COVID-19 Policy, Vaccination, and Outcome Dynamics in U.S. States

Sivani Dronamraju

Introduction

This analysis examines how COVID-19 policy strictness and vaccination uptake evolved across U.S. states and how those trajectories aligned with observed case and mortality patterns over time. The objective is not to attribute causality, but to surface meaningful state-level differences and temporal alignments under heterogeneous policy regimes and reporting constraints.

The focus is deliberately narrow: large U.S. states with high cumulative case counts, where policy variation, population scale, and healthcare capacity create materially different outcome paths. The analysis prioritizes interpretability and judgment over model complexity.

What this analysis does not attempt:

- It does not estimate causal policy effects.
- It does not evaluate policy "effectiveness" in isolation.
- It does not produce an operational forecasting system.

What it focuses on instead:

- How policy stringency and vaccination uptake co-evolved across states
- How outcome trajectories diverged and failed to reconverge
- Where intuitive narratives (e.g., stricter policy to fewer cases) break down

Data Description
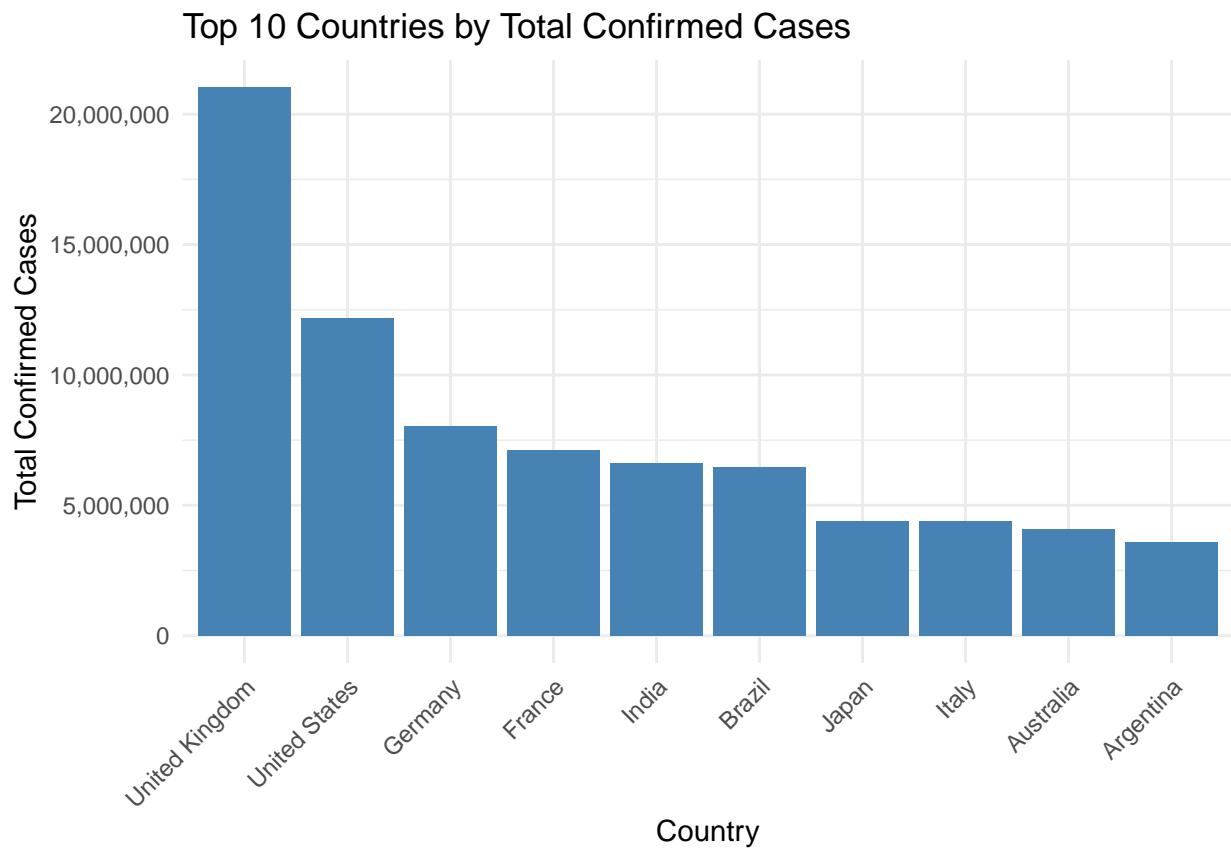
The dataset, extracted with the COVID19 R package, includes state-level data (level = 2) from the COVID-19 Data Hub. A brief summary of the dataset's covariates can be found in the COVID19 Data Hub documentation.
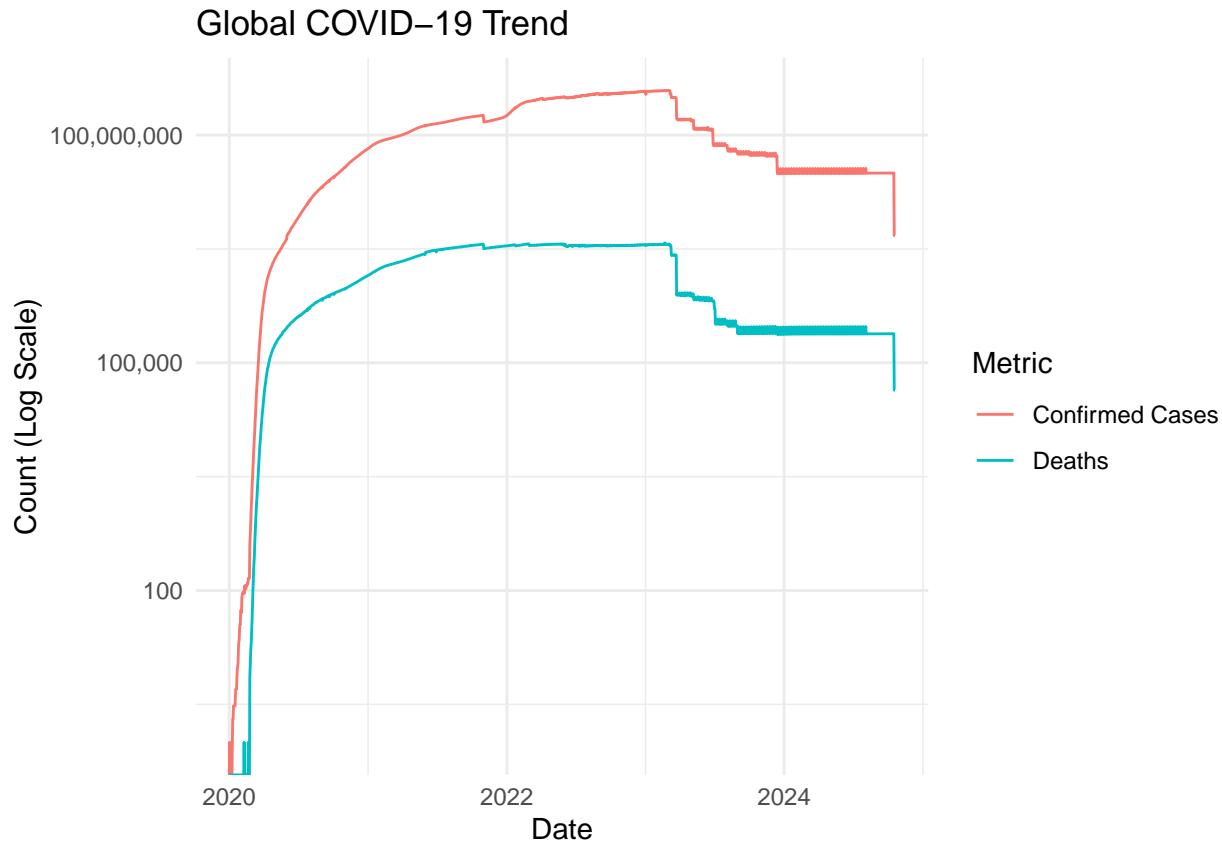
The key features include:

- Epidemiological outcomes: confirmed cases, deaths, hospitalizations, ICU occupancy
- Vaccination metrics: people vaccinated, people fully vaccinated
- Policy indices: stringency, containment & health, economic support
- Demographics: population, geographic identifiers

State-level granularity was chosen to balance policy heterogeneity with data availability, while avoiding over-interpretation of localized reporting noise.

To get an overview of the global situation, I identified the top 10 countries by total confirmed cases

Top 10 Countries by Total Confirmed Cases

Global COVID−19 Trend

Scope Selection: U.S. State-Level Focus

This analysis intentionally bypasses global comparisons and focuses directly on U.S. state-level dynamics. While global trends are well-documented, they add limited value for understanding how heterogeneous policy decisions, vaccination uptake, and healthcare capacity interacted within a single federal system.

Restricting the scope to U.S. states allows for clearer comparisons under shared reporting standards, vaccine availability timelines, and broadly similar institutional constraints.

Data Preprocessing

Preprocessing focused on isolating analysis-relevant signals while minimizing noise introduced by reporting gaps, jurisdictional inconsistencies, and redundant variables. The steps below reflect deliberate scoping decisions rather than exhaustive data cleaning.

```
##                              id                        date
##                        0.000000                    0.000000
##                       confirmed                      deaths
##                       28.575711                   28.575711
##                       recovered                       tests
##                       86.159541                    6.932337
##                        vaccines           people_vaccinated
##                       43.305198                   43.305198
##           people_fully_vaccinated                        hosp
```

```
##                                  43.305198                                14.278631
##                                        icu                                      vent
##                                  14.860938                                89.476961
##                             school_closing                         workplace_closing
##                                  38.104792                                38.104792
##                               cancel_events                    gatherings_restrictions
##                                  38.104792                                38.104792
##                            transport_closing                     stay_home_restrictions
##                                  38.104792                                38.104792
##              internal_movement_restrictions  international_movement_restrictions
##                                  38.104792                                38.104792
##                        information_campaigns                            testing_policy
##                                  38.104792                                38.104792
##                             contact_tracing                           facial_coverings
##                                  38.104792                                38.104792
##                           vaccination_policy                  elderly_people_protection
##                                  38.104792                                38.104792
##                     government_response_index                          stringency_index
##                                  38.104792                                38.104792
##                     containment_health_index                    economic_support_index
##                                  38.104792                                38.104792
##                      administrative_area_level               administrative_area_level_1
##                                   0.000000                                 0.000000
##                    administrative_area_level_2               administrative_area_level_3
##                                   0.000000                               100.000000
##                                    latitude                                 longitude
##                                   0.000000                                 0.000000
##                                  population                               iso_alpha_3
##                                   0.000000                                 0.000000
##                                 iso_alpha_2                               iso_numeric
##                                   0.000000                                 0.000000
##                                iso_currency                                 key_local
##                                   0.000000                                 0.000000
##                         key_google_mobility                          key_apple_mobility
##                                   6.601402                                 0.000000
##                                key_jhu_csse                                  key_nuts
##                                   0.000000                               100.000000
##                                    key_gadm
##                                   0.000000
```

Variables with structurally high missingness (e.g., early hospitalization and vaccination metrics) were retained only where temporal coverage was sufficient to support comparative analysis.

- To focus solely on U.S. states, I removed territories and the District of Columbia

This preprocessing yields a state-level panel aligned to the core analytical questions like how policy strictness, vaccination uptake, and healthcare strain evolved together over time.

Handling Missing Values

COVID-19 data is incomplete by nature, reporting practices changed over time, metrics were introduced gradually, and not all states reported the same fields consistently. Rather than attempting to reconstruct

missing values precisely, the goal here was to maintain continuity where trends matter, while avoiding artificial precision.

For cumulative counts (e.g., cases, deaths, vaccinations), missing values were filled forward and backward within each state. This preserves monotonic trends without implying new events occurred on missing days.

Policy indices were imputed using median values to avoid breaking time series while preserving relative differences between states. These values are treated as stabilizers, not as exact estimates.

Hospitalization and ICU data were handled conservatively. When a state reported some values, its own median was used; when a state never reported the metric, an overall median was used to avoid structural gaps in downstream analysis.
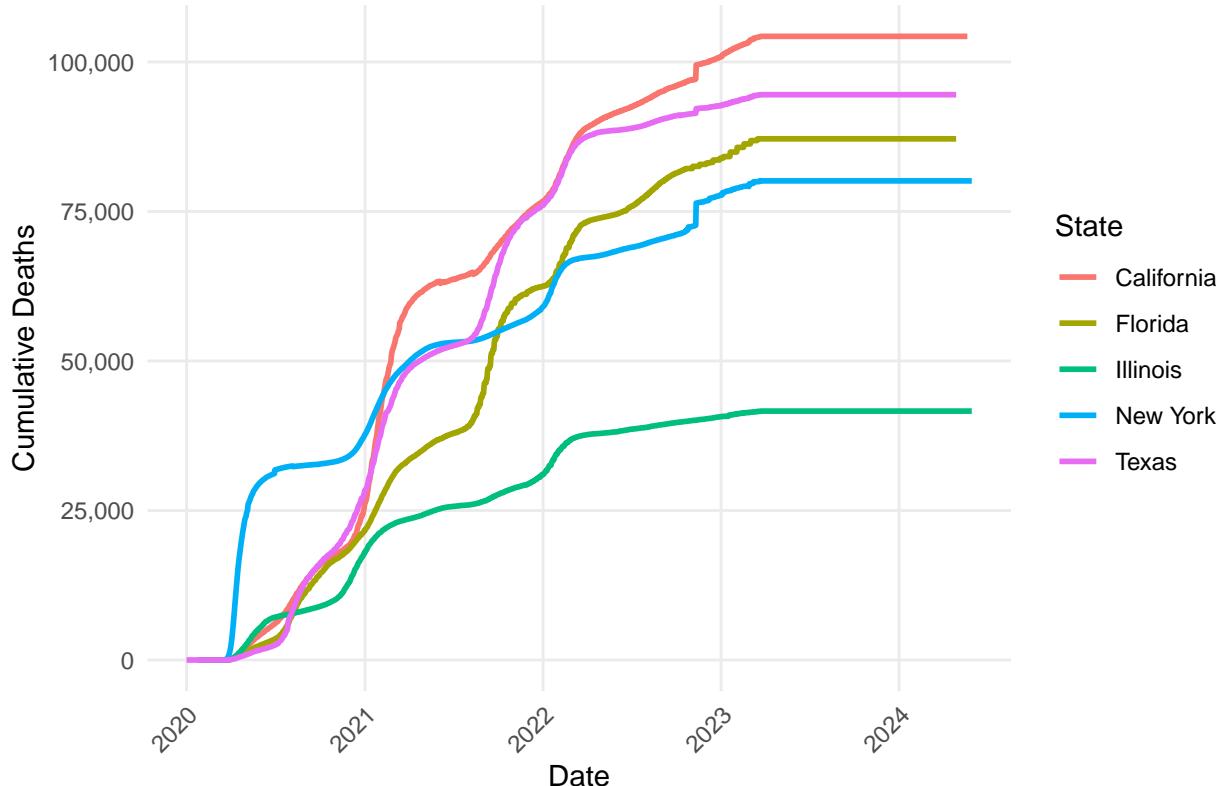
Policy variables were treated as ordinal measures rather than continuous quantities. To make comparisons more interpretable, each policy was decomposed into a simple notion of scope (general vs targeted) and intensity, without assuming linear effects.

These choices prioritize interpretability and continuity over completeness. Missing-value handling here supports comparative reasoning, not precise estimation, and downstream results are interpreted with this limitation in mind.

Outcome Trajectories in High-Impact States

Rather than aggregating outcomes into broad regions, the analysis focuses on the five U.S. states with the highest cumulative case counts. These states capture a wide range of population sizes, outbreak timing, and policy approaches, while keeping comparisons interpretable.
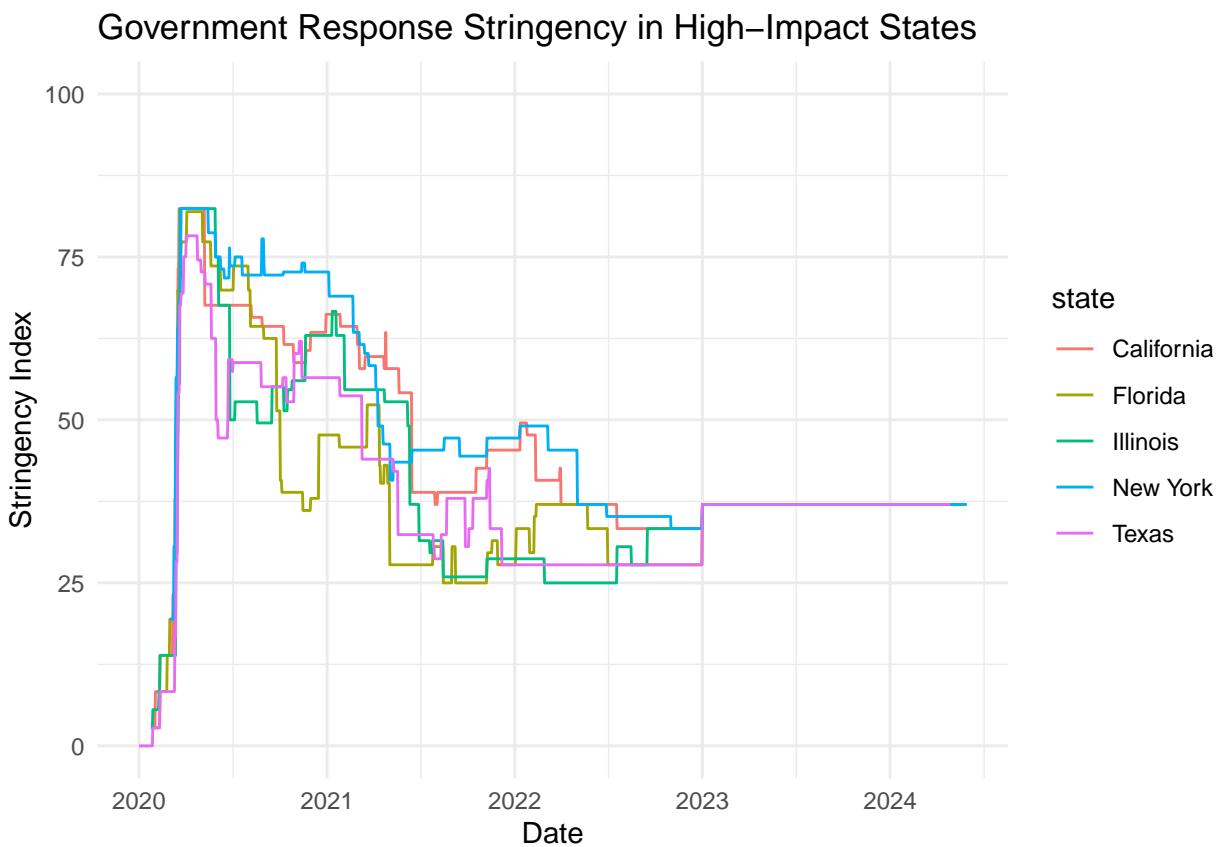

New York shows a sharp early increase in deaths, consistent with its role as an initial epicenter. In contrast, states such as Texas and Florida exhibit more gradual but sustained increases over a longer period. Cal-

ifornia's trajectory reflects both early exposure and population scale, with multiple periods of accelerated growth corresponding to major infection waves.

Looking beyond raw counts, population-normalized metrics highlight a different pattern. States with the highest infection reach are not always those with the highest mortality burden. Smaller states such as Rhode Island and Alaska rank high in percentage affected, while mortality rates are more concentrated in parts of the South and Southwest, pointing to the role of demographic and healthcare factors beyond case volume alone.
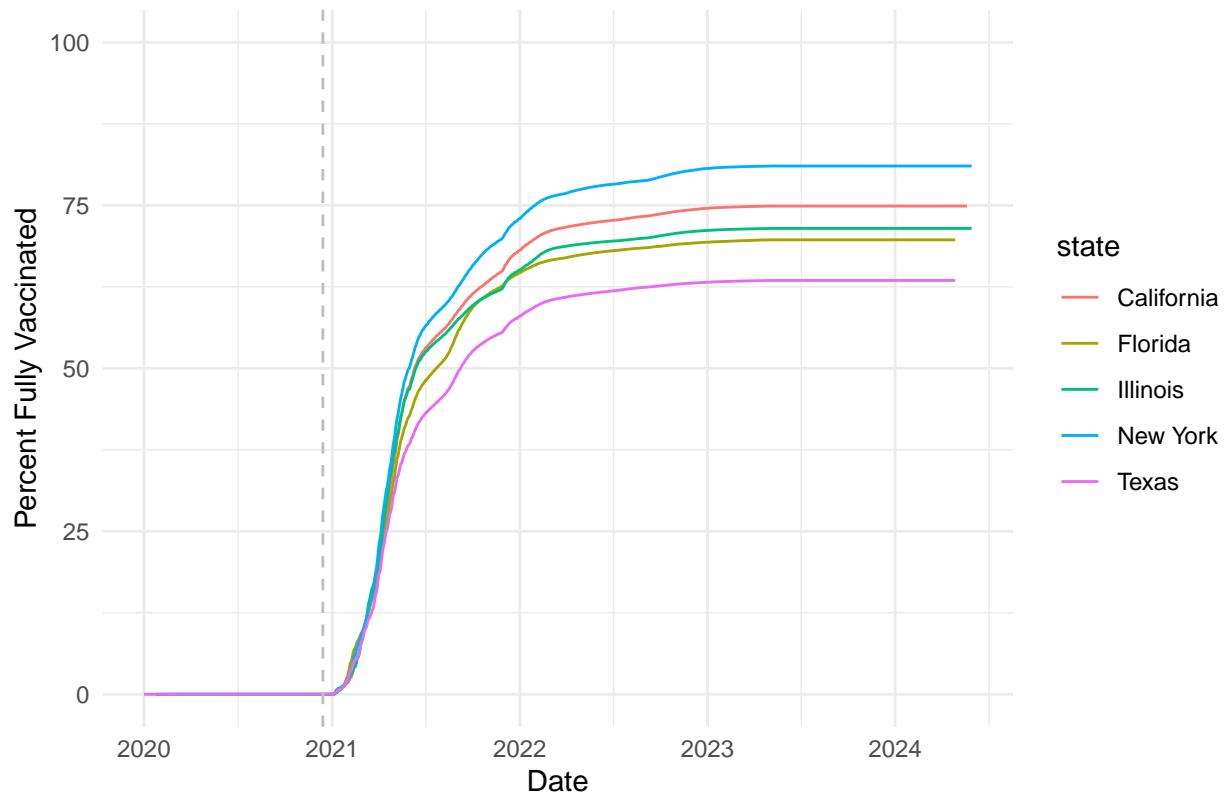
Policy Strictness, Vaccination Uptake, and Outcome Alignment

To understand how policy choices and vaccination rollout evolved alongside COVID-19 outcomes, the analysis focuses on the five U.S. states with the highest cumulative case counts. These states capture meaningful variation in timing, population scale, and policy posture under a shared national context.

### Government Response Stringency in High−Impact States
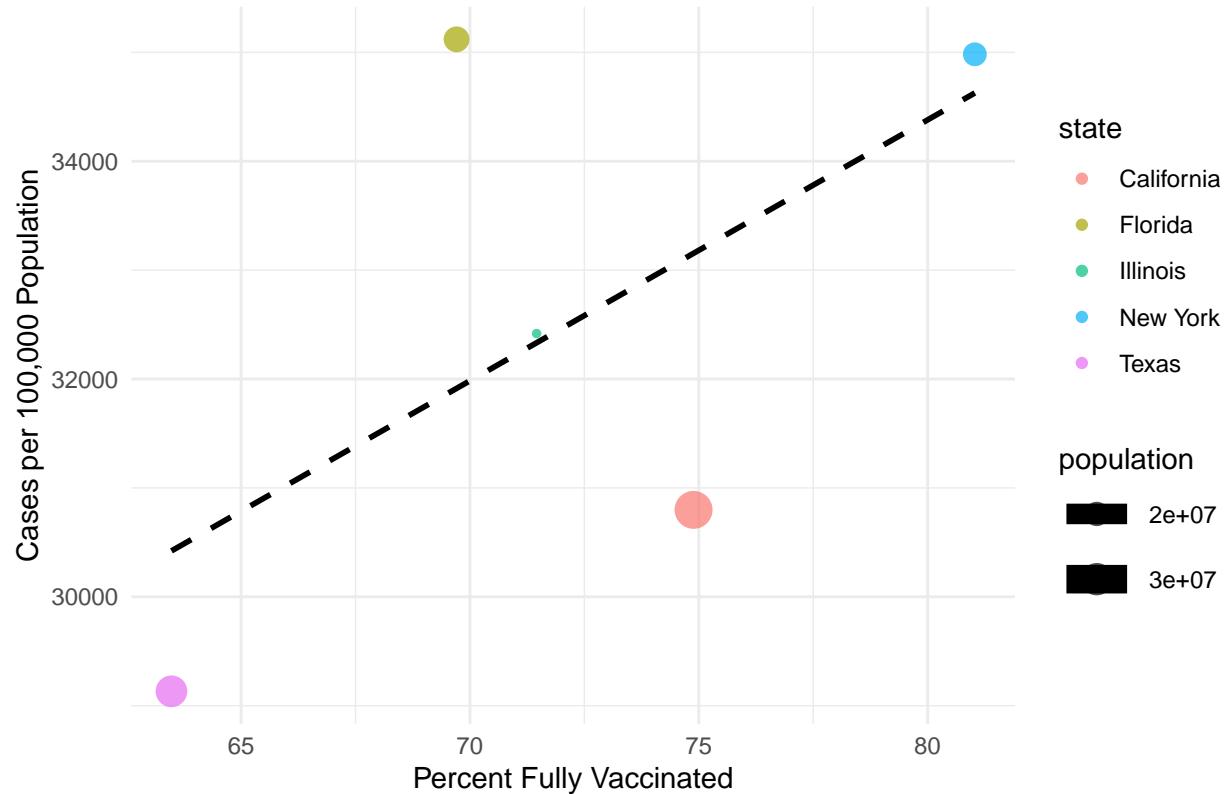
All five states implemented similarly strict measures during the initial phase of the pandemic in early 2020. Over time, policy paths diverged: New York and California maintained higher levels of restriction for longer periods, while Florida and Texas relaxed measures earlier. By late 2021, stringency levels converged across states, reflecting a broader shift toward managing COVID-19 under less restrictive policies.

Vaccination Uptake in High−Impact States

Vaccination uptake accelerated rapidly following vaccine availability in late 2020, but states diverged in their eventual coverage levels. New York and California reached higher vaccination plateaus than Florida and Texas, while Illinois followed an intermediate path. These trajectories reflect differences in timing, public response, and local implementation rather than a single explanatory factor.

## Vaccination Coverage vs. Cumulative Case Burden
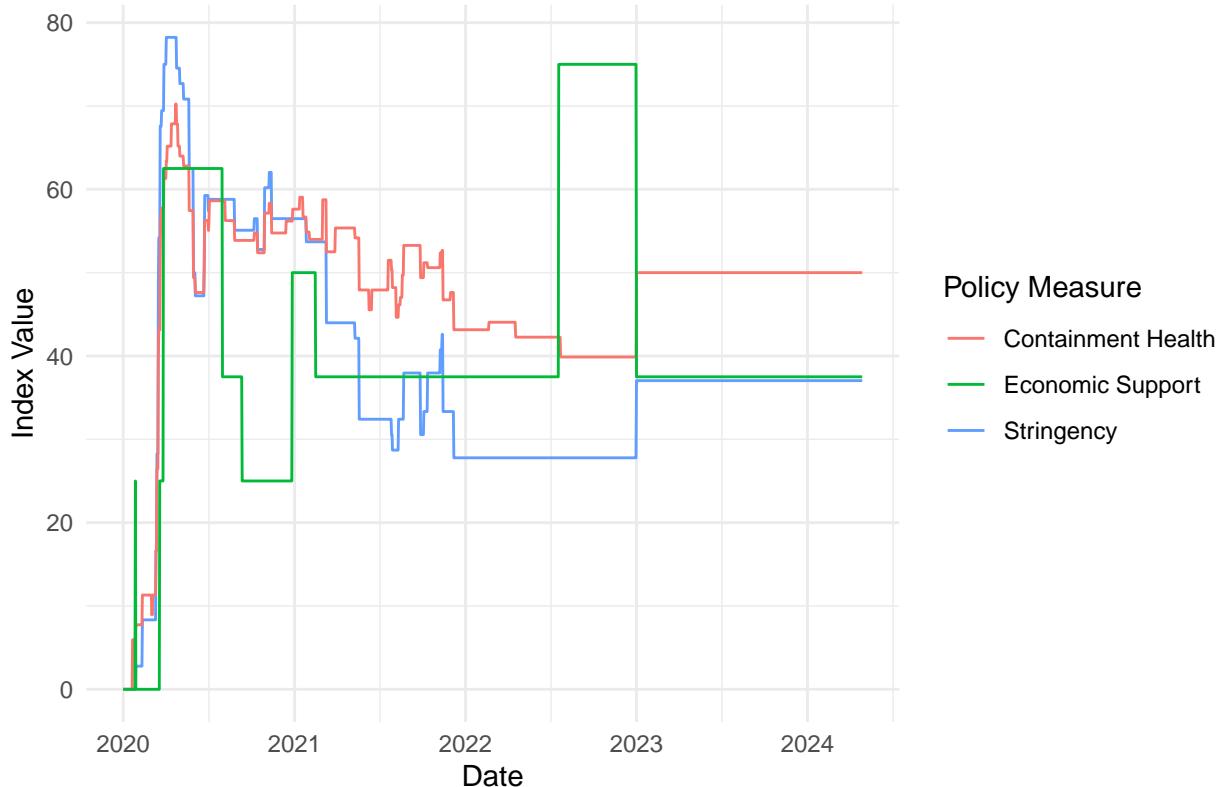
Higher vaccination coverage does not correspond to lower cumulative case burden in this cross-sectional view. States with the highest vaccination rates also experienced substantial early exposure and multiple infection waves. This reinforces the importance of timing, population density, and early outbreak dynamics when interpreting aggregate outcomes, and highlights the limits of simple policy–outcome narratives.
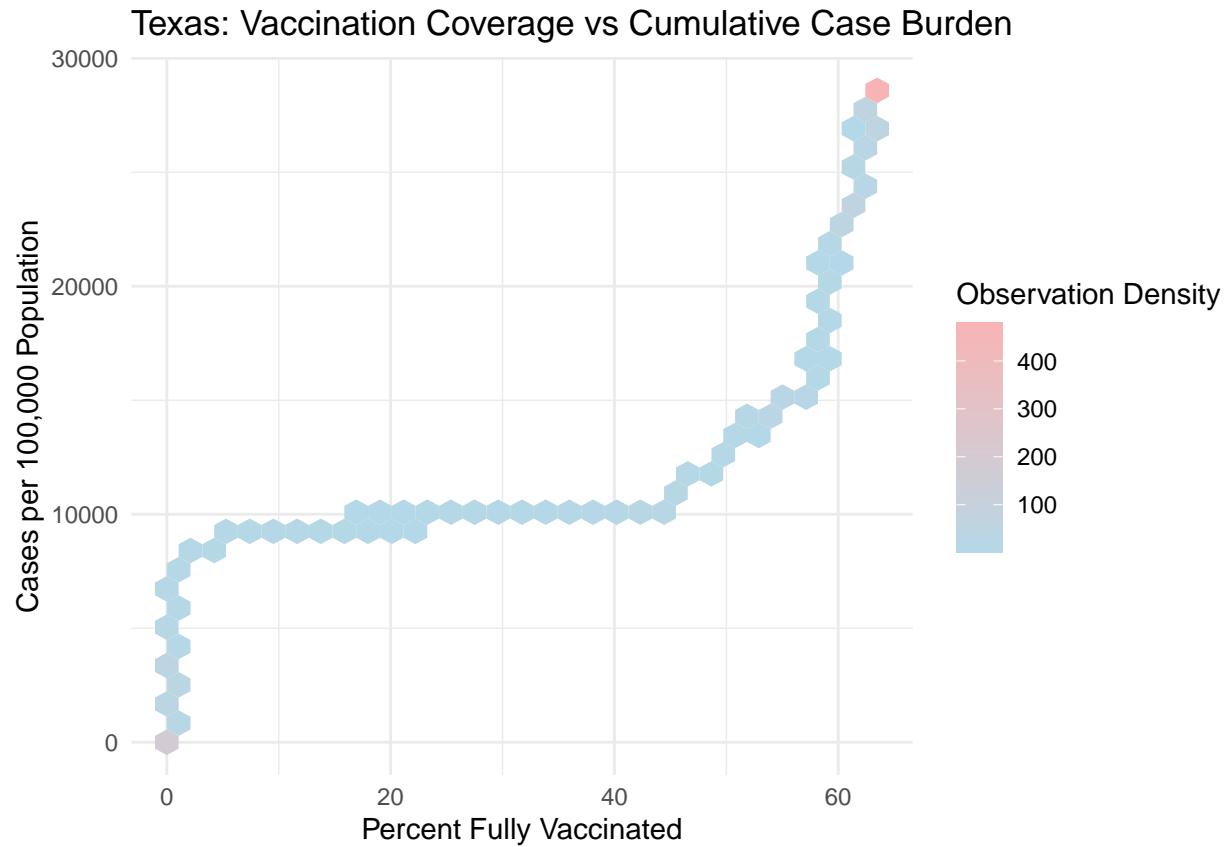
Texas as a Case Study in Policy Timing and Outcome Interpretation

Texas provides a useful contrast among high-impact states due to its relatively early relaxation of policy measures and distinct vaccination trajectory. Rather than treating Texas as an outlier to be explained away, this section uses it as a case study to highlight the limits of simple policy–outcome interpretations.

Texas Policy Measures Over Time

Texas followed a similar initial policy response to other large states in early 2020, but reduced restrictions earlier and maintained lower levels of stringency through much of 2021 and beyond. Economic support measures show a brief increase in early 2023, but overall policy indices remained comparatively stable during later pandemic phases.

Texas: Vaccination Coverage vs Cumulative Case Burden

The bimodal structure reflects the cumulative nature of both vaccination rollout and case reporting. Higher vaccination rates coincide with higher cumulative case counts not because vaccination increases transmission, but because both metrics increase over time. Sparse observations at higher vaccination levels correspond to later stages of the pandemic, when cumulative case totals had already risen substantially.

How to Interpret these Results

All findings in this analysis are descriptive and comparative. Observed alignments between policy measures, vaccination uptake, and outcomes reflect co-evolution over time rather than isolated effects. Differences across states should be interpreted as the result of interacting factors, including timing, population structure, healthcare capacity, and behavioral response, rather than as evidence for or against any single intervention.

Conclusion

This analysis highlights how policy strictness, vaccination uptake, and COVID-19 outcomes evolved together, while also illustrating the limits of drawing simple causal conclusions from aggregate data. States with higher vaccination coverage and prolonged policy measures often experienced higher cumulative case counts due to early exposure, population scale, and repeated infection waves.

Rather than pointing to a single dominant factor, the results underscore the importance of timing, baseline conditions, and behavioral responses when interpreting pandemic outcomes. These findings emphasize the need for careful framing when using observational data to evaluate policy effectiveness.Sharp declines in reported cases and deaths toward the end of the timeline likely reflect reporting changes rather than true epidemiological shifts.