal{C}_{it}$. A physician can remain in o_{it} or move to another county.

Location choices are highly persistent, so the model gives the incumbent county a stay premium in addition to its general destination attractiveness and moving-distance advantage. This premium includes a persistent physician-specific component, allowing physicians with different mobility histories to have different degrees of location attachment. The legal environment may affect the common attractiveness of a destination and the additional value of staying for an established incumbent. I estimate these margins separately and hold their policy slopes common across physicians in the baseline model.

The choice model conditions on physicians observed practicing in consecutive years and explains where these active physicians locate. Retirement, entry into practice, and specialty switching are outside the modeled choice set. The counterfactuals therefore trace geographic redistribution within the active physician population.

The annual choice problem is static: a physician evaluates current location payoffs, the incumbent location, and a persistent stay preference. In the counterfactual simulation, each year's chosen county becomes the incumbent location in the following year, producing a sequence of location choices with accumulated geographic effects.

5.2 Utility

Physician i receives the following utility from practicing in county j in year t:

$$U_{ijt} = V_{ijt} + \varepsilon_{ijt}, \tag{6}$$

Here ε_{ijt} is an idiosyncratic type-I extreme-value taste shock, and V_{ijt} collects the destination and individual-specific utility components. Conditional on the persistent stay taste, this utility index is

$$\begin{aligned} V_{ijt} = \delta_{jt} + (\alpha + \sigma_S \nu_i)\, S_{ijt} + \gamma_I P_{s(j)t} S_{ijt} + \gamma_B B_{s(j)} S_{ijt} \\ + \beta_d \log(1 + d_{ij}) + \beta_h H_{ij}, \end{aligned} \tag{7}$$

Here $S_{ijt} = \mathbf{1}\{j = o_{it}\}$ indicates that the county is the physician's incumbent location. The policy variable $P_{s(j)t}$ is standardized cumulative permissiveness change, while $B_{s(j)}$ is the state's 2015 legal environment, standardized by its 2015 cross-state dispersion rather than by the cumulative-change normalization used for P_{st}. Distance d_{ij} measures kilometers between the incumbent and candidate counties. The indicator H_{ij} captures the additional utility of remaining within the incumbent state, beyond the influence of distance. The

persistent physician-specific stay taste satisfies

$$\nu_i \sim N(0,1) \tag{8}$$

and is held fixed across all of physician i's observed annual choices.

The county-year destination effect δ_{jt} captures the common attractiveness of county j in year t to physicians who consider that destination, including practice opportunities, demand, earnings, amenities, and local conditions. The first stage estimates these effects without restricting their relationship to policy. The second stage then relates the recovered destination utilities to legal permissiveness and fixed effects. This separates destination attractiveness from the additional payoff to physicians who already practice in a location.

The mean stay premium α is the extra utility from remaining in the incumbent county after accounting for its destination attributes, distance, and same-state advantage. The random term $\sigma_S \nu_i$ permits persistent differences in this attachment across physicians. The incumbent interaction $\gamma_I P_{s(j)t} S_{ijt}$ measures how this additional stay payoff varies with changes in the legal environment. The baseline interaction $\gamma_B B_{s(j)} S_{ijt}$ allows incumbent attachment to differ with the state's initial legal regime, accounting for pre-existing geographic sorting.

5.3 Identification

The choice model identifies incumbent attachment by comparing physicians who evaluate the same county-year as their current practice location or as an alternative destination. The county-year effect δ_{jt} captures the attractiveness shared by all physicians considering that destination, while the stay indicator identifies the additional value of remaining in the incumbent county, conditional on distance and same-state mobility costs.

The incumbent-policy coefficient γ_I is identified by how the incremental stay value varies with legal permissiveness across locations and years. In the same county-year, some physicians evaluate the county as an incumbent and others as a prospective destination; the interaction distinguishes these valuations. The baseline interaction γ_B allows the incremental stay value to vary with the 2015 legal environment.

Repeated physician mobility histories identify the dispersion in persistent incumbent attachment, σ_S. Physicians who remain in place consistently and those who move more often provide information about the distribution of stay preferences, beyond the mean stay premium α.

The destination-policy margin is estimated from the county-year utilities recovered

in the first stage. Since legal permissiveness is common to counties within a state-year, its direct contribution cancels between two nonincumbent destinations in the same state. Relative attractiveness across states can nevertheless change, and the resulting choice probabilities may also redistribute physicians among counties within states. At the population level, write destination utility as $\delta^0_{jt} = \mu_j + \tau_t + \gamma_D P_{s(j)t} + \xi_{jt}$, where ξ_{jt} captures time-varying unobserved destination characteristics. The identifying condition is that residualized policy variation is orthogonal to ξ_{jt} after county and year effects. The empirical second stage uses estimated destination effects $\widehat{\delta}_{jt}$, centered within year, with the year-specific centering absorbed by the year fixed effects, and takes the form

$$\widehat{\delta}_{jt} = \mu_j + \tau_t + \gamma_D P_{s(j)t} + u_{jt}, \tag{9}$$

where μ_j and τ_t are county and year fixed effects. The destination parameter γ_D measures the change in common destination attractiveness associated with legal permissiveness, net of permanent county and common annual differences. A positive γ_D raises the relative utility of counties whose legal environments become more permissive. The residual u_{jt} combines time-varying unobserved destination characteristics, ξ_{jt}, with estimation error in the recovered county-year destination utilities.

The baseline second-stage specification includes policy, county fixed effects, and year fixed effects. I also examine specifications adding local earnings, hospital characteristics, OB bed capacity, fertility, and contemporaneous physician density to assess the stability of the destination-policy gradient. The additional controls are presented with the robustness results; some contemporaneous outcomes may themselves respond to policy and physician sorting.

The location-choice model has two policy margins: γ_D shifts the common utility of a county as a destination, while γ_I shifts the additional value of that county to a physician already practicing there. Their combined incumbent utility slope is $\gamma_D + \gamma_I$, although actual changes in stay probabilities also depend on the utilities of competing destinations. The estimates and their implications are discussed separately in Section 7.

5.4 Choice probabilities

Conditional on physician i's persistent stay preference ν_i, the probability of choosing county j follows the standard multinomial logit form (McFadden, 1974):

$$Pr(j_{it} = j \mid \nu_i) = \frac{\exp\{V_{ijt}(\nu_i)\}}{\sum_{k \in \mathcal{C}_{it}} \exp\{V_{ikt}(\nu_i)\}} \tag{10}$$

Because ν_i is persistent, the physician likelihood integrates the product of annual choice probabilities,

$$L_i = \int \prod_{t \in \mathcal{T}_i} \Pr(j_{it} = j_{it}^{\text{obs}} \mid \nu) \phi(\nu) \, d\nu. \tag{11}$$

This panel random-coefficients logit (mixed logit) follows the discrete-choice framework of McFadden and Train (2000) and the treatment in Train (2009). The persistent random stay coefficient allows repeated location decisions to reveal individual heterogeneity in mobility preferences.

6 Estimation

6.1 Choice occasions and support

The structural sample consists of 225,108 annual choices by 41,066 OB/GYNs observed in consecutive years during 2017–2024. The geographic universe contains 3,135 counties, excluding Connecticut in the common estimation geography.

The model identifies county-year destination utilities for 11,789 county-year cells, representing 47.0 percent of all county-years in the panel but 99.87 percent of observed OB/GYN stock and 91.5 percent of the population. County-years outside this set are typically very thin markets in which a separate destination effect cannot be estimated reliably. Structural destination effects and counterfactual predictions are therefore conditional on identified destination support. Appendix C.1 describes how off-support alternatives are handled in the sampled estimation sets.

Each observed choice lies on the identified county-year destination support. When the previous year's county has no identified destination effect in the current year, the stay alternative is excluded for that occasion and the physician chooses among the remaining supported locations.

6.2 Sampling of alternatives

The structural likelihood evaluates county alternatives for 225,108 physician-year choices. Evaluating all 3,135 counties would entail roughly 706 million individual-alternative evaluations per pass, before repeated optimization and integration over persistent stay preferences. I therefore sample 30 counties per annual choice, following the sampling-of-alternatives approach of McFadden (1978). This makes repeated likelihood evaluations feasible while retaining the observed choices and incumbent alternatives.

The observed chosen county is always retained. When the incumbent county is distinct, it is also retained. Remaining alternatives are drawn uniformly without replacement from the full county universe. Since the chosen and incumbent alternatives are included with different probabilities from the other sampled counties, I adjust the sampled-set likelihood using the conditional sampling correction

$$q_{ijt} = S_{ijt} \log\left(\frac{R-1}{M_t - 1}\right).$$

The offset adjusts the incumbent alternative for its sampling probability relative to a nonincumbent destination. The full derivation and the treatment of sampled off-support counties appear in Appendix C.1. The prediction and counterfactual calculations use the full identified destination support, without the estimation-only sampling correction.

I measure geographic distance between county centroids in kilometers, with $d_{ii} = 0$ for the incumbent county. Its contribution to utility is $\log(1 + d_{ij})$.

6.3 First-stage mixed logit

I estimate the first-stage panel random-coefficients logit by maximum likelihood, using the physician-level likelihood in equation (11). The parameter vector contains destination effects δ_{jt}, the mean stay premium α, persistent stay-preference heterogeneity σ_S, incumbent-policy interactions γ_I and γ_B, distance disutility β_d, and the same-state premium β_h. The likelihood is built from each physician's sequence of annual choices. Since the same stay taste carries across years, repeated stays and moves provide information about both the average stay premium and heterogeneity in attachment. I integrate over the distribution of that taste while recovering an unrestricted destination effect for each supported county-year.

I evaluate the random-coefficient integral with 11-point Gauss–Hermite quadrature in the baseline specification and jointly recover the county-year destination effects and behavioral parameters. The estimated model rejects the restriction $\sigma_S = 0$, supporting persistent differences across physicians in their preference for staying. Appendix C.1 reports estimation support, numerical stability, and quadrature details.

Model fit is assessed by comparing model-implied mobility patterns with their empirical counterparts. I focus on the annual probability of remaining in the incumbent county and the distribution of move distances, since these moments summarize the persistence and geographic mobility margins that are central to the location-choice model.

6.4 Second-stage destination attractiveness

The first stage recovers a county-year destination utility after accounting for incumbent attachment and geographic mobility costs. I use these recovered utilities to estimate equation (9) by unweighted ordinary least squares on county-year cells for which $\widehat{\delta}_{jt}$ is identified. County fixed effects account for persistent differences in the attractiveness of a location, and year fixed effects absorb changes common to all locations. The remaining within-county changes in permissiveness identify γ_D. Standard errors are clustered at the state level to allow for correlation across counties in the same policy jurisdiction (Cameron and Miller, 2015).

The baseline second-stage model includes legal permissiveness with county and year fixed effects. The coefficient γ_D describes how the recovered utility of a destination varies with permissiveness after accounting for those fixed effects. I assess robustness by adding market characteristics, including earnings, hospital share, OB bed capacity, fertility, and physician density; the estimates are reported in Appendix C.2.

6.5 Inference

State policy varies by state and year, whereas the reduced-form data contain many counties per state and the structural data contain repeated choices for physicians within states. I therefore cluster inference at the state level, allowing errors to be dependent within states. The number of state clusters is 50 jurisdictions (49 states and the District of Columbia), so I also use restricted state-level Rademacher wild-cluster bootstrap-t inference with 9,999 draws for the reduced-form regressions and the second-stage OLS regression. This checks the finite-cluster approximation without changing the underlying point estimates.

The first-stage mixed logit requires a different procedure because it estimates behavioral coefficients together with thousands of county-year destination effects. I calculate model-based standard errors from the profile Hessian and use state-clustered numerically profiled scores for the main behavioral-parameter standard errors. This follows the profile-likelihood treatment of nuisance parameters (**?**) while allowing state-level dependence in the profiled scores. Complete physician histories are assigned to the initial incumbent state when constructing state scores. A separate 9,999-draw Rademacher score-multiplier simulation reproduces the first-order cluster covariance without fully reestimating the mixed logit in each draw. It is a linearized covariance diagnostic, not the restricted wild-bootstrap-t test used in the regressions. First-stage significance stars use state-clustered normal-reference p-values, except that inference for the boundary hypothesis $\sigma_S = 0$ uses a boundary-

appropriate likelihood-ratio test. The preferred specification retains γ_B; setting $\gamma_B = 0$ is a robustness check.

The second-stage regression uses recovered county-year utilities and reports state-clustered inference conditional on those first-stage estimates. Appendix C.1 documents the numerical profile-score checks, including the accepted half-step perturbation test covering all 15 high-influence states.

7 Structural Results

7.1 Location preferences and incumbent attachment

Table 5 reports the behavioral parameters from the panel random-coefficients logit. The estimated mean stay premium, $\widehat{\alpha} = 1.0693$, raises the conditional relative odds of remaining in the incumbent county by a factor of about 2.9, beyond the destination's own utility, zero moving distance, and same-state advantage. The standard deviation of the persistent stay preference, $\widehat{\sigma}_S = 1.3482$, indicates substantial heterogeneity in location attachment across physicians. Geographic mobility is also costly: the distance coefficient $\widehat{\beta}_d = -1.3050$ implies that doubling $1 + d$ lowers relative destination-choice odds by about 60 percent. The same-state coefficient $\widehat{\beta}_h = 1.8305$ implies more than six times higher relative odds for an otherwise comparable destination within the incumbent state than outside it. Together, these estimates explain why physician practice locations are highly persistent and why most observed moves remain within state.

The estimated destination-policy parameter is $\widehat{\gamma}_D = 0.0610$ (state-cluster SE 0.0208; wild-bootstrap $p = 0.011$). A one-standard-deviation increase in permissiveness therefore raises the relative odds of choosing a county as a destination by about 6.3 percent, holding the other components of the location-choice utility index fixed. This is the destination-sorting margin of the structural model.

The estimated incumbent-policy interaction, $\widehat{\gamma}_I = -0.1372$ (state-cluster SE 0.0584; $p = 0.0187$), indicates that the additional utility of remaining in the current county is greater in more restrictive legal environments. This incumbent-specific margin operates alongside the county's general destination attractiveness. It is consistent with existing practices, patient relationships, local professional capital, and relocation costs making established physicians less mobile even when legal conditions change. The separate 2015-baseline interaction is $\widehat{\gamma}_B = -0.1212$ (SE 0.0759); it permits incumbent attachment to vary with the initial legal environment and absorbs part of the pre-existing geographic

sorting component.

For the incumbent county, the two estimated policy terms sum to $\widehat{\gamma}_D + \widehat{\gamma}_I = -0.0762$. This is the conditional utility slope for the incumbent alternative, while its effect on the probability of staying depends on the full set of competing destinations. The two policy coefficients therefore describe distinct components of the location-choice response and are interpreted separately; I do not estimate a joint two-stage standard error for their sum.

Together, the results show that more permissive environments attract movers, while more restrictive environments are associated with stronger incumbent attachment. Because most physicians stay in place each year, the destination and incumbent margins allow practice-location flows to adjust before large changes emerge in physician stocks.

Table 5: Structural location-choice estimates

Parameter	Estimate	Profile SE	State-cluster SE
Panel A: mixed-logit location parameters			
Mean stay coefficient, α	1.0693^{***}	(0.0781)	(0.4025)
Incumbent policy interaction, γ_I	-0.1372^{**}	(0.0177)	(0.0584)
2015 baseline × stay, γ_B	-0.1212	(0.0175)	(0.0759)
$\log(1 + \text{distance km})$, β_d	-1.3050^{***}	(0.0177)	(0.1053)
Same-state indicator, β_h	1.8305^{***}	(0.0344)	(0.1292)
SD of persistent stay taste, σ_S	1.3482^{***}	(0.0257)	(0.0545)
Panel B: recovered destination attractiveness			
Permissiveness-change slope, γ_D	0.0610^{**}	–	(0.0208)
Incumbent policy slope, $\gamma_D + \gamma_I$	-0.0762	–	–
Identified county-year destination cells		11,789	
Choice occasions		225,108	
Physicians		41,066	
LR statistic for $\sigma_S = 0$		1,165.2	
State-cluster p-value for γ_I		0.0187	
Wild-cluster p-value for γ_D		0.0111	

Notes: Panel A reports the accepted NQ=11 mixed-logit estimates. "Profile SE" is the model-based profile-Hessian standard error obtained by re-profiling the county-year destination effects at each low-dimensional perturbation. "State-cluster SE" is based on state-level numerically profiled efficient scores, with complete physician histories assigned to the physician's initial incumbent state. Significance stars in Panel A use state-cluster normal-reference p-values, except σ_S, whose star uses the boundary-adjusted likelihood-ratio test. Panel B reports the second-stage estimate of γ_D; its standard error is clustered by state, and its significance star uses the wild-cluster-bootstrap p-value. γ_I and γ_D correspond to a one-standard-deviation change in the cumulative restriction index, while γ_B uses the separately standardized 2015 level. No joint standard error is reported for $\gamma_D + \gamma_I$. $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.10$.

The distance coefficient provides a useful scale for the destination-policy gradient. The

estimate $\widehat{\gamma}_D = 0.0610$ is equivalent in utility terms to a 0.0467 reduction in $\log(1 + d)$. Evaluated at the median observed move distance of 116.4 kilometers, this is approximately the utility gain from shortening a destination by 5.4 kilometers. The same-state coefficient is much larger. At the median move distance, losing the same-state advantage is equivalent to increasing distance from about 116 kilometers to about 476 kilometers. Legal permissiveness is therefore economically meaningful at the margin of destination choice, while distance and state-specific frictions remain larger determinants of geographic mobility.

As a robustness check, I re-estimate the model without allowing incumbent attachment to depend on the initial legal environment. The destination-policy coefficient increases from 0.0610 to 0.0825, implying an increase in relative destination-choice odds from about 6.3 percent to 8.6 percent for a one-standard-deviation increase in permissiveness. The incumbent-policy interaction becomes more negative, from -0.1372 to -0.1868. The 2015 baseline severity and subsequent cumulative permissiveness change are correlated (0.381 across 400 state-year observations), so omitting the separate baseline-state stay interaction alters the variation attributed to the current-policy interaction. The coefficient changes show that retaining this baseline control is substantively important, even though its own estimated slope is imprecise. The magnitudes change, but the signs and economic interpretation of both policy margins remain the same. I therefore retain the richer specification as the preferred model and use it for the main inference and counterfactual exercises. Appendix C reports the restricted estimates.

7.2 Model fit

Table 6 compares model-implied mobility patterns under the observed policy path with their empirical counterparts for the preferred specification used in the counterfactual simulations. The model closely reproduces the central persistence margin: the simulated annual stay rate is 93.0 percent compared with 94.6 percent in the data. It also reproduces the center of the mover-distance distribution, with a simulated median move of 117.2 kilometers compared with 116.4 kilometers in the data. The correlation between simulated and observed county-year physician stocks is 0.9984. Because county stocks are highly concentrated and many county-years contain zero physicians, I treat this correlation as a coarse aggregate fit statistic rather than as a discriminating measure of fit in thin markets.

Fit is weaker for long-distance and interstate moves. Among movers, the model predicts that 30.6 percent cross state lines, compared with 38.3 percent in the data. The simulated mean move distance is 348.6 kilometers compared with 542.7 kilometers in the data, and the simulated 90th percentile is 976 kilometers compared with 1,716 kilometers. The model

therefore captures persistence and typical move distances more closely than the upper tail of geographic mobility. I therefore interpret counterfactual geographic composition cautiously; the direction of any resulting bias in aggregate reallocation is not determined.

Table 6: Model fit under the observed policy path

Moment	Data	Model
Stay rate	94.59%	93.03%
Within-state share among movers	61.74%	69.42%
Cross-state share among movers	38.26%	30.58%
Mean move distance	542.7 km	348.6 km
Median move distance	116.4 km	117.2 km
90th percentile move distance	1,715.5 km	976.3 km
County-year stock correlation	–	0.9984

8 Counterfactual Policy Paths

8.1 Counterfactual design

The counterfactuals ask how the geographic allocation of an existing OB/GYN workforce changes under alternative state legal environments. I consider five policy paths. The narrowest holds each state's permissiveness at its 2021 level after *Dobbs*; a second freezes the 2019 level from 2020 onward. A third partially restores restrictive states toward their 2015 baseline by up to one standard deviation. The main restriction-specific experiment removes all net tightening since 2015 while preserving observed liberalizations. The broadest path returns every state to its 2015 legal environment. These comparisons distinguish the consequences of recent post-*Dobbs* changes from the cumulative effects of earlier restrictive provisions. Monte Carlo standard errors reported in this section and the appendix measure simulation noise conditional on the estimated structural parameters; they do not include parameter-estimation uncertainty.

I use the estimated model to examine how the geographic allocation of the observed OB/GYN workforce would differ under alternative state-level policy permissiveness paths. The counterfactual is simulated year by year. For each physician, the county chosen in year t becomes that physician's incumbent county in year $t+1$. The physician-specific stay preference is held fixed over time, so persistent differences in attachment to the incumbent location carry forward across simulated years. The annual choice rule remains myopic, i.e., there is no forward-looking value function. The simulator evaluates all county-year

destinations on the identified support rather than the 30 sampled alternatives used for estimation. In the final structural geography, that support contains 11,789 of 25,080 county-years, or 47.0 percent, covering 99.87 percent of observed OB/GYN stock and 91.5 percent of population. The counterfactual therefore reallocates physicians among destinations whose county-year effects are identified in the estimation data and does not extrapolate choice probabilities to county-years without identified destination effects.

A counterfactual permissiveness path changes utility through both estimated policy margins. For destination attractiveness, I shift the recovered county-year mean utility by

$$\Delta\delta_{jt} = \widehat{\gamma}_D \left(P^{CF}_{s(j)t} - P^{OBS}_{s(j)t} \right). \tag{12}$$

For an incumbent alternative, the counterfactual policy change also enters through the incumbent-policy interaction, $\widehat{\gamma}_I \left(P^{CF}_{s(j)t} - P^{OBS}_{s(j)t} \right)$. The 2015 baseline-policy component is held fixed, and all remaining county-year destination utility components are held at their estimated values under the observed policy path. The counterfactual therefore changes only the policy terms while keeping the estimated nonpolicy determinants of destination attractiveness fixed. It should be interpreted as a partial-equilibrium reallocation exercise based on the estimated relationship between the legal environment and physician location choice.

The main outcome measures the difference between simulated geographic physician allocations under each policy path and the observed-policy baseline. For year t, define

$$\mathcal{R}_t = \frac{1}{2} \sum_j \left| N^{CF}_{jt} - N^{OBS}_{jt} \right|, \tag{13}$$

where N_{jt} denotes simulated physician stock in county j. The half-L_1 measure $\mathcal{R}_t$ reports the number of physician-equivalents of geographic mass reassigned across counties under the counterfactual. To express its economic scale relative to the mobile portion of the workforce, I compare it with the number of physicians who change counties in a typical year. The resulting percentage is a mover-volume equivalence, since $\mathcal{R}_t$ is a difference between simulated county-stock distributions rather than a count of distinct physicians changing their individual decisions.

Both the observed-policy arm and each counterfactual arm are simulated using common random numbers from a common simulated history. Before a policy path diverges, the arms face the same simulated environment and therefore share the same simulated location history; after divergence, the chosen county in each arm becomes the next year's incumbent.

Accordingly, N_{jt}^{OBS} in equation (13) is the stock simulated under the observed-policy path, not the realized stock in the raw data.

8.2 2021 policy benchmark

I first consider a post-2021 benchmark in which each state's legal environment remains at its observed 2021 level from 2022 onward. Formally,

$$P_{st}^{CF} = \begin{cases} P_{st}^{OBS}, & t \leq 2021, \\ P_{s,2021}^{OBS}, & t \geq 2022. \end{cases} \tag{14}$$

This counterfactual replaces the observed policy changes from 2022–2024 with each state's 2021 policy level. It therefore removes both subsequent tightening and subsequent liberalization.

The 2021 policy benchmark reallocates about 130 OB/GYN physician-equivalents across counties by 2024. It changes the simulated legal environment in 21 of the 50 jurisdictions, yielding a geographically dispersed adjustment concentrated in the location choices of the existing workforce. Appendix D.3 provides state-level changes and simulation uncertainty.

The implied state-level changes are individually small and geographically dispersed, so I treat the aggregate county-level reallocation as the main outcome. Figure 4 reports the state-level changes, and Appendix D.3 reports the corresponding simulation uncertainty.

8.3 Broader policy counterfactuals

The broader policy paths restore increasing portions of the cumulative tightening that began before *Dobbs*. Table 7 reports the resulting geographic reallocation in 2024 for all five experiments. These scenarios are defined in Section 8.1, before the simulation equations.

The no-net-tightening experiment reallocates 427.28 OB/GYN physician-equivalents across counties by 2024. At the observed annual move rate of 5.41 percent, a workforce of 27,181 physicians would generate approximately 1,470 geographic moves in a typical year. The 2024 reallocation represents 1.57 percent of the active simulated workforce and is *equivalent in magnitude to about 29 percent of typical annual mover volume.* This is a comparison of an end-of-horizon geographic stock difference with a typical single-year flow. It highlights the scale of geographic adjustment among a largely immobile workforce

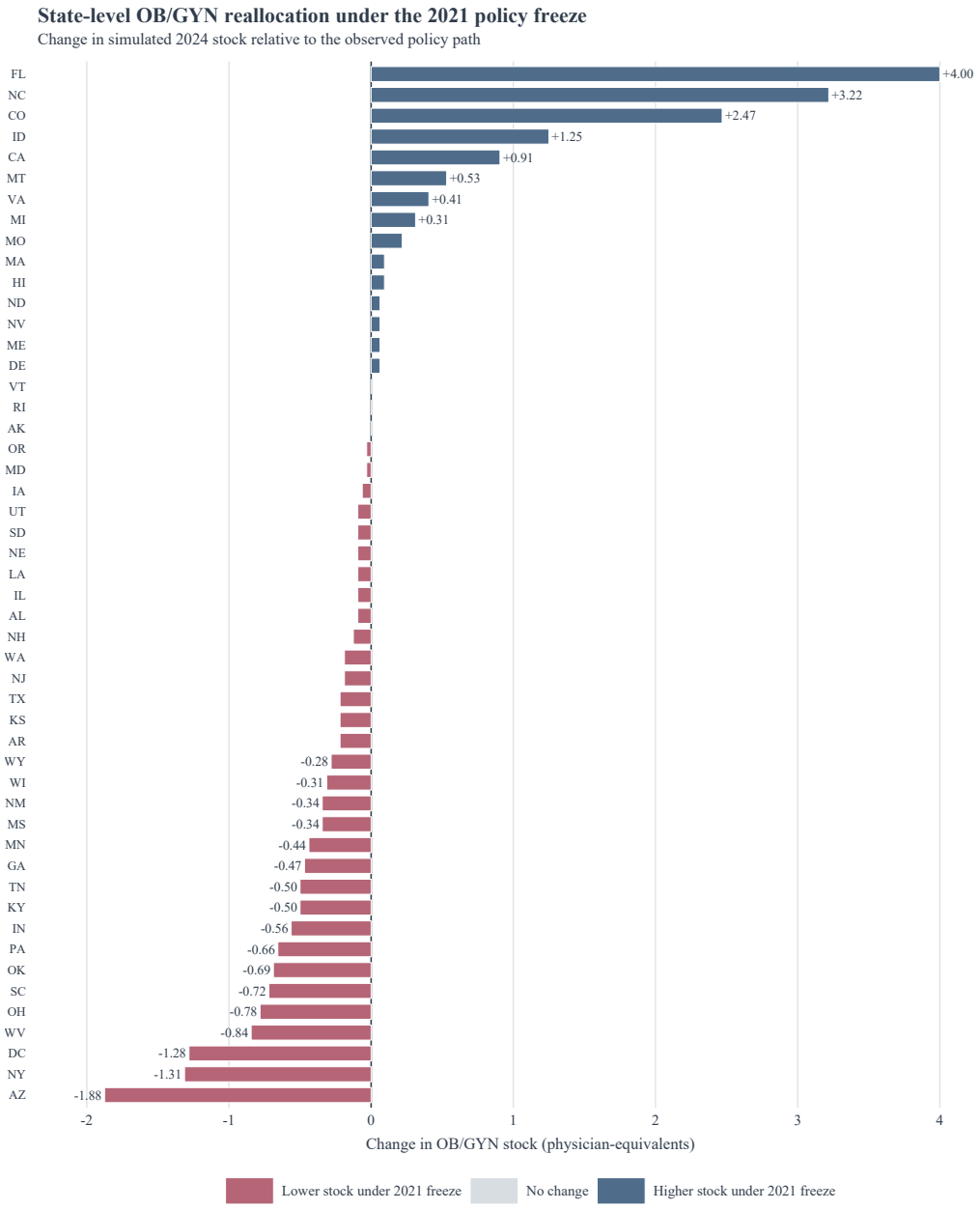


Figure 4: State-level mean changes in simulated 2024 OB/GYN stock under the 2021 policy benchmark. Positive values indicate more physicians under the counterfactual than under the observed policy path. Appendix D.3 reports simulation uncertainty for the largest state-level changes.

but does not mean that 29 percent of individual movers change their chosen county. The narrower 2021 and 2019 policy freezes reallocate about 130 and 194 physician-equivalents, respectively. A one-standard-deviation restoration toward 2015 reallocates about 291, and a full return to the 2015 legal environment reallocates about 431.

I use the no-net-tightening experiment as the main restriction-specific counterfactual because it removes measured tightening while preserving observed liberalizations. The similarity between the 427- and 431-physician estimates reflects the fact that the two policy paths differ in only five state-years: most movement in the primary policy index after 2015 comes from tightening rather than liberalization. Appendix D.1 reports the corresponding destination-only simulation.

I also evaluate how the main counterfactual redistributes physicians across counties with different initial OB/GYN availability. Using observed 2021 physician supply, I group counties into zero-supply, lower-density, and higher-density markets and hold these classifications fixed across policy paths. In the no-net-tightening experiment, the simulated 2024 allocation shifts 31.25 physician-equivalents toward counties with zero or below-median initial OB/GYN availability, including 25.91 to lower-density counties and 5.34 to counties without an observed OB/GYN in 2021. The lower- and zero-supply group's stock rises by approximately 0.54 percent under the experiment. This net shift across supply groups accompanies the larger geographic reallocation across individual counties. Appendix D.2 reports the comparison across all five policy paths.

Because each year's simulated destination becomes the next year's incumbent location, differences in choices accumulate across the policy path. The 2024 reallocation therefore summarizes changes in the geographic stock arising from successive annual mobility decisions.

Table 7: Counterfactual geographic reallocation in 2024

Policy path	Reallocation (physician-equivalents)	Mover-volume equivalent (typical annual moves)
Hold permissiveness at 2021 level	130	8.9%
Hold permissiveness at 2019 level	194	13.2%
One SD toward 2015 baseline	291	19.8%
Remove all net tightening since 2015	427	29.1%
Return to 2015 permissiveness	431	29.3%

Notes: Reallocation is the county-level half-L_1 difference between simulated county-stock distributions in equation (13). The mover-volume equivalent divides each reallocation by approximately 1,470 typical annual geographic moves, calculated as 27,181 active simulated physicians times the observed 5.41% annual move rate. It is a scale benchmark, not the share of unique movers whose decisions change. The main 427-physician-equivalent reallocation is also 1.57% of the active simulated workforce. "Remove all net tightening since 2015" restores negative cumulative policy changes to the 2015 baseline while preserving observed liberalizations; the one-SD path restores up to 2.8638 raw index units. Exact simulated outcomes and conditional Monte Carlo standard errors appear in Appendix D.1.

9 Robustness and Scope

9.1 Policy measurement and model specifications

The reduced-form arrivals pattern is not tied to the severity weights used in the main index. In the levels DDD specification used for the index-construction sensitivity exercise, the arrivals coefficient is 0.0503 under the main severity weights, 0.0529 with equal weights, 0.0482 with physician-risk weights, and 0.0477 using the first principal component of the underlying policy provisions. A joint specification separating total-ban changes from a standardized count of non-ban restrictions estimates −0.0338 (SE 0.0513) for total-ban changes and 0.0463 (SE 0.0211) for the non-ban margin. The total-ban estimate is too imprecise to establish that its effect differs from the non-ban effect, so I treat the component comparison as suggestive rather than a test of distinct causal effects. These alternatives support the use of the severity-weighted index as the primary measure while showing that the motivating arrivals relationship is robust to index construction. Full results are reported in Appendix B.4.

The reduced-form associations and structural parameters have been examined under several alternative specifications. For the reduced form, the arrivals coefficient remains positive when small counties are excluded and when the outcome is winsorized. In the preferred first-difference timing diagnostics, the one- and two-year future policy changes are not statistically significant, either separately or jointly, and the current-policy estimate remains stable. The levels specification has a significant two-year future-policy coefficient, so I report both designs as timing diagnostics rather than definitive pre-trend tests; Appendix B.3 and Appendix B.6 report these exercises. For the structural model, I also re-estimate the specification without allowing incumbent attachment to depend on the 2015 legal environment. The signs and qualitative interpretation of both policy margins are unchanged, although the coefficient magnitudes differ. Appendix C reports the restricted estimates.

I also examine heterogeneity specifications by career stage and quality of medical-school training on both the destination and incumbent margins. The destination point estimates imply roughly 6.3 percent and 6.4 percent higher relative destination odds among experienced and early-career physicians, but the screen standard error of 0.0541 cannot rule out economically meaningful differences. Among physicians with matched training ranks, the corresponding point estimates are 5.5 percent for non-Top-50 and 10.3 percent for Top-50 training; their difference of 0.0447 has a standard error of 0.0763 and $p = 0.560$. These imprecise screens cannot establish equal slopes or rule out substantial

heterogeneity. Neither dimension on either structural margin survives the pre-specified multiplicity adjustments or meets the pre-specified headline rule. I therefore treat the subgroup estimates as descriptive rather than as evidence of differential policy responses. Appendix C.3 reports all four screens and the multiplicity adjustments.

9.2 Scope and limitations

The design uses continuous changes in the legal environment before and after *Dobbs*. The reduced-form estimates compare OB/GYN geographic outcomes with those of general surgery as the legal-policy index evolves. The design relies on the two specialties sharing local healthcare-market shocks after conditioning on fixed effects, rather than on general surgery being entirely unaffected by the legal environment. The timing diagnostics show no statistically detectable future-policy-change coefficients in the preferred first-difference design, while the levels-gap specification has a significant two-year future-policy coefficient. I interpret these as distinct specification-dependent timing patterns rather than direct evidence establishing or ruling out anticipation. The static structural model describes current location incentives and recursively updates incumbent locations in the simulations.

The primary policy index measures the stock and severity of restrictive provisions. It does not mechanically increase when protective statutes are enacted. I therefore interpret the continuous coefficient as a gradient with respect to measured restriction severity, not as the effect of a universal cardinal measure of abortion access. The alternative index constructions and external calibration exercises are intended to assess whether the results depend on the specific coding choices.

The annual structural choice rule is myopic. It does not model expectations about future legal changes. If physicians relocate in anticipation of a future policy change, that anticipatory response is not separately identified from contemporaneous destination attractiveness. The 2021 policy benchmark accordingly inherits any anticipatory sorting already embodied in 2021 locations.

The counterfactuals are restricted to identified destination support. The estimated destination effects cover 11,789 county-years, 47.0 percent of the county-year grid but 99.87 percent of observed OB/GYN stock and 91.5 percent of population. The exercises therefore describe reallocation among destinations whose attractiveness is identified from observed choices; they do not extrapolate the model to every county-year.

The medical-school training heterogeneity screen uses 21,321 physicians with matched training records, or 51.9 percent of the structural sample. The comparisons are interpreted within that matched group.

The structural model conditions on observed participation. It does not model retirement, entry into practice, changes in specialty, or changes in the total number of active OB/GYNs. The counterfactuals therefore reallocate a fixed active physician population across counties. They are location-choice counterfactuals, not workforce-supply counterfactuals.

The model matches the median move distance closely but underpredicts the upper tail of long-distance moves. Counterfactual geographic composition should therefore be interpreted conditional on the estimated mobility kernel; the direction of any resulting bias in aggregate reallocation is not determined.

The physician data cover Medicare-enrolled clinicians observed in the CMS Doctors and Clinicians file. This provides a large national panel with consistent identifiers and practice locations, but it need not include every practicing OB/GYN. Reduced-form stocks and flows and the structural decision sample should therefore be interpreted within that sampling frame.

Second-stage inference conditions on the recovered destination effects. The counterfactual section separately states the scope of simulation uncertainty.

10 Conclusion

This paper studies how local legal environments shape the geographic allocation of physicians. I combine a novel state-year panel of abortion-related legal provisions with longitudinal physician-location data and use variation in a continuous legal permissiveness measure before and after *Dobbs*. The reduced-form evidence is strongest on the arrival margin. In the primary first-difference specification, a one-standard-deviation increase in permissiveness is associated with a 0.089 increase in the annual OB/GYN-minus-general-surgery arrivals gap per 100,000 residents, a magnitude benchmark equivalent to about 26 percent of the 2017–2021 mean annual OB/GYN arrival rate. Departures respond little and contemporaneous stock effects are imprecise, consistent with the high persistence of physician locations.

The structural model provides a behavioral explanation for this pattern. More permissive legal environments are more attractive destinations: a one-standard-deviation increase in permissiveness is associated with about 6.3 percent higher relative odds of choosing a county as a destination. At the same time, physicians are strongly and heterogeneously attached to their incumbent locations. I find that incumbent persistence is greater in more restrictive legal environments. The model therefore separates two margins that aggregate stock data can obscure, where physicians choose to move and where incumbent physicians

remain, and shows how both contribute to gradual geographic adjustment.

The counterfactuals give this mechanism an economically meaningful scale relative to the physicians who actually make geographic choices. Removing all measured net tightening since 2015 reallocates about 427 physician-equivalents across counties by 2024, or 1.57 percent of the active simulated workforce, or an amount equivalent to approximately 29 percent of a typical year's geographic mover volume for a workforce of this size. The same experiment shifts about 31 physician-equivalents toward counties with lower initial OB/GYN availability. A narrower post-2021 experiment reallocates about 130 physician-equivalents. The mover-volume comparison scales a difference in county distributions against annual mobility and is not a percentage of unique physicians whose decisions changed. The simulations hold national specialty supply fixed and measure where physicians practice.

Taken together, the results show that local legal environments shape physician location through both where physicians move and where they remain. In a workforce with strong location persistence, changes in practice-related legal conditions can alter recruitment and geographic allocation well before large stock changes become visible. Geographic distribution is therefore an important margin of adjustment even when aggregate physician supply changes only gradually.

The policy implication is that changes in the scope, liability, and conditions of clinical practice can operate through physician recruitment and destination choice, even when most established physicians remain in their existing locations. In a workforce with high annual persistence, the relevant short-run adjustment is concentrated among prospective movers. Geographic inflows can therefore change well before substantial movements in the local stock become apparent, with implications for the future distribution of access to specialized care.

More broadly, the results highlight the importance of modeling physician location as a dynamic choice with substantial persistence. Similar mechanisms may arise in other settings in which professional rules vary across jurisdictions and relocation is costly. The framework can be further extended to study additional legal channels, including provider liability, shield protections, and other scope-of-practice regulations, and to estimate heterogeneity across physician career stages and market types with greater precision than the present subgroup screens allow. Incorporating patient-side behavior in future iterations would provide a more complete picture of how physician reallocation translates into healthcare access and welfare across local markets.

References

Barrett, Tyson, Matt Dowle, Arun Srinivasan, Jan Gorecki, Michael Chirico, Toby Hocking, Benjamin Schwendinger, and Ivan Krylov, "data.table: Extension of data.frame," R package 2026.

Bergé, Laurent R., Kyle Butts, and Grant McDermott, "Fast and User-Friendly Econometrics Estimations: The R Package fixest," *arXiv preprint arXiv:2601.21749*, 2026.

Bolduc, Denis, Bernard Fortin, and Marc-André Fournier, "The Effect of Incentive Policies on the Practice Location of Doctors: A Multinomial Probit Analysis," *Journal of Labor Economics*, 1996, *14* (4), 703–732.

Cameron, A. Colin and Douglas L. Miller, "A Practitioner's Guide to Cluster-Robust Inference," *Journal of Human Resources*, 2015, *50* (2), 317–372.

Centers for Medicare & Medicaid Services, "Doctors and Clinicians: National Downloadable File," Provider Data Catalog 2026. Dataset identifier mj5m-pzi6; documentation accessed October 2026.

Chou, Chiu-Fang and Anthony T. Lo Sasso, "Practice Location Choice by New Physicians: The Importance of Malpractice Premiums, Damage Caps, and Health Professional Shortage Area Designation," *Health Services Research*, 2009, *44* (4), 1271–1289.

Felix, Mabel, Laurie Sobel, and Alina Salganicoff, "A Review of Exceptions in State Abortion Bans: Implications for the Provision of Abortion Services," KFF June 2024. Issue brief, updated and expanded from the original May 2023 version.

Finkelstein, Amy, Matthew Gentzkow, and Heidi Williams, "Sources of Geographic Variation in Health Care: Evidence from Patient Migration," *Quarterly Journal of Economics*, 2016, *131* (4), 1681–1726.

Ganguly, Anisha P., Anirban Basu, and Anna M. Morenz, "State-Level Disparities in Residency Applications After Dobbs v Jackson Women's Health Organization," *JAMA Network Open*, 2026, *9* (3), e260286.

Gottlieb, Joshua D. and Sean Nicholson, "Physician Competition: Entry and Substitution," *Journal of Economic Perspectives*, 2026, *40* (2), 117–142.

Gottlieb, Joshua D., Maria Polyakova, Kevin Rinz, Hugh Shiplett, and Victoria Udalova, "The Earnings and Labor Supply of U.S. Physicians," *Quarterly Journal of Economics*, 2025, *140* (2), 1243–1298.

Guttmacher Institute, "Hostile and Supportive Abortion Laws as of January 2020," Guttmacher Institute January 2020. State policy-score table updated January 24, 2020.

Guttmacher Institute, "Abortion After Roe: New Comprehensive Map Tracks Abortion Policies and Statistics for Each State," Guttmacher Institute June 2022. June 9, 2022.

Hammoud, Maya M., Helen K. Morgan, Karen George, Arthur T. Ollendorff, John L. Dalrymple, Dana Dunleavy, Min Zhu, Erika Banks, Bukky Ajagbe Akingbola, and AnnaMarie Connolly, "Trends in Obstetrics and Gynecology Residency Applications in the Year After Abortion Access Changes," *JAMA Network Open*, 2024, *7* (2), e2355017.

Health Resources and Services Administration, "Area Health Resources Files (AHRF)," Bureau of Health Workforce, National Center for Health Workforce Analysis 2026. Annual county-, state-, and national-level health workforce and market data.

Holmes, George M., "Increasing Physician Supply in Medically Underserved Areas," *Labour Economics*, 2005, *12* (5), 697–725.

Hurley, Jeremiah E., "Physicians' Choices of Specialty, Location, and Mode: A Reexamination within an Interdependent Decision Framework," *Journal of Human Resources*, 1991, *26* (1), 47–71.

Jones, Rachel K., Meghan Ingerick, and Jenna Jerman, "Differences in Abortion Service Delivery in Hostile, Middle-ground, and Supportive States in 2014," *Women's Health Issues*, 2018, *28* (3), 212–218.

Khoury, Stephanie, Jonathan M. Leganza, and Alex Masucci, "Health Professional Shortage Areas and Physician Location Decisions," *American Journal of Health Economics*, 2025, *11* (3), 487–527.

McFadden, Daniel, "Conditional Logit Analysis of Qualitative Choice Behavior," in Paul Zarembka, ed., *Frontiers in Econometrics*, New York: Academic Press, 1974, pp. 105–142.

McFadden, Daniel, "Modeling the Choice of Residential Location," in Anders Karlqvist, Lars Lundqvist, Folke Snickars, and Jorgen W. Weibull, eds., *Spatial Interaction Theory and Planning Models*, Amsterdam: North-Holland, 1978, pp. 75–96.

McFadden, Daniel and Kenneth Train, "Mixed MNL Models for Discrete Response," *Journal of Applied Econometrics*, 2000, *15* (5), 447–470.

Mermin-Bunnell, Kellen, Ariana M. Traub, Kelly Wang, Bryan Aaron, Louise Perkins King, and Jennifer Kawwass, "Abortion Restrictions and Medical Residency Applications," *Journal of Medical Ethics*, 2025, *51* (2), 79–86.

Nash, Elizabeth, "State Abortion Policy Landscape: From Hostile to Supportive," Guttmacher Institute August 2019. Historical seven-category policy framework; counts on the maintained page refer to laws in effect as of December 30, 2020.

Nash, Elizabeth and Sophia Naide, "State Policy Trends at Midyear 2021: Already the Worst Legislative Year Ever for U.S. Abortion Rights," Guttmacher Institute July 2021.

Nash, Elizabeth, Lizamarie Mohammed, Olivia Cappello, and Sophia Naide, "State Policy Trends 2019: A Wave of Abortion Bans, But Some States Are Fighting Back," Guttmacher Institute December 2019.

Nash, Elizabeth, Lizamarie Mohammed, Olivia Cappello, and Sophia Naide, "State Policy Trends 2020: Reproductive Health and Rights in a Year Like No Other," Guttmacher Institute December 2020.

Nash, Elizabeth, Rachel Benson Gold, Zohra Ansari-Thomas, Olivia Cappello, and Lizamarie Mohammed, "Policy Trends in the States: 2016," Guttmacher Institute January 2017.

R Core Team, "R: A Language and Environment for Statistical Computing," R Foundation for Statistical Computing, Vienna, Austria 2026.

Small, Kenneth A. and Harvey S. Rosen, "Applied Welfare Economics with Discrete Choice Models," *Econometrica*, 1981, *49* (1), 105–130.

Staiger, Becky, Valentin Bolotnyy, Sonya Borrero, Maya Rossin-Slater, Jessica Van Parys, and Caitlin Myers, "Obstetrician and Gynecologist Physicians' Practice Locations Before and After the Dobbs Decision," *JAMA Network Open*, 2025, *8* (4), e251608.

StatsAmerica, "Population Estimates and Decennial Census," Download Center, Indiana Business Research Center, Indiana University 2025. County population download used in the project; StatsAmerica identifies the U.S. Census Bureau as the source.

Supreme Court of the United States, "Dobbs v. Jackson Women's Health Organization," 2022. 597 U.S. 215 (2022), decided June 24, 2022.

Train, Kenneth E., *Discrete Choice Methods with Simulation*, 2 ed., Cambridge: Cambridge University Press, 2009.

U.S. Census Bureau, "Annual Estimates of the Resident Population for Counties in the United States: April 1, 2020 to July 1, 2024," Population Estimates Program, Vintage 2024 March 2025. County total population estimates; later Census vintages revise the post-2020 time series.

Wickham, Hadley, Mara Averick, Jennifer Bryan, Winston Chang, Lucy D'Agostino McGowan, Romain François, Garrett Grolemund, Alex Hayes, Lionel Henry, Jim Hester, Max Kuhn, Thomas Lin Pedersen, Evan Miller, Stephan Milton Bache, Kirill Müller, Jeroen Ooms, David Robinson, Dana Paige Seidel, Vitalie Spinu, Kohske Takahashi, Davis Vaughan, Claus Wilke, Kara Woo, and Hiroaki Yutani, "Welcome to the Tidyverse," *Journal of Open Source Software*, 2019, *4* (43), 1686.

A Data Construction and Policy Measurement

A.1 Sources, annual coding, and timing sensitivity

The original legal panel combines a hand-coded event ledger (`compiled_unsorted.csv`) with the historical workbook `abortion_policy_panel_2015_2025.xlsx`. The event ledger classifies dated changes in legal provisions by state and direction. In the original construction, the recorded event year is extracted and the most recent event for a provision is carried forward to subsequent years. This is an event-year convention; a recorded enactment or policy event does not necessarily coincide with the year of sustained legal enforcement. The historical workbook supplies the remaining five restrictive flags and overrides their event-ledger counterparts where applicable. The annual panel contains 51 jurisdictions for 2015–2025, including Washington, DC. Protective provisions are tracked separately, but they do not receive positive points in the main index.

The provision definitions, source classifications, and severity weights are given in Table 1. In the original historical workbook, the R2–R6 restrictive columns supply waiting periods, Medicaid funding restrictions, telehealth restrictions, parental involvement, and provider regulations (TRAP), respectively.

The source audit reproduces all 561 original state-year weighted-index observations. All 4,488 event-derived restrictive provision-year entries agree with the locally used original event classification under the recorded-year carry-forward rule, and all 2,805 workbook-derived entries match the original workbook. These checks establish source consistency, not complete legal accuracy. Where enactment, activation, injunction, and overlapping-law histories differ, a single annual binary indicator requires judgment. I therefore evaluate exact-cell alternative timing choices as sensitivity exercises without altering the baseline index or estimation sample.

A.2 Physician-location assignment and geographic support

The physician panel assigns one county to each physician-year from the CMS Doctors and Clinicians National Downloadable File. Geographic moves are defined only for physicians observed in consecutive years. Panel entry, panel exit, and nonconsecutive reappearance are therefore not classified as moves. The reduced-form county universe contains 3,135 counties. Connecticut is excluded from the final common structural geography because of the change in its county classification during the sample period. The final structural sample contains 41,066 physicians and 225,108 annual choice occasions.

Table A.1: Selected legal-timing and category sensitivity checks: first-difference arrivals DDD

Policy coding scenario	Estimate	State-cluster SE	p (asymptotic)
Original index	0.08923	0.02883	0.00325
Seven-state pre-enforcement timing	0.10289	0.02728	0.00044
Seven-state timing plus IA/FL	0.10883	0.02616	0.00013
Prior category corrections only	0.09034	0.03161	0.00624
MS/IN/ND overlap scenario (exploratory)	0.07409	0.03043	0.01857

Notes: All specifications use the same 21,945 county-year differences, 50 state clusters, original policy normalization, and first-difference regression. The seven-state scenario changes identified pre-enforcement coding for AL, AR, GA, ID, KY, MO, and TN. The expanded scenario also adjusts selected IA and FL implementation years. Prior category corrections concern CO, NC, NE, and NH. The overlap scenario involving MS, IN, and ND is exploratory because partial-year enforcement and overlapping provisions complicate annual classification. The reported p-values in this table are state-clustered asymptotic values, not wild-cluster p-values. Separately computed 9,999-draw restricted wild-cluster p-values in the fixed-seed audit are 0.0037 for the seven-state scenario, 0.0021 for the expanded IA/FL scenario, and 0.0391 for the exploratory MS/IN/ND scenario. These tests examine the stated correction scenarios, not the legal accuracy of all index cells. The baseline 9,999-draw wild-bootstrap p-value for arrivals is 0.0140. These tests establish robustness to specified coding alternatives, not full state-by-state legal enforceability certification.

The structural model identifies 11,789 county-year destination effects. These cells represent 47.0 percent of the 25,080 county-years in the 2017–2024 grid, but cover 91.5 percent of population and 99.87 percent of observed OB/GYN stock. Every observed chosen destination in the structural sample lies on identified support. Table A.2 summarizes the support used for estimation and counterfactual simulation.

Table A.2: Structural sample and destination support

Statistic	Value
Physicians	41,066
Annual choice occasions	225,108
County universe	3,135
County-years, 2017–2024	25,080
Identified county-year destination cells	11,789
Share of county-year grid on identified support	47.0%
Share of population on identified support	91.5%
Share of observed OB/GYN stock on identified support	99.87%

Notes: Choice occasions require consecutive physician-year observations. The estimation sample uses the identified destination support described in Section 6. Counterfactual predictions evaluate the full identified support rather than the sampled estimation alternatives.

A.3 Policy-index calibration and categorical support

I use two distinct calibration exercises for different purposes. First, I compare the rebuilt absolute 2022 restriction index with Guttmacher's state policy classification (Guttmacher Institute, 2022). I code states whose Guttmacher category contains "hostile" as restrictive and search cutoffs in 0.5-index-point increments over the observed 2022 range. The best-fitting cutoff is -6.0 and agrees with the external classification for 50 of 51 jurisdictions, or 98.0 percent; Michigan is the only disagreement. Because the cutoff is selected to maximize agreement on this same 2022 cross section, this is a scale calibration rather than an out-of-sample validation. It is not used to estimate the continuous reduced-form or structural specifications.

Second, the exploratory three-way categorical reduced-form design uses thresholds calibrated to the Guttmacher Institute's 2020 historical policy classification (Nash, 2019; Guttmacher Institute, 2020). Index values at or below -4.75 are classified as restrictive, values at or above -1.00 as permissive, and values between those cutoffs as intermediate. On the 49 jurisdictions common to the 2020 benchmark and the final estimation geography, this rule has 85.7 percent agreement, balanced accuracy of 0.846, macro-F1 of 0.794, and a Spearman correlation of 0.831. The -6.0 binary calibration above is not used for these three-way regressions.

The jurisdiction counts refer to different objects: 51 jurisdictions enter the 2022 binary calibration, 49 enter the 2020 three-way calibration overlap, 50 enter the final estimation geography because Connecticut is excluded for geographic reasons, and Table A.4 displays all 50 states plus DC using the full policy-only panel. The three-way categorical design has only 13 state transitions over 2018–2023, and the permissive group shrinks from 10 states in 2021 to 5 in 2023–2024. More importantly, the restriction index records statutes on the books rather than effective legal availability. I therefore retain the three-way categories only as an exploratory description of the index and do not use their coefficients as headline evidence.

A.4 State policy index and category in 2024

Table A.4 reports the 2024 absolute policy-index value and corresponding restrictive, intermediate, or permissive category for all 50 states and the District of Columbia. The values are generated from the same policy source used in estimation. Connecticut is recovered from the full 51-jurisdiction policy panel because it is excluded from the common estimation geography for reasons unrelated to policy measurement.

Table A.3: Support for the absolute policy categories

Year	Restrictive	Intermediate	Permissive	Category transitions
2021	29	11	10	2
2022	30	13	7	4
2023	29	16	5	3
2024	29	16	5	0

Notes: Counts refer to the final 50-jurisdiction estimation geography (49 states and the District of Columbia). Across the full 2018–2023 period there are 13 state category transitions. "Restrictive" is the paper presentation label for the label in the original reduced-form construction.

Table A.4: State policy index and category, 2024

State	Abbrev.	Absolute index	Category
Alabama	AL	-14.00	Restrictive
Alaska	AK	-3.00	Intermediate
Arizona	AZ	-13.00	Restrictive
Arkansas	AR	-14.50	Restrictive
California	CA	-2.50	Intermediate
Colorado	CO	-4.50	Intermediate
Connecticut	CT	-2.00	Intermediate
Delaware	DE	-2.00	Intermediate
District of Columbia	DC	-4.50	Intermediate
Florida	FL	-11.50	Restrictive
Georgia	GA	-11.50	Restrictive
Hawaii	HI	-2.50	Intermediate
Idaho	ID	-18.50	Restrictive
Illinois	IL	0.00	Permissive
Indiana	IN	-16.00	Restrictive
Iowa	IA	-12.00	Restrictive
Kansas	KS	-10.50	Restrictive
Kentucky	KY	-19.00	Restrictive
Louisiana	LA	-15.50	Restrictive
Maine	ME	-5.00	Restrictive
Maryland	MD	-4.00	Intermediate
Massachusetts	MA	-3.00	Intermediate

Continued on next page

Table A.4 – continued

State	Abbrev.	Absolute index	Category
Michigan	MI	-3.50	Intermediate
Minnesota	MN	-4.00	Intermediate
Mississippi	MS	-14.00	Restrictive
Missouri	MO	-14.00	Restrictive
Montana	MT	-6.50	Restrictive
Nebraska	NE	-14.00	Restrictive
Nevada	NV	-3.00	Intermediate
New Hampshire	NH	-5.50	Restrictive
New Jersey	NJ	0.00	Permissive
New Mexico	NM	0.00	Permissive
New York	NY	-3.50	Intermediate
North Carolina	NC	-11.50	Restrictive
North Dakota	ND	-12.00	Restrictive
Ohio	OH	-13.00	Restrictive
Oklahoma	OK	-15.00	Restrictive
Oregon	OR	-2.50	Intermediate
Pennsylvania	PA	-6.00	Restrictive
Rhode Island	RI	-2.00	Intermediate
South Carolina	SC	-11.50	Restrictive
South Dakota	SD	-16.50	Restrictive
Tennessee	TN	-16.00	Restrictive
Texas	TX	-21.00	Restrictive
Utah	UT	-10.50	Restrictive
Vermont	VT	0.00	Permissive
Virginia	VA	-2.00	Intermediate
Washington	WA	0.00	Permissive
West Virginia	WV	-14.00	Restrictive
Wisconsin	WI	-7.00	Restrictive
Wyoming	WY	-2.00	Intermediate

Notes: The absolute index is the 2024 value of the paper's restriction-severity index, $p^{level}_{s,2024}$, constructed from the same underlying policy panel used in the structural and reduced-form analyses. Lower values indicate more restrictive legal environments. Categories use the pre-specified calibrated cutoffs: restrictive if the index is ≤ -4.75, permissive if the index is ≥ -1.00, and intermediate otherwise. The common reduced-form and structural estimation geography excludes Connecticut, but Connecticut's displayed index is recovered directly from the same 51-jurisdiction policy source (`policy_panel_consolidated.rds`); no external value or imputation is used.

B Additional Reduced-Form Results

B.1 Exploratory categorical estimates

Table B.1 reports the pre-specified categorical DDD estimates for transparency. These estimates use the $-4.75/-1.00$ cutoffs described in Appendix A.3. They are not treated as primary evidence because the categories have limited within-state switching and do not map cleanly to effective legal availability in the middle of the index distribution. The reported standard errors use asymptotic state-clustered inference and are not adjusted for multiplicity.

Table B.1: Exploratory absolute-level policy categories and county-level outcomes

	Restrictive vs. intermediate	Permissive vs. intermediate	Restrictive vs. nonrestrictive	Restrictive vs. permissive
Physicians per 100,000	−0.2010	0.2649	−0.2111	−0.4658
	(0.0836)	(0.3053)	(0.0802)	(0.2966)
Geographic arrivals per 100,000	−0.2359	0.0281	−0.2370	−0.2640
	(0.0786)	(0.0563)	(0.0786)	(0.0999)
Geographic departures per 100,000	−0.1093	−0.0574	−0.1071	−0.0519
	(0.0514)	(0.0480)	(0.0516)	(0.0648)
Net geographic flow per 100,000	−0.1266	0.0855	−0.1299	−0.2121
	(0.0642)	(0.0484)	(0.0640)	(0.0780)

Notes: Pre-specified categorical DDD coefficients; standard errors clustered by state are in parentheses. These contrasts are correlated and no multiple-testing adjustment is applied. The table is exploratory for the classification reasons discussed in Appendix A.3.

B.2 Supporting outcomes

Table B.2 reports supporting stock and flow outcomes from the continuous DDD specification that are not used as headline estimates.

Table B.3 expands the restrictive-versus-intermediate categorical comparison to the extensive and cross-state margins. The any-present and cross-state-arrival estimates move in the same direction as the main recruitment results.

Table B.2: Continuous permissiveness change: supporting reduced-form outcomes

Outcome	OB/GYN only	OB/GYN relative to general surgery
Physicians per 100,000	−0.0296	−0.0197
	(0.0349)	(0.0495)
Any specialty physician present	0.0041	0.0048
	(0.0029)	(0.0039)
Geographic arrivals per 100,000	0.0170	0.0503
	(0.0180)	(0.0261)
Geographic departures per 100,000	−0.0111	0.0041
	(0.0174)	(0.0217)
Net geographic flow per 100,000	0.0281*	0.0463*
	(0.0165)	(0.0213)
Cross-state arrivals per 100,000	−0.0013	−0.0019
	(0.0062)	(0.0097)

Notes: Coefficients correspond to one standard deviation of the continuous restriction-index change; the observed identifying variation is predominantly differential tightening. The DDD specification includes county-by-year, county-by-specialty, and specialty-by-year fixed effects. Standard errors are clustered by state. Stars in the DDD column use the 9,999-draw restricted Rademacher wild-cluster bootstrap reported in Appendix Table B.9; stars in the OB/GYN-only column use state-clustered asymptotic p-values. Supporting outcomes in this table are reported for completeness and are not used to define the paper's main mechanism.

Table B.3: Restrictive versus intermediate environments: supporting DDD outcomes

Outcome	Restrictive vs. intermediate
Physicians per 100,000	−0.2010
	(0.0836)
Any specialty physician present	−0.0188
	(0.0051)
Geographic arrivals per 100,000	−0.2359
	(0.0786)
Geographic departures per 100,000	−0.1093
	(0.0514)
Net geographic flow per 100,000	−0.1266
	(0.0642)
Cross-state arrivals per 100,000	−0.0575
	(0.0235)

Notes: Estimates come from the county-level OB/GYN–general-surgery triple-difference design. Intermediate environments are the omitted category. Standard errors are clustered by state using asymptotic cluster-robust inference. These exploratory contrasts are correlated and no multiplicity adjustment is applied.

B.3 Weighting and first-difference robustness

The preferred first-difference DDD coefficient can be decomposed into separate specialty regressions estimated on exactly the same 21,945 county-year differences. The OB/GYN-only arrivals slope is 0.0312, and the general-surgery-only slope is −0.0581; their difference is the preferred DDD estimate of 0.0892. Thus the positive specialty difference includes both a positive OB/GYN point estimate and a negative general-surgery estimate. Table B.4 reports their separate state-clustered inference. This decomposition describes the DDD comparison rather than substituting either single-specialty estimate for it.

Table B.4: Specialty-specific components of the preferred FD arrivals DDD

FD arrivals per 100,000	Coefficient	State-cluster SE	Wild p
OB/GYN only	0.0312	(0.0259)	0.2985
General surgery only	−0.0581	(0.0244)	0.0243
OB/GYN minus general surgery	0.0892	(0.0288)	0.0140
County-year differences		21,945	
Jurisdiction clusters		50	

Notes: Each specialty-specific specification regresses the annual first difference of its county arrivals rate on the annual change in the original standardized permissiveness index, with year fixed effects and jurisdiction-clustered standard errors. The difference of the two coefficients reproduces the preferred DDD point estimate. Wild p-values use separately seeded 9,999-draw restricted state-level Rademacher tests. The specialty-specific tests are descriptive diagnostics, not independent validation of the DDD identification assumption.

Table B.5: First-difference DDD by county weighting

	Equal county weight		2017-population weight	
Outcome per 100,000	Coefficient (SE)	Wild p	Coefficient (SE)	Wild p
Arrivals	0.0892 (0.0288)	0.0140	0.0420 (0.0161)	0.0220
Departures	0.0098 (0.0257)	0.7063	0.0349 (0.0162)	0.0673
Net flow	0.0795 (0.0385)	0.0666	0.0071 (0.0164)	0.6726
Cross-state arrivals	0.0202 (0.0130)	0.1127	0.0159 (0.0127)	0.2234
County-year first differences		21,945		
Jurisdiction clusters		50		

Notes: Both specifications use the same original fixed-sample annual differenced OB/GYN-minus-general-surgery gaps and policy changes. Population weights are fixed at 2017 county population. State-clustered standard errors are in parentheses. Restricted Rademacher wild-cluster-t p-values use 9,999 draws, with each outcome's seed predetermined in the fixed-seed computational audit. The weighting schemes target different regression objectives.

Table B.6 collects the principal weighting and first-difference checks discussed in

Section 9. The arrivals estimate is larger in first differences. Net flow is shown as a supporting outcome.

Table B.6: Reduced-form robustness across weighting and differencing choices

	Baseline	2017-population weighted	First difference
Panel A. Continuous DDD: +1 SD permissiveness			
Arrivals per 100,000	0.0503	–	0.0892**
	(0.0261)		(0.0288)
Net flow per 100,000	0.0463*	0.0185**	0.0795*
	(0.0213)	(0.0083)	(0.0385)
Panel B. Restrictive versus permissive DDD			
Arrivals per 100,000	−0.2640	−0.1088	−0.4160
	(0.0999)	(0.0419)	(0.0932)
Net flow per 100,000	−0.2121	−0.0946	−0.3810
	(0.0780)	(0.0373)	(0.1476)

Notes: Baseline models use the level of the county outcome. Population-weighted models use fixed 2017 county population weights. First-difference models use annual changes in the outcome and policy index. For the baseline and first-difference continuous DDD columns, stars use 9,999-draw state-level wild-cluster p-values; the corresponding arrivals p-values are 0.136 and 0.0140, and the net-flow p-values are 0.063 and 0.0666. The 9,999-draw first-difference wild-cluster tests use the fixed-seed V8 audit with unchanged point estimates. The population-weighted levels-arrivals cell is not available in the frozen table and should not be inferred from the weighted first-difference regression; that separate result is reported in Table B.5. The population-weighted and exploratory categorical columns in this table retain state-clustered asymptotic inference. Panel B is exploratory because category-transition support is thin and the absolute categories do not map cleanly to effective legal availability.

Table B.7 focuses on the arrivals outcome across county-size and support restrictions. The levels magnitude falls when very small counties are removed. The first-difference coefficient remains positive across these restrictions and after winsorization, although precision decreases under the population thresholds. The structural-support row compares the reduced-form result on the counties where the structural destination effects are identified; it is a sample-correspondence exercise, not a conventional exogenous sample-robustness check.

B.4 Policy-index construction sensitivity

Table B.8 re-estimates the two principal flow outcomes using alternative continuous summaries of the same underlying policy panel. The coefficients are similar across the baseline severity weights, equal weights, physician-risk weights, and the first principal component. The final two rows separate changes in total-ban status from a standardized

Table B.7: Arrival estimates across county-size and support restrictions

Sample or outcome treatment	Levels DDD	First-difference DDD
Baseline	0.0503	0.0892 [0.0140]
Population $\geq$ 10,000	0.0255	0.0725 [0.115]
Population $\geq$ 20,000	0.0114	0.0865 [0.160]
Structural-support county-years	0.0948	0.1253 [0.075]
Winsorized per-capita outcome	0.0204	0.0473 [0.032]

Notes: Entries are OB/GYN arrival coefficients per 100,000 residents for a one-standard-deviation increase in permissiveness relative to general surgery. Bracketed values in the first-difference column are 9,999-draw state-level wild-cluster p-values. The structural-support row is a sample-correspondence comparison and not an exogenous robustness restriction: support is partly determined by observed OB/GYN choices. These estimates are sensitivity and comparability checks rather than alternative headline specifications.

count of non-ban restrictions, allowing the data to estimate these two margins separately rather than imposing one additive scale. The two terms have little collinearity after the fixed effects are absorbed: their residualized correlation is -0.195, the VIF is 1.04, and 26 states contribute changes in total-ban status.

Table B.8: Reduced-form sensitivity to policy-index construction

Policy measure	Arrivals per 100,000	Net flow per 100,000
Main severity weights	0.0503	0.0463
	(0.0261)	(0.0213)
Equal-weight restriction count	0.0529	0.0485
	(0.0233)	(0.0185)
Physician-risk weights	0.0482	0.0456
	(0.0260)	(0.0219)
First principal component	0.0477	0.0398
	(0.0184)	(0.0149)
Change in total-ban indicator	-0.0338	0.0064
	(0.0513)	(0.0562)
1-SD non-ban restriction-count change	0.0463	0.0512
	(0.0211)	(0.0188)

Notes: Each coefficient is the OB/GYN interaction from the continuous county-level triple-difference specification with county-by-year, county-by-specialty, and specialty-by-year fixed effects. Standard errors are clustered by state. The first four rows enter one standardized index at a time. The last two rows come from a joint specification containing a change in total-ban status and a standardized count of non-ban restrictions. The joint Wald p-value for the two terms is 0.069 for arrivals and 0.018 for net flow; the total-ban coefficient itself is imprecisely estimated.

B.5 Finite-cluster inference

Table B.9 applies a common restricted Rademacher wild-cluster bootstrap-t procedure at the state level to the six headline continuous DDD outcomes. The point estimates are unchanged. In levels, net geographic flow has the smaller wild-cluster p-value, while in first differences arrivals is the strongest flow result.

Table B.9: State-level wild-cluster inference for continuous DDD outcomes

Outcome	Estimate	Asymptotic p	Wild p
Physicians per 100,000	−0.0197	0.692	0.701
Any specialty physician present	0.0048	0.230	0.269
Arrivals per 100,000	0.0503	0.060	0.136
Departures per 100,000	0.0041	0.852	0.859
Net geographic flow per 100,000	0.0463	0.035	0.063
Cross-state arrivals per 100,000	−0.0019	0.842	0.841

Notes: The wild p-values use a restricted Rademacher wild-cluster bootstrap-t with clustering by state and 9,999 draws. The bootstrap is applied to the same continuous DDD specification reported in the main text. The procedure changes inference, not the coefficient estimates.

B.6 Continuous-treatment timing diagnostics

I evaluate timing by adding one- and two-year future policy values to the continuous specification and by estimating the relationship over 2017–2019 alone. Table B.10 reports both the contemporaneous and future-policy coefficients using the same 9,999-draw state-level wild-cluster procedure. The contemporaneous arrivals and net-flow coefficients remain positive when either future policy value is included. The two-year arrivals lead is also negative and statistically different from zero, while the net-flow leads are close to zero. Part of the change in the contemporaneous arrivals coefficient relative to the full-sample headline estimate comes from the shorter samples required by the leads: the same-window arrivals baseline is 0.0631 for 2017–2023 and 0.0627 for 2017–2022, before the corresponding future policy term is added. The two-year arrivals lead is −0.0499 in the levels-gap regression and is statistically different from zero under the previously reported wild-cluster test. A separate state-clustered covariance calculation for that regression yields $0.0698 - 0.0499 = 0.0199$ (SE 0.0278; asymptotic $p = 0.474$) for the sum of the current and future policy coefficients. This reparameterization is a statement about the levels regression with future policy, not the preferred first-difference estimand, and does not resolve the significant lead. The 2017–2019 arrivals estimate is also positive (0.1306)

but imprecise (wild-cluster $p = 0.166$) over that short window. I therefore treat these levels exercises as timing diagnostics in an evolving policy environment rather than as a sharp pre-trend test around a single event date.

Table B.10: Timing diagnostics for the continuous DDD specification

Outcome	Current, 1-year model	1-year lead	Current, 2-year model	2-year lead	2017–2019
Physicians per 100,000	0.0362 [0.474]	−0.0267 [0.631]	0.0507 [0.454]	−0.0501 [0.559]	−0.0443 [0.564]
Any specialty physician present	0.0034 [0.422]	0.0039 [0.270]	0.0034 [0.375]	−0.0016 [0.700]	0.0025 [0.552]
Arrivals per 100,000	0.0745 [0.042]	−0.0211 [0.399]	0.0698 [0.045]	−0.0499 [0.040]	0.1306 [0.166]
Departures per 100,000	0.0036 [0.911]	−0.0080 [0.854]	0.0036 [0.870]	−0.0575 [0.125]	0.0019 [0.933]
Net geographic flow per 100,000	0.0710 [0.044]	−0.0131 [0.725]	0.0662 [0.047]	0.0075 [0.865]	0.1287 [0.168]
Cross-state arrivals per 100,000	0.0207 [0.063]	−0.0324 [0.034]	0.0018 [0.853]	−0.0264 [0.057]	0.0210 [0.349]

Notes: Entries are coefficients with 9,999-draw state-level restricted Rademacher wild-cluster p-values in brackets. The first two columns come from regressions containing the contemporaneous standardized policy index and its one-year future value; the next two use the contemporaneous index and its two-year future value. All lead specifications include county and year fixed effects in the OB/GYN–general-surgery gap representation. The final column uses 2017–2019 only. The cross-state arrivals estimate is −0.0019 in the baseline levels DDD (Table B.9), +0.0207 in the one-year-lead-adjusted levels specification here, and +0.0202 in the preferred FD DDD (Table 3); these are estimates under different specifications and sample windows.

B.7 First-difference future-policy-change diagnostics

For the preferred first-difference specialty-gap specification, I add the one-year and two-year *future annual changes* in the same standardized permissiveness index used in the headline regression. Each diagnostic includes year fixed effects and jurisdiction-clustered inference. I compare the contemporaneous coefficient with the baseline estimated on the identical available sample, rather than with a full-window estimate based on more years. The one-year test covers 2018–2023 (18,810 county-year differences); the two-year and joint tests cover 2018–2022 (15,675 differences). Table B.11 reports the individual tests. The joint null that both future-policy-change coefficients are zero has a state-level wild-cluster Wald $F = 0.2426$ and $p = 0.8064$ (9,999 draws). The contemporaneous coefficient remains near 0.09 across the diagnostics, while neither lead is statistically distinguishable from zero. This is a timing check for the preferred FD design, not proof of parallel trends or an explanation that renders the significant levels-regression two-year lead spurious.

C Structural Estimation Details and Robustness

C.1 Alternative-set sampling and likelihood evaluation

For each annual physician location choice, the estimation file contains 30 sampled counties. The chosen county is always included. The previous year's county is also included when

Table B.11: Preferred first-difference DDD: future annual policy-change tests

Specification	Term	Coefficient	Cluster SE	Wild p	County-years
1-year: matched baseline	ΔP_t	+0.0892	(0.0288)	0.0136	18,810
1-year: with future change	ΔP_t	+0.0899	(0.0284)	0.0129	18,810
	ΔP_{t+1}	+0.0229	(0.0350)	0.5287	18,810
2-year: matched baseline	ΔP_t	+0.0912	(0.0325)	0.0251	15,675
2-year: with future change	ΔP_t	+0.0911	(0.0325)	0.0234	15,675
	ΔP_{t+2}	−0.0070	(0.0416)	0.8765	15,675
Joint: both future changes	ΔP_t	+0.0919	(0.0321)	0.0259	15,675
	ΔP_{t+1}	+0.0226	(0.0334)	0.5125	15,675
	ΔP_{t+2}	−0.0043	(0.0407)	0.9201	15,675

Notes: The dependent variable is the annual first difference in the county-level OB/GYN-minus-general-surgery arrivals rate per 100,000 residents. ΔP_t is the contemporaneous change in the original fixed-standardization permissiveness index; future changes are ΔP_{t+1} and ΔP_{t+2}. All models include year fixed effects, 50 jurisdiction clusters, and 9,999-draw null-imposed restricted Rademacher wild-cluster tests. One-year and two-year baselines use the respective identical samples. In the frozen panel, the index has no measured state-policy changes in 2024; thus the full-window and 2018–2023 baseline slopes and clustered SEs coincide even though the sample sizes differ. In the joint model, the joint test of the two future coefficients is $F = 0.2426$, wild $p = 0.8064$. The tests assess future policy-change associations in the FD specification; they do not establish causal identification or resolve the levels-model timing pattern.

distinct and when it is on identified support. Sampling occurs from the full geographic county universe, before applying the restriction to county-years with identified destination effects. Counties outside this identified support remain in the sampled set with zero modeled choice probability rather than being removed after sampling.

There are 6,753,240 sampled county alternatives across the 225,108 annual choices. Of these, 51.1 percent are outside the identified destination support in the relevant year. The median sampled choice set therefore contains 15 counties with identified destination utilities (10th and 90th percentiles: 11 and 18). These are the alternatives with nonzero modeled probability, not counties literally open or closed for medical practice. When the incumbent county lacks an identified destination effect, the stay alternative is unavailable; this occurs for 280 observed moves (0.124 percent of 225,108 annual choices). The chosen county remains on the identified support.

The incumbent-only sampling correction follows the chosen-and-incumbent inclusion rule described in Section 6.2. Conditional on the sampled set containing the observed chosen county, the set probability is $\binom{M_t-1}{R-1}^{-1}$ when the physician stays and $\binom{M_t-2}{R-2}^{-1}$ when the physician moves and the distinct incumbent is retained. Their ratio is $\binom{M_t-2}{R-2}/\binom{M_t-1}{R-1} = (R-1)/(M_t-1)$, which gives $\log[(R-1)/(M_t-1)] = -4.6828$ for $R = 30$ and $M_t = 3,135$.

This correction is applied to estimation utilities only. The prediction step uses the full identified choice set, so no sampling correction is required. The preferred mixed-logit likelihood uses 11-point Gauss–Hermite quadrature.

The state-profile inference diagnostic uses 50 state clusters. The estimated covariance matches the saved profile-score sandwich exactly and the multiplier-draw SD differs from its sandwich counterpart by at most 1.49 percent. All 50 state profiles satisfy the implemented nuisance-gradient threshold of 0.02, although 44 have nonzero optimizer termination codes. A targeted follow-up tested the union of the ten largest absolute influence states for α and the ten largest for γ_I (15 distinct state profiles), using both full-size and half-size perturbations. All 15 passed the 5-percent numerical acceptance threshold; the largest discrepancy was 2.353 percent of the aggregate influence scale across parameters, and the largest discrepancy for γ_I was 1.328 percent. Re-profiled full-step contributions matched their saved counterparts exactly. All profiles in this targeted check met the nuisance-gradient threshold, although the majority of optimizer calls returned nonzero codes under the short-chunk L-BFGS-B implementation. This evidence supports the finite-difference stability of the influential profile scores without claiming conventional optimizer convergence for every call. The second-stage inference remains conditional on the recovered destination effects.

The main behavioral estimates and their uncertainty are reported in Table 5. The appendix focuses on estimation support, sampling probabilities, numerical inference checks, and the parameter restriction used in the robustness specification.

C.2 Second-stage controls and fixed versus mixed logit

Table C.2 compares the preferred structural specification with the parsimonious $\gamma_B = 0$ robustness and selected second-stage conditioning checks. The destination-policy coefficient remains between 0.0593 and 0.0611 across the reported second-stage conditioning variants. It is 0.0825 when γ_B is excluded and 0.0710 under the fixed-logit first stage; these are specification comparisons, not bounds on the causal parameter. The coefficient changes little when earnings, hospital share, and OB bed capacity are added. Fertility and contemporaneous OB/GYN density may respond to policy and physician location, so the specification including them is an over-control check rather than a preferred estimate.

Table C.1: Parsimonious robustness: setting $\gamma_B = 0$

Parameter	Preferred model	$\gamma_B = 0$ robustness
Mean stay, α	1.0693	0.9644
Policy × stay, γ_I	−0.1372	−0.1868
2015 baseline × stay, γ_B	−0.1212	0
Log distance, β_d	−1.3050	−1.3166
Same-state destination, β_h	1.8305	1.8141
Random-stay SD, σ_S	1.3482	1.3501
Stage-2 permissiveness slope, γ_D	0.0610	0.0825
Log likelihood	−42,004.07	−42,022.61

Notes: The second column is the preferred specification reported in Table 5. The third column imposes $\gamma_B = 0$ and re-estimates the exact NQ=11 likelihood. The restricted Stage-2 destination coefficient has a state-clustered standard error of 0.0207 and $p = 0.00023$. The restricted first-stage re-estimation is used only as a point-estimate robustness check; the main first-stage inference comes from the preferred specification. A one-standard-deviation increase in permissiveness corresponds to about 6.3 percent higher relative destination-choice odds in the preferred model and 8.6 percent in the restricted model.

Table C.2: Structural destination-policy robustness

Specification	$\widehat{\gamma}_D$	Additional result
Preferred specification	0.0610	SE = 0.0208
$\gamma_B = 0$ robustness	0.0825	SE = 0.0207
Policy only, common 2017–2023 sample	0.0611	8,270 cells
+ earnings, hospital share, OB bed capacity	0.0604	
+ fertility and contemporaneous OB/GYN density	0.0593	
Fixed-logit first stage	0.0710	

Notes: All controlled specifications use the common complete-case sample of 1,416 counties over 2017–2023. Fertility and contemporaneous OB/GYN density may themselves adjust to policy and are therefore treated as sensitivity controls.

C.3 Pre-defined structural heterogeneity screens

I examine two pre-defined heterogeneity dimensions, career stage and medical-school training, on both the destination-policy and incumbent-policy margins. The career-stage sample covers 41,043 physicians and 225,017 choice occasions. The medical-school training sample covers 21,321 physicians (51.9 percent of the 41,066-physician structural sample) and 118,486 choice occasions with matched medical-school rank. Results are conditional on successful training-rank matching; they need not generalize to physicians without matched training information. Both subgroup samples include all 50 jurisdictions in the structural geography. The four comparisons have large standard errors relative to the differences, so nonrejection is not evidence of equal policy slopes. The screens were evaluated under a pre-specified decision rule rather than by searching across alternative cutoffs. Neither dimension on either margin met the headline criterion.

Table C.3: Pre-defined heterogeneity screens across structural margins

Margin	Dimension	Reference slope	Subgroup slope	Difference	Screen SE	Raw p	Holm p	Four-test Bonferroni p
Destination	Early career ($\leq$ 10 years)	0.0608	0.0620	0.0012	0.0541	0.982	1.000	1.000
Destination	Top-50 medical-school training	0.0535	0.0982	0.0447	0.0763	0.560	1.000	1.000
Incumbent	Early career ($\leq$ 10 years)	-0.1380	-0.1159	0.0221	0.0359	0.542	0.542	1.000
Incumbent	Top-50 medical-school training	-0.1291	-0.2014	-0.0723	0.0601	0.235	0.470	0.941

Notes: For career stage, the reference group is physicians with more than 10 years since medical-school graduation; the subgroup is physicians with 10 or fewer years. For medical-school training, the reference group is physicians matched to non-Top-50 medical schools and the subgroup is physicians matched to Top-50 schools. Destination slopes are changes in destination utility for a one-standard-deviation increase in permissiveness; the corresponding relative-odds changes are about 6.3 versus 6.4 percent by career stage and 5.5 versus 10.3 percent by medical-school training. Incumbent slopes are type-specific deviations in the current-policy incumbent interaction. The destination screen SE is the largest of physician-, incumbent-state-, and destination-state-clustered standard errors; the incumbent screen uses state-clustered inference. Holm adjustments are applied within each two-test margin family. The final column applies a Bonferroni adjustment to the four raw tests across both margins. The pre-specified decision rule selected none of the four tests for headline interpretation. Joint-model and mover/transition checks do not overturn this conclusion.

D Counterfactual Details

D.1 Simulation algorithm and uncertainty

The counterfactual simulation is recursive in location. For each physician, the simulated county chosen in year t becomes the incumbent county in year $t + 1$, while the physician-specific stay taste is held fixed across years. Policy changes enter both destination attractiveness and, for the incumbent alternative, the policy-dependent stay term. All other recovered county-year destination components are held at their observed-path

estimates. The active physician population is fixed, so the counterfactuals describe geographic reallocation rather than entry into or exit from the specialty.

Reported Monte Carlo standard errors summarize simulation noise conditional on the estimated parameters. They do not propagate parameter-estimation uncertainty through the mixed-logit and second-stage estimation pipeline. The exact 2024 reallocations and conditional MCSEs are 130.47 (1.42) for the 2021 benchmark, 194.06 (1.68) for the 2019 benchmark, 290.78 (2.13) for the one-standard-deviation experiment, 427.28 (2.60) for no net tightening, and 431.16 (2.83) for the 2015 benchmark. These MCSEs measure numerical simulation precision, not structural confidence intervals.

For the no-net-tightening path, a separate destination-channel-only arm shuts down the current-policy incumbent interaction while retaining the destination-policy shift. It reallocates 165.56 physician-equivalents in 2024 (MCSE 1.81), compared with 427.28 in the full model. This is not an additive causal decomposition because policy changes alter choice probabilities nonlinearly and simulated incumbent locations evolve recursively. I report it only to show the scale of the destination channel in a separate simulation.

D.2 Baseline physician supply and counterfactual reallocation

To characterize where the counterfactual physician mass is reallocated, I classify counties using OB/GYN availability measured in the observed 2021 data, so the groups are identical across all counterfactual arms. Counties with zero OB/GYNs in 2021 form a separate group. Among counties with positive stock, I calculate the median number of OB/GYNs per 100,000 residents in the observed 2021 county panel, equal to 7.919, and classify counties below that median as low density and counties at or above it as higher density. For the two-group summary, zero-stock and low-density counties are combined as lower/zero-supply markets. This classification is based only on baseline OB/GYN availability and should not be interpreted as a formal designation of local access or shortage status.

The supply-group statistics use the mean simulated county stocks stored for each counterfactual. They therefore answer whether the *mean 2024 allocation* shifts toward lower- or higher-supply markets. They are not components of the official reallocation statistic, which first computes the county-level half-L_1 distance within each simulation replication and then averages across replications. Because absolute values are taken before versus after simulation averaging, the two objects need not coincide.

Under no net tightening, the 31.25 physician-equivalent gain in lower/zero-supply markets comprises 25.91 in low-density counties and 5.34 in counties with no OB/GYNs in 2021. The latter group begins from a very small baseline stock, so its mean simulated

Table D.1: Counterfactual shifts across baseline OB/GYN-supply groups in 2024

Policy path	Net shift to lower/zero supply markets (physician-equivalents)	Lower/zero change (percent)	Higher-supply change (percent)
Hold permissiveness at 2021 level	3.50	0.060	−0.016
Hold permissiveness at 2019 level	6.16	0.106	−0.029
One SD toward 2015 baseline	11.88	0.204	−0.056
Remove all net tightening since 2015	31.25	0.537	−0.146
Return to 2015 permissiveness	30.53	0.525	−0.143

Notes: The first column reports the between-group shift in the mean simulated 2024 allocation. With two exhaustive supply groups and fixed national physician mass, it equals the net gain of lower/zero-supply counties and the absolute net loss of higher-supply counties. Baseline groups are defined using 2021 OB/GYN availability: zero-stock counties are combined with positive-stock counties below the 2021 median density of 7.919 OB/GYNs per 100,000; positive-stock counties at or above that median form the higher-supply group. Percentage changes use each group's simulated observed-policy 2024 stock as the denominator. These quantities are descriptive decompositions of mean simulated stocks and are distinct from the replication-level half-L_1 reallocation in Table 7.

2024 stock is 8.17 percent higher under this counterfactual. The signed net gain of 31.25 physician-equivalents is numerically 7.3 percent of the 427.28 county-level half-L_1 reallocation, but the signed group shift and unsigned county-level reallocation are different accounting objects, not additive components. I treat the pattern as a descriptive shift in physician availability rather than as a patient-access or welfare effect.

D.3 State-level changes under the 2021 policy benchmark

The largest positive state-level net stock changes in 2024 are Florida (+4.00), North Carolina (+3.22), Colorado (+2.47), Idaho (+1.25), and California (+0.91) physician-equivalents. The largest negative changes are Arizona (−1.88), New York (−1.31), the District of Columbia (−1.28), West Virginia (−0.84), and Ohio (−0.78). These are net state stock changes after county-level resorting. They are distinct from the national county-level half-L_1 reallocation statistic. The half-L_1 distance computed from the mean state changes is 13.75 physician-equivalents, compared with 130.47 at the county level. Their ratio is not the share of interstate moves because absolute values and simulation averaging are applied at different aggregation levels.

State-level policy can generate within-state county resorting in this model even though counties within a state share the same policy index. The destination term $\gamma_D P_{st}$ shifts all counties in the state together, while the incumbent county additionally receives the policy-dependent state-dependence term $\gamma_I P_{st} S_{ijt}$. The same-state coefficient then makes another county in the incumbent state a relatively attractive alternative. This is a property

of the estimated choice specification, not a claim that the data identify a separate primitive mechanism behind γ_I.

Table D.2: Largest state-level net stock changes under the 2021 policy benchmark

Largest gains	Change	MCSE	Largest losses	Change	MCSE
Florida	4.00	0.49	Arizona	−1.88	0.26
North Carolina	3.22	0.65	New York	−1.31	0.28
Colorado	2.47	0.44	District of Columbia	−1.28	0.28
Idaho	1.25	0.37	West Virginia	−0.84	0.35
California	0.91	0.37	Ohio	−0.78	0.27

Notes: Changes are counterfactual minus observed simulated 2024 state stocks, in physician-equivalents. MCSEs use the 16 antithetic simulation pairs and measure simulation uncertainty conditional on the estimated parameters. State changes are descriptive; the main counterfactual object remains county-level geographic reallocation.

D.4 Location-choice inclusive value

For completeness, the model also produces the standard discrete-choice inclusive value, or expected maximum utility over practice locations (Small and Rosen, 1981). Under the 2021 policy benchmark, the mean 2024 inclusive value is 0.0276 utility units lower than under the observed policy path, equivalent to approximately 0.0211 units of $\log(1 + d)$ at the estimated distance coefficient. The negative sign is compatible with the model: the estimated current-policy incumbent interaction is negative, so increasing permissiveness can lower the direct stay component of utility even as the destination component rises. Because staying is overwhelmingly common, the incumbent component can dominate the simulated logsum. I treat this quantity as a descriptive location-choice logsum rather than as a social-welfare measure. In particular, the incumbent-policy interaction is not given a primitive welfare interpretation, and the current two-step inference does not provide a joint confidence interval for the full counterfactual logsum.

E Replication and Numerical Notes

All data construction, reduced-form estimation, structural estimation, and counterfactual simulation are implemented in R (R Core Team, 2026). The workflow uses `data.table` for large-panel manipulation (Barrett et al., 2026), packages in the tidyverse ecosystem for data import and supporting transformations (Wickham et al., 2019), and `fixest` for high-dimensional fixed-effects regressions (Bergé et al., 2026). The mixed-logit likelihood, quadrature, sampled-choice correction, and counterfactual simulator are implemented directly in the project code. Exact R and package versions are recorded by `sessionInfo()`

in the frozen replication objects. The underlying CMS practice-location files are publicly distributed; the state-year legal-provision panel was assembled for this study. The replication files are not currently posted in a public repository.