Medicaid Spillovers on Commercially Insured Patients: Evidence from Postpartum Depression Screening

Sobin Lee*

October 17, 2025

Abstract

In January 2014, Colorado Medicaid began reimbursing pediatric providers for postpartum depression (PPD) screening during infant well-child visits. This paper examines how this targeted reimbursement affected screening practices across payer types. Using linked birth records and All-Payer Claims data from 2012–2019, I find that practices with greater pre-policy Medicaid exposure increased screening more after the policy, primarily among commercially insured patients. Screening for Medicaid patients rose broadly across all practices, consistent with practice-wide adoption rather than payer-specific targeting. Leveraging physician moves across practices, I show that about two-thirds of physicians adjust their screening behavior toward the norms of their new practice, indicating that organizational systems play a key role in shaping provider behavior. The results highlight how payer-specific incentives can influence care delivery beyond the targeted payer through practice-level mechanisms.

*

1 Introduction

Understanding spillovers from public to private insurance is central to evaluating the broader impact of health policy. When providers respond to payment incentives for publicly insured patients, their behavior may extend to privately insured patients as well, amplifying the reach of targeted programs. Prior work has documented such spillovers in hospital spending (Baicker et al., 2013), physician pricing (Clemens and Gottlieb, 2017), and bundled payment reforms (Einav et al., 2020; Barnett et al., 2020). Yet little is known about whether similar effects arise in mental health care, where detection and treatment gaps remain large.

Postpartum depression (PPD) provides a particularly useful case for studying these dynamics. PPD affects 10–20% of mothers in the United States (Khadka et al., 2024), yet often goes undiagnosed in 50–70% of cases and untreated in nearly 85% (Dagher and Bair-Merritt, 2021). Untreated PPD harms maternal well-being and can lead to developmental delays and higher healthcare utilization in early childhood (Slomian et al., 2019). Improving detection and treatment has therefore become a national public health priority. Medicaid, one of the largest payers of maternity care, offers a natural policy lever for addressing these gaps. Because Medicaid covers only a subset of the population, policies targeting its enrollees may leave out the commercially insured. Spillovers through provider behavior can bridge this divide: if practices adopt new screening protocols in response to Medicaid incentives, these workflows may extend to all patients, regardless of payer.

This paper investigates how pediatric practices responded to a targeted Medicaid reimbursement policy for postpartum depression screening and whether these changes affected care for commercially insured mothers. In January 2014, Colorado implemented a policy allowing pediatric providers to bill Medicaid for screening mothers for PPD during infant well-child visits. While the policy aimed to increase screening among Medicaid-enrolled mothers, many pediatric practices serve both Medicaid and commercially insured patients. The policy thus created heterogeneous incentives across providers, depending on their baseline Medicaid patient share, and introduced the possibility that practices exposed to stronger Medicaid incentives would adopt screening workflows that spill over to commercial patients.

Provider behavior offers a key mechanism for these spillovers. Physicians and practices often face multiple payer arrangements and may find it administratively simpler to standardize care delivery rather than tailor services by insurance type. Standardization can reduce complexity, improve efficiency, and ensure compliance with quality metrics. When new financial incentives encourage screening for Medicaid patients, providers may find it more cost-effective to apply screening protocols uniformly across their entire patient panel, generating spillovers for privately insured mothers. Pediatric care provides an especially relevant setting for this process, given the frequent and structured schedule of infant well-child visits during the first year of life. By embedding screening into this routine, the Medicaid policy created an opportunity for broader

PPD detection among mothers.

To test these hypotheses, I use the Colorado All-Payer Claims Database (APCD) linked to statewide birth records from 2012–2019. A key feature of the analysis is the use of practice-level variation in Medicaid patient share to measure financial incentives directly. I define each practice's pre-policy Medicaid exposure as the average share of patients enrolled in Medicaid in 2012–2013 and implement a difference-in-differences design using this exposure as a continuous treatment variable. I restrict the analytic sample to practices with at least one pediatric or primary care physician in the pre-policy period. To assess whether practices applied screening uniformly across patient groups, I estimate a triple-difference model comparing Medicaid and commercially insured mothers within the same practices. Finally, to distinguish physician behavior from practice-level effects, I implement a physician-movers design following the same clinician as their primary practice changes over time. I identify moves using NPPES primary-practice listings, validate timing with billed well-child and screening claims, and estimate an event study around the move to test whether a physician's screening behavior converges toward the destination practice's environment.

The results show that the 2014 Medicaid policy led to substantial increases in postpartum depression screening overall, driven primarily by gains among commercially insured mothers. Among commercially insured mothers, practices with full pre-policy Medicaid exposure experienced large increases in both the share screened and the share of well-child visits that included screening, relative to practices with little or no Medicaid exposure. Screening adoption for both patient groups was concentrated in high-Medicaid-share practices early in the policy period and persisted thereafter. Triple-difference estimates confirm that providers applied screening protocols uniformly across Medicaid and commercial patients, consistent with practice-level workflow changes rather than payer-targeted responses. The physician-movers analysis shows that when a clinician moves to a higher-Medicaid-share practice, their screening behavior shifts toward the destination practice's screening environment by roughly 70%, similar to other estimates of place effects in health care (Molitor, 2018). These findings suggest that financial incentives at the practice level, rather than fixed physician propensities, are the main drivers of post-policy adoption.

This paper contributes to three strands of research. First, I contribute to the literature on spillovers from public to private insurance. A large body of work shows that changes in public payment systems can influence care delivered to privately insured patients through shared providers and market interactions. For example, Baicker et al. (2013) show that greater Medicare Advantage penetration reduces hospital spending for commercially insured patients; Clemens et al. (2025) demonstrate that the introduction of new Medicare billing codes led to gradual but heterogeneous take-up across physicians, highlighting how payment design and administrative frictions shape provider adoption; Clemens and Gottlieb (2017) document that Medicare fee changes spill over to private-sector physician prices; and Einav et al. (2020) and Barnett et al. (2020) find that Medicare bundled payment reforms altered treatment intensity beyond the targeted population. I extend this literature by focusing on mental health care, an area rarely studied in this context, and by using a direct, practice-level measure of Medicaid exposure that captures the financial incentives faced by providers rather than relying on geographic or population-level proxies. This approach allows me to identify how targeted Medicaid reimbursement can influence practice-wide adoption of screening protocols that reach commercially insured patients as well.

Second, I contribute to work on the organizational drivers of care delivery and the diffusion of clinical practices across payers. Providers operating under multiple payer arrangements often standardize workflows to reduce administrative burden and ensure compliance with quality metrics (Glied and Zivin, 2002; Baicker and Robbins, 2015; Barnett et al., 2020). By documenting uniform adoption of postpartum depression screening across payer types following a Medicaid reimbursement change, I show how organizational routines can translate payer-specific incentives into system-wide improvements in care delivery. This perspective highlights the importance of practice-level decision-making and offers a broader interpretation of how targeted public incentives can reshape clinical standards.

Finally, I contribute to the literature on physician behavior and the role of place versus person effects. Prior studies separate individual provider preferences from local practice environments by tracking patients or physicians who move across regions or institutions (Finkelstein et al., 2016; Molitor, 2018). These studies show that much of observed variation in treatment intensity reflects local environments rather than fixed provider propensities. I adapt this framework to a new setting, maternal mental health screening in pediatric care, and demonstrate that roughly half of a physician's screening behavior adjusts to the destination practice after a move. This provides novel evidence that organizational context, rather than stable physician traits, drives much of the observed response to financial incentives.

The rest of the paper proceeds as follows. Section 2 provides institutional background on the Colorado Medicaid reimbursement policy and the clinical context for postpartum depression screening. Section 3 describes the linked Colorado All-Payer Claims Database and Vital Records data, outlines the construction of key variables, and summarizes sample characteristics. Section 4 examines the effect of Medicaid reimbursement on screening rates, first estimating the overall impact using a difference-in-differences framework and then extending the analysis to assess spillovers across payers through a triple-difference design that compares Medicaid and commercially insured patients within the same practices. Section 5 investigates mechanisms of adoption using a physician-movers design that distinguishes physician behavior from practice-level effects. Section 6 concludes with policy implications and directions for future research.

2 Background

Postpartum depression (PPD) is a prevalent and serious maternal health condition that often goes unrecognized in routine care. Obstetric care typically concludes after a single postpartum visit around six weeks after delivery, and as many as 40% of women do not attend that visit (American College of Obstetricians and Gynecologists, 2018). By contrast, mothers are far more likely to accompany their infants to regularly scheduled pediatric visits, creating repeated opportunities for early identification of depressive symptoms (Liberto, 2012; Earls et al., 2010). In recognition of this, the American Academy of Pediatrics (AAP) recommends that pediatricians screen for maternal depression during early infant well-child visits (Earls et al., 2010). Yet despite these guidelines, fewer than half of mothers with postpartum depressive symptoms are identified in clinical settings, with major barriers including lack of reimbursement, limited referral pathways, and uncertainty about billing procedures (Yawn et al., 2012; Liberto, 2012).

In January 2014, Colorado implemented a policy authorizing pediatric and family medicine providers to bill Medicaid for postpartum depression screening conducted during infant well-child visits. Providers could bill under either the mother's or the infant's Medicaid identification number, and state guidance outlined validated screening tools and referral expectations. The policy aimed to increase screening and follow-up treatment among Medicaid-enrolled mothers, who face particularly high barriers to accessing postpartum mental health services. A recent evaluation by Gordon et al. (2025) found that the policy raised screening among Medicaid mothers by 9.6 percentage points relative to commercially insured mothers and also increased diagnosis and treatment rates. Similarly, Currie and Malinovskaya (2025) show that a comparable policy adopted in Michigan in 2018 roughly doubled screening rates, though treatment gains were concentrated in higher-income areas.

At the national level, the U.S. Preventive Services Task Force (USPSTF) issued a Grade B recommendation in 2016 for universal depression screening, including postpartum women. Under the Affordable Care Act, this required most commercial insurers to cover screening without cost sharing beginning in 2017. Because this change may have independently affected screening rates, I later verify that my results are robust when restricting the analysis to the pre-2017 period.

Colorado's 2014 Medicaid policy thus provides a valuable setting to examine whether targeted reimbursement can generate broader changes in provider behavior. Once screening tools such as the Edinburgh Postnatal Depression Scale (EPDS) are integrated into practice workflows, the marginal cost of extending them to all patients is low, especially in practices serving both Medicaid and commercially insured families. This institutional feature allows for a clean test of whether Medicaid-specific payment incentives can reshape practice-wide screening routines and produce spillovers across payer types.[1]

[1] Appendix Figure A1 reproduces the standard Edinburgh Postnatal Depression Scale (EPDS) screening form used in pediatric and primary care settings.

3 Data

3.1 Data and Variable Definition

This study links multiple administrative and survey-based data sources to construct a comprehensive mother–infant panel spanning 2012–2019. Each dataset contributes distinct information on screening behavior, provider characteristics, and the socioeconomic environment of practice locations. The Colorado All-Payer Claims Database (APCD) serves as the primary source of information on utilization and billing, including detailed medical and pharmacy claims, enrollment, and payment data across Medicaid and commercial insurers. I use it to measure infant well-child visits, postpartum depression (PPD) screening claims, and maternal insurance coverage. The Colorado Vital Records provide official mother–infant linkages and detailed maternal demographics and birth characteristics, allowing claims to be matched to maternal profiles. The National Plan and Provider Enumeration System (NPPES) supplies information on each provider's National Provider Identifier (NPI), primary practice address, and organizational affiliation, which I use to define consistent practice identifiers and track providers over time. Finally, I link area-level socioeconomic and demographic characteristics from the American Community Survey (ACS) to each practice location using ZIP Code Tabulation Areas (ZCTAs) to account for differences in neighborhood context.

Each claim in the APCD lists both a billing and a service provider National Provider Identifier (NPI). I link these NPIs to the NPPES primary practice location file and assign a consistent practice identifier based on the billing provider's primary address, which best reflects the locus of financial incentives and billing decisions. When multiple billing NPIs share the same address, I treat them as a single practice. The billing and service provider practice locations coincide for approximately 75% of well-child visits; in the remaining cases, billing is often centralized through larger networks or affiliated systems. To capture changes in provider location over time, I use the last-month NPPES file from each quarter to extract quarterly practice addresses. If a provider's practice location is missing in a given quarter, I impute it using the closest available quarter's address. As a robustness check, I replicate key results using the service provider's practice location, which yields similar estimates. To characterize provider types, I link taxonomy codes reported in the APCD to the National Uniform Claim Committee (NUCC) taxonomy crosswalk, which maps each NPI to a standardized specialty category. This classification identifies pediatric and primary care providers (internal medicine, family medicine, or general practice) and distinguishes organization types such as hospitals, ambulatory facilities, nursing facilities, agencies, and managed care organizations. The analytic sample is restricted to practices that include at least one pediatric or primary care provider in the pre-policy period.

The unit of analysis is a practice–year–quarter cell (p, t). I examine two main outcomes. The first, Any screening, is an indicator equal to one if at least one postpartum depression screening was billed at practice p in quarter t. The second, Percent of visits screened, is the percentage of eligible well-child visits at practice p in quarter t with a billed screening claim. Eligible visits are those occurring within the infant's first year of life, consistent with the policy's coverage window. The Any screening outcome is analyzed without weights to capture adoption at the practice level, while the Percent of visits screened outcome is weighted by the number of well-child visits in the corresponding practice–quarter to reflect visit volume. For descriptive analyses and robustness checks, I also use visit-level data to visualize adoption dynamics.

A key variable in the analysis is each practice's pre-policy Medicaid exposure, denoted MedicaidShare_p. This measure captures the extent of financial incentives each practice faced when Colorado introduced Medicaid reimbursement for postpartum depression screening in 2014.

To construct MedicaidShare_p, I use all claims from 2012 and 2013, the two years prior to the policy. For each practice p, identified by its billing practice ID, I calculate the share of claims paid by Medicaid among all claims paid by either Medicaid or commercial insurance. Claims covered by other insurance types, such as self-pay, are excluded. I then average these shares across 2012 and 2013 to obtain a time-invariant measure of baseline Medicaid exposure for each practice.

This measure reflects the proportion of a practice's patient volume financed by Medicaid prior to the policy and therefore its relative financial exposure to the new reimbursement. Practices with higher pre-policy Medicaid shares faced stronger incentives to adopt screening once reimbursement became available. In all regression specifications, I control for the total number of infant well-child visits billed by each practice in a given year to account for differences in practice size and patient volume that may correlate with both payer mix and adoption behavior. For robustness, I confirm that the results are consistent when using the service provider's practice identifier instead of the billing provider's and that the two measures yield a highly correlated distribution of Medicaid shares.

Maternal insurance status is derived from the APCD enrollment file and classified as Medicaid or commercial based on the majority of coverage during the calendar year. This variable is used for subgroup analyses and as the Medicaid indicator in the triple-difference specification. The key treatment variable, MedicaidShare_p, measures each practice's payer mix prior to the 2014 policy. Using claims from 2012–2013, I compute the share of unique patients with Medicaid coverage among those with Medicaid or commercial insurance, averaging across the two pre-policy years to create a time-invariant measure of exposure. The post-policy indicator Post_t equals 1 for 2014–2019 and 0 for 2012–2013.

Covariates include both patient- and practice-level characteristics. Maternal covariates (aggregated to the practice–quarter level as visit-weighted means) include age, education, race and ethnicity, marital status, nativity, and indicators for chronic or pregnancy-related conditions (e.g., diabetes, hypertension, obesity, cesarean delivery, or preterm birth). Practice-level covariates include the number of physicians, total patient volume, and counts of pediatric, primary care, and OB/GYN physicians, as well as indicators for practice type. Details on the CPT and HCPCS codes used to identify PPD screening and well-child visits are provided in Appendix ??.

To capture neighborhood socioeconomic context, I link the practice file to the American Community Survey (ACS) 5-year estimates (2012–2019). Practice ZIP codes from NPPES are mapped to ZIP Code Tabulation Areas (ZCTAs) using a ZIP–ZCTA crosswalk, and year-specific ACS tables are merged to each practice record. From ACS DP02, DP05, and S1901 tables, I extract measures of educational attainment, household income, language proficiency, nativity, fertility, and racial and ethnic composition. These area-level covariates enter regressions as time-varying controls to account for socioeconomic and demographic differences in practice surroundings.

Unless otherwise noted, regressions are estimated at the practice–year–quarter level without weights. Observations missing outcomes, insurance assignment, or key identifiers are excluded from the analytic sample.

3.2 Sample and Summary Statistics

I form practice–payer–quarter cells and compute two outcomes: an indicator for any PPD screening in the cell and the share of well-child visits with a billed screen. The baseline uses all practice–quarter observations. As summarized in Table 1, the analytic sample retains practices observed both pre-policy (2012–2013) and post-policy (2014–2019), yielding 658 practices, 5,838 pre-policy and 18,117 post-policy practice–quarter cells, and a total of 1,088,807 well-child visits. The share of practice–quarters with any PPD screening increased substantially after the policy, from 4.4 to 18.5 percent among Medicaid practices, and also rose among commercial practices, from 1.3 to 14.5 percent, despite no contemporaneous changes in federal or state reimbursement guidelines for commercial plans. This parallel rise motivates the analysis of potential spillover effects in subsequent sections.

Table 1: Sample and Summary Statistics

	All		**Commercial**		**Medicaid**	
	Pre	Post	Pre	Post	Pre	Post
Practices (count)	658	658	536	536	532	532
Practice–quarter observations	5,838	18,117	2,768	8,526	3,004	8,500
Well-child visits	219,102	869,705	52,244	257,665	166,746	603,919
Any screening (%)	2.9	15.9	1.3	14.5	4.4	18.5
	(16.8)	(36.5)	(11.5)	(35.2)	(20.6)	(38.8)
% WCVs screened (mean)	0.3	7.6	0.1	4.3	0.3	9.1
	(1.8)	(14.9)	(1.9)	(10.7)	(1.8)	(16.2)

Notes: Pre = 2012–2013; Post = 2014–2019. "Any screening" is the share of practice–payer–quarter cells with at least one PPD screen. "% WCVs screened" is the visit-weighted mean share of well-child visits with a billed PPD screen. Practices are restricted to those observed in both periods. Standard deviations are in parentheses below the mean.

Pre-policy Medicaid exposure. Figure 1 plots the distribution of pre-policy Medicaid share across practices, averaged over 2012–2013. Practices vary widely in the proportion of their patients covered by Medicaid: some serve predominantly commercially insured families, while others draw mostly from Medicaid, with the median practice having roughly half of its patients enrolled in Medicaid. This variation captures meaningful differences in financial exposure to the 2014 policy, which reimbursed postpartum depression screening only for Medicaid-covered mothers. The dispersion in pre-policy Medicaid share provides the identifying source of cross-practice variation in the difference-in-differences design, where practices with higher baseline Medicaid shares faced stronger incentives to adopt screening.

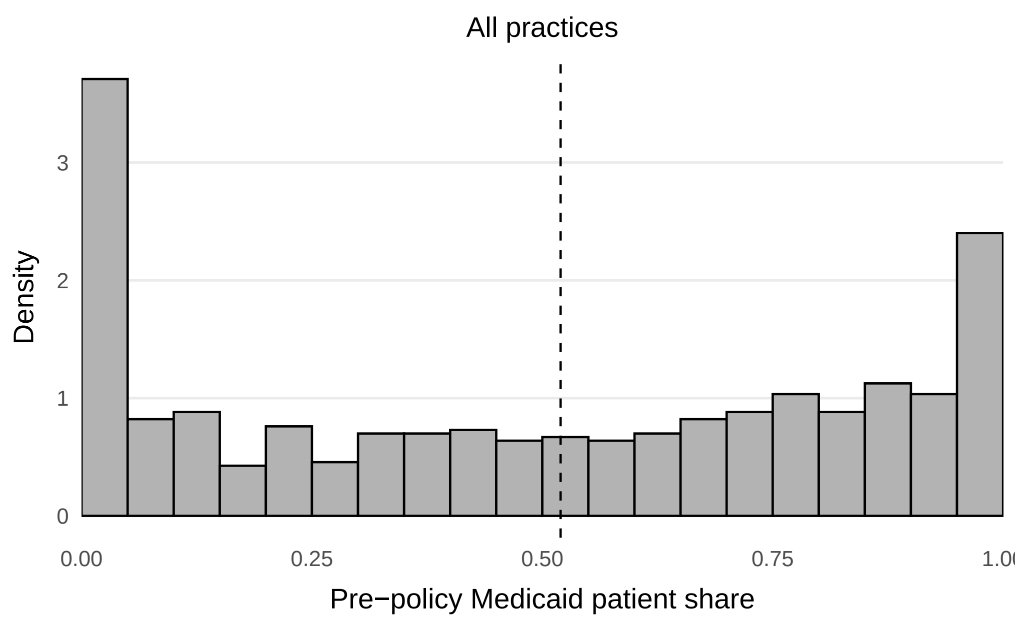


Figure 1: Distribution of Pre-Policy Medicaid Share.

Notes: Each bar shows the density of practices by their average share of Medicaid patients across 2012–2013. The measure is based on all claims at the billing-practice level and reflects the proportion of patients with Medicaid coverage among those with either Medicaid or commercial insurance. The dashed line marks the median practice.

Figure 2 presents the same distribution separately for mothers covered by commercial insurance and Medicaid, based on their infants' visits to the same set of pediatric practices. Practices serving commercially insured mothers tend to have lower Medicaid shares on average, while those serving Medicaid mothers have higher shares, yet there is substantial overlap across the two distributions. This overlap is central for identification: within the same statewide policy environment, both groups of mothers are seen at practices with differing levels of pre-policy Medicaid exposure, allowing comparisons of how screening behavior evolved across payer types while holding practice characteristics and broader policy shocks constant.

Some commercially insured mothers appear in high–Medicaid-share practices, and conversely, some Medicaid-enrolled mothers visit low–Medicaid-share practices. This pattern arises because a mother's insurance classification is based on her majority enrollment during the calendar year, whereas a practice's Medicaid share is defined using the share of total billed claims across all patients seen in that practice in a given year. The two measures therefore capture related but distinct aspects of coverage. Insurance coverage is also not always consistent over time: mothers

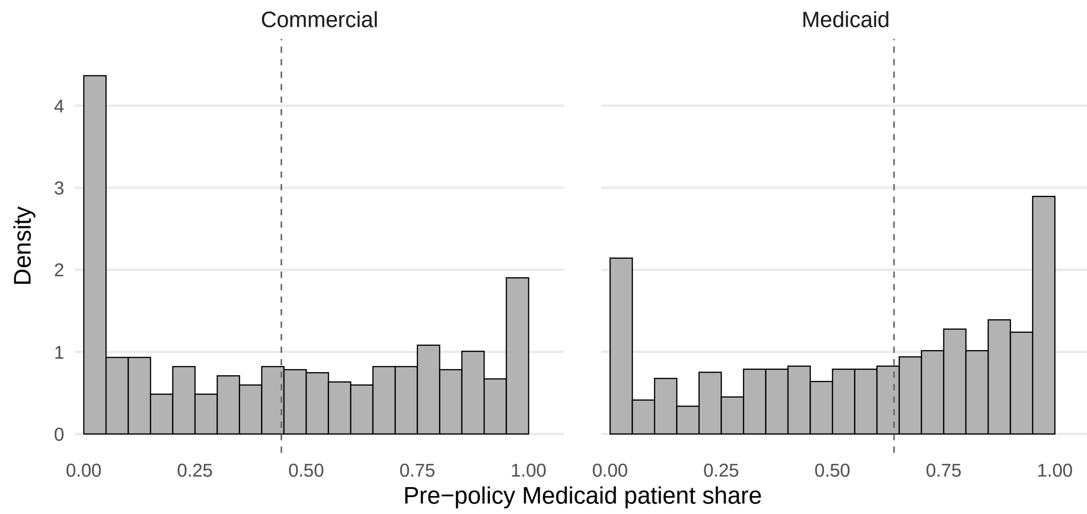


Figure 2: Distribution of Pre-Policy Medicaid Share by Payer Group.

Notes: Each panel shows the distribution of practices by pre-policy Medicaid patient share for visits billed to commercially insured (left) and Medicaid (right) mothers. Both panels use the same horizontal and vertical scales. Dashed lines indicate the median practice within each group. Distributions overlap substantially, supporting within-state comparisons across payer types.

can hold both Medicaid and commercial coverage during the postpartum year, and I classify each mother based on her majority coverage during that period. In addition, infant and maternal coverage can diverge. Infants born under Medicaid are automatically enrolled in Medicaid for their first year of life, but maternal eligibility changes shortly after delivery. During the study period, pregnant individuals in Colorado qualified for Medicaid up to 265 percent of the federal poverty level (FPL), while eligibility for low-income adults was limited to 138 percent FPL under the state's Medicaid expansion in 2014.

4 Effect of Medicaid Reimbursement on Screening Rates

4.1 Impact on Overall Screening Rate

4.1.1 Empirical Strategy

This section estimates how the 2014 Medicaid reimbursement policy for postpartum depression (PPD) screening affected screening rates at pediatric and primary care practices in Colorado. The empirical strategy exploits cross-practice variation in pre-policy Medicaid exposure as a continuous measure of treatment intensity under a single statewide policy change.

I examine two outcomes. The first is a binary indicator for whether a practice–payer–quarter cell recorded any PPD screening (Any screening). The second is the percentage of well-child visits in that cell that included a billed PPD screen (Percent screened). The first outcome captures the extensive margin of adoption—whether a practice ever screens—while the second measures the intensity of screening among practices that do.

I estimate the following difference-in-differences (DiD) model:

$$Y_{pt} = \beta_0 + \beta_1 \text{Post}_t + \beta_2 \text{MedicaidShare}_p + \beta_3 \big(\text{Post}_t \times \text{MedicaidShare}_p\big) + X_{pt}\gamma + \epsilon_{pt}. \quad (1)$$

Here, Y_{pt} represents one of the two outcomes for practice p in year–quarter t; Post_t equals 1 for quarters after January 2014 and 0 otherwise; and MedicaidShare_p is the practice's pre-policy Medicaid patient share, defined using all patient visits in 2012–2013. The interaction term $\text{Post}_t \times \text{MedicaidShare}_p$ captures the differential post-policy change in screening rates across practices with higher versus lower baseline Medicaid exposure.

The vector X_{pt} includes both quarterly and annual practice-level covariates. Quarterly covariates summarize the average patient characteristics for well-child visits in that practice–quarter, including maternal age, education, and area-level income and rurality. Annual covariates include practice-level characteristics constructed from provider taxonomies and patient volume, such as the total number of physicians, the counts of pediatric, primary care, OBGYN, and mental health providers, and the number of infants under age one visiting the practice in that year. ZIP-code–level characteristics from the American Community Survey are also merged annually by practice location to capture local socioeconomic and demographic context, including income, education, race, and household composition. These covariates vary by year but are constant within each year–quarter. ϵ_{pt} is the error term.

The coefficient of interest, β_3, measures how screening outcomes changed in the post-policy period as a function of a practice's pre-policy Medicaid share. A positive β_3 indicates that practices with a higher share of Medicaid patients experienced larger increases in screening following the policy. Coefficients are interpreted as the effect of a one-unit (100 percentage point) increase in pre-policy Medicaid share on the post-policy change in screening. Standard errors are clustered at the practice level to account for serial correlation, and all regressions are restricted to practices observed in both the pre- and post-policy periods. I estimate the model separately for Medicaid and commercially insured visits to assess whether the increase in screening extended beyond the directly targeted population.

4.1.2 Identification

The identification strategy relies on two conditions. First, parallel trends: absent the policy, screening rates for practices with different pre-policy Medicaid shares would have evolved similarly. Second, the absence of sorting or compositional changes: the mix of patients visiting a given practice should not change systematically in ways correlated with pre-policy Medicaid exposure after the policy was introduced. Both conditions are supported by the institutional setting. The policy was implemented statewide with no phase-in or differential eligibility across practices.

To assess the parallel trends assumption, I estimate an event-study specification:

$$Y_{pt} = \alpha_p + \lambda_t + \sum_{k\in\mathcal{K},\, k\neq -1} \beta_k \big(D_k(t) \times \text{MedicaidShare}_p\big) + X_{pt}\gamma + \epsilon_{pt}. \tag{2}$$

Here, $D_k(t) = \mathbb{1}\{t - t_0 = k\}$ are event-time indicators relative to the policy quarter $t_0 =$ 2014Q1, with $k = -1$ (2013Q4) omitted as the reference period. The coefficients β_k trace the dynamic evolution of screening rates before and after the policy by pre-policy Medicaid share. Practice fixed effects α_p absorb time-invariant differences across practices, while year–quarter fixed effects λ_t capture aggregate time shocks common to all practices. X_{pt} includes both quarterly and annual time-varying covariates as described above, and standard errors are clustered at the practice level.

Under the identifying assumptions, a positive and statistically significant β_3 in the DiD specification—or upward shifts in post-2014 coefficients β_k in the event study—indicate that Medicaid reimbursement increased PPD screening, with stronger effects among practices more financially exposed to Medicaid prior to the policy.

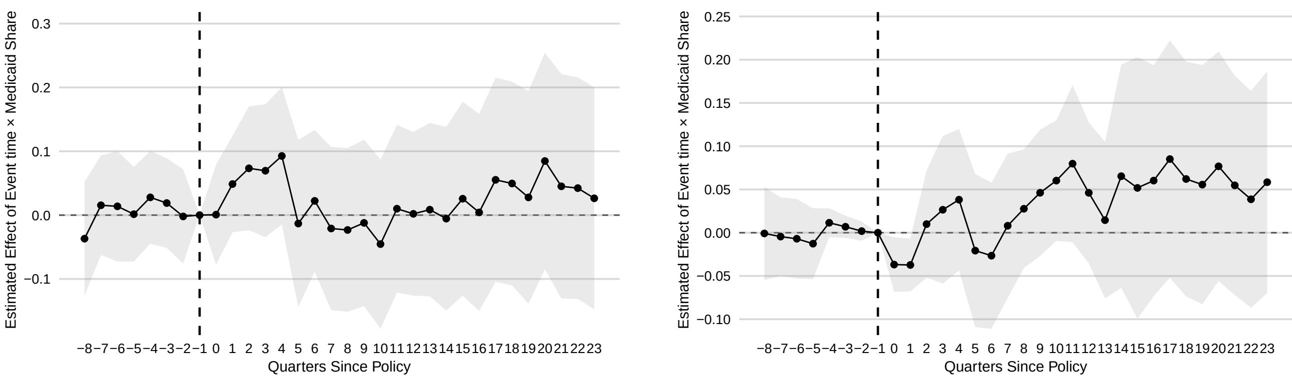
To examine the dynamic effects of the policy and assess the identifying assumption of parallel trends, I estimate Eq. (2) separately for Medicaid and commercially insured patients. The outcomes are 1) an indicator for any screening during well-child visits and 2) the percentage of well-child visits screened. The first captures the extensive margin of adoption, while the second reflects screening intensity among visits. The coefficients $\widehat{\beta}_k$ trace how screening changed over time in practices with higher pre-policy Medicaid shares, relative to those with lower shares.

Joint pre-trend tests. For each insurance group $g \in \{\text{Medicaid}, \text{Commercial}\}$, I jointly test that all pre-policy event-time coefficients are zero,

$$H_0 : \beta_k = 0 \quad \text{for all } k \in \{-8, -7, -6, -5, -4, -3, -2\},$$

with $k = -1$ (2013Q4) omitted. Cluster-robust Wald tests (clustered by practice) fail to reject the joint null of no pre-policy effects at conventional 5% levels. The results are: Medicaid, any screening: $F(7, 769) = 1.50,\ p = 0.163$; Medicaid, percent screened: $F(7, 769) = 1.12,\ p = 0.346$; Commercial, percent screened: $F(7, 747) = 1.30,\ p = 0.248$; Commercial, any screening: $F(7, 747) = 1.79,\ p = 0.086$.

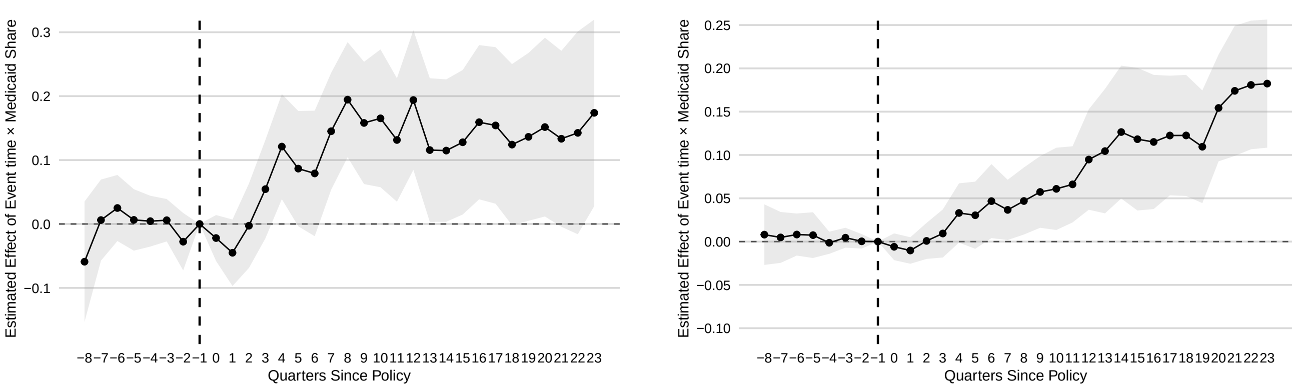
Across both insurance groups, pre-policy coefficients are small and jointly insignificant, indicating no systematic differences in screening trends across practices with different baseline Medicaid exposure prior to the policy. After 2014, screening rates among commercially insured patients rise more in practices with higher pre-policy Medicaid shares, and the event-study coefficients become positive and statistically distinguishable from zero in the commercial sample for both outcomes. In the Medicaid sample, post-policy coefficients are generally positive but imprecise and not statistically different from zero, consistent with the payer-specific DiD estimates that show economically meaningful but statistically insignificant gradients for Medicaid. Taken together, the event studies support parallel pre-trends for both payer groups and indicate that the strongest post-policy divergence occurs for commercial patients, consistent with practice-wide workflow adoption that spills over to commercially insured mothers rather than a Medicaid-only response especially for practices with higher pre-policy Medicaid patient share.



(A) Any screening **(B) Percentage of well-child visits screened**

Figure 3: Event-study coefficients for Medicaid patients

Notes: Points plot $\widehat{\beta}_k$ from Eq. (2), where $D_k(t)$ are event-time indicators relative to 2014Q1 and $k = -1$ (2013Q4) is omitted. The gray shaded area shows the 95% confidence intervals clustered by practice. Specification includes practice and year–quarter fixed effects, ZIP-code and practice-level covariates X_{pt}, and interactions $D_k(t) \times \text{MedicaidShare}_p$.



(A) Any screening **(B) Percentage of well-child visits screened**

Figure 4: Event-study coefficients for Commercial patients

Notes: Defined as in Fig. 3. 95% confidence intervals clustered by practice. Same sample, controls, and fixed effects as the Medicaid panel.

Compositional change by pre-policy Medicaid exposure. To assess whether observable maternal characteristics changed differentially across practices with varying pre-policy Medicaid exposure, I estimate difference-in-differences regressions at the practice–quarter level. Each regression interacts the post-policy indicator (2014–2019) with the practice's pre-policy Medicaid share, estimated separately for the Medicaid and commercial samples. All models include practice and year–quarter fixed effects, cluster standard errors at the practice level, and weight by the number of eligible well-child visits in each cell. To adjust for multiple hypothesis testing across covariates, I report Benjamini–Hochberg false discovery rate (FDR) q-values within payer. Tables 2 and 3 report, for each covariate and payer, the estimated coefficient on the interaction term, with its standard error shown in parentheses below, and the corresponding p-value and FDR-adjusted q-value in separate columns.

After accounting for multiple testing, most maternal covariates show no differential change by pre-policy Medicaid exposure between the pre- and post-policy periods. In the commercial sample (Table 2), two covariates remain statistically significant after FDR adjustment: the share of mothers with income \$25,000–34,999 (coefficient 0.092, s.e. 0.028, $q = 0.023$) and the preterm birth rate (0.17, s.e. 0.056, $q = 0.024$). In the Medicaid sample (Table 3), only racial composition changes significantly: the share of White mothers declines (coefficient -0.37, s.e. 0.11, $q = 0.036$) while the share of mothers classified as Other race increases (0.35, s.e. 0.12, $q = 0.039$). The magnitudes of these coefficients are modest, indicating that the composition of patients served by each practice remained largely stable over time.

These results suggest that the estimated effects of the policy are not driven by differing compositional changes across practices with varying pre-policy Medicaid shares. Combined with the event-study evidence of parallel pre-trends, this supports the validity of the identification strategy used in the main difference-in-differences analysis.

4.1.3 Results

Table 4 reports the difference in differences estimates for the impact of the Medicaid reimbursement policy on postpartum depression screening, estimated separately for Medicaid and commercially insured patients. Each specification includes practice and year–quarter fixed effects, with standard errors clustered at the practice level.

Screening rates increased more in practices with higher pre-policy Medicaid exposure, but the pattern differs by payer. In the Medicaid sample, the post-policy slope in pre-policy Medicaid share is small and statistically indistinguishable from zero: 0.019 (s.e. 0.042) for any screening and 0.023 (s.e. 0.033) for the percent screened (columns 2 and 4). This pattern indicates that screening among Medicaid patients rose substantially across practices after the policy, with little variation by baseline Medicaid share. In other words, once reimbursement became available, practices appear to have adopted screening for their Medicaid population broadly, regardless of how many Medicaid patients they served before the policy.

By contrast, the commercial sample shows clear and statistically significant gradients. The coefficients are 0.109 (s.e. 0.033, $p < 0.01$) for any screening and 0.070 (s.e. 0.018, $p < 0.01$) for the percent screened (columns 6 and 8), implying increases of roughly 1.1 and 0.7 percentage points, respectively, for every 10 percentage point higher pre-policy Medicaid share. Because baseline pre-policy screening rate was near zero, these effects represent meaningful post-policy gains concentrated among practices more exposed to Medicaid reimbursement incentives.

The evidence suggests that practices serving more Medicaid patients before the policy were more likely to implement screening after reimbursement became available, and that this im-

Table 2: Maternal Covariate DiD by Pre-Policy Medicaid Share: Commercial Sample

Covariate	Coefficient (s.e.)	p-value	FDR-adjusted q
Age, mean (years)	4.3 (1.8)	0.02	0.12
Number of prenatal visits, mean	0.94 (0.82)	0.25	0.47
Income: \$0–14,999	0.13 (0.086)	0.12	0.34
Income: \$15,000–24,999	-0.11 (0.090)	0.21	0.46
Income: \$25,000–34,999	0.092 (0.028)	0.0012	0.023
Income: \$35,000–49,999	-0.24 (0.17)	0.15	0.34
Income: \$50,000–74,999	0.15 (0.073)	0.042	0.17
Income: \$75,000+	0.046 (0.31)	0.88	0.91
Urban	-0.0095 (0.065)	0.88	0.91
Rural	-0.078 (0.083)	0.35	0.56
Rural (isolated)	0.0032 (0.026)	0.90	0.91
Diabetes	-0.0080 (0.014)	0.58	0.74
Hypertension	0.00074 (0.0068)	0.91	0.91
Any chronic condition	0.16 (0.11)	0.14	0.34
Cesarean-section births	0.086 (0.098)	0.38	0.65
Pregnancy and delivery complications	0.059 (0.030)	0.050	0.18
Hispanic	0.045 (0.12)	0.70	0.82
Education: High school	-0.061 (0.10)	0.56	0.74
Education: College	0.18 (0.080)	0.029	0.13
White	-0.031 (0.042)	0.46	0.64
Black	0.045 (0.019)	0.021	0.12
Asian	0.037 (0.041)	0.37	0.56
Other race	-0.051 (0.054)	0.35	0.56
Married	-0.12 (0.10)	0.24	0.47
Prenatal care initiated in first 3 months	0.12 (0.076)	0.12	0.34
Preterm birth	0.17 (0.056)	0.0026	0.024
Born outside US	0.030 (0.081)	0.71	0.83

Notes: Each row reports a separate difference-in-differences regression of the listed maternal covariate on the interaction Post (2014–2019) × pre-policy Medicaid share, estimated for the commercial sample. Coefficients are reported with standard errors in parentheses below. Except for age and number of prenatal visits (reported in their natural units), all covariates are measured as proportions (0–1), and coefficients are expressed in percentage points. Each coefficient reflects the change in the outcome from the pre- to post-policy period associated with a 100% (one-unit) increase in a practice's pre-policy Medicaid share. All models include practice and year–quarter fixed effects, are weighted by the number of eligible well-child visits, and cluster standard errors by practice. p-values are unadjusted; FDR-adjusted q values follow the Benjamini–Hochberg procedure.

Table 3: Maternal Covariate DiD by Pre-Policy Medicaid Share: Medicaid Sample

Covariate	Coefficient (s.e.)	p-value	FDR-adjusted q
Age, mean (years)	0.064 (0.45)	0.89	0.97
Number of prenatal visits, mean	-0.61 (0.38)	0.11	0.46
Income: $0–14,999	-0.011 (0.022)	0.62	0.83
Income: $15,000–24,999	0.024 (0.012)	0.046	0.17
Income: $25,000–34,999	0.031 (0.053)	0.56	0.96
Income: $35,000–49,999	0.0010 (0.0090)	0.97	0.97
Income: $50,000–74,999	-0.022 (0.025)	0.38	0.82
Income: $75,000+	-0.011 (0.010)	0.31	0.72
Urban	-0.00084 (0.0041)	0.84	0.90
Rural	0.023 (0.019)	0.23	0.71
Rural (isolated)	-0.099 (0.090)	0.27	0.72
Diabetes	-0.011 (0.010)	0.31	0.72
Hypertension	-0.022 (0.025)	0.38	0.82
Any chronic condition	0.029 (0.037)	0.45	0.84
Cesarean-section births	0.063 (0.040)	0.12	0.46
Pregnancy and delivery complications	-0.046 (0.025)	0.076	0.46
Hispanic	-0.023 (0.022)	0.30	0.72
Education: High school	0.090 (0.053)	0.091	0.46
Education: College	-0.022 (0.028)	0.43	0.84
White	-0.37 (0.11)	0.0013	0.036
Black	0.048 (0.038)	0.22	0.71
Asian	-0.023 (0.022)	0.30	0.72
Other race	0.35 (0.12)	0.0028	0.039
Married	0.0017 (0.0076)	0.82	0.89
Prenatal care initiated in first 3 months	-0.099 (0.090)	0.27	0.72
Preterm birth	-0.0041 (0.0069)	0.55	0.80
Born outside US	-0.011 (0.010)	0.31	0.72

Notes: Each row reports a separate difference-in-differences regression of the listed maternal covariate on the interaction Post (2014–2019) × pre-policy Medicaid share, estimated for the Medicaid sample. Coefficients are reported with standard errors in parentheses below. Except for age and number of prenatal visits (reported in their natural units), all covariates are measured as proportions (0–1), and coefficients are expressed in percentage points. Each coefficient reflects the change in the outcome from the pre- to post-policy period associated with a 100% (one-unit) increase in a practice's pre-policy Medicaid share. All models include practice and year–quarter fixed effects, are weighted by the number of eligible well-child visits, and cluster standard errors by practice. p-values are unadjusted; FDR-adjusted q values follow the Benjamini–Hochberg procedure.

Table 4: Difference-in-Differences Estimates by Payer

	Medicaid				Commercial			
	Any Screening		% WCVs Screened		Any Screening		% WCVs Screened	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Post-policy (2014+)	0.0856***		0.0334		0.0305*		-0.0144*	
	(0.0274)		(0.0226)		(0.0156)		(0.00765)	
Pre-policy Medicaid share	0.101***		0.0667**		0.0576**		0.0102	
	(0.0324)		(0.0296)		(0.0250)		(0.0149)	
Post × Medicaid share (DiD)	0.0124	0.0192	-0.0206	0.0226	0.0767**	0.109***	0.0628***	0.0696***
	(0.0412)	(0.0415)	(0.0301)	(0.0329)	(0.0320)	(0.0331)	(0.0180)	(0.0177)
Practice fixed effects	No	Yes	No	Yes	No	Yes	No	Yes
Year–quarter fixed effects	No	Yes	No	Yes	No	Yes	No	Yes
Observations	11,336	11,332	11,336	11,332	11,144	11,143	11,146	11,145

Notes: Each column reports a separate regression at the practice–year–quarter level using billing-practice IDs. "Any screening" equals 1 if at least one well-child visit at the practice in quarter t included a PPD screen. "% WCVs Screened" is the share of well-child visits with a PPD screen. Standard errors in parentheses, clustered at the practice level. Percent-screening regressions are weighted by the number of well-child visits at that practice-quarter for the corresponding payer. $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.10$.

plementation extended to commercially insured visits within the same practices. The absence of a differential gradient among Medicaid visits indicates that screening for Medicaid patients increased broadly across practices, while commercial patients benefited more in high Medicaid share settings. This pattern supports the interpretation that the policy induced practice-level adoption rather than payer-specific behavior, consistent with the event study evidence of parallel pre-trends and post-policy divergence.

4.1.4 Robustness

I assess the sensitivity of results to practice definition and sample period. The All-Payer Claims Database records both billing and service provider identifiers, which align for roughly 75 percent of well-child visits based on the NPPES practice address. The main analysis uses billing identifiers, reflecting the locus of financial decision-making and reimbursement incentives. As a robustness check, I replicate the difference-in-differences model using the service-provider identifier, defined by the service provider's primary practice address in the NPPES. This alternative definition more directly reflects where care is delivered, rather than where claims are processed or aggregated for billing purposes.

Table 5 reports the corresponding estimates. The results are highly consistent across specifications: practices with higher pre-policy Medicaid exposure continue to exhibit larger post-policy increases in screening, particularly among commercially insured patients. The estimated coefficients are similar in magnitude to those from the billing-provider specification, and the patterns remain statistically significant for the commercial sample but small and imprecise for Medicaid. This pattern suggests that Medicaid screening increased broadly across practices after the policy, while diffusion to commercial patients was concentrated among practices more exposed to Medicaid incentives. The consistency of these findings under both definitions indicates that the results are not driven by billing structure or attribution but reflect true behavioral responses within practices to the introduction of Medicaid reimbursement.

Table 5: Difference-in-Differences Estimates by Payer (Service-Provider ID)

	Medicaid				Commercial			
	Any screening		% WCVs screened		Any screening		% WCVs screened	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Post-policy (2014+)	0.0757***		0.0334		0.0444**		-0.00521	
	(0.0254)		(0.0226)		(0.0190)		(0.00807)	
Pre-policy Medicaid share	0.0524**		0.0667**		0.0205		0.0189	
	(0.0243)		(0.0296)		(0.0224)		(0.0138)	
Post × Medicaid share (DiD)	0.0580	0.0651	-0.0206	0.0384	0.111***	0.116***	0.0697***	0.0696***
	(0.0431)	(0.0430)	(0.0301)	(0.0320)	(0.0366)	(0.0379)	(0.0185)	(0.0177)
Practice fixed effects	No	Yes	No	Yes	No	Yes	No	Yes
Year–quarter fixed effects	No	Yes	No	Yes	No	Yes	No	Yes
Observations	12,918	12,917	12,918	12,917	11,296	11,296	11,296	11,296

Notes: Each column reports a separate regression at the practice–year–quarter level using service-provider practice IDs. Any screening equals 1 if at least one well-child visit at the practice in quarter t included a PPD screen. % WCVs screened is the share of well-child visits with a PPD screen; those regressions are weighted by the number of eligible visits at that practice–quarter for the corresponding payer. Practice and year–quarter fixed effects absorb the main effects of Post and pre-policy Medicaid share in FE specifications, so those rows are blank in FE columns by design. Standard errors in parentheses, clustered at the practice level. $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.10$.

I also re-estimate the difference-in-differences models restricting the sample to the 2014–2016 period, before the USPSTF recommendation took effect in 2017. As discussed in Section 2, this update required commercial insurers to cover postpartum depression screening without cost sharing, potentially diluting the identifying variation in later years. The pre-2017 restriction yields similar results: the post-by-Medicaid-share coefficient remains positive and significant for commercially insured patients, with magnitudes comparable to the baseline estimates (Table 6). If commercial plans began reimbursing screening after 2017, increased screening among low–Medicaid-share practices would bias the DiD estimates toward zero, implying that the main estimates are, if anything, conservative. The persistence of the effect in the pre-2017 sample supports the interpretation that the observed increases reflect behavioral responses to the Medicaid reimbursement policy rather than subsequent documentation or coverage changes.

Table 6: Difference-in-Differences Estimates by Payer (Pre-2017 Sample)

	Medicaid				Commercial			
	Any screening		% WCVs screened		Any screening		% WCVs screened	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Post-policy (2014+)	0.0759***		0.0279		0.0314***		-0.00553*	
	(0.0184)		(0.0173)		(0.0108)		(0.00388)	
Pre-policy Medicaid share	0.0329		0.0195		0.0278		0.000766	
	(0.0219)		(0.0160)		(0.0200)		(0.00540)	
Post × Medicaid share (DiD)	0.0449	0.0374	-0.00227	0.00950	0.0670***	0.0844***	0.0372***	0.0301***
	(0.0338)	(0.0343)	(0.0237)	(0.0149)	(0.0245)	(0.0265)	(0.0115)	(0.00792)
Practice fixed effects	No	Yes	No	Yes	No	Yes	No	Yes
Year–quarter fixed effects	No	Yes	No	Yes	No	Yes	No	Yes
Observations	7,592	7,590	7,593	7,590	7,123	7,122	7,130	7,122
R^2	0.008	0.399	0.017	0.611	0.011	0.387	0.024	0.431

Notes: Each column reports a separate regression at the practice–year–quarter level using billing-practice identifiers. Any screening equals 1 if at least one well-child visit at the practice in quarter t included a PPD screen. % WCVs screened is the share of well-child visits with a PPD screen; those regressions are weighted by the number of eligible visits at that practice–quarter for the corresponding payer. Fixed effects columns omit the main effects of Post and pre-policy Medicaid share by construction. Standard errors in parentheses, clustered at the practice level. $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.10$.

4.2 Impact Across Payers

4.2.1 Empirical Strategy

In the previous section, I examined how screening rates changed across practices with differing baseline Medicaid exposure using a difference-in-differences framework. Those results showed that screening rose more in high–Medicaid-share practices, particularly among commercially insured patients, suggesting that the policy's influence extended beyond directly affected Medicaid claims. To test whether this pattern reflects payer-targeted responses or practice-wide adoption, I estimate a triple-difference (DDD) model.

This specification compares changes in screening for Medicaid versus commercially insured patients within the same practice, before and after the policy, and across practices with different pre-policy Medicaid shares:

$$\begin{aligned} Y_{pgt} = {} & \beta_0+\beta_1\text{Post}_t+\beta_2\text{MedicaidShare}_p+\beta_3\text{Medicaid}_g+\beta_4(\text{Post}_t\times\text{MedicaidShare}_p)+\beta_5(\text{Post}_t\times\text{Medicaid}_g) \\ & + \beta_6(\text{MedicaidShare}_p \times \text{Medicaid}_g) + \beta_7(\text{Post}_t \times \text{MedicaidShare}_p \times \text{Medicaid}_g) + X_{pgt}\gamma + \epsilon_{pgt}. \end{aligned} \tag{3}$$

Here, Y_{pgt} denotes the percentage of well-child visits at practice p for insurance group g (Medicaid or commercial) in quarter t that included a PPD screening. Medicaid_g equals 1 for Medicaid and 0 for commercial. Post_t and MedicaidShare_p are defined as before, and X_{pgt} includes time-varying practice-level covariates and patient composition controls.

The coefficient of interest, β_7, measures whether practices with higher baseline Medicaid exposure experienced larger post-policy increases in screening for Medicaid patients relative to commercially insured patients. A positive and significant β_7 would indicate a payer-targeted response, where financial incentives led practices to expand screening primarily for Medicaid patients. Conversely, an insignificant β_7 coupled with a positive β_4 (the DiD interaction) would imply practice-wide adoption—once screening was introduced in response to Medicaid reimbursement, practices applied it broadly across payer groups. This design thus distinguishes financial targeting from organizational spillovers within practices.

4.2.2 Results

I extend the difference-in-differences analysis by comparing Medicaid and commercially insured patients within the same practice. The goal is to test whether practices with higher pre-policy Medicaid exposure expanded screening primarily for Medicaid patients (a payer-targeted response) or more uniformly across all patients (a practice-level adoption).

Table 7 compares within-practice changes for Medicaid versus commercially insured patients and allows the post-policy effect to vary with pre-policy Medicaid share. In the fixed-effects specification, the triple interaction $\text{Post}_t \times \text{MedicaidShare}_p \times \text{Medicaid}_g$ is negative and statistically significant for the extensive margin (any screening), approximately -0.080 with a standard error of about 0.031 (column 2), but is close to zero and statistically insignificant for the intensive margin (percent of well-child visits screened), roughly -0.0049 with a standard error of about 0.0089 (column 4). The negative extensive-margin estimate indicates that, in higher–Medicaid-share practices, the post-policy increase in the probability that a quarter has any screening is smaller for Medicaid patients relative to commercial patients. Combined with the near-zero intensive-margin estimate, this pattern is consistent with practices adopting uniform workflows once screening is in place: the share of visits screened within screening-active quarters moves similarly across payers, while the expansion at the extensive margin is relatively larger for commercial patients in higher–Medicaid-share practices. This aligns with the DiD findings that show stronger post-policy gradients for commercial patients and supports a practice-level adoption story rather than payer-targeted implementation.

Table 7: Triple Difference-in-Differences Estimates

	Any Screening		% WCVs Screened	
	(1)	(2)	(3)	(4)
Post-policy (2014+)	0.0314**		-0.00384	
	(0.0156)		(0.0117)	
Pre-policy Medicaid share	0.0307		0.0283	
	(0.0240)		(0.0185)	
Post × Medicaid share	0.0709**	0.110***	0.0218	0.0570**
	(0.0302)	(0.0324)	(0.0195)	(0.0224)
Medicaid × Medicaid share	0.0751***	0.182***	0.0171*	0.0244***
	(0.0223)	(0.0254)	(0.00931)	(0.00884)
Post × Medicaid	0.0385**	0.0520***	0.0154**	0.0203***
	(0.0160)	(0.0173)	(0.00632)	(0.00638)
Post × Medicaid × Medicaid share	-0.0422	-0.0795**	-0.00174	-0.00488
	(0.0287)	(0.0312)	(0.00936)	(0.00886)
Practice fixed effects	No	Yes	No	Yes
Year–quarter fixed effects	No	Yes	No	Yes
Observations	23,589	23,586	23,589	23,586

Notes: Each column reports a separate regression at the practice–year–quarter level. "Any screening" equals 1 if at least one well-child visit at the practice in quarter t included a PPD screen. "% WCVs screened" is the share of well-child visits with a PPD screen; these regressions are weighted by the number of eligible visits at that practice in that quarter. Standard errors in parentheses, clustered at the practice level. $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.10$.

Across specifications, the evidence points to a practice-level behavioral response to Medicaid reimbursement rather than a payer-specific one. Practices with higher pre-policy Medicaid exposure were the most likely to increase screening following the introduction of reimbursement, and these gains extended to commercially insured patients. The absence of a positive triple-difference effect and the similar pre-policy trends across payers suggest that screening adoption reflected a change in clinical workflow rather than differential targeting of Medicaid patients. The results imply that financial incentives for one payer can influence broader provider behavior, generating cross-payer spillovers that improve care delivery beyond the directly affected population.

4.2.3 Robustness

As a robustness check, I re-estimate the triple-difference model using the practice location linked to the service provider's NPI rather than the billing provider's address. This approach attributes screening activity to the physical site of care, capturing potential diffusion through co-located providers rather than shared billing systems.

Table 8 reports the results. The triple interaction term, $\text{Post}_t \times \text{MedicaidShare}_p \times \text{Medicaid}_g$, remains negative and statistically significant for the any-screening outcome in the specification with fixed effects (-0.069, s.e. 0.029), but close to zero and insignificant for the percent-screening outcome. The direction and magnitude of the coefficients closely match the billing-based estimates, indicating that the observed pattern is not sensitive to how practices are defined. The results again suggest that high–Medicaid-share practices did not disproportionately expand screening for Medicaid patients relative to commercially insured patients, consistent with practice-wide adoption and diffusion of screening protocols once reimbursement was introduced.

Table 8: Triple Difference-in-Differences Estimates (Service-Provider ID)

	Any Screening		% WCVs Screened	
	(1)	(2)	(3)	(4)
Post-policy (2014+)	0.0174		0.00630	
	(0.0163)		(0.0127)	
Pre-policy Medicaid share	0.0129		0.0296*	
	(0.0226)		(0.0166)	
Post × Medicaid share	0.0681**	0.0984***	0.0268	0.0420*
	(0.0318)	(0.0330)	(0.0228)	(0.0251)
Medicaid × Medicaid share	0.0743***	0.153***	-0.00631	0.0233***
	(0.0219)	(0.0241)	(0.00979)	(0.00691)
Post × Medicaid	0.0362**	0.0431**	0.00189	0.0143**
	(0.0172)	(0.0183)	(0.00745)	(0.00659)
Post × Medicaid × Medicaid share	-0.0451*	-0.0692**	0.0215*	0.00894
	(0.0264)	(0.0291)	(0.0113)	(0.0103)
Practice fixed effects	No	Yes	No	Yes
Year–quarter fixed effects	No	Yes	No	Yes
Observations	25,134	25,132	25,134	25,132

Notes: Each column reports a separate regression at the practice–year–quarter level using service-provider practice identifiers. "Any screening" equals 1 if at least one well-child visit at the practice in quarter t included a PPD screen. "% WCVs screened" is the share of well-child visits with a PPD screen; these regressions are weighted by the number of eligible visits at that practice in that quarter. Standard errors in parentheses, clustered at the practice level. $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.10$.

Restricting the sample to the pre-2017 period, before the USPSTF's B-grade recommendation for postpartum depression screening took effect, produces consistent results. The triple-difference estimates in Table 9 mirror the main findings: the triple interaction term, $\text{Post}_t \times \text{MedicaidShare}_p \times \text{Medicaid}_g$, remains negative and statistically significant for the any-screening outcome in the specification with fixed effects (-0.073, s.e. 0.031) and close to zero for the percent-screening outcome. These results imply that, as in the main analysis, practices with greater Medicaid exposure continued to expand screening broadly across payer types, rather than disproportionately for Medicaid patients.

Because the USPSTF update in 2016 required commercial insurers to cover postpartum depression screening without cost sharing starting in 2017, any resulting spillover would bias the difference-in-differences and triple-difference estimates toward zero. The persistence of the effects when restricting to 2014–2016 therefore reinforces that the observed screening increases were driven by the 2014 Medicaid reimbursement policy rather than by later coverage mandates or changes in coding and documentation practices.

Table 9: Triple Difference-in-Differences Estimates (Pre-2017 Sample)

	Any Screening		% WCVs Screened	
	(1)	(2)	(3)	(4)
Post-policy (2014+)	0.0341**		0.00844	
	(0.0150)		(0.00578)	
Pre-policy Medicaid share	0.0326*		0.000118	
	(0.0188)		(0.0108)	
Post × Medicaid share	0.0669**	0.109***	0.0128	0.0196***
	(0.0315)	(0.0332)	(0.0110)	(0.00706)
Medicaid × Medicaid share	0.0682**	0.175***	0.0180***	0.00915**
	(0.0274)	(0.0298)	(0.00587)	(0.00393)
Post × Medicaid	0.0392**	0.0543***	0.0122***	0.0123***
	(0.0168)	(0.0180)	(0.00404)	(0.00361)
Post × Medicaid × Medicaid share	-0.0395	-0.0725**	-0.000709	0.00151
	(0.0295)	(0.0308)	(0.00620)	(0.00543)
Practice fixed effects	No	Yes	No	Yes
Year–quarter fixed effects	No	Yes	No	Yes
Observations	15,460	15,458	15,460	15,458
R-squared	0.194	0.501	0.194	0.501

Notes: Each column reports a separate regression at the practice–year–quarter level using billing-practice IDs, restricted to 2014–2016. "Any screening" equals 1 if at least one well-child visit at the practice in quarter t included a PPD screen. "% WCVs screened" is the share of well-child visits with a PPD screen; these regressions are weighted by the number of eligible visits at that practice in that quarter. Standard errors in parentheses, clustered at the practice level. $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.10$.

5 Mechanism for Screening Adoption: Individual Physicians versus Practices

5.1 Empirical Strategy

The preceding results show that the Medicaid reimbursement led to broad increases in postpartum depression (PPD) screening across patients, consistent with practice-wide adoption. To further assess whether this diffusion reflects changes in individual physician behavior or institutional systems at the practice level, I adapt a movers framework following Finkelstein et al. (2016) and Molitor (2018). The analysis exploits physicians who switch billing practices within Colorado, allowing comparison of screening behavior before and after the move.

The intuition follows migration-style research designs that distinguish individual from place effects. If screening adoption reflects practice-level systems, such as EMR prompts, nurse-administered checklists, or standardized protocols for well-child visits, then a physician who moves from a low-screening to a high-screening billing practice should increase their screening rate after the switch. If screening reflects persistent physician-specific preferences or beliefs, the rate should remain unchanged. Observing how screening behavior changes around a move therefore helps identify whether adoption is driven by institutional incentives or individual persistence.

For each move episode m of physician j, I compute two measures of the difference in screening environments between the destination and origin billing practices:

$$\Delta^{\text{pre}}_{jm} = \text{Rate}^{(-j)}_{\text{dest}(jm),\text{pre}} - \text{Rate}^{(-j)}_{\text{orig}(jm),\text{pre}}, \quad \Delta^{\text{post}}_{jm} = \text{Rate}^{(-j)}_{\text{dest}(jm),\text{post}} - \text{Rate}^{(-j)}_{\text{orig}(jm),\text{post}},$$

where each rate is averaged over four quarters before and after the switch, excluding the moving physician (leave-one-out). These two measures capture the pre- and post-move differences in practice-level screening intensity, providing a concise summary of the screening environments experienced by movers. Calculating these practice rates in a leave-one-out manner ensures that a physician's own screening behavior does not mechanically affect the measured screening rate of the origin or destination practice, isolating the practice-level environment they are exposed to.

The distribution of $\Delta^{\text{post}}_{jm}$ is centered near zero, with substantial variation in both directions: many movers transition to practices with modestly higher or lower screening rates (for example, $|\Delta^{\text{post}}| \leq 0.10$), while others experience much larger changes. The presence of both positive and negative gaps indicates that movers are exposed to a wide range of screening environments, providing identifying variation in cross-practice exposure.

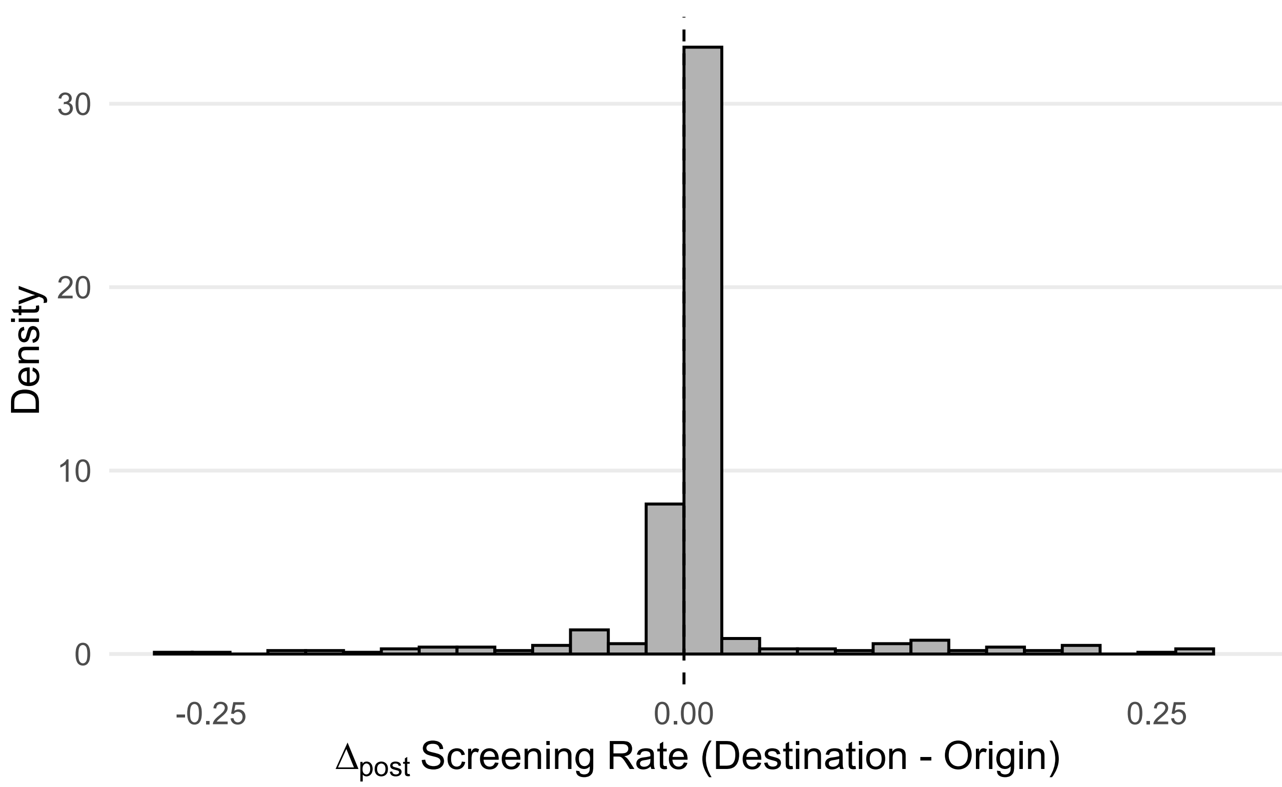


Figure 5: Distribution of Post-Move Differences in Screening Rates Across Moves

I estimate a stacked event study at the provider–practice–quarter level, controlling for the same practice-level, ZIP-code, and maternal covariates used in the main difference-in-differences analysis:

$$Y_{jpmt} = \eta_{o(jm)} + \lambda_t + \sum_{k \in \mathcal{K},\, k \neq -1} \theta_k \left[\mathbf{1}\{\tau_{jmt} = k\} \times \Delta_{jm}^{\text{pre/post}} \right] + X_{pt}\gamma + \varepsilon_{jpmt}. \tag{4}$$

Here, Y_{jpmt} is the share of well-child visits with a PPD screen for physician j at practice p in quarter t. The interaction term uses Δ_{jm}^{pre} for quarters before the move ($\tau_{jmt} < 0$) and $\Delta_{jm}^{\text{post}}$ for quarters after the move ($\tau_{jmt} \geq 0$), corresponding to the leave-one-out difference in average screening rates between the destination and origin billing practices during the pre- and post-switch periods. The vector X_{pt} includes the same practice-, ZIP-code-, and maternal-level covariates as in the main DiD specification. All regressions are weighted by the number of well-child visits in each physician–practice–quarter cell. Origin-practice fixed effects $\eta_{o(jm)}$ absorb baseline differences across billing systems, while λ_t controls for quarter-specific shocks common to all physicians. Standard errors are clustered by both physician and practice. This is the main specification. I also show the event study results are similar when using provider fixed effects in the Appendix.

The coefficients θ_k trace how physician screening behavior responds to differences in screening intensity between destination and origin practices. Pre-move coefficients ($k \leq -2$) should be close to zero if movers are not differentially trending before the switch. Positive post-move coefficients ($k \geq 0$) indicate convergence toward the destination practice's screening rate per unit of $\Delta_{jm}^{\text{post}}$, consistent with institution-driven changes in behavior. To summarize these effects, I also estimate a single difference-in-differences specification that interacts a post-move indicator with $\Delta_{jm}^{\text{post}}$ ("Post $\times \Delta^{\text{post}}$").

The movers sample includes physicians with at least two consecutive quarters at their origin billing practice and two at their destination, ensuring sufficient pre- and post-move observations. Moves are identified using the NPPES primary billing practice address, which reflects the financial and administrative entity responsible for claims submission. The analysis is restricted to allopathic and osteopathic physicians whose primary specialty is pediatrics or primary care, as these providers are eligible to bill for postpartum depression screening under Medicaid guidelines. This design provides a direct test of whether the diffusion of screening practices occurred through billing-level organizational systems or through individual physician persistence.

Table 10 compares physicians who switched primary billing practices ("movers") to those who remained at the same practice ("non-movers") during the post-policy period (2014–2019). For movers, practice characteristics are measured based on their origin billing practice prior to the switch. For each physician, I compute the yearly average ZIP-code and practice-level characteristics of their primary billing location between 2014 and 2019, then take the mean across years to obtain post-policy averages. These physician-level averages are subsequently averaged across movers and non-movers. Because practices with more pediatric or primary care

Table 10: Movers vs. Non-Movers: Post-Policy Practice and ZIP Characteristics (2014–2019)

	Non-movers	Movers
ZIP-level characteristics (%)		
High school or higher	91.42	90.85
	(6.83)	(6.34)
Bachelor's degree or higher	42.06	41.22
	(17.35)	(17.61)
Foreign-born	10.13	10.05
	(6.45)	(5.39)
English only	83.09	82.48
	(10.17)	(9.05)
Spanish spoken	10.85	12.05
	(9.27)	(8.92)
Non-English spoken	16.91	17.52
	(10.17)	(9.05)
White	83.30	84.36
	(10.21)	(9.78)
Black	4.45	3.57
	(6.44)	(5.61)
Asian	3.33	2.98
	(2.90)	(2.61)
Other race	4.22	4.44
	(3.49)	(3.74)
Female	50.48	50.42
	(2.95)	(2.49)
Age 20–64	62.72	62.31
	(7.47)	(7.07)
Household and population measures		
Total households	11,177	11,561
	(5,401)	(5,472)
Average household size	2.44	2.47
	(0.35)	(0.34)
Total population	28,144	29,574
	(14,583)	(14,610)
Mean household income ($)	82,765	83,363
	(26,896)	(26,737)
Median household income ($)	63,655	63,370
	(21,760)	(19,479)
Practice characteristics		
Number of infants	1,549	1,006
	(3,610)	(2,995)
Total physicians	61.75	35.80
	(115.58)	(74.65)
Hospital (%)	11.8	6.8
	(32.3)	(25.2)
Ambulatory facility (%)	33.2	33.6
	(47.1)	(47.3)
Agency (%)	1.0	1.5
	(10.1)	(12.1)
Managed care organization (%)	6.6	2.2
	(24.8)	(14.5)
Nursing facility (%)	0.2	0.0
	(4.4)	(0.0)
Number of physicians	3,739	404

Notes: Means with standard deviations in parentheses. ZIP and practice characteristics are averaged over 2014–2019 (post-policy period). Each observation corresponds to a physician's primary billing practice; for movers, characteristics are measured using their origin billing practice prior to the switch. Movers are physicians who switched primary billing practices during 2014–2019 with at least two consecutive quarters observed at both the origin and destination. Non-movers remained in the same practice over the same period.

physicians contribute more individual observations, they implicitly receive greater weight in the comparison.

Movers and non-movers exhibit broadly similar ZIP-level and household characteristics, including education, income, and racial composition, suggesting that mobility is not systematically concentrated in distinct socioeconomic areas. This similarity supports the credibility of the movers framework by showing that differences in community context are unlikely to bias within-physician estimates of behavioral change.

The main differences appear at the practice level. Movers tend to come from smaller practices, with fewer total physicians and infants served on average, and are somewhat less likely to be affiliated with hospitals or managed care organizations. This pattern suggests that mobility is more common among physicians in smaller or independent practices, where structural changes and practice consolidation occur more frequently. While movers may not represent all physicians statewide, the event study isolates within-physician adjustments in screening behavior, holding constant individual characteristics through fixed effects. The comparison therefore supports that the observed convergence toward destination practice behavior reflects place-based adaptation rather than selection into systematically different environments.

5.2 Results: Physician Movers

This section examines whether the observed practice-level adoption of postpartum depression (PPD) screening reflects physician-specific behavior or adaptation to practice-level systems. Figure 6 plots the event-study coefficients θ_k from equation (4), which interact relative event time with the difference in screening intensity between the physician's destination and origin practices. The estimates are centered at the quarter of the move ($\tau = 0$), with $\tau = -1$ omitted.

The event-study results show little movement in screening behavior prior to the switch: preperiod coefficients are close to zero and statistically indistinguishable from each other, suggesting that movers and non-movers followed similar trajectories before changing practices. Immediately after the switch, screening rates among movers begin to converge toward the destination practice's screening level. The increase stabilizes within a few quarters, implying that adaptation occurs relatively quickly once physicians are exposed to new practice systems. This pattern supports the view that institutional features—such as EMR prompts, billing templates, or staff-administered checklists—shape how often screening is conducted, independent of physician-specific preferences.

Table 11 presents the corresponding DiD results. Across all specifications, the interaction term between post-switch and the destination–origin screening gap is positive and significant, with an estimated coefficient around 0.7. This implies that when a physician moves to a practice with a 10–percentage-point higher screening rate than their origin (measured excluding their own behavior), their own screening rate rises by roughly 7 percentage points. The estimates are stable across specifications that include origin-practice, physician, and year–quarter fixed effects.

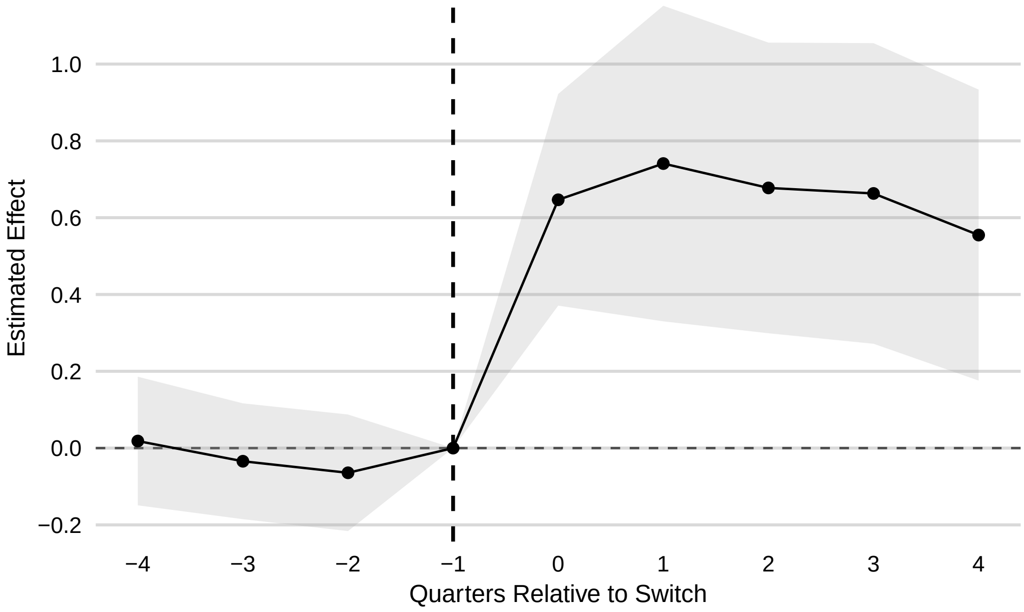


Figure 6: Physician Movers Event Study: 4-Quarter Window

Notes: Points plot coefficients from a stacked event study regression of the physician–practice–quarter screening rate on event-time indicators interacted with the leave-one-out difference in average screening rates between the destination and origin billing practices, using Δ_{jm}^{pre} for quarters before and $\Delta_{jm}^{\text{post}}$ for quarters after the move. The regression includes origin-practice and quarter fixed effects and controls for the same practice-, ZIP-code-, and maternal-level covariates as in the main difference-in-differences specification. Standard errors are clustered by physician and practice. Shaded areas show 95% confidence intervals. The -1 bin is omitted by construction. Joint pre-trend test (all $k \leq -2$): $p = 0.486$.

Overall, the results indicate that the diffusion of screening practices operates partly through physician adaptation to new organizational environments. When physicians move to practices with higher baseline screening rates, their own screening behavior rises proportionally, suggesting that institutional factors such as billing protocols, EMR prompts, or standardized workflows play a substantial role in shaping individual practice patterns. The magnitude of the response, roughly 70 percent convergence to the destination practice's screening level, points to an important role for practice-level systems in driving adoption, rather than screening behavior being solely determined by physician-specific preferences or beliefs.

Table 11: Physician Movers: Difference-in-Differences Estimates

	(1)	(2)	(3)
Post-switch	-0.006	-0.016*	-0.010
	(0.007)	(0.009)	(0.009)
ΔScreening (dest $-$ origin)	0.021	-0.238	-0.472***
	(0.118)	(0.194)	(0.147)
Post-switch $\times$ ΔScreening	0.677***	0.700***	0.695***
	(0.163)	(0.158)	(0.165)
Origin practice FE	Yes		Yes
Provider FE		Yes	Yes
Year–quarter FE	Yes	Yes	Yes
Observations	3,314	3,301	3,301

Notes: Outcome is the share of well-child visits with a billed PPD screen in the physician–practice–quarter cell. "Post-switch" equals 1 for quarters after the physician changes billing practices. ΔScreening is the leave-one-out difference in average screening rates between destination and origin billing practices in the post period. Models include the same maternal, practice-level, and ZIP-code covariates as in the main specification. Standard errors are clustered by physician and practice. $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.10$.

5.3 Robustness

As a robustness check, I extend the event-study window to include eight quarters before and after each physician move. This specification allows for a longer adjustment period and provides an additional test of pre-trend stability. The estimation follows the same specification as in equation (4), interacting event-time indicators with the corresponding pre- and post-move leave-one-out screening gaps. For this version, Δ_{jm}^{pre} and $\Delta_{jm}^{\text{post}}$ are computed as the average differences in screening rates between the destination and origin billing practices over the eight quarters before and after each move, respectively. The regression includes the same set of practice-, ZIP-code-, and maternal-level covariates and fixed effects as in the main specification.

The results, shown in Figure 7, confirm the main findings. Coefficients in the pre-period remain close to zero, suggesting no anticipatory changes in screening behavior prior to the switch. Post-move coefficients rise sharply in the first few quarters after the move and then stabilize, consistent with convergence toward the destination practice's screening rate. The magnitude of the post-move response is somewhat larger than in the four-quarter specification, likely reflecting the use of longer averaging windows that smooth quarter-to-quarter noise and capture more sustained changes in physician behavior. Overall, the extended event window reinforces the conclusion that the adoption of postpartum depression screening is largely driven by practice-level systems that shape physician behavior after switching billing environments, rather than by persistent physician-specific screening preferences.

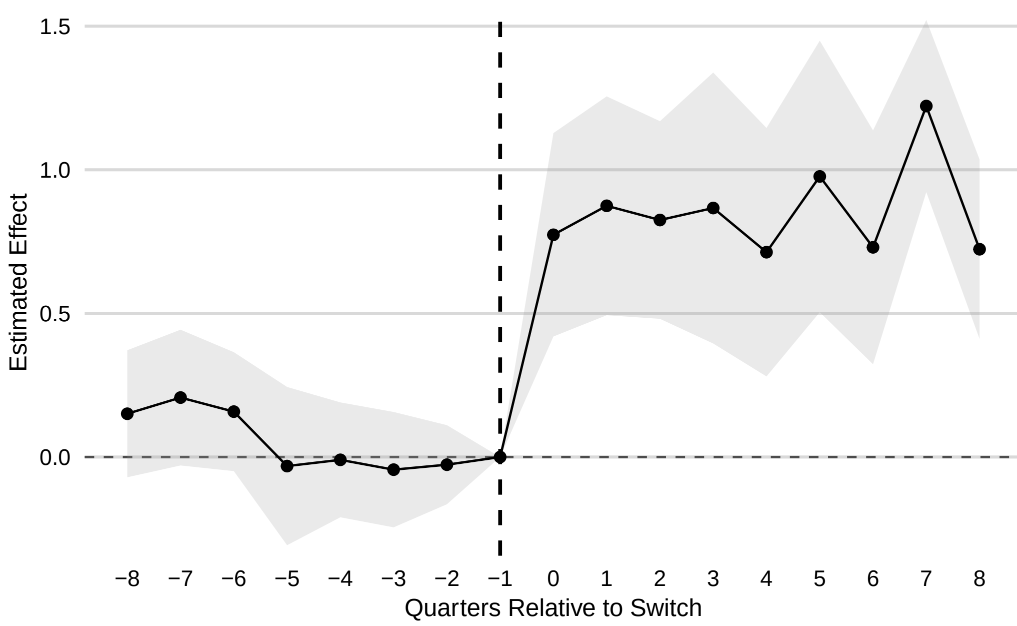


Figure 7: Physician Movers Event Study: Extended 8-Quarter Window

Notes: Points plot coefficients from a stacked event study regression of the physician–practice–quarter screening rate on event-time indicators interacted with the leave-one-out difference in average screening rates between the destination and origin billing practices, using Δ_{jm}^{pre} for quarters before and $\Delta_{jm}^{\text{post}}$ for quarters after the move. The regression includes origin-practice and quarter fixed effects and controls for the same practice-, ZIP-code-, and maternal-level covariates as in the main difference-in-differences specification. Standard errors are clustered by physician and practice. Shaded areas show 95% confidence intervals. The −1 bin is omitted by construction. Joint pre-trend test (all $k \leq -2$): $p = 0.372$.

Overall, the extended event window reinforces the conclusion that the adoption of postpartum depression screening is largely driven by practice-level systems that influence physician behavior after switching billing environments, rather than by persistent physician-specific screening preferences.

6 Conclusion

This paper examines how targeted Medicaid reimbursement for postpartum depression (PPD) screening shaped provider behavior and generated cross-payer spillovers. Using linked birth records and the Colorado All-Payer Claims Database from 2012–2019, I find that the introduction of Medicaid reimbursement for an existing screening code led to broad increases in PPD screening among Medicaid patients across all practices, consistent with widespread adoption of reimbursable screening. Consequently, the difference-in-differences estimates for the Medicaid group are not statistically significant, reflecting uniform improvements across practices regardless of their baseline Medicaid exposure. In contrast, screening among commercially insured mothers rose more sharply in high–Medicaid-share practices, indicating that providers in Medicaid-exposed settings adopted standardized screening workflows and applied them to all patients, not just those covered by Medicaid. The triple-difference estimates further confirm this interpretation: conditional on a practice's baseline Medicaid share, post-policy gains were not disproportionately concentrated among Medicaid patients, consistent with payer-neutral, practice-wide adoption.

The physician movers analysis provides complementary evidence on mechanisms. When physicians switched billing practices, their screening behavior adjusted sharply toward the destination practice's norms, converging by roughly 70 percent within the first year after the move. The extended eight-quarter event study confirms this pattern: no pre-trends are observed, and screening rates rise immediately following the move, consistent with physicians adopting the institutional routines of their new billing environment. These findings suggest that practice-level systems, rather than fixed physician preferences, drive much of the observed diffusion in screening behavior.

The results have several policy implications. First, targeted public payment reforms can reshape provider behavior more broadly than intended when they alter organizational routines that affect all patients. Designing reimbursement incentives with practice-wide implementation in mind may thus amplify their total impact. Second, because physicians adapt rapidly to new organizational environments, interventions that reach provider groups or practices, rather than individual clinicians, may be more effective in achieving sustained behavioral change. Finally, in maternal mental health care, where screening can be easily integrated into existing well-child workflows, such practice-level diffusion offers a scalable mechanism for improving detection and treatment of postpartum depression.

Overall, the evidence shows that targeted Medicaid reimbursement not only increased screening among publicly insured mothers but also elevated screening standards across payer types. In this setting, a modest payment policy catalyzed organization-wide behavioral change, illustrating how public insurance reforms can diffuse through shared delivery systems to enhance maternal and child health at a population level.

References

American College of Obstetricians and Gynecologists (2018). Optimizing postpartum care. Obstetrics & Gynecology, 131(5):e140–e150.

Baicker, K., Chernew, M. E., and Robbins, J. A. (2013). Medicare advantage and hospital utilization. The American Journal of Managed Care, 19(10):e350–e357.

Baicker, K. and Robbins, J. A. (2015). The spillover effects of medicare managed care: Medicare advantage and hospital utilization. Journal of Public Economics, 125:1–12.

Barnett, M. L., Mehrotra, A., and Landon, B. E. (2020). Effects of physician payment reform on patient care: Evidence from the medicare bundled payments for care improvement initiative. Health Affairs, 39(10):1762–1770.

Clemens, J. and Gottlieb, J. D. (2017). Do physicians' financial incentives affect medical treatment and patient health? American Economic Review, 107(5):983–1019.

Clemens, J., Leganza, J. M., and Masucci, A. (2025). Plugging gaps in payment systems: Evidence from the take-up of new medicare billing codes. Contemporary Economic Policy, 43(1):1–26.

Currie, J. and Malinovskaya, A. (2025). Can pediatric well-child visits improve maternal mental health? evidence from michigan's postpartum depression screening policy. Health Economics Review, 15(1):1–22.

Dagher, R. K. and Bair-Merritt, M. H. (2021). Trends and disparities in postpartum depression. Maternal and Child Health Journal, 25(2):186–194.

Earls, M. F., Yogman, M. W., Mattson, G., and Rafferty, J. (2010). Incorporating recognition and management of perinatal and postpartum depression into pediatric practice. Pediatrics, 126(5):1032–1039.

Einav, L., Finkelstein, A., and Mahoney, N. (2020). Spillover effects of bundled payment programs: Evidence from medicare. American Economic Review, 110(7):1951–1989.

Finkelstein, A., Gentzkow, M., and Williams, H. (2016). Sources of geographic variation in health care: Evidence from patient migration. The Quarterly Journal of Economics, 131(4):1681–1726.

Glied, S. A. and Zivin, J. G. (2002). How do doctors behave when some (but not all) of their patients are in managed care? Journal of Health Economics, 21(2):337–353.

Gordon, S. H., Andrews, C., Beharie, N., and Others (2025). Medicaid reimbursement and postpartum depression screening: Evidence from colorado. JAMA Pediatrics, 179(2):143–152.

Khadka, R., Smith, J., and Thompson, L. (2024). Postpartum depression: Epidemiology and recent updates. Journal of Maternal Health, 10:50–62.

Liberto, T. L. (2012). Screening for depression in pregnant and postpartum women: a review of the evidence. Journal of the American Academy of Nurse Practitioners, 24(12):754–760.

Molitor, D. (2018). The evolution of physician practice styles: Evidence from cardiologist migration. American Economic Journal: Economic Policy, 10(1):326–356.

Slomian, J., Honvo, G., Emonts, P., Reginster, J.-Y., and Bruyere, O. (2019). Consequences of maternal postpartum depression: A systematic review of maternal and infant outcomes. Women's Health, 15:1–12.

Yawn, B. P., Olson, A. L., Bertram, S., Pace, W., Wollan, P., and Dietrich, A. J. (2012). Postpartum depression: Screening, diagnosis, and management programs 2000 through 2010. Depression Research and Treatment, 2012:363964.

Note: An online appendix with additional figures, tables, and robustness checks will be made available in a future version.