

Examining Auto Loss Trends through the COVID Pandemic

AUGUST | 2022





Examining Auto Loss Trends through the COVID Pandemic

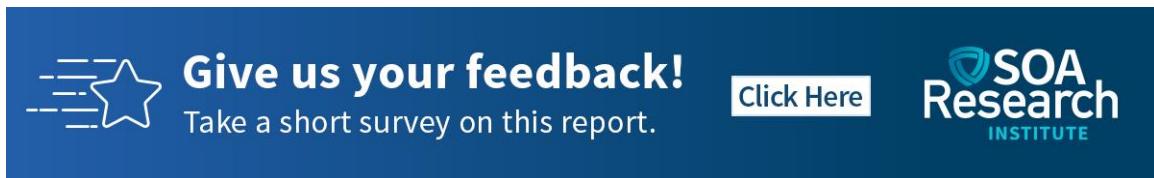
AUTHORS Robert Richardson, ASA, PhD

Brian Hartman, ASA, PhD

SPONSOR Society of Actuaries Research Institute

American Property Casualty Insurance Association

Casualty Actuarial Society

A dark blue horizontal banner with white text and icons. On the left is a white star icon with horizontal lines extending from its left side. To the right of the star is the text "Give us your feedback!" in a large, bold, white sans-serif font. Below this is the text "Take a short survey on this report." in a smaller white font. To the right of this text is a white rectangular button with the text "Click Here" in a dark blue font. On the far right of the banner is the SOA Research Institute logo in white.

Caveat and Disclaimer

The opinions expressed and conclusions reached by the authors are their own and do not represent any official position or opinion of the Society of Actuaries Research Institute, the Society of Actuaries or its members. The Society of Actuaries Research Institute makes no representation or warranty to the accuracy of the information.

Copyright © 2022 by the Society of Actuaries Research Institute. All rights reserved.

CONTENTS

Executive Summary	4
Section 1: Introduction	5
Section 2: Data.....	5
Section 3: Methods.....	6
3.1 Dynamic Linear Models	6
3.2 Multivariate Dynamic Linear Model	7
Section 4: Results.....	8
4.1 Countrywide Results.....	8
4.2 State-specific Results.....	10
Section 5: Conclusions	13
Section 6: Acknowledgments	14
References.....	15
About The Society of Actuaries Research Institute	16

Examining Auto Loss Trends through the COVID Pandemic

Executive Summary

The COVID-19 pandemic affected the auto insurance industry in many ways. In this report, we will examine the effects of the pandemic on auto losses using dynamic linear models. We examine which predictors best describe the trends, and the impact the pandemic has on predicted future claims.

The model results showed a few consistent themes. To follow intuition, miles driven positively affects frequency and supply costs positively relate to severity. However, it is interesting to note that rental car prices were a much more significant predictor for severity than car part prices. Also, for loss cost, miles driven was more significant than either of the car part or rental car prices.

The pandemic had an impact on the future predictions of frequency (lower), severity (higher), and loss cost (mixed).

In addition to this report, we include an appendix with plots of all of the state models to help readers understand the impact in their location and some R code to quickly tabulate the monthly miles driven by state.



Give us your feedback!

Take a short survey on this report.

[Click Here](#)

SOA
Research
INSTITUTE

Section 1: Introduction

The COVID-19 pandemic has affected nearly all areas of modern life and auto insurance is no exception. In March 2020 widespread shutdowns greatly reduced traffic (and therefore accidents and claims) across the country. Supply chain issues and increased demand for products led to the highest inflation in decades.

In this report we examine the auto loss trends through the pandemic. We describe the impact of the pandemic through exploratory data analysis. We also compare multiple potential predictors and see which ones perform the best and where they perform well. We also incorporate the pandemic into our future predictions to examine its impacts.

Section 2: Data

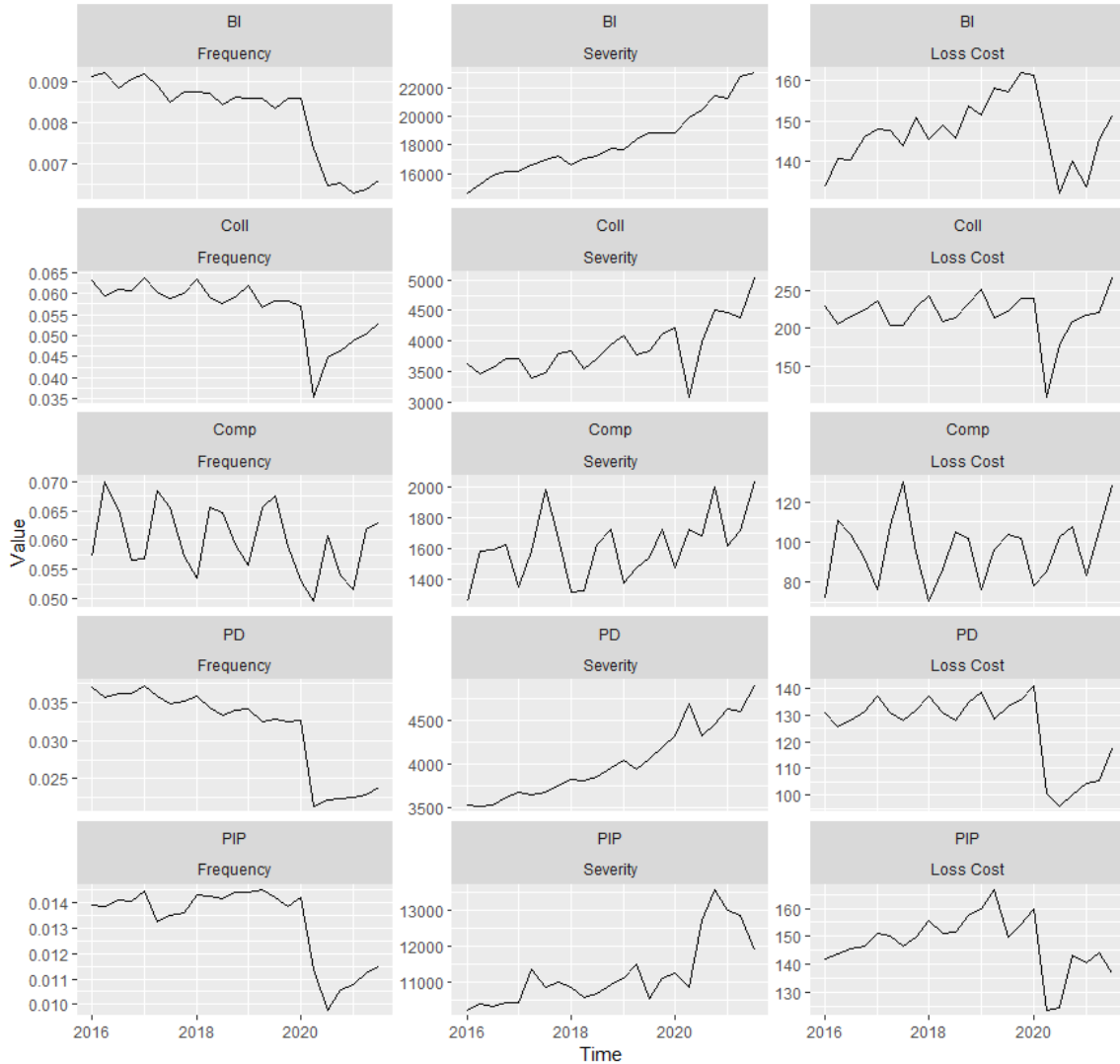
The loss data is gathered from the Fast Track Plus database (Independent Statistical Service, 2022). We obtained quarterly loss amounts, claims, and earned car years for each state and coverage (BI, PD, Comp, Coll, PIP, and property protection (PPI)) from Q1 2016 through 2021 (the most recent quarter available at the time we did our analysis). We then calculated frequency (claims/earned car years), severity (loss amount/claims), and loss cost (loss amount/earned car years). For the remainder of this report, we focus on those three metrics. Figure 1 shows the countrywide frequency, severity, and loss cost for all quarters in our dataset. Five different coverages are considered: Bodily Injury (BI), Collision (Coll), Comprehensive (Comp), Property Damage (PD), and Personal Injury Protection (PIP). The pandemic drop is obvious in Q2 2020 for essentially all the coverage frequencies. Comprehensive seems less dramatic because of its seasonal pattern, but it was still affected. Collision also had a momentary drop in severity and property damage had a momentary increase.

In addition to the loss data, we gathered some covariates. Monthly miles traveled were scraped from the Federal Highway Administration website (2022). The R code which scrapes those values can be found on the web page which hosts this report, as well as on the GitHub repository <https://github.com/Society-of-actuaries-research-institute/GIP125-Auto-Loss-Cost-Reports>. To account for population size, miles traveled was scaled by individual state averages, and so will be referred to as Scaled Miles for this report. The supply variables we felt were most representative of the effect of the supply shortage were car replacement parts, used car prices, and passenger rental car costs. These variables are measured as part of the Producer Price Index (Bureau of Labor Statistics, 2022). They are also highly correlated and so we will have to account for that in the modeling. Not all coverage types will include all three of these supply variables directly. However, we will consider all three supply variables for each coverage type to allow for indirect correlations that better explain the trends.

The final variable that was used is Covid-19 numbers. Again, these are highly correlated with the other variables, so to account for that we use the log of the proportion of a state with Covid-19 divided by the proportion of the country with Covid-19. This will hopefully capture lingering effects from areas with higher or lower Covid-19 values than the rest of the country.

There are other factors that have contributed to the auto insurance trends during the pandemic. Thefts have risen causing more claims to be made. Even though fewer miles have been driven, the number of fatal accidents has not decreased at the same rate, suggesting riskier driving. Rental car rates are represented in our model, but the duration of rental periods has also increased. These all would affect insurance claims and costs, but data for these variables were not available during the domain of the study.

Figure 1
COUNTRYWIDE METRICS FOR Q1 2016 – Q3 2021



Section 3: Methods

A dynamic linear model is a model for time series data that is often considered to be more realistic for real-world data than other time series models because it allows for relationships between key factors in the model to change slightly over time. We use a dynamic linear model that allows for a linear trend in the time series and regression variables chosen from the predictors mentioned above. We fit the model on the data for the country as well as for the individual state time series.

3.1 DYNAMIC LINEAR MODELS

The dynamic linear model used for the individual time series can be broken into two individual components. The first component is a polynomial model of order 2, which is essentially a random walk with a linear trend. The second component is a regression model where the coefficients of the model evolve

over time according to a random walk. This part of the model is called the observation equation which is shown here with only one regressor for simplicity, but there can be many included.

$$Y_t = \theta_t + \beta_t X_t + \varepsilon_t,$$

where Y_t represents the time series at time t and X_t represents the value of a single predictor at time t . The error term, ε_t is included to show that the data has normally distributed errors. The rest of the model describes how the parameters vary over time. They are often written in matrix form but can be decomposed into several different formulae known as the process equations.

$$\theta_t = \theta_{t-1} + \gamma_{t-1} + \omega_{1,t}$$

$$\gamma_t = \gamma_{t-1} + \omega_{2,t}$$

$$\beta_t = \beta_{t-1} + \omega_{3,t}$$

The variables θ_t and γ_t represent unknown latent variables that provide the temporal structure for the data. The error terms, $\omega_{i,t}$ are normal.

The latent variables and regression coefficients are estimated via Kalman Filtering equations, which is essentially using the process equations as a Bayesian prior for the unknown parameters and then updating them via the observation equation. This is a very stable and effective estimation procedure. The only other unknown parameters are the variances of the processes and time 0 priors, which are either fixed at reasonable values or estimated using maximum likelihood.

One method that can be used to compare model fits is using one-step ahead predictions. This can be done at the end of a time series for a holdout group, or it can be done in-sample. For in-sample one-step ahead predictions, the data at time $t+1$ is predicted using the filtering equations through time t , which essentially simulates holding out the time point. This is done for all time points it is possible to predict ahead. These can be summarized through a mean squared error of the one-step ahead predictions. We will use the MSE for in sample one-step ahead predictions to compare model fits to determine which supply variable to include in the model.

3.2 MULTIVARIATE DYNAMIC LINEAR MODEL

A multivariate dynamic linear model can be used to model several different time series at once. Essentially this looks very similar to the individual time series, but the observation level is a matrix equation

$$\mathbf{Y}_t = \mathbf{F}_t \boldsymbol{\theta}_t + \boldsymbol{\varepsilon}_t$$

where \mathbf{F}_t contains both information mapping the vector of latent variables, $\boldsymbol{\theta}_t$, to the data vector, \mathbf{Y}_t , as well as covariate information. The vector $\boldsymbol{\theta}_t$ is augmented to contain both the latent variables and the coefficients. The process level looks very similar to the individual time series,

$$\boldsymbol{\theta}_t = \mathbf{G}_t \boldsymbol{\theta}_{t-1} + \boldsymbol{\omega}_t$$

where \mathbf{G}_t is a block diagonal where there is one block for each value in \mathbf{Y}_t that contains information for the polynomial trend of order 2 and then a block that is just a diagonal matrix with a dimension equal to the number of covariates. The benefit of a multivariate model is that the regression component will be fit using all the data instead of the individual time series, thus increasing the overall amount of information that is

used to estimate parameter coefficients. We will use both individual time series to detect trends for individual states and then a multivariate model with each state represented to determine a general trend for the full country.

Section 4: Results

We fit the data to the dynamic linear model using the approach detailed in Section 3.

4.1 COUNTRYWIDE RESULTS

The data we use is through Q4 2021. From this point we can project forward several steps. However, the data is not entirely available for these time points. To account for this, we make projections under two scenarios. The Pre-Covid scenario uses values from Q4 2018 and the first two quarters in 2019. Some of the variables have a seasonal effect, and so it is important that we match the predicted quarters, and these would be the most recent representative stretch. Post-covid values mimic variables from Q4 2020 and the first two quarters of 2021.

Figure 2 shows that projections for frequency for comprehensive coverage change significantly under the two scenarios. The main cause of this is the lower miles driven during 2020 and 2021.

Figure 2
COUNTRYWIDE COMPREHENSIVE FREQUENCY PROJECTIONS

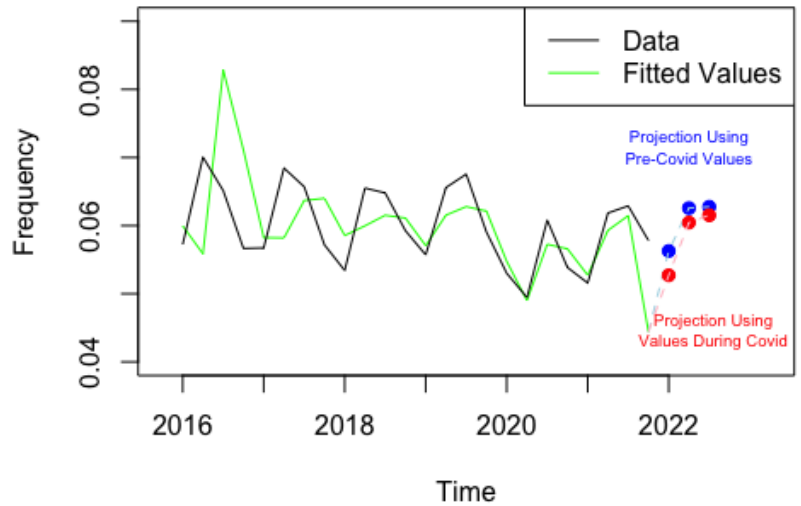
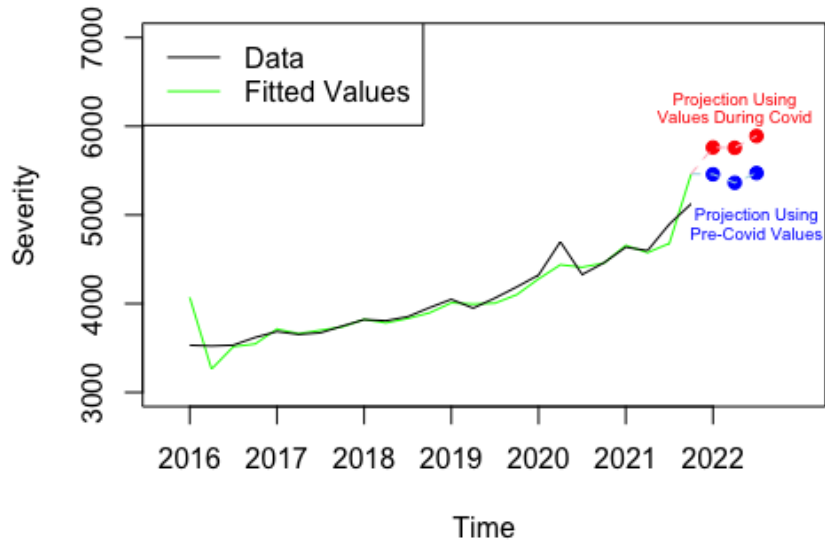


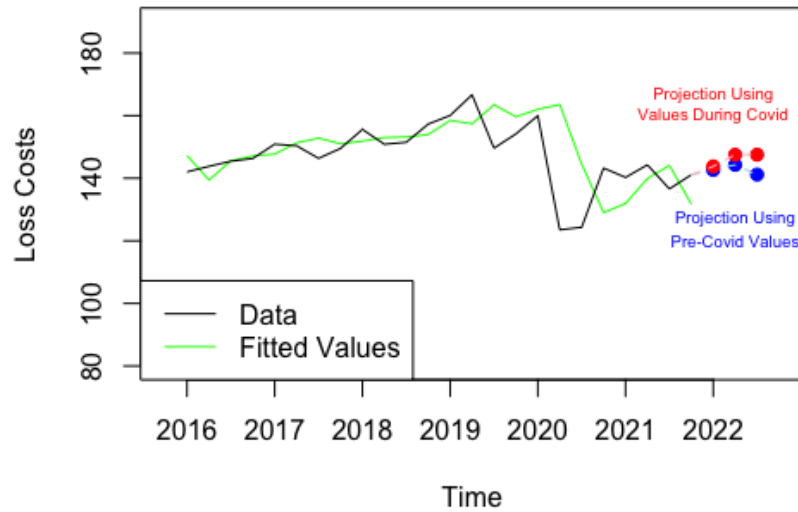
Figure 3 shows that post-covid projections are higher than projections using pre-covid values, showing the effect of increased supply prices in 2021.

Figure 3
COUNTRYWIDE PROPERTY DAMAGE SEVERITY PROJECTIONS



Finally, loss cost will depend highly on both miles driven and supply variables. We see in Figure 4 that the supply costs outweigh the effect of miles driven to cause projections using post-covid values to be higher than pre-covid values.

Figure 4
COUNTRYWIDE LOSS COST PIP PROJECTIONS



Figures 2 through 4 also show the overall model fit of the different models. To be more accurate in predictions, the insignificant features were not included in the model fit when making these plots. There are essentially three paradigms for the shape of the curve for the national series. There is a consistent

upward trend, a seasonal trend that is either relatively flat or upward sloping, and then a curve that dips significantly starting in 2020. The model fits are very good for the first two of these paradigms whereas the model is not able to capture the sharp dip of the third paradigm. This could be because the predictors are slightly delayed or because there may yet be a significant predictor that explains this data that we do not have in the model.

The multivariate model is able to break apart the individual time series for each state and build a model that can determine the effect of these variables accounting for each state's values. Here we report the effect of each of the significant covariates on several of the different series. The values are given in terms of the scaled data, meaning that the coefficient given is the expected number of standard deviations the target variable changes with a one standard deviation change in the respective variable. We fit several models from each series to find the best set of variables. The best model is shown with the estimated covariates. Supply variables were not used in the frequency models and only one was ever used at a time in the severity and loss cost models because they were so highly correlated. Variables that were not included in the best model do not have the coefficients reported. The best supply variable for the models is used cars (UC), rentals (R) or replacement parts (P).

Table 1
COUNTRYWIDE PARAMETER ESTIMATES

Target	Coverage	Miles Driven	Covid-19 above national average	Supply Variable
Frequency	Comprehensive	0.63	0.23	---
Frequency	Collision	0.24	---	---
Frequency	Property Damage	0.21	---	---
Frequency	Bodily Injury	0.09	---	---
Frequency	PIP	---	1.21	---
Severity	Comprehensive	-0.09	---	0.24 (R)
Severity	Collision	-0.04	---	0.22 (R)
Severity	Property Damage	---	---	0.09 (UC)
Severity	Bodily Injury	0.03	0.11	0.08 (R)
Severity	PIP	---	---	0.05 (UC)
Loss Cost	Comprehensive	0.43	---	0.88 (R)
Loss Cost	Collision	0.09	---	0.74 (R)
Loss Cost	Property Damage	---	---	0.39 (UC)
Loss Cost	Bodily Injury	0.01	-0.05	0.53 (R)
Loss Cost	PIP	---	---	0.46 (UC)

The best supply variable was consistent between Severity and Loss Cost models. Specifically, rentals was the best variable for comprehensive, collision, and bodily injury coverages while used cars was the best variable for property damage and PIP coverages. Also, the size of the coefficients suggests the loss costs target variable is most drastically affected by the supply variables while severity was not as affected.

4.2 STATE-SPECIFIC RESULTS

Models were also run individually for each state. Each frequency model uses the miles driven and the Covid-19 proportion difference from the rest of the country as covariates. The significance of these variables is measured and recorded. For the severity and loss cost model we also add a supply variable. Car part prices and rental prices are so highly correlated that we did not want to use both in the model. We use

one step ahead prediction error to determine which variable is better and then use that in the model. The chosen supply variable for each state and coverage is summarized in Tables 2 and 3.

Table 2
OPTIMAL SUPPLY VARIABLE SUMMARY FOR SEVERITY

	BI	Coll	Comp	PD	PIP
Parts	12	3	8	0	1
Rentals	33	41	31	18	7
Used Cars	6	7	12	33	11

Table 3
OPTIMAL SUPPLY VARIABLE SUMMARY FOR LOSS COST

	BI	Coll	Comp	PD	PIP
Parts	12	20	8	8	4
Rentals	21	27	27	21	13
Used Cars	18	4	16	22	2

This chosen supply variables follow the nationwide model showing that rental prices is the best for BI, comp, and collision with used car prices being the best for PD and PIP (though rentals were the best in the majority of PIP loss cost models). Maps of the significant variables are available in the appendix. The maps in the appendix only show the significant variables, sometimes the optimal variable is not statistically significant. As mentioned above, it seems in all these cases that the supply variables are surrogates for broader market conditions.

Figures 5-7 show whether the other covariates are positively or negatively significant for a select subset of coverages and metrics. All the maps are available in the appendix. States without data for a particular coverage are colored dark gray.

In Figure 5 we see that in almost every state, scaled miles traveled is positively correlated with comprehensive frequency. That makes sense because as the number of miles driven decreases, the number of claims should also decrease. Also, both comprehensive loss cost and mile traveled have a strong seasonal trend which is higher in the summer (because of road trips and hail prevalence) and lower in the winter.

Figure 5
SIGNIFICANCE OF SCALED MILES TRAVELED IN COMPREHENSIVE FREQUENCY

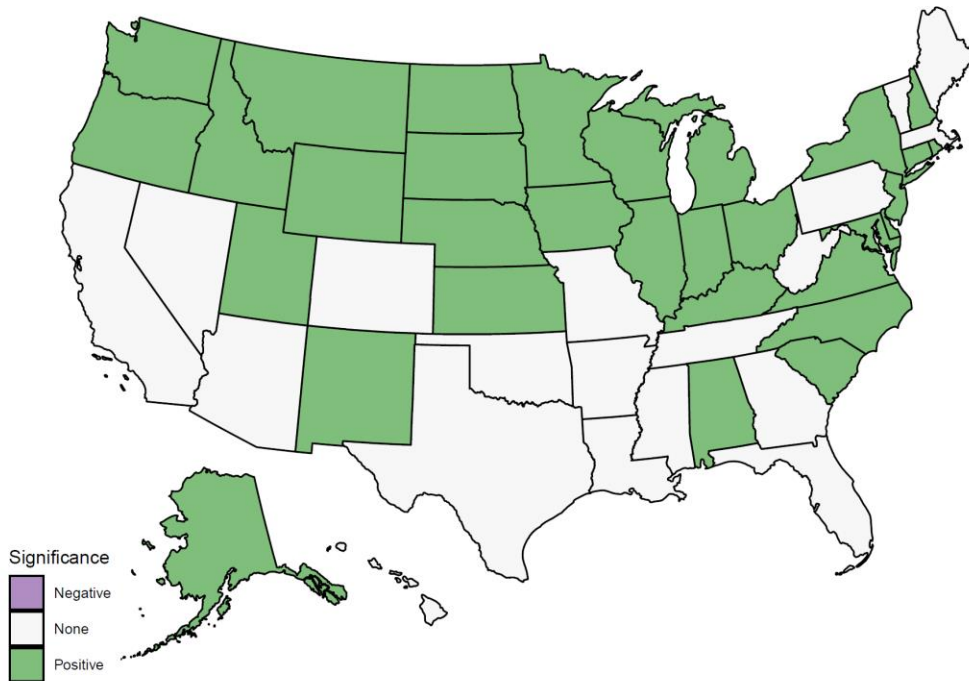
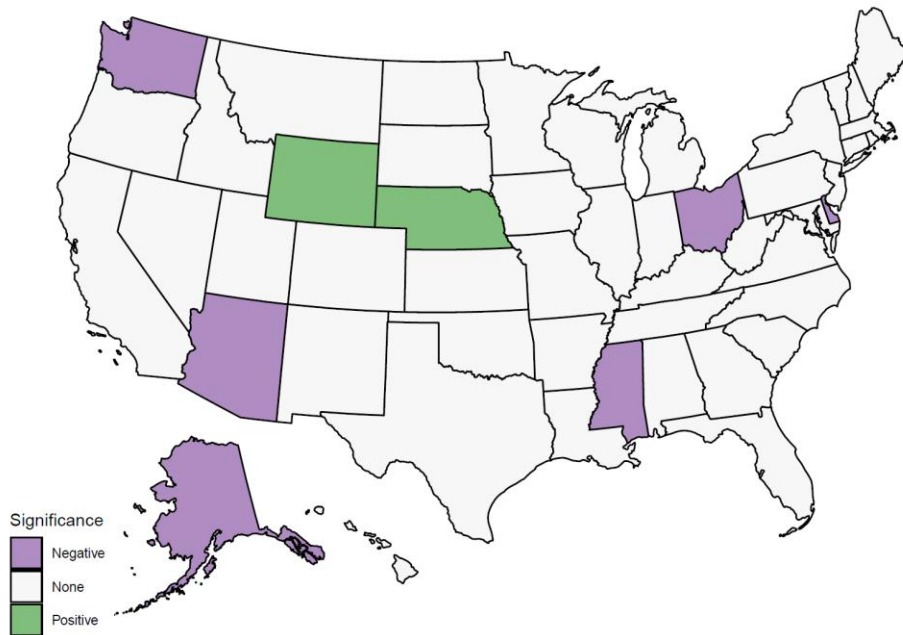


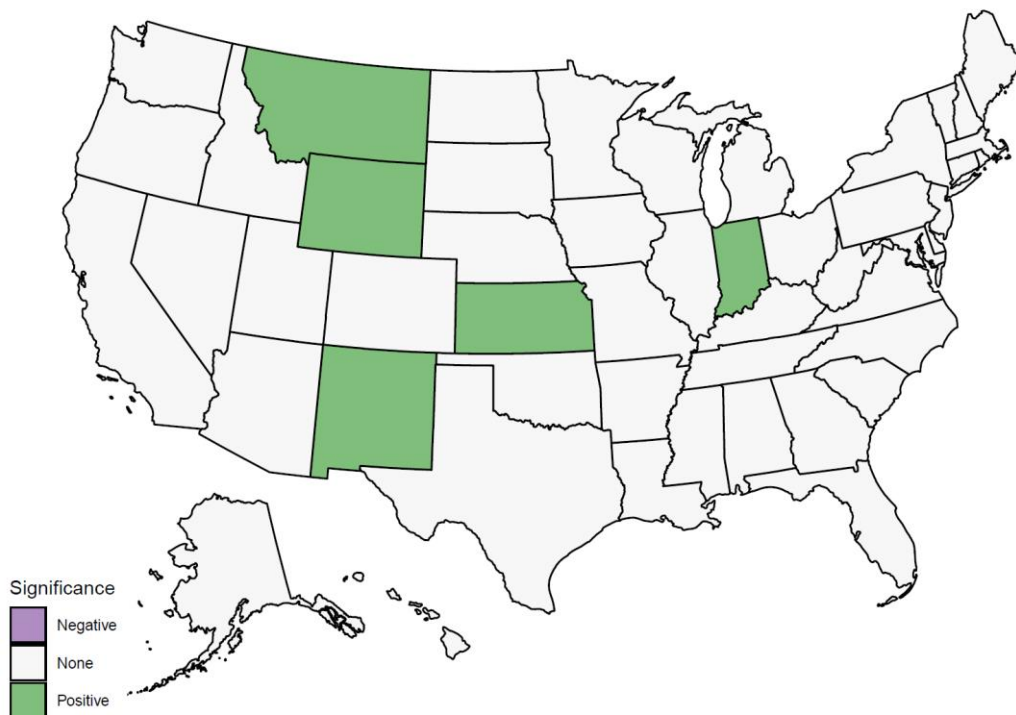
Figure 6 shows that scaled miles traveled doesn't not have much of a correlation with comprehensive severity.

Figure 6
SIGNIFICANCE OF SCALED MILES TRAVELED IN COMPREHENSIVE SEVERITY



Because loss cost is the product of frequency and severity, the correlation between loss cost and miles traveled is somewhat muted, though definitely still positive (Figure 7).

Figure 7
SIGNIFICANCE OF SCALED MILES TRAVELED IN COMPREHENSIVE LOSS COST



Section 5: Conclusions

Using the Dynamic Linear Model structure, we built several models to try to determine which variables are most important for various insurance coverages. Using both countrywide and state-specific data, we see how miles traveled has a significant impact on frequency and loss cost, although we suspect the impact of miles traveled on severity is only present due to certain features of the data and not practically significant. Using Covid-19 numbers beyond the other predictors that were affected by Covid-19 did not consistently help improve the model in predicting frequency, severity, or loss cost.

Rental costs were the best supply variable for bodily injury, comprehensive, and collision, while used car prices were the best supply variable for property damage and PIP. The projections also confirm this as projected loss cost under pre-covid conditions was higher than projections for post-covid conditions, suggesting that the effect of fewer miles driven was stronger than the effect of increased supply costs.

Supply costs have continued to rise while miles driven has since reverted to pre-covid levels, so additional data may be insightful as we compare more thoroughly how the pandemic will affect auto insurance in years to come.



Give us your feedback!

Take a short survey on this report.

[Click Here](#)

SOA
Research
INSTITUTE

Section 6: Acknowledgments

The researchers' deepest gratitude goes to those without whose efforts this project could not have come to fruition: the Project Oversight Group for their diligent work overseeing, reviewing, and editing this report for accuracy and relevance.

Project Oversight Group members:

David R. Clark, FCAS, MAAA

Brian Fannin ACAS CSPA MAAA

Newton Jennings, FCAS, MAAA

Thomas G. Myers, FCAS, MAAA

Norman Niami FCAS, MAAA, Affiliate IFoA

Robert C. Passmore, CPCU

At the Society of Actuaries Research Institute:

Dale Hall, FSA, MAAA, CERA, CFA – Managing Director of Research

Rob Montgomery, ASA, MAAA, FLMI, Consultant – Research Project Manager

References

Bureau of Labor Statistics. (2022) *Producer Price Indexes*. <https://www.bls.gov/ppi/>

Federal Highway Administration. (2022) *Travel Monitoring: Traffic Volume Trends*.
https://www.fhwa.dot.gov/policyinformation/travel_monitoring/tvt.cfm

Independent Statistical Service. (2022) *Fast Track Plus Data*. https://www.iss-statistical.net/docs/default-source/zmag-links/iss_fasttrackplus.pdf

About The Society of Actuaries Research Institute

Serving as the research arm of the Society of Actuaries (SOA), the SOA Research Institute provides objective, data-driven research bringing together tried and true practices and future-focused approaches to address societal challenges and your business needs. The Institute provides trusted knowledge, extensive experience and new technologies to help effectively identify, predict and manage risks.

Representing the thousands of actuaries who help conduct critical research, the SOA Research Institute provides clarity and solutions on risks and societal challenges. The Institute connects actuaries, academics, employers, the insurance industry, regulators, research partners, foundations and research institutions, sponsors and non-governmental organizations, building an effective network which provides support, knowledge and expertise regarding the management of risk to benefit the industry and the public.

Managed by experienced actuaries and research experts from a broad range of industries, the SOA Research Institute creates, funds, develops and distributes research to elevate actuaries as leaders in measuring and managing risk. These efforts include studies, essay collections, webcasts, research papers, survey reports, and original research on topics impacting society.

Harnessing its peer-reviewed research, leading-edge technologies, new data tools and innovative practices, the Institute seeks to understand the underlying causes of risk and the possible outcomes. The Institute develops objective research spanning a variety of topics with its [strategic research programs](#): aging and retirement; actuarial innovation and technology; mortality and longevity; diversity, equity and inclusion; health care cost trends; and catastrophe and climate risk. The Institute has a large volume of [topical research available](#), including an expanding collection of international and market-specific research, experience studies, models and timely research.

Society of Actuaries Research Institute
475 N. Martingale Road, Suite 600
Schaumburg, Illinois 60173
www.SOA.org