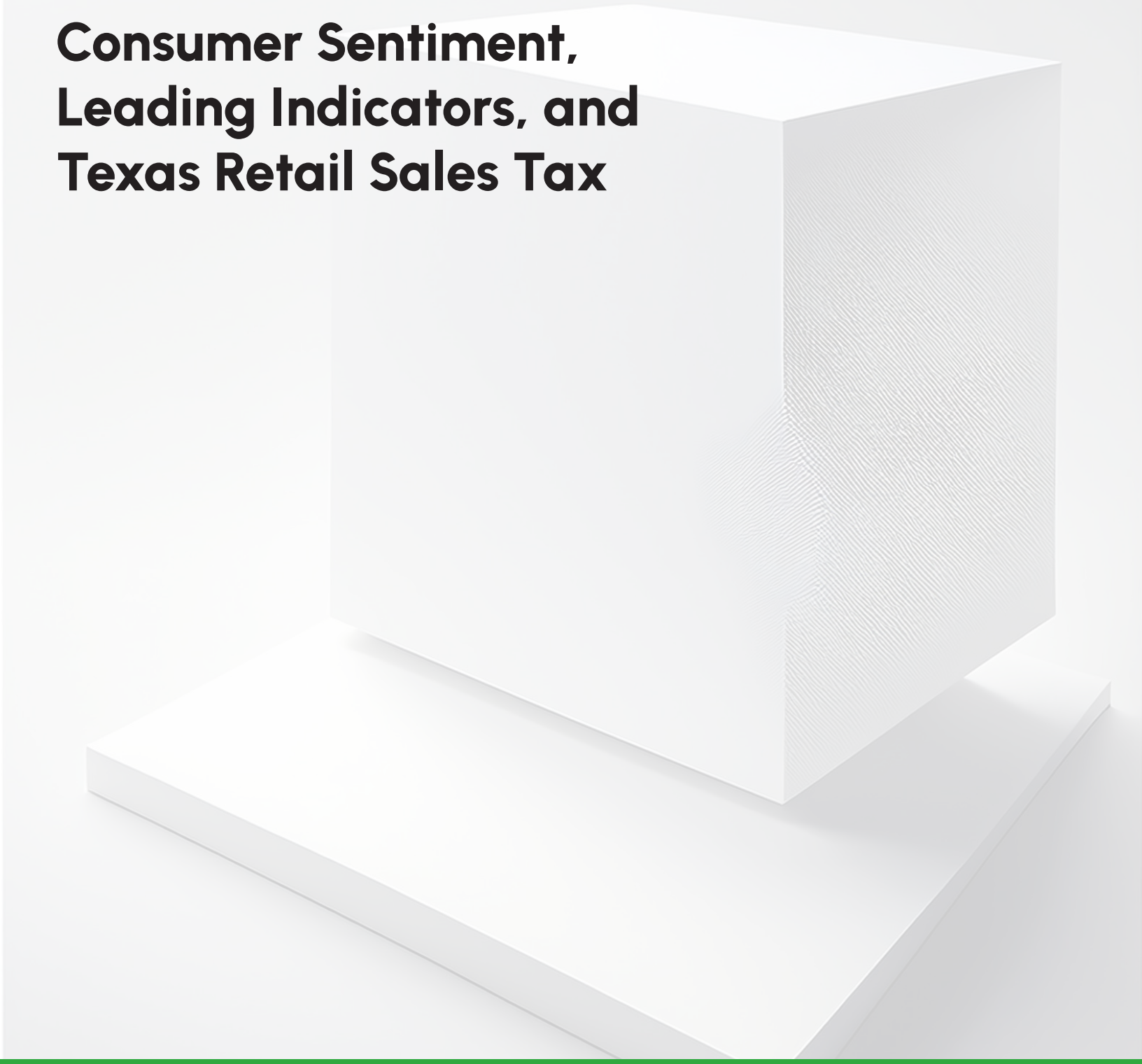
zactax research

Consumer Sentiment, Leading Indicators, and Texas Retail Sales Tax



zactax

Consumer Sentiment, Leading Indicators, and Texas Retail Sales Tax

This analysis investigates the correlation and predictive abilities between consumer sentiment indices and state retail sales taxes, with a focus on Texas. Using a range of sentiment and confidence measures - including the University of Michigan Consumer Sentiment Index (CSI), OECD Consumer Confidence and Composite Leading Indicators (CCC, CLI), Comptroller Reported Texas and US Consumer Confidence Indices (TX CCI, US CCI), and Dallas Federal Reserve Reported Texas Leading Indicator (TLI) - we evaluate their correlation and predictive power for quarterly gross retail sales tax collections in Texas.

In our analysis, we consider our retail sales tax as time series data, which in many models requires data transformation for careful analysis. We first assess baseline relationships after applying differencing and normalization techniques to our sales tax and sentiment data, ensuring numerical stability.

We then apply cross-correlation analysis, Granger causality tests, and ordinary least squares (OLS) regressions that suggest that multiple sentiment indices exhibit lagged predictive relationships with sales tax growth, as well as a specific indicator showing significant predictability after accounting for the autoregressive nature of retail sales tax.

In our final model, we identify that the TLI at a 3-quarter lag is a robust and statistically significant predictor of future sales tax performance. Other sentiment variables that appear significant in simpler models without autoregression include the UMich CSI and TLI at lag 1 quarter and the US CCI at lag 3 quarters.

These findings highlight the importance of accounting for both internal temporal dynamics and the predictive value of external economic indicators. These results can aid city managers in providing more informed projections of their sales tax revenue.

Introduction

Forecasting public revenue is a central challenge for state and local governments, especially when revenues depend on volatile economic activity such as consumer retail spending. This report explores whether consumer sentiment indices, widely used indicators of public economic outlook and confidence, can meaningfully improve short-to-medium forecasts of retail sales tax collections in Texas.

Specifically, we will test whether movements in several sentiment indicators have correlative and predictive power over future sales tax collections, including:

- The University of Michigan: Consumer Sentiment Index (CSI)
- The Organization for Economic Cooperation and Development: Composite Consumer Confidence and Leading Indicators (CCC, CLI)
- The Texas Comptroller: Consumer Confidence Indices for Texas and the United States (TX CCI, US CCI)
- The Dallas Federal Reserve: Texas Leading Index (TLI)

Methodology

Data Sources and Processing

Our retail sales tax data is sourced directly from the Quarterly Sales Tax Historical Data by Industry from the Texas Comptroller of Public Accounts. However, our sentiment indices were reported monthly. To remedy, we transformed each of our sentiment indices to a quarterly average, aligning with the frequency of our sales tax data.

Gross sales tax allocations and sentiment indicators were treated as time series. To meet the assumptions required for correlation and predictive modeling testing (e.g. stationary time series), we applied differencing techniques to our series and validated using Augmented Dickey-Fuller tests on our differenced series as well as visual inspections of the autocorrelation (ACF) and partial autocorrelation (PACF) plots. All sentiment indices were either already stationary or needed first-differencing, while the retail sales tax series required second-order differencing.

After the differencing process, all series were standardized (z-scored) to address:

- Differences in scale between variables (e.g. taxes in billions versus sentiment indices around 100)
- Multicollinearity and ill-conditioning in regression and ARIMA-based models
- Comparability of coefficients across models

We apply normalization after differencing to ensure that scaling reflects stationary variance, not trend-induced distortion.

Correlation and Lead-Lag Analysis

To evaluate the correlative and lagged relationships, we used:

- Cross Correlation (CCF) analysis, after prewhitening our sales tax and sentiment series to remove shared autocorrelation structure
- Granger causality tests up to 6 quarter lags to determine whether lagged sentiment values improved prediction of future retail sales tax growth

This analysis surfaced candidate lags for each sentiment variable, giving us clues as to sentiments that may contain statistically significant predictive power.

Regression and Predictive Modeling

Ordinary Least Squares (OLS) Regression
We constructed a linear regression model of second-differenced sales tax using its own lagged values (chosen from PACF) and selected sentiment indices at candidate lags (from CCF/Granger tests and minimizing multicollinearity via the Variance Inflation Factor).

A backward-elimination procedure removed statistically insignificant predictors at the 95% level of significance, yielding a parsimonious OLS model with high explanatory power (Adjusted R-squared value of roughly .94).

ARIMAX Modeling
To properly account for autocorrelation in the dependent variable, we implemented a (3,2,0) Autoregressive Integrated Moving Average with Exogenous Regressors (ARIMAX) model of order (3,2,0) based on:

- Residual diagnostics (ACF, PACF)
- Stationarity assessmen
- Ljung-Box tests to confirm white noise residuals

This model was able to retain statistically significant exogenous predictors, while others (previously significant in OLS) became insignificant once autoregressive structure was correctly accounted for.

Residual diagnostics confirmed that the ARIMAX(3,2,0) model fully removed autocorrelation, producing white noise residuals, and accurately isolated the exogenous signal from our sentiments.

Results

Exploratory Findings
Initial differenced and normalized series reveal that multiple consumer sentiment indices exhibit meaningful correlation with sales tax changes:

- Cross Correlation (CCF) analysis on prewhitened series show that several sentiment indices lead gross retail sales tax by 1 to 3 quarters.
- Most notably, the Texas Leading Index (TLI) and UMich Consumer Sentiment Index (CSI) demonstrated statistically significant positive correlations at lags 1 and 3.
- Granger causality tests confirmed these lead-lag relationships.
- TLI at lag 3 and CSI at lag 1-3 were statistically significant under tests for Granger causality.

L*inear Regression with Autoregressive Terms*
An OLS regression model that included three autoregressive lags of sales tax and lagged sentiment predictors produced strong in-sample performance:

Sentiment	Lag	p-value
TLI	1	0.000
TLI	3	0.000
OECD CC	3	0.032
UMich CSI	3	0.023
US CCI	3	0.001

Our OLS model has an adjusted R-squared value of 0.939, which is considered to be a good fit of our sales tax data. All autoregressive terms of the dependent variable were highly significant ($p < .001$), confirming temporal dependence in sales tax differences.

To reduce multicollinearity and improve interpretability, we applied:

- Variance Inflation Factor (VIF) filtering to remove sentiment features with comparatively high scores
- Resulting variables show very low VIF scores, ensuring we have minimal multicollinearity:

Feature	VIF
UMich CSI_lag1	1.088520
TLI_lag1	1.471547
US CCI_lag3	1.121485
TLI_lag3	1.452633

These findings support the idea that both sentiment and autoregressive history are required to explain sales tax variation. However, a linear regression model is not as robust at modeling autoregressive time series as a model such as ARIMA, which we will showcase the results for next.

ARIMAX Modeling

Our final model used an ARIMAX(3,2,0) specification to account for second-order differencing and AR(3) behavior in the sales tax series. Key results include:

1) A significant exogenous predictor
- TLI at lag 3 remained significant ($p < .001$) even after controlling for autoregression and differencing.
- Other predictors, inducing those significant in OLS – became statistically insignificant once autocorrelation was properly modeled.

2) Model fit and diagnostics:
- Model residuals pass the Ljung-Box test, meaning there is no remaining significant autocorrelation.
- ACF/PACF of residuals also show no remaining structure, supporting proper model specification.
- Numeric stability is showcased, supporting the prior z-score normalization of utilized features.

Insights and Commentary

After controlling for autocorrelation and stabilizing the data, only the Dallas Fed's Texas Leading Index at lag 3 showed robust, statistically significant predictive power for retail sales tax growth in Texas. This suggests that while several sentiment indicators are correlated with sales tax changes, most do not offer unique predictive value beyond the lagged sales tax itself, with the exception of TLI.

Given that the Texas Leading Index is designed to project the future of the state's economy, composed of indicators that typically change direction before the overall economy, we certainly would expect this indicator to have some predictive power around 6-9 months (or 2-3 quarters, in this case).

Note that correlations and Granger causality tests do not imply a physical cause-effect relationship between any sentiment and the retail sales tax, and the relationships we conclude are meant to showcase the aid that consumer sentiment can provide when projecting future sales tax earnings.

Sources

"Allocation - Quarterly Sales Tax Historical Data." *state.tx.us*, 2015, mycpa.cpa.state.tx.us/allocation/HistSales.

Federal Reserve Bank of Dallas. "Texas Leading Index." *dallasfed.org*, 2024, www.dallasfed.org/research/econdata/tli.

OECD. "Composite Leading Indicators." Oecd.org, 2018, data-explorer.oecd.org/vis?lc=en&pg=0&snb=1&vw=tb&df.

Texas Open Data Portal. "Key Economic Indicators | Open Data Portal." *data.texas.gov*, data.texas.gov/dataset/Key-Economic-Indicators/karz-jr5v/about_data.

University of Michigan. "Table 1: The Index of Consumer Sentiment." *data.sca.isr.umich.edu*, data.sca.isr.umich.edu/data-archive/mine.php.

zactax