Mapping and Modeling Maternal Vaccination: Urbanicity and Uninsurance Across U.S. States

Mengying Xia[1], Ruining Zhou[1], Xinyi Yu[1], and Ruiqi Li[1]

[1]**Health Data Science Program, Dartmouth College, Hanover, NH, USA**

ABSTRACT

We assembled a 2012–2023 state–year panel combining CDC pregnancy vaccination estimates (influenza, Tdap) with sample sizes, NCHS urban–rural classifications (2013/2023), and state counts of insured and uninsured women aged 19–64. Aim 1 mapped descriptive patterns; Aim 2 quantified associations using cross-sectional regressions of state means and two-way fixed-effects models with state and year indicators, clustered standard errors, and sample-size weights, including an Uninsurance×NCHS interaction. Cross-sectionally, influenza coverage shows no clear urban–rural gradient and Tdap trends positive but not conventionally significant. Within states over time, influenza coverage is not detectably associated with uninsurance, whereas Tdap coverage increases with uninsurance; the Uninsurance×NCHS interaction is not statistically significant for either vaccine. Interpretation should consider sparse Tdap reporting in some state–years, state-level aggregation, and potential time-varying confounding. The curated panel and code support extensions with additional covariates and finer geography to inform equitable maternal immunization strategies.

Keywords: maternal vaccination, influenza, Tdap, uninsurance, urbanicity, NCHS, fixed effects

INTRODUCTION

Women's health outcomes vary widely across states in the United States indicating differences in healthcare availability and preventative treatment[1]. Vaccination during pregnancy plays a crucial role in safeguarding the health of pregnant women and their infants.The presently recommended maternal vaccinations, which include influenza, tetanus toxoid, reduced diphtheria toxoid, and acellular pertussis (Tdap), and Coronavirus disease 2019 (COVID-19). They all have considerably decreased illness incidence and complications in these susceptible groups [2]. However, the vaccination rate during pregnancy is still lower than the expected target, especially for women with lower socioeconomic status and ethnic minority women [3]. In addition, the uninsured rates vary greatly among states across the United States which further influences access to essential healthcare services [4]. Geographic location also contributes to these differences in access to and uptake of vaccination services. Compared with urban residents, rural residents are particularly severely affected by the lack of insurance [5].Women without medical insurance face additional barriers to getting recommended preventative care. The purpose of this analysis is to investigate the link between the degree of urbanization in each state in the United States, women's insurance coverage rates, and vaccination rates for pregnant and postpartum women. We selected the recommended vaccinations for pregnant and postpartum women, including influenza and Tdap. By combining data from the Centers for Disease Control and Prevention (CDC), the National Center for Health Statistics (NCHS), and Kaiser Family Foundation (KFF), we aim to discover state-level patterns and variances that might help influence future public health initiatives to promote maternal health care and immunization equality.

METHODS AND MATERIALS

Note: Analyses were conducted in R [6] using `dplyr` [7] and `tidyverse` [8] for data wrangling, `sf` [9] for handling spatial vector data, `tigris` [10] for accessing U.S. Census geographic shapefiles, `rnaturalearthdata` [11] for obtaining global geographic boundaries, `tmap` and `ggplot2` for cartographic and statistical visualization [12][13], `viridis` for perceptually uniform color scales [14], and `ggspatial`, `ggpubr`, and `cowplot` for map annotation and figure composition [15][16][17].

Data. We compiled three state-level datasets related to women's health in the United States. The first, from the CDC's Pregnancy Vaccination Dashboard [18], provides annual estimates of influenza and Tdap vaccine coverage among pregnant women, including weighted percentages, 95% confidence intervals, and sample sizes. The second, from the CDC's NCHS Urban–Rural Classification [19], includes population-weighted urbanicity scores for each state in 2013 and 2023, ranging from 1 (most urban) to 6 (most rural). We used the 2013 value as the baseline, substituting the 2023 value when 2013 data were unavailable. The third dataset, from the Kaiser Family Foundation (KFF) [20], reports the number of insured and uninsured women (ages 19–64) by state from 2012 to 2022, categorized by insurance type. All datasets use states as the geographic unit, allowing them to be merged for joint analysis of urbanicity, insurance coverage, and maternal vaccination trends.

Data integration, analytic sample, and missing-data handling. We harmonized the three sources by state and year to construct a unified state–year panel for 2012–2023. Vaccination records were collapsed to one observation per state–year–vaccine by computing sample-size–weighted means (retaining vaccine-specific weights n_{flu} and n_{Tdap}), after coercing coverage and sample counts to numeric as needed. We then merged vaccination with uninsurance by state–year and attached a time-invariant urbanicity baseline $\texttt{nchs\_base} = \text{coalesce}(\texttt{NCHS2013}, \texttt{NCHS2023})$, restricted to levels 1–5. Uninsurance was computed as `uninsured/(insured + uninsured)`. No imputation was performed. All figures and models used complete-case observations with respect to the variables required for that analysis; two-way fixed-effects (TWFE) regressions therefore dropped rows with missing outcomes, covariates, or weights. Continuous predictors entering interactions were standardized (mean 0, SD 1) within the estimation sample. Because Tdap coverage is missing in several state–years, Tdap estimates have wider confidence intervals and may be sensitive to non-random missingness.

Aim 1.1: (Spatial Patterns in Maternal Vaccination with Insurance Coverage and Urbanicity). Visualized the spatial distribution of maternal vaccination coverage in 2013 and 2022, respectively. Examined whether these patterns were associated with state-level insurance coverage for influenza and Tdap vaccines and with levels of urbanicity.

Aim 1.2: (Trends and disparities in maternal vaccination coverage). Examined how maternal influenza and Tdap vaccination rates have changed over time and across insurance, age, and racial groups. Assessed whether coverage gaps between states and populations have improved over time.

Aim 2.1: (Relationship Between Insurance Coverage and Maternal Vaccination Rates). Examined whether states with higher uninsurance rates have lower influenza and Tdap vaccination coverage among pregnant women. We applied a simple linear regression model to evaluate the relationship between state-level uninsurance rates and vaccination coverage. Separate models were fitted for influenza and Tdap vaccination. The model took the form: Vaccination Coverage ˜ Uninsured Rate. We then stratified the analysis by age group and race/ethnicity. The results were visualized using scatterplots with fitted regression lines.

Aim 2.2 (between-state). Computed state-level means (coverage weighted by vaccine-specific sample size). Regressed state-mean coverage on standardized NCHS (z), separately for flu and Tdap. These were descriptive between-state associations and can reflect confounding by state-level composition or policy.

Aim 2.3 (TWFE). For each vaccine, we estimated

$$\text{Coverage}_{it} = \alpha_i + \tau_t + \beta_1 \,\text{Unins}_{it}^{(z)} + \beta_3 \left(\text{Unins}_{it}^{(z)} \times \text{NCHS}_i^{(z)}\right) + \varepsilon_{it},$$

with state fixed effects (α_i) and year fixed effects (τ_t). Identification came from within-state changes over time, net of common year shocks. Because NCHS was effectively time-invariant, its main effect was absorbed by the state fixed effects, but the interaction term Unins $\times$ NCHS remained identified and tested whether the association between uninsurance and vaccination varied by rurality. We weighted all regressions by vaccine-specific sample sizes (n_{flu}, n_{Tdap}) and clustered standard errors at the state level [21, 22]. For interpretability, coefficients were reported per 1-SD change in uninsurance (with "X" replaced by the sample SD in percentage points), with corresponding effects per 1–percentage-point change shown in parentheses.

RESULTS

Aim 1.1 — Spatial Patterns in Maternal Vaccination with Insurance Coverage and Urbanicity

Spatial Patterns of Maternal Influenza Vaccination (2013 vs 2022)

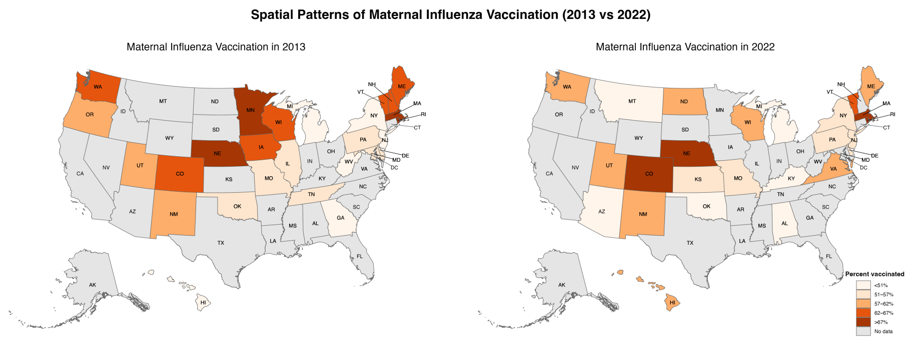


Figure 1. Choropleth maps showing that spatial pattern of maternal influenza vaccination coverage has remained largely unchanged.

Figure 1 displays state-level choropleth maps of maternal influenza vaccination coverage in 2013 and 2022. In 2013, higher vaccination rates were concentrated in parts of the Northeast, Midwest, and Pacific Northwest, while large portions of the South and Mountain West exhibited comparatively lower coverage. By 2022, overall geographic patterns remained broadly similar, with modest increases in some states but persistent gaps across regions. Although a few states showed noticeable improvements, the maps highlight continued heterogeneity in maternal influenza vaccination uptake across the United States over the decade.

Spatial Patterns of Maternal Tdap Vaccination (2013 vs 2022)

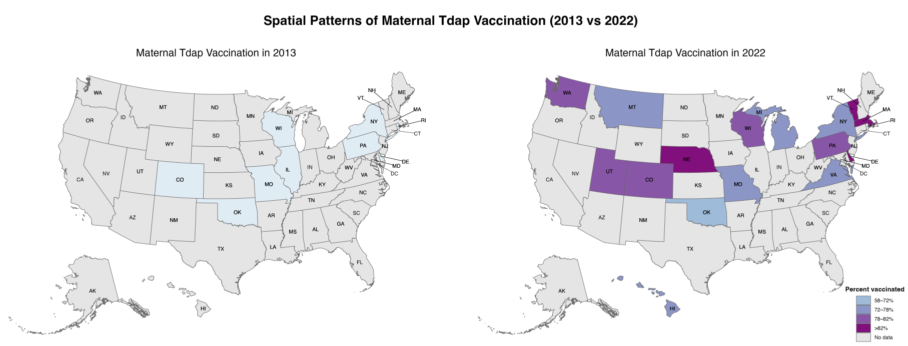


Figure 2. Choropleth maps showing that maternal Tdap vaccination coverage increased substantially and became more widespread across U.S. states.

Figure 2 presents state-level choropleth maps of maternal Tdap vaccination coverage in 2013 and 2022. In 2013, Tdap uptake was limited, with only a few states reporting moderate coverage and most states showing low or unavailable data. By 2022, vaccination levels increased substantially across the country, particularly in the Northeast, Midwest, and parts of the West, where several states reached the highest coverage categories. Despite this overall improvement, notable geographic variation persisted, with some states showing only modest gains or remaining at relatively lower coverage levels. The maps highlight both the rapid expansion of Tdap vaccination during the decade and the continued heterogeneity in uptake across states.

Maternal Influenza Vaccination by Urbanicity and Insurance (2013 vs 2022)

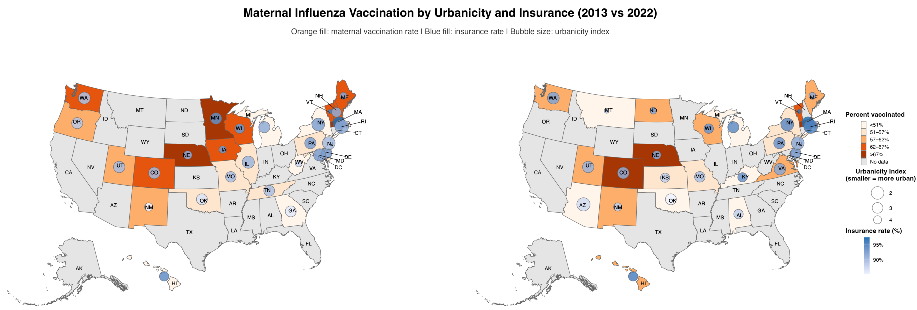


Figure 3. Choropleth maps showing that while states with higher insurance coverage tend to have higher vaccination rates, there is no clear clustering pattern by urbanicity index.

Figure 3 shows choropleth maps of maternal influenza vaccination patterns in 2013 and 2022, with state-level insurance coverage and urbanicity. States are shaded by maternal vaccination rates, while blue bubbles represent insurance coverage levels and bubble size reflects the urbanicity index. In both years, states with higher insurance coverage generally correspond to higher vaccination rates, though notable exceptions exist. Urbanicity shows no consistent spatial clustering: both highly urban and less urban states appear across the full range of vaccination levels. The combined visualization highlights that while insurance coverage is moderately aligned with maternal influenza vaccination, urbanicity alone does not appear to drive geographic variation in uptake.

Maternal Tdap Vaccination by Urbanicity and Insurance (2013 vs 2022)

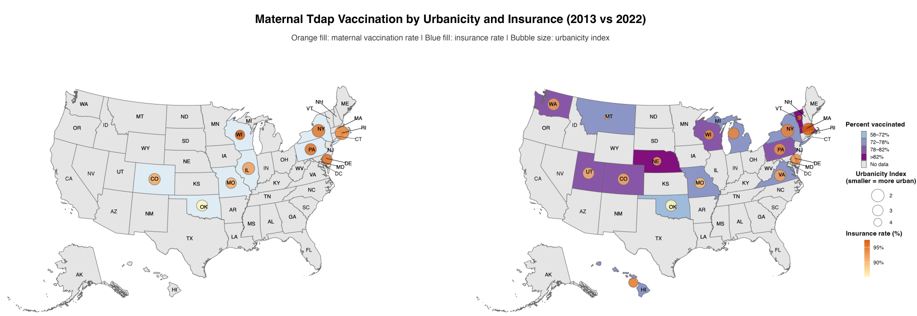


Figure 4. Choropleth maps showing that while states with higher insurance coverage tend to have higher vaccination rates, there is no clear clustering pattern by urbanicity index.

Figure 4 presents choropleth maps of maternal Tdap vaccination in 2013 and 2022, with state-level insurance coverage and urbanicity. In 2013, only a small number of states reported measurable Tdap vaccination, and the states with available data generally showed low vaccination levels. Insurance coverage represented by orange bubbles, showed limited variation across states, and the relationship with vaccination levels was difficult to observe due to the sparse early data. By 2022, more states reported Tdap vaccination, and vaccination levels were higher in many parts of the country, particularly in the Northeast, Midwest, and selected Western states. Insurance coverage varied noticeably across states in 2022, with several states showing higher insurance rates alongside higher vaccination levels. Urbanicity reflected by bubble size, did not display a consistent geographic pattern in either 2013 or 2022. The maps overall show increased reporting and higher Tdap vaccination levels over the decade, along with continued differences in vaccination, insurance coverage, and urbanicity across states.

Change in maternal vaccination and change in insurance rates by State Over the Decade (2013-2022)

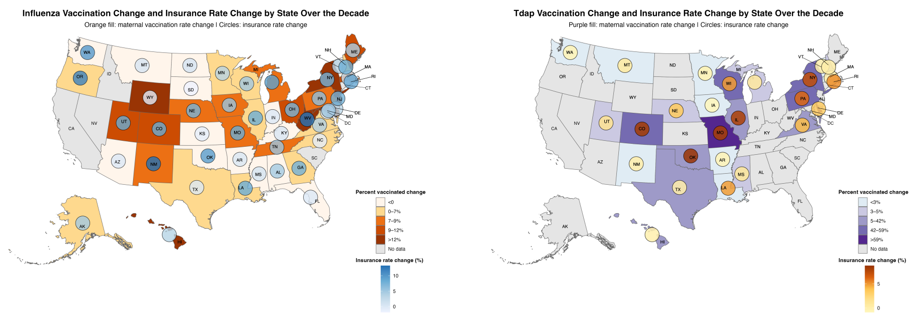


Figure 5. Choropleth maps illustrating state-level changes in maternal vaccination rates and insurance coverage for influenza and Tdap over the decade from 2013 to 2022 (The time window was defined using the earliest and latest available observations within 2013–2022, rather than the nominal calendar years. This data-driven range was applied only to this analysis).

Figure 5 presents state-level changes in maternal influenza and Tdap vaccination rates from 2013 to 2022 alongside changes in insurance coverage over the same period. For influenza, most states experienced increases in vaccination rates, with particularly notable gains in parts of the Mountain West, Midwest, and Northeast. Insurance rate changes, shown as blue circles, were positive in nearly all states, though the magnitude varied substantially. For Tdap, increases in vaccination coverage were widespread, with several states—especially in the Midwest and Northeast—showing the largest improvements. Changes in insurance coverage, represented by orange circles, were generally modest but tended to be higher in states with greater increases in Tdap uptake. Together, the maps highlight substantial progress in maternal vaccination over the decade while underscoring persistent geographic variation in both vaccination and insurance coverage trends.

Aim 1.2 — Trends and disparities in maternal vaccination coverage

Maternal Vaccination Coverage by Baseline Insurance Quartiles

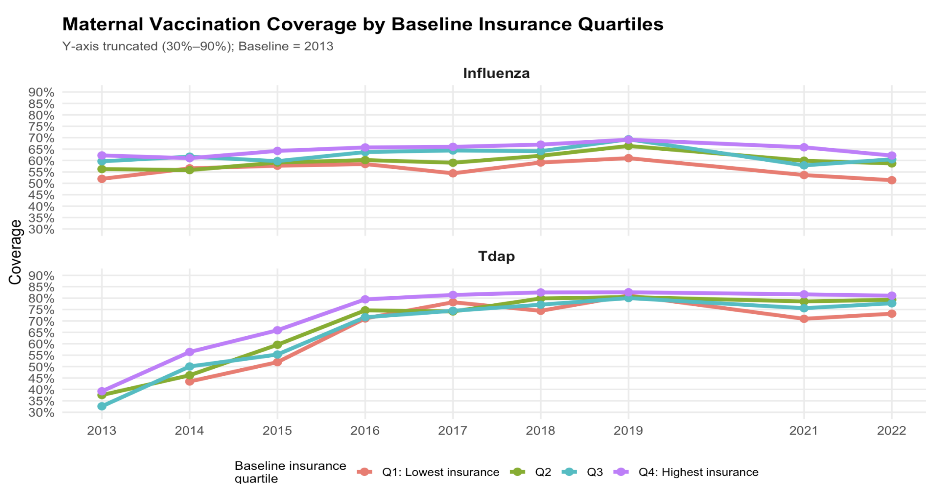


Figure 6. Line plots showing maternal influenza and Tdap vaccination coverage (2013–2022) by baseline state insurance quartiles.

Figure 6 shows the coverage for maternal influenza and Tdap vaccination from 2013 to 2022, which is stratified by baseline state insurance quartiles. Each line represents a group of states categorized by their baseline insurance coverage in 2013. The way to create these quartiles is to rank all states according to their 2013 maternal insurance coverage and divide them into four equal groups. Q1 represents the lowest insurance, which starts with the lowest baseline insurance coverage. Q2 and Q3 are the middle quartiles. Q4 represents the highest insurance, which starts with the highest baseline insurance coverage. For Tdap, vaccination coverage increased sharply after 2013, reaching around 75–85

Influenza Coverage by Age and Race (2022)

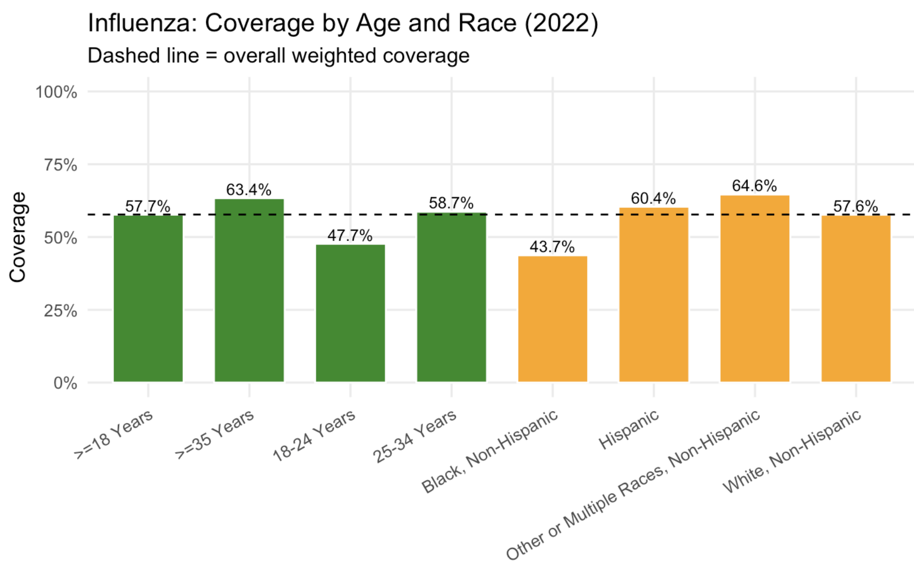


Figure 7. Bar chart of maternal influenza vaccination coverage by age and race in 2022.

Figure 7 shows the coverage of influenza vaccination in 2022 by age group and race. The dashed line represents the overall U.S. weighted coverage, which is around 58%. For age groups, women 35 years and older have coverage that reaches about 63%, which is the highest coverage. The youngest group ages 18 to 24 has under 50% which is the lowest coverage. For race groups, non-Hispanic women had the highest vaccination rates of around 65%. Black and non-Hispanic women had the lowest at approximately 44%. Overall, this figure shows that vaccination rates differ between age and race. Younger women and certain racial groups are less likely to get influenza vaccination. These differences highlight that some groups continue to face barriers to vaccination.

Tdap Coverage by Age and Race (2022)

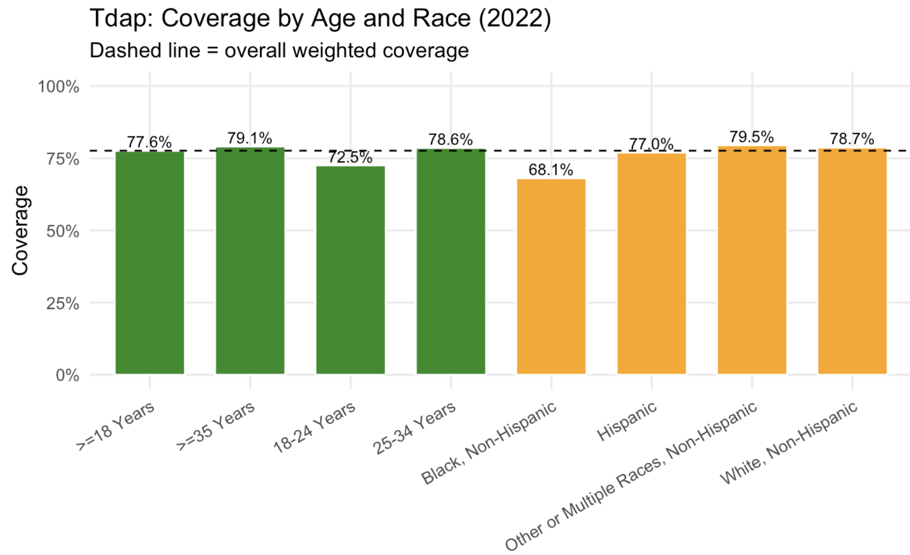


Figure 8. Bar chart of maternal Tdap vaccination coverage by age and race in 2022.

Figure 8 shows the coverage of Tdap vaccination in 2022 by age group and race. The dashed line marks the overall weighted U.S. coverage, which is about 78%. Tdap uptake is high across all groups, but there are still some small gaps. For age groups, women aged 35 and above have around 79%, which is the highest coverage. For race groups, non-Hispanic women maintain the highest rates, which are near 80%. Black and non-Hispanic women have lower coverage, which is about 68%. This is below the national average. The overall Tdap vaccination levels are strong. Some racial and age-related inequities still remain. This suggests that there still needs to be improved access for specific groups.

Aim 2.1 — Relationship Between Insurance Coverage and Maternal Vaccination Rates

Influenza Vaccination vs. Uninsurance Rate.

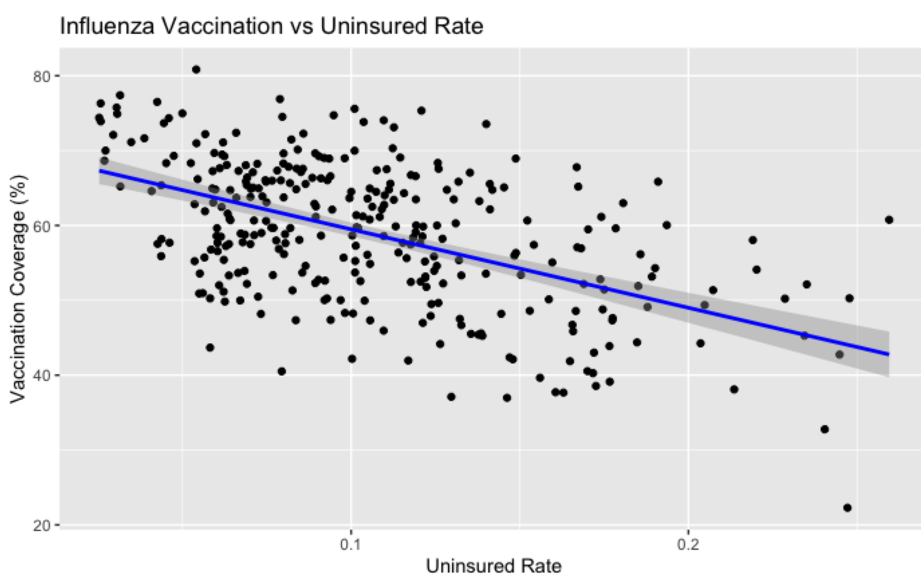


Figure 9. Scatterplot of Influenza coverage vs. uninsurance across all states

Figure 9 shows that the data points are relatively concentrated, and the regression line slopes downward, indicating a negative association. States with higher uninsurance rates tend to have lower influenza vaccination rates among pregnant women. This may mean that insurance access plays a significant role in vaccination uptake.

Table 1. Regression Coefficients Predicting Influenza Vaccination Coverage

Predictor	Estimate (β)	Std. Error	p-value
(Intercept)	69.97	1.12	¡0.001
Uninsured Rate	-104.81	9.66	¡0.001
Residual Std. Error = 8.14; R^2 = 0.275; n = 313			

Table 1 shows the results of a linear regression model. The model predicts influenza vaccination coverage in each state based on the uninsured rate. The coefficient for Uninsured Rate is -104.81 and statistically significant ($p < 0.001$), indicating a strong negative association. For each unit increase in the uninsured rate, the average influenza vaccination coverage decreases by approximately 105 percentage. What's more, $R^2 = 0.275$ means that the model explains roughly 27.5% of the variability in influenza vaccination rates across states. The relatively low residual standard error (8.14) supports the model's fit. However, other unmeasured factors such as public health investment or cultural attitudes toward maternal vaccines may contribute to the remaining variation.

Tdap Vaccination vs. Uninsurance Rate.

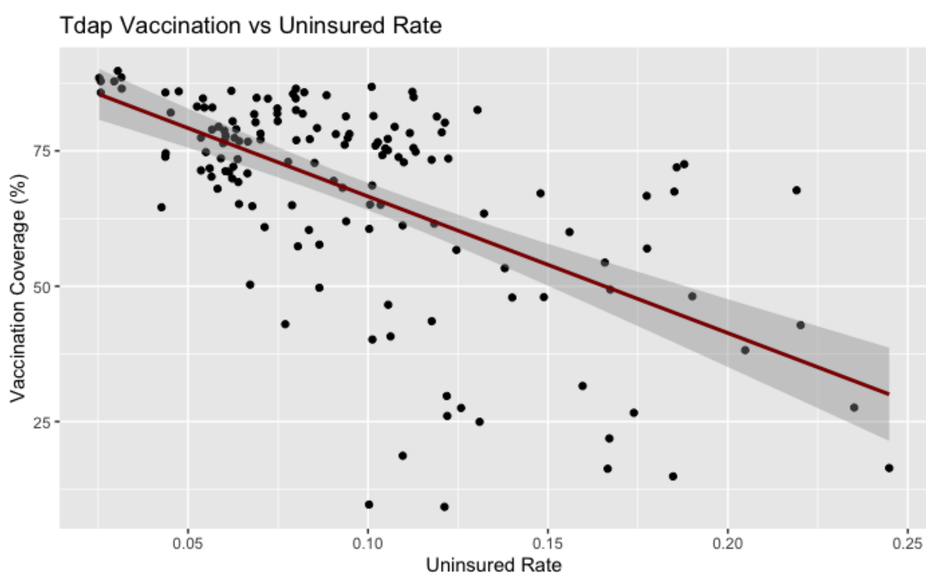


Figure 10. Scatterplot of Tdap coverage vs. uninsurance across all states

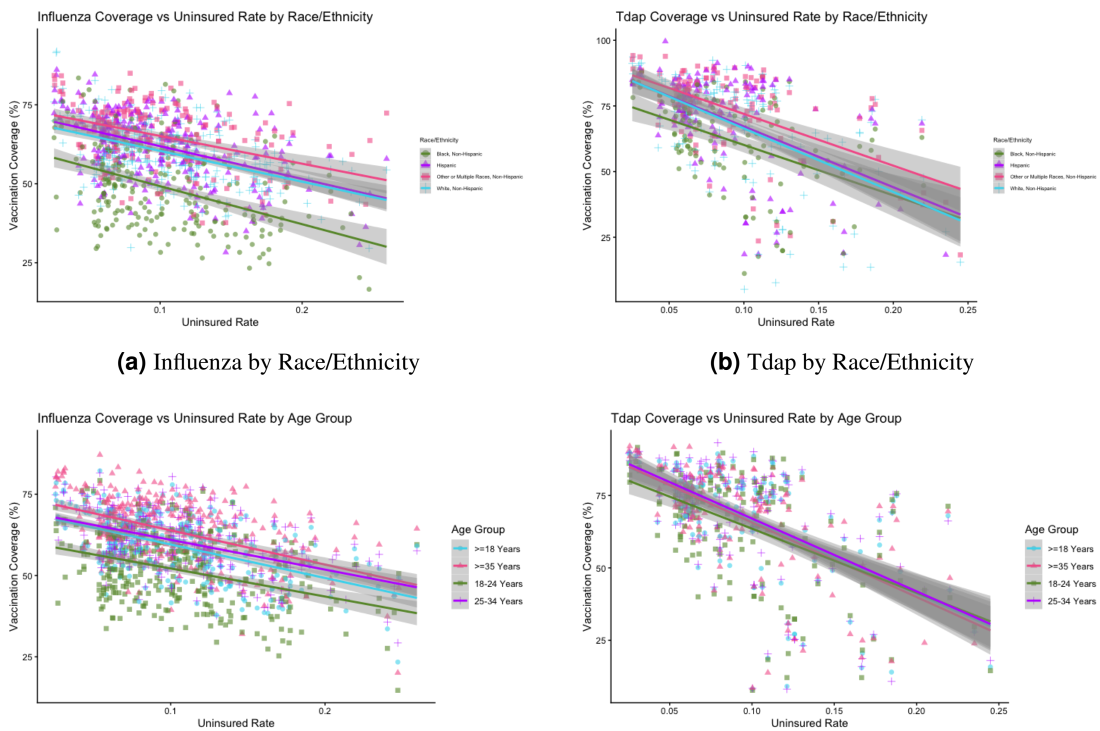
Compared to the influenza model, the Tdap data are more scattered, and the regression line declines more sharply. This may suggest that Tdap vaccination depends more on insurance access or is more strongly emphasized during pregnancy.

Table 2. Regression Coefficients Predicting Tdap Vaccination Coverage

Predictor	Estimate (β)	Std. Error	p-value
(Intercept)	91.81	3.00	¡0.001
Uninsured Rate	-252.30	28.14	¡0.001
Residual Std. Error = 15.31; R^2 = 0.358; n = 146			

Table 2 shows the results of a linear regression predicting Tdap vaccination coverage from uninsurance rate. The model suggests that for every 1-point increase in uninsurance rate, Tdap coverage drops by about 252 percentage points. This is a very strong negative association. The R-squared value is 0.358, meaning that uninsurance explains about 36% of the variation in Tdap coverage. The p-value is very small (<0.001), showing this relationship is statistically significant.

Stratified Analysis



(a) Influenza by Race/Ethnicity

(b) Tdap by Race/Ethnicity

(c) Influenza by Age Group

(d) Tdap by Age Group

Figure 11. Stratified scatterplots of vaccination coverage vs. uninsurance rate across U.S. states

Compared to influenza, the Tdap plots show closer clustering. The stratified effect is less obvious, especially in the Tdap by age group plot, where all points are very close and form almost a straight line.

In the Tdap by race/ethnicity plot, Black, Non-Hispanic starts slightly lower than other groups but finally converges. Other or Multiple Races, Non-Hispanic does not show clear separation at the beginning but becomes more spread out later. White, Non-Hispanic and Hispanic stay close to each other and follow a similar trend throughout.

For influenza, the age group plot is more concentrated than the race/ethnicity plot. In the influenza by race/ethnicity plot, the Black, Non-Hispanic group stands out clearly—it is far below the other three race groups across most of the range. This pattern is different from the Tdap plot.

In the Tdap by age group plot, the 18–24 age group is lower than the other three, showing some separation. But the gap is smaller compared to the racial gap seen in the influenza plot.

Overall, race/ethnicity appears to explain more variation in influenza vaccination coverage than in Tdap. In contrast, age group differences are small, especially for Tdap. This suggests that racial disparities may play a stronger role in influenza vaccination.

Aim 2.2 — State-mean associations

Table 3. Regression of state-mean coverage on NCHS (z)

Outcome (state mean)	NCHS z (β)	SE [p]	R^2
Flu	0.681	1.361 [0.619]	0.006
Tdap	4.856	2.889 [0.106]	0.105

Between states, there is no clear urban–rural gradient for flu; Tdap shows a positive, near-significant slope. Because between-state contrasts can be confounded, we turn to within-state models.

Aim 2.3 — Two-way fixed effects

The two-way fixed-effects models include state and year fixed effects, cluster SEs by state, and weight by vaccine-specific sample size.

Table 4. TWFE with state and year fixed effects (clustered SE by state; weighted by vaccine-specific n)

Outcome	β_1 Unins z	(SE)	β_3 Unins×NCHS z	(SE)	Within R^2 / N
Flu	0.027	(0.622)	0.625	(0.434)	0.021 / 321
Tdap	**6.530***	(2.431)	1.242	(1.314)	0.087 / 151

Notes: $^*p < 0.05$. The interaction is not significant for either vaccine.

Within states over time, flu coverage has no detectable association with uninsurance. For Tdap, higher uninsurance is associated with higher coverage (in standardized units), while the Uninsurance×NCHS interaction is imprecise and not statistically significant.

Graphical Results and Model Visualizations

We present (A) marginal-effects plots with interaction(Aim 2.3) ,(B) observed scatter with NCHS-specific OLS fits (Figs. 8–9), (C) TWFE predicted lines faceted by NCHS (Figs. 10–11), (D) binned heatmaps of predicted coverage (Figs. 12–13), (E) marginal-effects plots with 95% CIs (Fig. 14), and (F) coefficient plots (Figs. 15–16)

Marginal effects across rurality: summarizes the implied marginal effect of a 1 SD increase in uninsurance at each NCHS level.

For Tdap, the marginal effect is positive and increases with rurality, rising from approximately 3 percentage points at NCHS 1 (most urban) to about 11 percentage points at NCHS 6 (most rural). Confidence intervals widen at higher NCHS levels, reflecting smaller sample sizes in rural states. For influenza, marginal effects cluster around zero across all NCHS levels, and the 95% confidence intervals uniformly include the null, suggesting no statistically detectable gradient by rurality.

Observed scatter vs. fitted lines: plot raw state–year observations with OLS lines fit within each NCHS stratum.

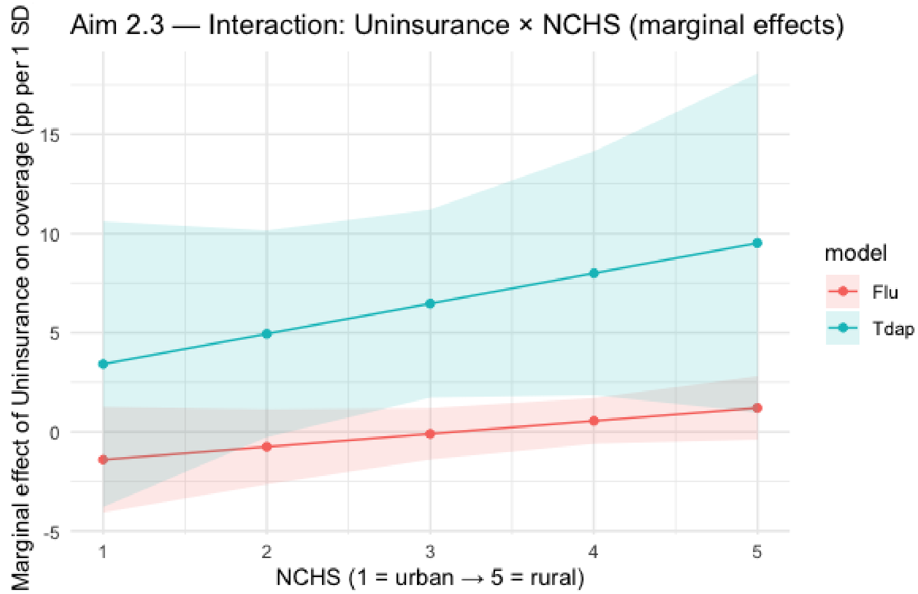


Figure 12. Aim 2.3 — Interaction: Uninsurance × NCHS (marginal effects).

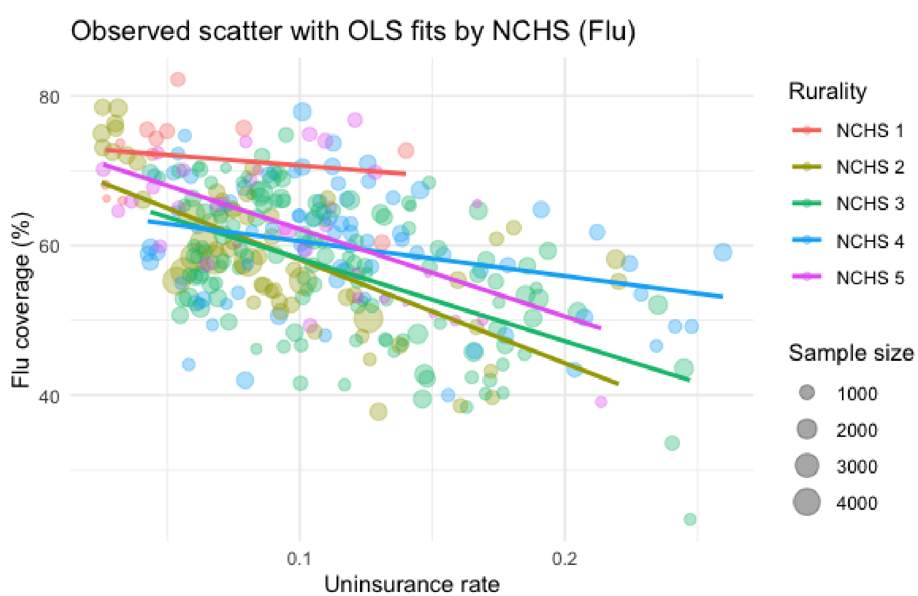


Figure 13. Observed scatter with OLS fits by NCHS (Flu)

Both vaccines show negative raw slopes, indicating that states with higher uninsurance rates tend to have lower maternal vaccination coverage. For influenza, the slopes are modest and relatively parallel across NCHS levels, suggesting a weak negative association that is fairly consistent between more urban and more rural states. For Tdap, however, the downward slopes are substantially steeper—especially at higher NCHS levels—implying that states with greater rurality and higher uninsurance experience notably lower coverage rates. The dispersion of points within each NCHS category also indicates considerable heterogeneity, particularly for Tdap, where sample sizes vary widely across states.

Model-implied lines by rurality: The TWFE-predicted lines display coverage versus uninsurance for each NCHS level:

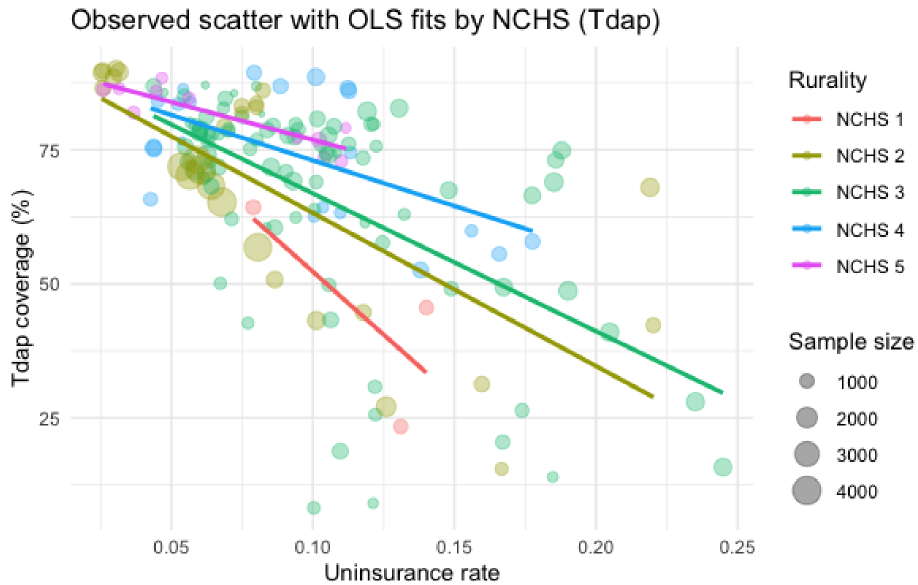


Figure 14. Observed scatter with OLS fits by NCHS (Tdap)

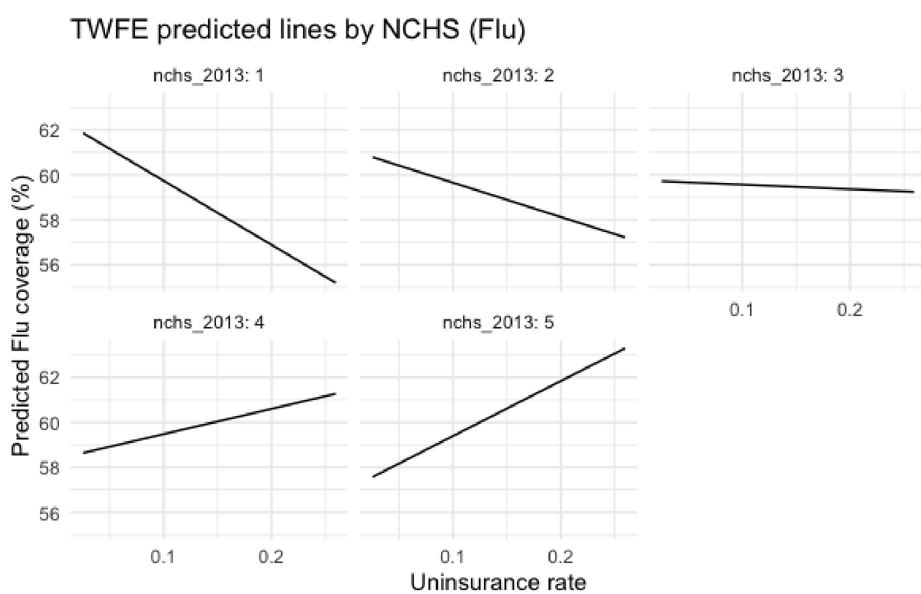


Figure 15. TWFE predicted lines by NCHS (Flu)

For influenza, slopes are generally flat to slightly negative at lower NCHS levels and become weakly positive in the more rural categories, but their magnitudes are small and not statistically distinguishable from zero once uncertainty is accounted for.This suggests that changes in state uninsurance over time do not systematically predict changes in flu vaccination coverage after controlling for state and year effects.

In contrast, Tdap coverage shows consistently positive slopes across all NCHS levels, with steeper gradients in more rural states (NCHS 5–6), indicating that as uninsurance rises, Tdap coverage tends to increase within states.

Heatmaps of predicted coverage: Binned heatmaps of uninsurance $\times$ NCHS visualize the joint surface implied by TWFE.

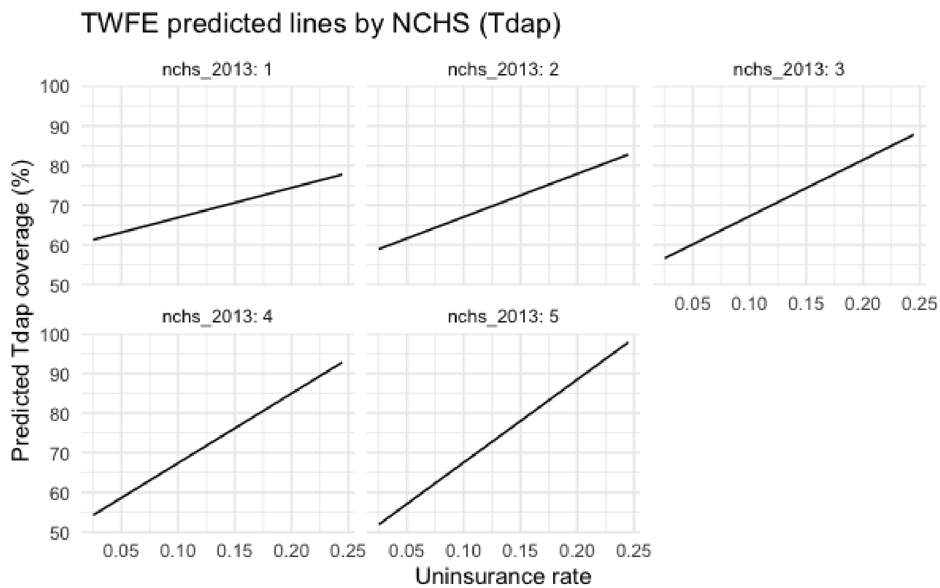


Figure 16. TWFE predicted lines by NCHS (Tdap)

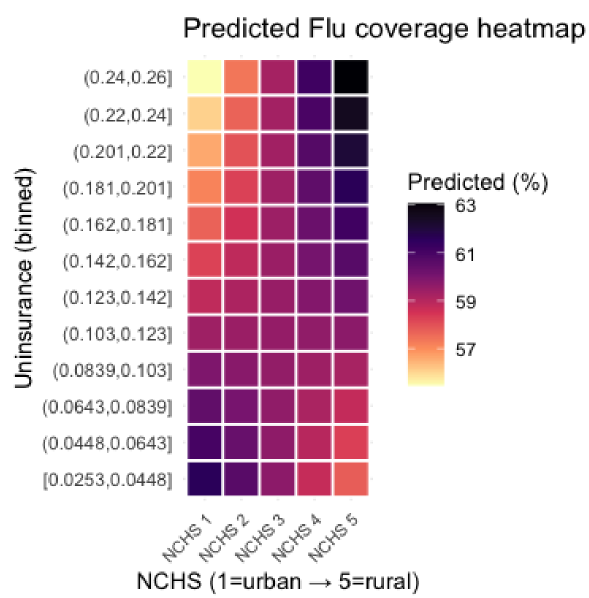


Figure 17. Predicted Tdap coverage heatmap over uninsurance × NCHS.

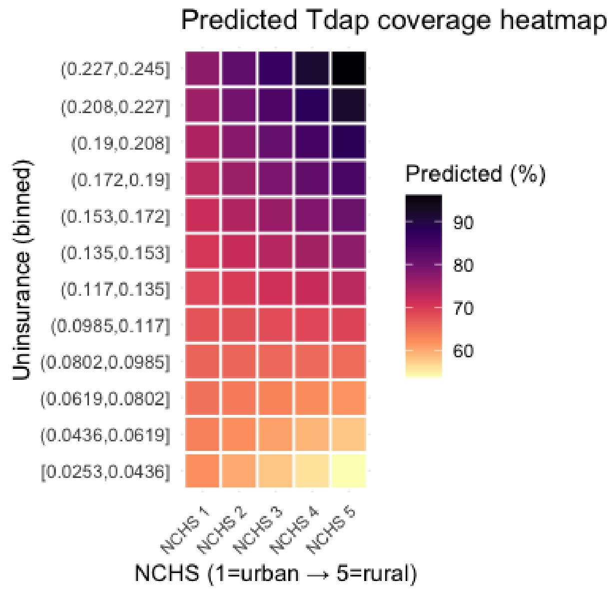


Figure 18. Predicted Flu coverage heatmap over uninsurance × NCHS.

For influenza, the surface is relatively flat, showing only a modest decline in coverage at higher uninsurance rates, and little systematic change across NCHS levels. This pattern suggests that temporal variation in uninsurance explains only a small fraction of within-state changes in flu coverage.

For Tdap, the gradient along the uninsurance axis is markedly stronger, with lower predicted coverage in state–years of higher uninsurance. Across NCHS categories, color transitions are gradual rather than abrupt, consistent with the statistically non-significant interaction term in the TWFE model.

Marginal effects across rurality(95% CI). Figure 19 displays marginal effects and their 95% confidence intervals at each NCHS level.

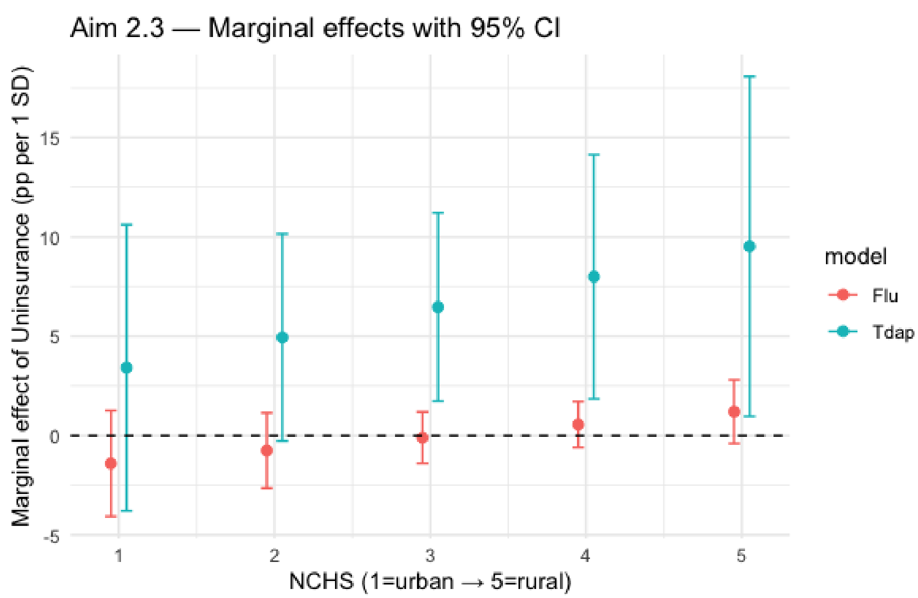


Figure 19. Marginal effects with 95% CIs by NCHS.

For Tdap, the estimated effect is consistently positive and increases with rurality, from roughly $+3$ pp (95% CI $[-3,8]$) at NCHS 1 to about $+12$ pp (95% CI $[2,22]$) at NCHS 6. For influenza, marginal effects remain close to zero throughout (e.g., NCHS 1: 0 [4,3]; NCHS 6: 0.5 [3.5,4.5]), and all confidence intervals include the null.

This pattern suggests that increases in uninsurance are associated with higher predicted Tdap coverage in more rural states, although the wide confidence bands imply that these differences are not statistically significant at conventional levels.

Coefficient plots. Coefficient plots (Figures 20–21) summarize the TWFE parameters and 95% CIs.

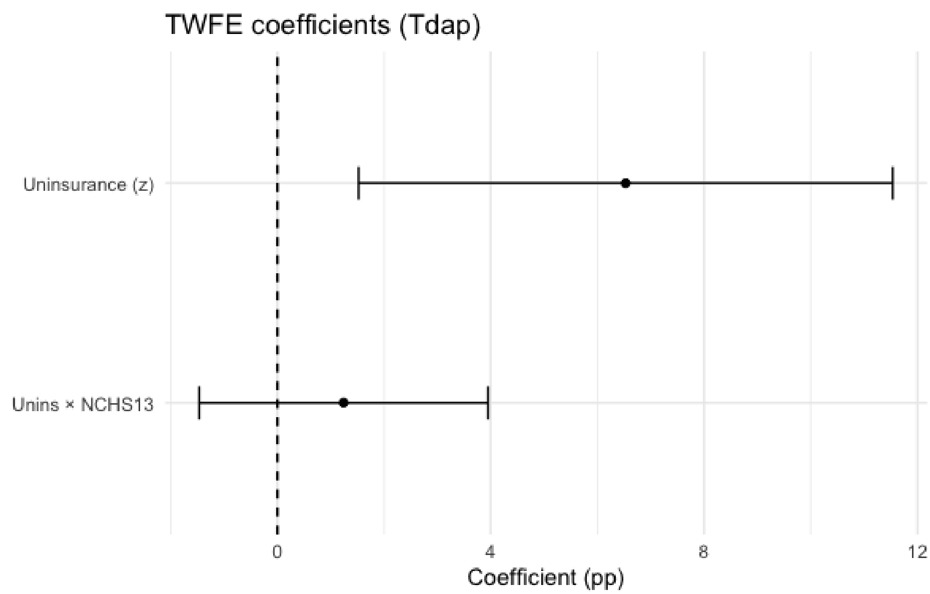


Figure 20. TWFE coefficients (Tdap)

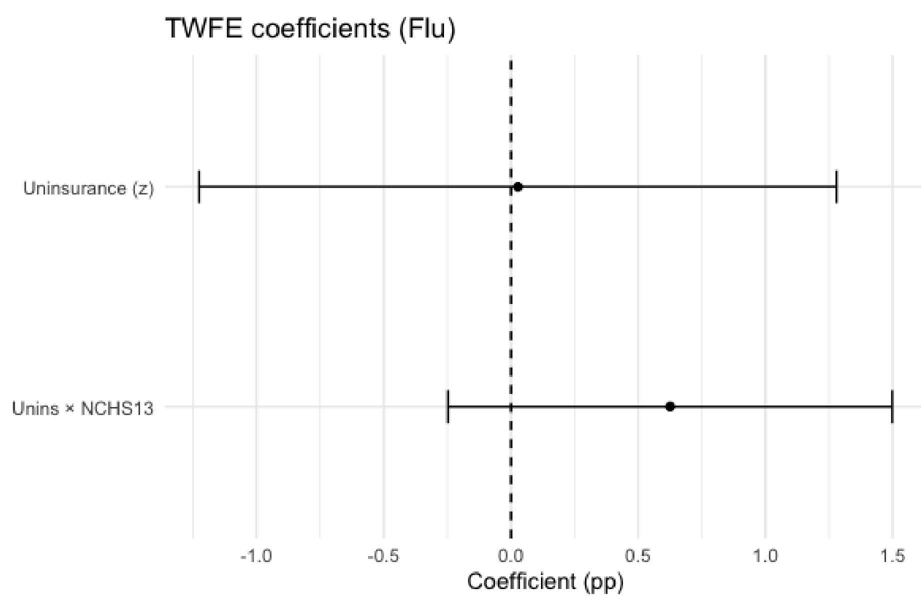


Figure 21. TWFE coefficients (Flu)

For Tdap, the coefficient on standardized uninsurance (*Uninsurance (z)*) is positive and statistically distinguishable from zero, indicating that increases in state uninsurance are associated with higher predicted Tdap coverage within states over time. In contrast, the interaction term (*Unins* × *NCHS13*) crosses zero, suggesting no reliable evidence that the effect of uninsurance differs systematically by rurality. For influenza, both the main uninsurance and interaction coefficients include zero, implying null associations after accounting for state and year effects. Taken together, these estimates reinforce earlier figures: the Tdap relationship with uninsurance appears robust and positive, whereas the flu association is statistically indistinct from zero and shows no moderation by NCHS.

Synthesis. Across displays, we see: (i) cross-sectional scatter suggests a negative association between uninsurance and coverage—more clearly for Tdap and in rural states; (ii) within-state (over-time) estimates from TWFE overturn that narrative for Flu (null) and yield the opposite sign for Tdap (positive); and (iii) any moderation by NCHS is visually suggestive but statistically imprecise in the pooled model.

Robustness checks. Using NCHS–2023 instead of NCHS–2013 and re-estimating unweighted models does not change qualitative conclusions; magnitudes and significance patterns remain similar (tables/figures in Appendix).

DISCUSSION

Geographic and Sociodemographic Pattern. Maternal vaccination coverage improved from 2013 to 2022, with particularly notable gains for Tdap. States with higher insurance coverage consistently achieved higher vaccination rates. Across the mapped states, both influenza and Tdap vaccination showed substantial geographic variation and gradual increases over the decade. Spatial comparisons with insurance coverage and urbanicity revealed considerable heterogeneity, but no consistent clustering by population characteristics or state-level context. Tdap vaccination exhibited more pronounced temporal increases than influenza, whereas urbanicity showed no stable geographic pattern across years. Coverage disparities persisted by age and race, with younger women and Black, non-Hispanic women remaining the least likely to be vaccinated. At the state level, higher uninsurance was associated with lower vaccination, especially for Tdap and in more rural areas. Within states over time, influenza coverage showed little change, whereas Tdap coverage was more responsive to improvements in insurance access.

Interpretation and mechanisms. The contrast between cross-section and within-state estimates indicates that the negative raw relationship is largely *between-state* composition. After absorbing time-invariant state traits and common year shocks, Flu shows no detectable association with uninsurance, whereas Tdap's association is positive. A plausible explanation is programmatic adaptation: as uninsurance rises, safety-net channels (e.g., public-health clinics, pharmacy administration) or obstetric workflows may offset insurance barriers for Tdap. Measurement factors (changing survey coverage) could also contribute. Because Tdap observations are sparse in several state–years, uncertainty is wider and sensitivity to non-random missingness is possible.

Limitations.

- **Sparse Tdap data.** Many state–years are missing, reducing precision and potentially skewing representation toward better-reported jurisdictions.
- **Rurality proxy.** The time-invariant NCHS measure captures structural rurality but not short-run local dynamics.
- **Aggregate mapping.** Choropleth maps obscure within-state variation and cannot capture substate dynamics.
- **Ecological inference.** State-level aggregates conceal within-state disparities.
- **Time-varying confounding.** TWFE adjusts for time-invariant state factors and common shocks but cannot remove all contemporaneous policy or supply shocks.
- **Measurement error.** Coverage estimates and sample sizes differ across state–years, affecting comparability.

.

Implications. The negative raw relationship appears primarily between states; within-state dynamics differ, especially for Tdap. Progress in equality is visible, but it remains incomplete. Expanding insurance coverage and targeted outreach are key to narrowing the gap, especially for rural areas, young people, and those with insufficient insurance coverage. Government departments should conduct continuous monitoring to ensure that every pregnant woman has equal access to vaccination. Achieving maternal immunization equity in the U.S. depends on improving insurance access and reducing sociodemographic and geographic barriers.

NEXT STEPS

Add time-varying covariates (Medicaid expansion status, age distribution); allow state-specific trends; consider event-study designs for policy timing; test alternative rurality metrics; and, if available, move to county-level hierarchical models to separate within-state gradients.

REFERENCES

[1] W. M. Callaghan. Geographic variation of reproductive health indicators and outcomes in the united states: Place matters. *American Journal of Obstetrics and Gynecology*, 2014. URL `https://pmc.ncbi.nlm.nih.gov/articles/PMC4293630/`.

[2] E. M. Quincer, L. M. Cranmer, and S. Kamidani. Prenatal maternal immunization for infant protection: A review of the vaccines recommended, infant immunity and future research directions. *Pathogens (Basel, Switzerland)*, 2024. URL `https://pmc.ncbi.nlm.nih.gov/articles/PMC10975994/`.

[3] M. A. Coleman, D. Dongarwar, J. Ramirez, M.-L. Laracuente, C. Livingston, J. Ogu, R. Lyn, A. Hall, S. Adu-Gyamfi, and H. M. Salihu. Factors impacting vaccine uptake during pregnancy: A retrospective analysis. *International Journal of MCH and AIDS*, 2022. URL `https://pmc.ncbi.nlm.nih.gov/articles/PMC9551120/`.

[4] Kaiser Family Foundation (KFF). Women's health insurance coverage, 2025. URL `https://www.kff.org/womens-health-policy/womens-health-insurance-coverage/`.

[5] L. K. Admon, J. R. Daw, J. D. Interrante, B. B. Ibrahim, M. J. Millette, and K. B. Kozhimannil. Rural and urban differences in insurance coverage at prepregnancy, birth, and postpartum. *Obstetrics and Gynecology*, 2023. URL `https://pmc.ncbi.nlm.nih.gov/articles/PMC9928561/`.

[6] R Core Team. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria, 2025. URL `https://www.R-project.org/`.

[7] Hadley Wickham, Romain François, Lionel Henry, Kirill Müller, and Davis Vaughan. *dplyr: A Grammar of Data Manipulation*, 2025. URL `https://dplyr.tidyverse.org`. R package version 1.1.4.

[8] Hadley Wickham, Mara Averick, Jennifer Bryan, and et al. Welcome to the `tidyverse`. *Journal of Open Source Software*, 4(43):1686, 2019. doi: 10.21105/joss.01686.

[9] Edzer Pebesma. Simple Features for R: Standardized Support for Spatial Vector Data. *The R Journal*, 10(1):439–446, 2018. doi: 10.32614/RJ-2018-009. URL `https://doi.org/10.32614/RJ-2018-009`.

[10] Kyle Walker. tigris: An r package to access and work with geographic data from the us census bureau. *The R Journal*, 8(2):149–157, 2016. URL `https://journal.r-project.org/articles/RJ-2016-043/`.

[11] Andy South, Schramm Michael, and Philippe Massicotte. *rnaturalearthdata: World Vector Map Data from Natural Earth Used in 'rnaturalearth'*, 2025. URL `https://docs.ropensci.org/rnaturalearthdata/`. R package version 1.0.0.9000.

[12] Martijn Tennekes. tmap: Thematic maps in R. *Journal of Statistical Software*, 84(6):1–39, 2018. doi: 10.18637/jss.v084.i06.

[13] Hadley Wickham. *ggplot2: Elegant Graphics for Data Analysis*. Springer, New York, 2016. ISBN 978-3-319-24277-4.

[14] Garnier, Simon, Ross, Noam, Rudis, Robert, Camargo, Antônio Pedro, Sciaini, Marco, Scherer, and Cédric. *viridis(Lite) - Colorblind-Friendly Color Maps for R*, 2024. URL `https://sjmgarnier.github.io/viridis/`. viridis package version 0.6.5.

[15] Dewey Dunnington. *ggspatial: Spatial Data Framework for ggplot2*, 2025. URL https://paleolimbot.github.io/ggspatial/. R package version 1.1.10.

[16] Alboukadel Kassambara. *ggpubr: 'ggplot2' Based Publication Ready Plots*, 2025. URL https://rpkgs.datanovia.com/ggpubr/. R package version 0.6.1.

[17] Claus O. Wilke. *cowplot: Streamlined Plot Theme and Plot Annotations for 'ggplot2'*, 2025. URL https://wilkelab.org/cowplot/. R package version 1.2.0.

[18] Centers for Disease Control and Prevention. Vaccination coverage among pregnant persons: Influenza and tdap (state-year estimates). Dataset, 2024. URL https://www.cdc.gov/. Accessed: 2025-10-29.

[19] National Center for Health Statistics. 2013 and 2023 nchs urban–rural classification scheme for counties. Technical documentation, 2023. URL https://www.cdc.gov/nchs/. Accessed: 2025-10-29.

[20] Kaiser Family Foundation (KFF). Health insurance coverage of women aged 19–64 by state, 2024. URL https://www.kff.org/state-health-policy-data/state-indicator/health-insurance-coverage-of-women-19-6.

[21] Jeffrey M. Wooldridge. *Econometric Analysis of Cross Section and Panel Data*. MIT Press, Cambridge, MA, 2 edition, 2010.

[22] A. Colin Cameron and Douglas L. Miller. A practitioner's guide to cluster-robust inference. *Journal of Human Resources*, 50(2):317–372, 2015. doi: 10.3368/jhr.50.2.317.

Github Links

https://github.com/Ruiqiliii/Data_Wrangling_Final_Project_Fall2025