Auto Loss Cost Drivers: Physical Damage

August 2018

Auto Loss Cost Drivers: Physical Damage

August 2018

Examining the drivers of collision and comprehensive frequency and severity.

In the latter half of 2013, personal auto insurance carriers began to notice an uptick in property damage liability and collision frequency. This marked the beginning of a new increasing frequency trend bucking over 25 years of falling crash rates. While the period of falling frequency preceding this increase was largely attributable to safety awareness, technology, and enforcement, explanations supporting increasing frequency needed further investigation.

In response, industry partners banded together to analyze these trends. Using publicly available data from the Federal Highway Administration, Bureau of Labor Statistics, the Census Bureau, and other sources, an analysis group is searching for explanatory variables. This paper represents some of their findings in collision and comprehensive frequency and severity. The group's goal is to provide an analytical basis for discussing and understanding auto insurance loss cost drivers that ultimately affect premiums. Please note that the collision frequency section of the report was first published in January 2018 under the title "Auto Loss Cost Trends Report."

The key findings from this work are:

- Congestion (through drivers per lane mile or average commute times) is positively related to collision frequency and comprehensive severity.
- Collision severity is largely driven by economic factors.
- Comprehensive frequency is impacted by the seasonal pattern of hailstorms and is generally higher in states with a zero-deductible windshield policy.
- Comprehensive severity is also impacted by natural disasters (like hurricane Sandy in our dataset).

Data Description

We mainly focus on two sources of data, the auto claims data and the economic variables which we will try to relate to them. We have quarterly frequency and severity data for each state (excluding DC and HI) from Q4 2011 through Q4 2015 from the FAST TRACK PLUS™ database. Additionally, we have the following state-level explanatory variables.

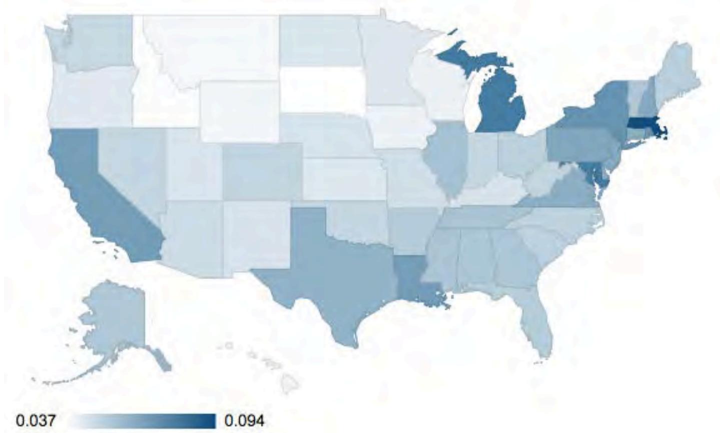
- UrbanVMTPercent: Percent of the vehicle miles traveled (VMT) in an urban area.
- LawyersPer1MillionCapita: Number of lawyers in the state per 1 million people.
- UrbanAvgCommuteTime: Average commute time in minutes for people in urban areas.
- RuralAvgCommuteTime: Average commute time in minutes for people in rural areas.
- MobileBroadbandPercent: Percent of population with access to mobile broadband
- InterstateGood: Percent of interstate miles rated as good
- DriversUnder20Percent: Percent of drivers under age 20
- DriversOver75Percent: Percent of drivers over age 75
- CommutePrivateVehiclePercent: Percent of people who commute by private vehicle
- AverageQuarterlyPrecipitation: Average quarterly precipitation in inches.
- BLSUnemployment: Unemployment rate from Bureau of Labor Statistics (BLS)
- UrbanVMTperLane: Urban vehicle miles traveled per urban lane mile.
- RuralVMTperLane: Rural vehicle miles traveled per rural lane mile.

- CapitalOutlayperVMT: Total transportation dollars spent on capital projects, per vehicle miles travelled.
- MaintenanceExpensesperVMT: Total transportation dollars spent on maintenance expenses, per VMT.
- PolicingExpensesperVMT: Total transportation dollars spent on policing expenses, per VMT.
- DUIs: Total DUIs per driver
- GasPricevsWage: Average gas price in dollars divided by average hourly wage in dollars.
- TortSystem: No-fault, optional no-fault, tort
- LicensedDrivers: Number of licensed drivers in the state.
- LaneMilesTotal: Total number of lane miles in the state.
- DriversperLaneMile: LicensedDrivers/LaneMilesTotal

Collision Frequency

National private passenger auto collision frequency rates from 4Q 2011 through 4Q 2015 are shown in Figure 1.[1] Massachusetts, Michigan, Maryland, and New York have the highest average collision frequency, while South Dakota, Idaho, Wyoming, and Montana have the lowest.

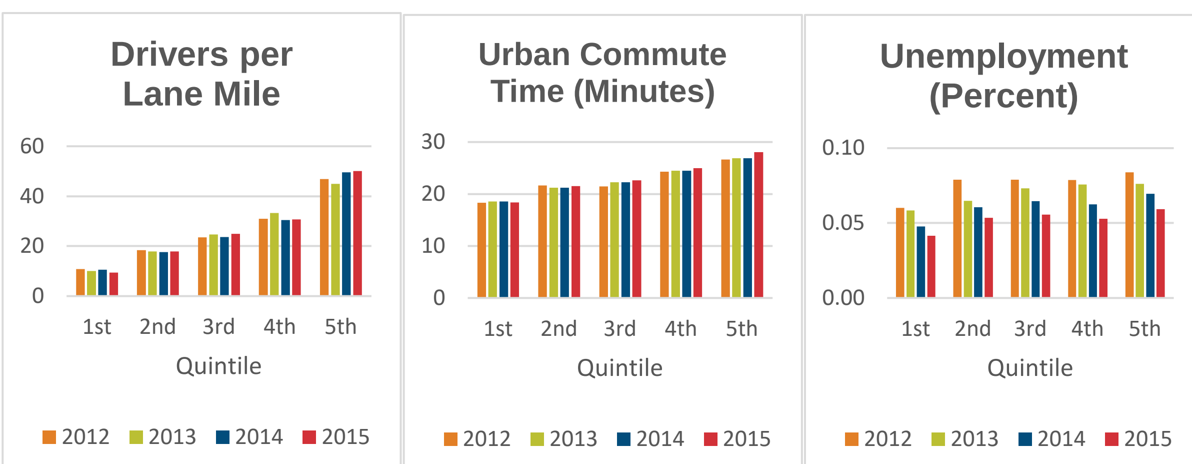
Figure 1
Average Collision Frequency (claims per car year), 4Q 2011 – 4Q 2015



To analyze collision frequencies, states were divided into quintiles each year based on their average collision frequency. Quintile averages were then plotted against a set of automotive and financial variables. A description of all the variables in the analysis is available at the end of this report. Sample plots are shown in Figure 2.

Drivers per lane mile has a strong positive relationship with collision frequency, and the relationship is very consistent across years. The same is true about many of the other congestion variables (rural/urban vehicle miles traveled (VMT) per lane, rural/urban commute time, number of licensed drivers, etc.).

Figure 2
Histograms of Factors Grouped by Quintile



[1] Private Passenger Auto Paid Collision Claim Frequency, FAST TRACK PLUS™, 06 October 2017

Other variables which are driven by national trends, like unemployment, don't show much of a pattern other than between the years within the quintiles.

Many of the variables which appear to be strongly correlated with collision frequency are related to congestion. To distinguish between the variables and find the ones which best predict collision frequency, a random forest was constructed to compare the importance of each variable to the model. Variables with more importance have the best predictive ability. As seen in Figure 3, five variables stand out: Drivers per Lane Mile, Urban Average Commute Time, Rural Average Commute Time, System, and Urban VMT.

Another way to examine the impact of a single variable is an added variable plot. These plots show the relationship between collision frequency and another variable after accounting for the impact of all the other variables in the model. The actual values don't matter as much as the general pattern. When looking at the added variable plot we see that the pattern is consistent for both commute time variables and for drivers per lane mile. Urban VMT became slightly negatively related, though the relationship is not terribly strong.

Figure 3
Variable Importance

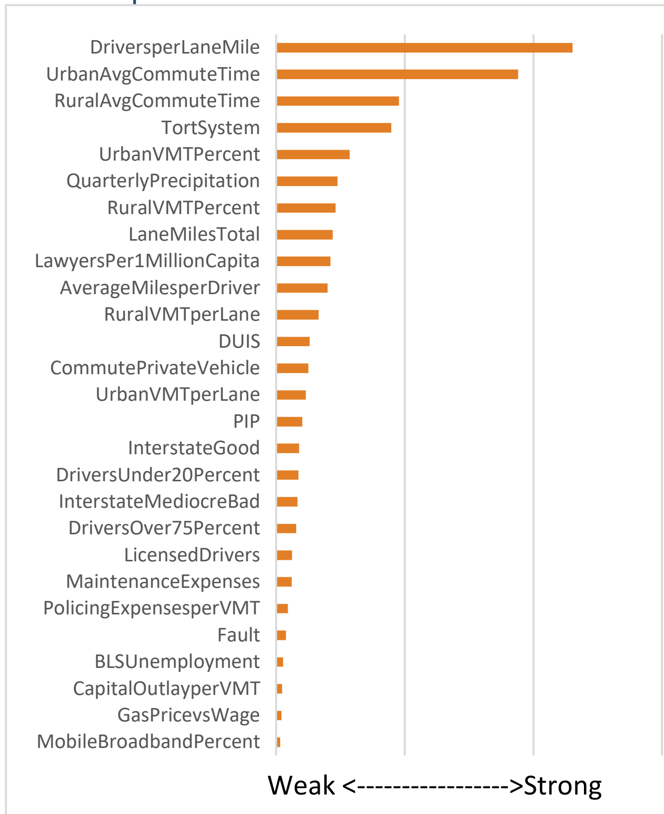
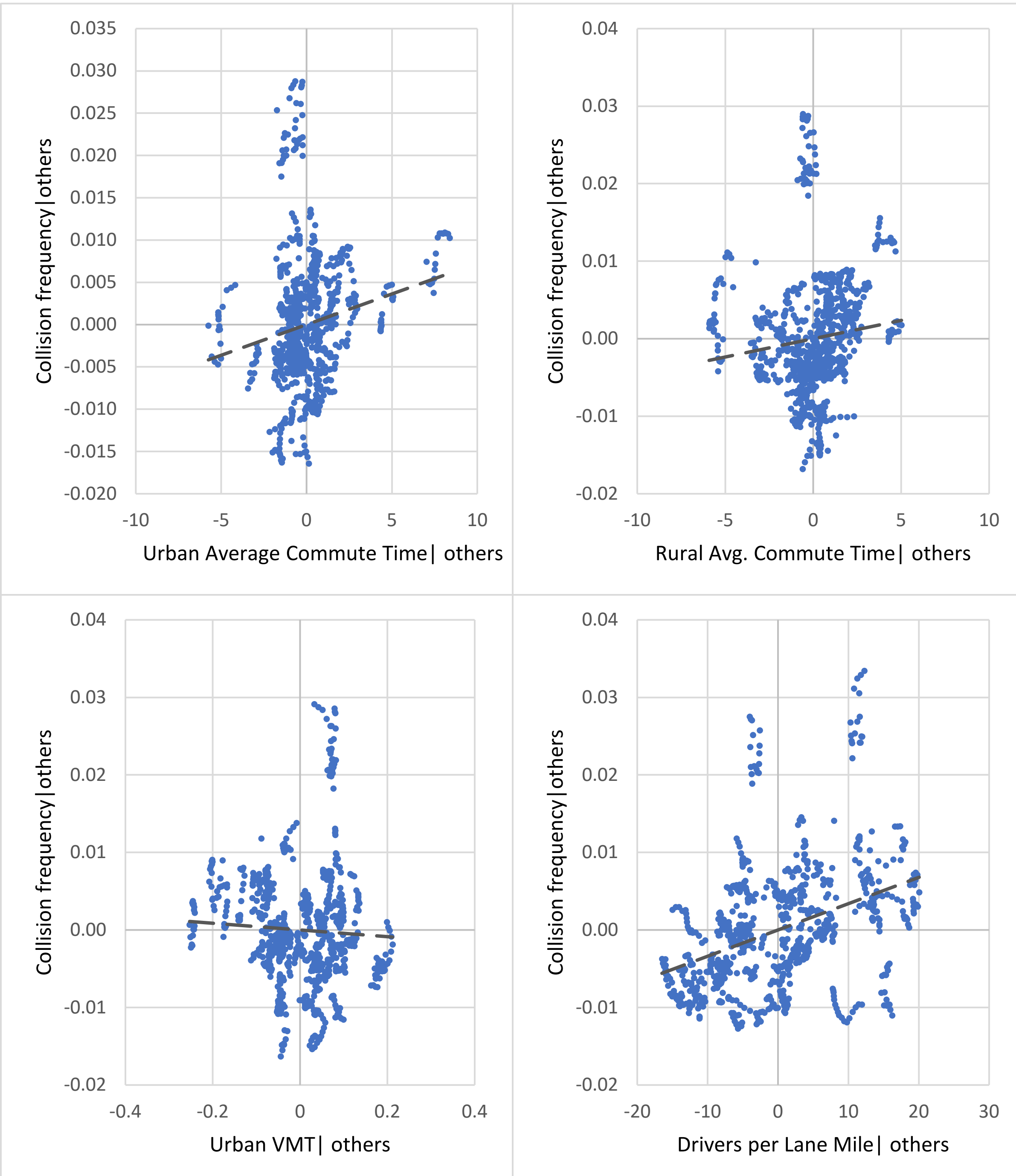
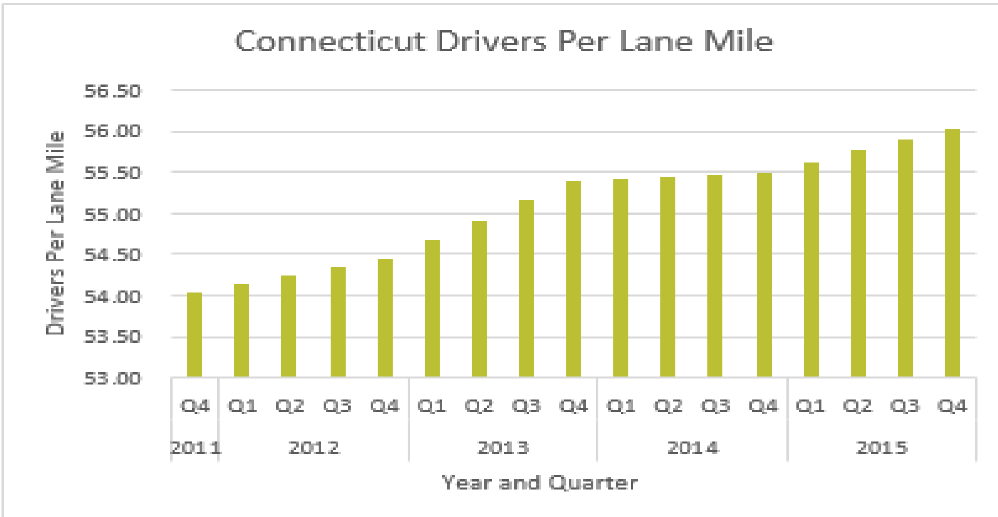


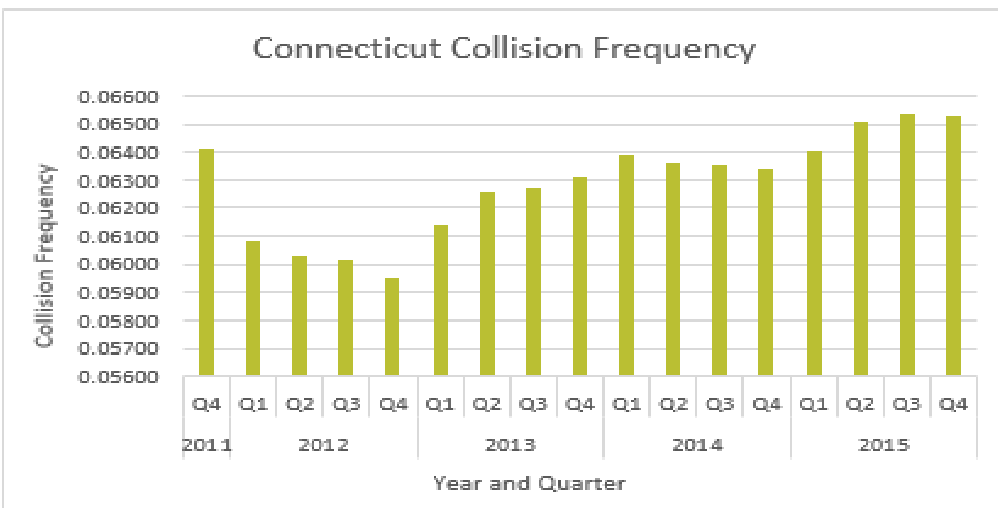
Figure 4
Added Variable Plots



Similar to the relationships seen in Figure 4, the state-specific relationships with collision frequency are relatively strong for all the congestion variables, especially drivers per lane mile. One notable exception is Connecticut, shown in Figure 5. It seems to follow the pattern relatively well except for 4Q 2011. During that quarter, a major storm caused excessive snowfall throughout the state.

Figure 5
Connecticut collision frequency and congestion





Other Interesting Relationships to be Further Explored

There are a few interesting relationships that we are still investigating. As we continue this analysis to more coverages and more years, we hope to gain a better understanding.

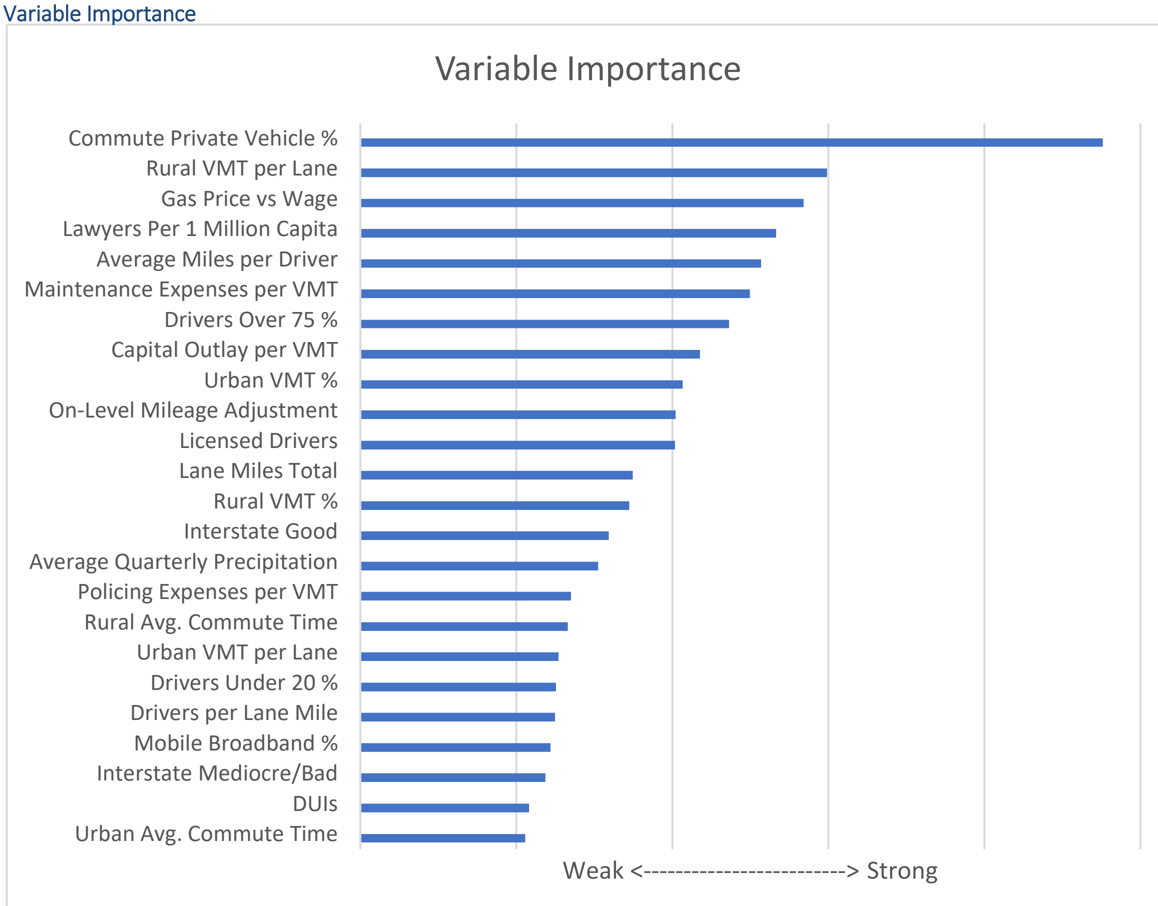
- DUIs appear to be negatively related to collision frequency, even after accounting for a few outlying states.
- Mobile broadband access (used as a proxy the likelihood that a driver may have a mobile device while driving) appears to have no impact on collision frequency. With all the current press around distracted driving, this was surprising (if it is a good proxy). Likely, we need to find a better proxy for distracted driving.
- The system (no-fault vs. tort) doesn't appear to impact the expected collision frequency, but has a big impact on the variance of the frequency.
- Both CA and WY increase in collision frequency, but decrease in drivers per lane mile each quarter.

Collision Severity

Important Variables

A random forest model was fit to determine which explanatory variables had the most predictive ability for collision severity. Variables that have the highest importance in a random forest model have the best predictive ability. A plot displaying the relative importance of each of the variables is shown in Figure 6. The top five explanatory variables include percentage of people commuting with private vehicles, rural vehicle miles traveled per lane, gas price vs. wage, average miles per driver, and percentage of vehicle miles traveled that are urban. Some other notable variables include percentage of drivers over 75, lawyers per 1 million capita, and maintenance expenses per vehicle mile traveled. Most of these variables are related to the wealth of an area, such as gas price vs. wage and percentage of private vehicles. One theory is that with more wealth in an area, vehicles tend to be more expensive, and thus the average collision claim costs more.

Figure 6
Variable Importance



Percentage of Commuters with Private Vehicles

The percentage of commuters with private vehicles in each state was a particularly important explanatory variable. However, when examining the plot of private vehicle percentages by collision severity quintile, a trend does not seem to be so obvious, which may mean that percentage of commuters with private vehicles could be a useful intermediary explanatory variable when coupled with others. New York's percentage of commuters with private vehicles is much lower than all other states. Almost half of the state population lives in New York City where many households do not own a car. Therefore, private vehicle percentage may largely be a surrogate, identifying New York. The added-variable plot of the model with respect to the percentage of commuters with private vehicles (excluding New York) is shown in Figure 7.

Figure 7

Private Vehicle Added Variable Plot

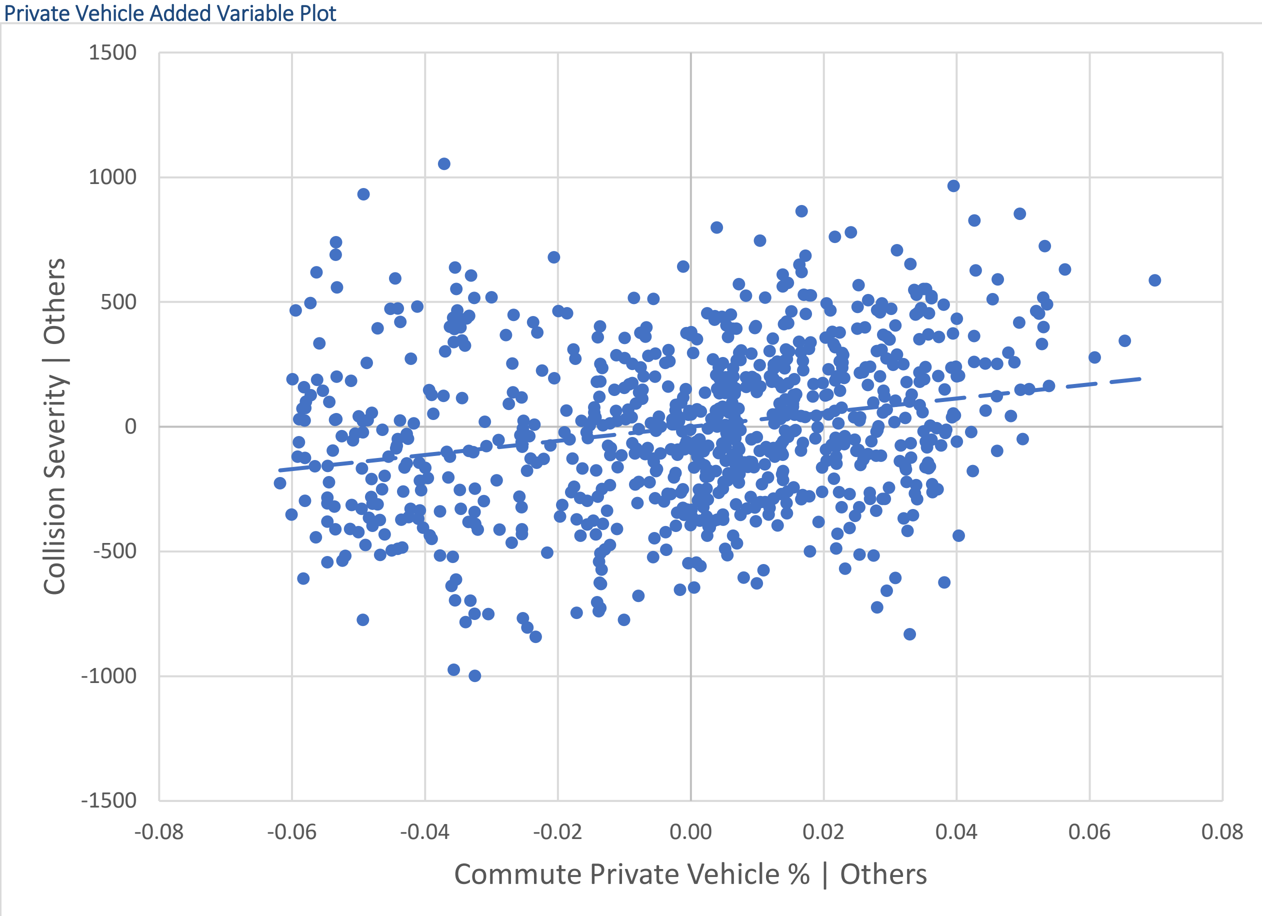


Figure 8
Rural VMT per lane quintile plot

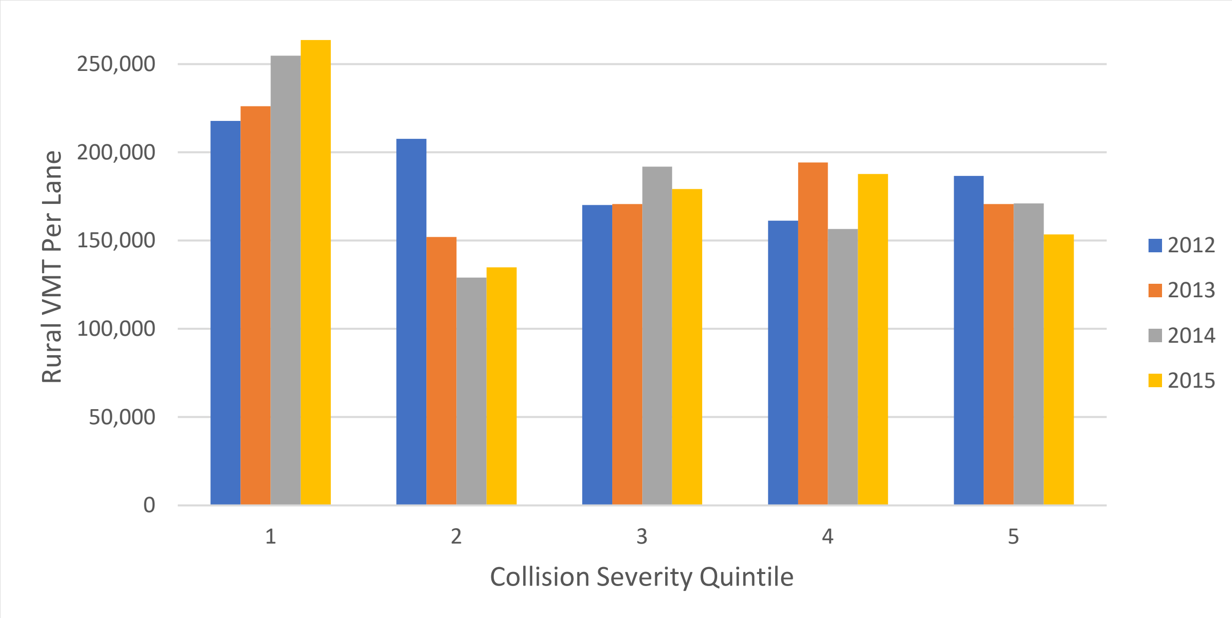
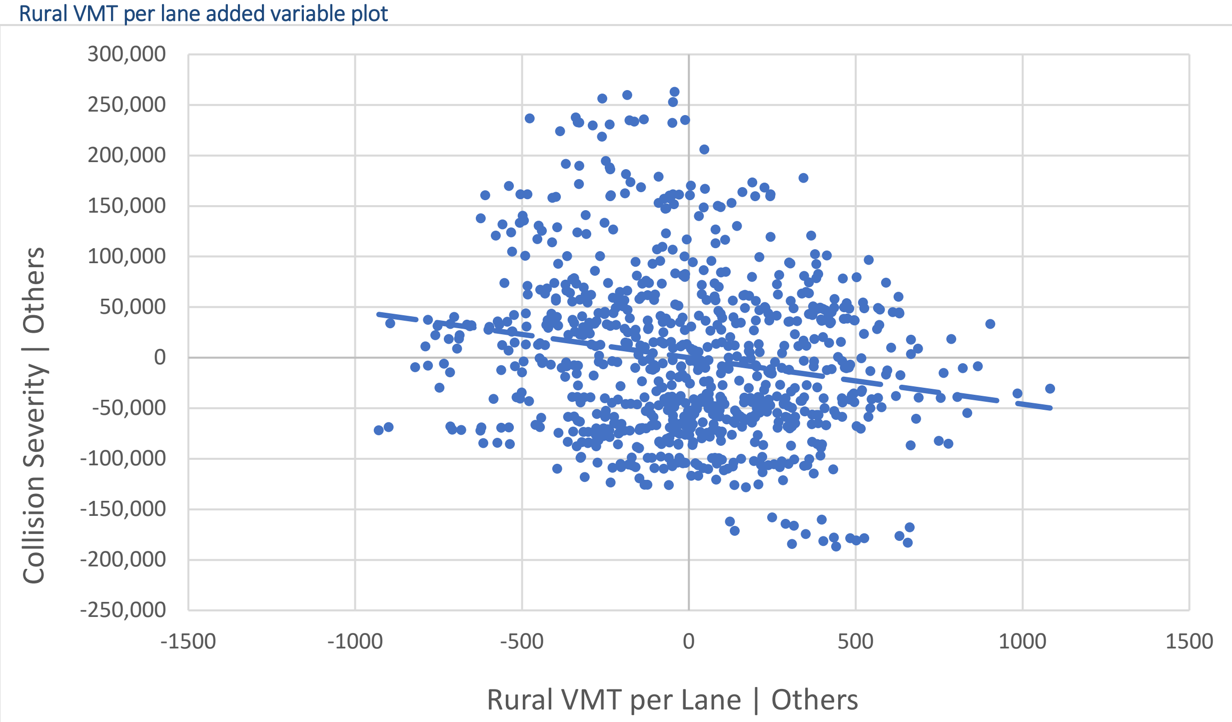


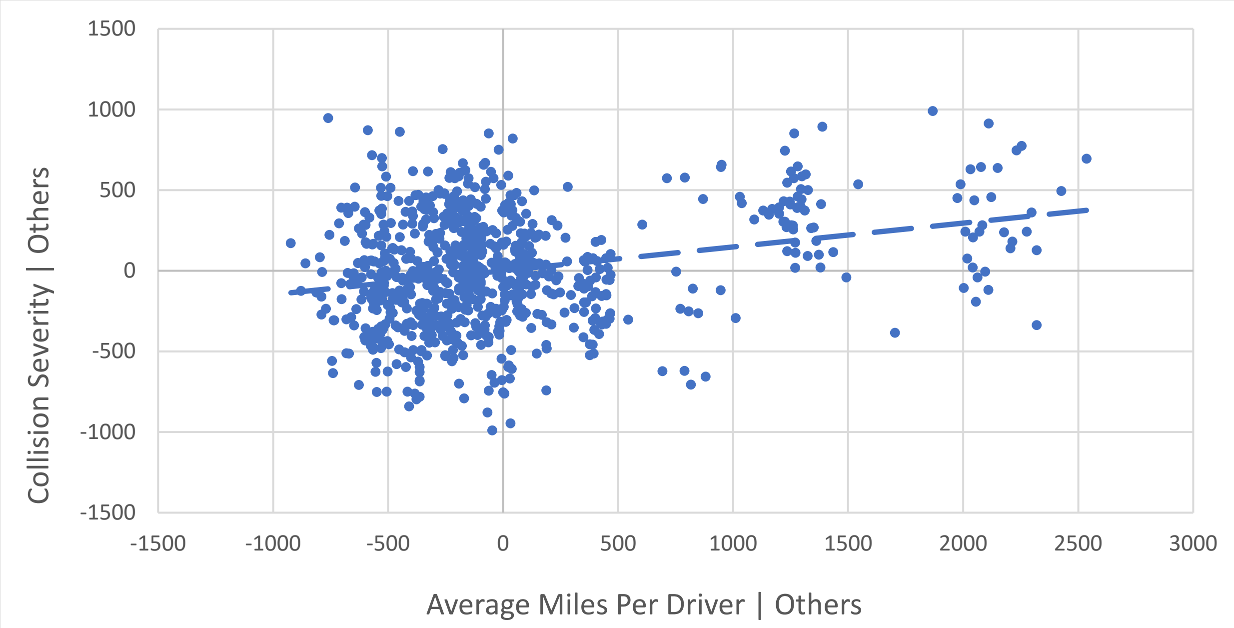
Figure 9
Rural VMT per lane added variable plot



Rural Vehicle Miles Traveled per Lane

Collision severity increases as rural vehicle miles traveled per lane decreases. This may be because driving speeds can be higher on rural roads that are less congested, leading to more expensive collisions. The quintile plot (Figure 8) shows that in every year the highest rural vehicle miles traveled per lane occurred in the states with the areas with the lowest collision severity. The added variable plot in Figure 9 displays this inverse trend clearly.

Figure 10
Average miles per driver added variable plot



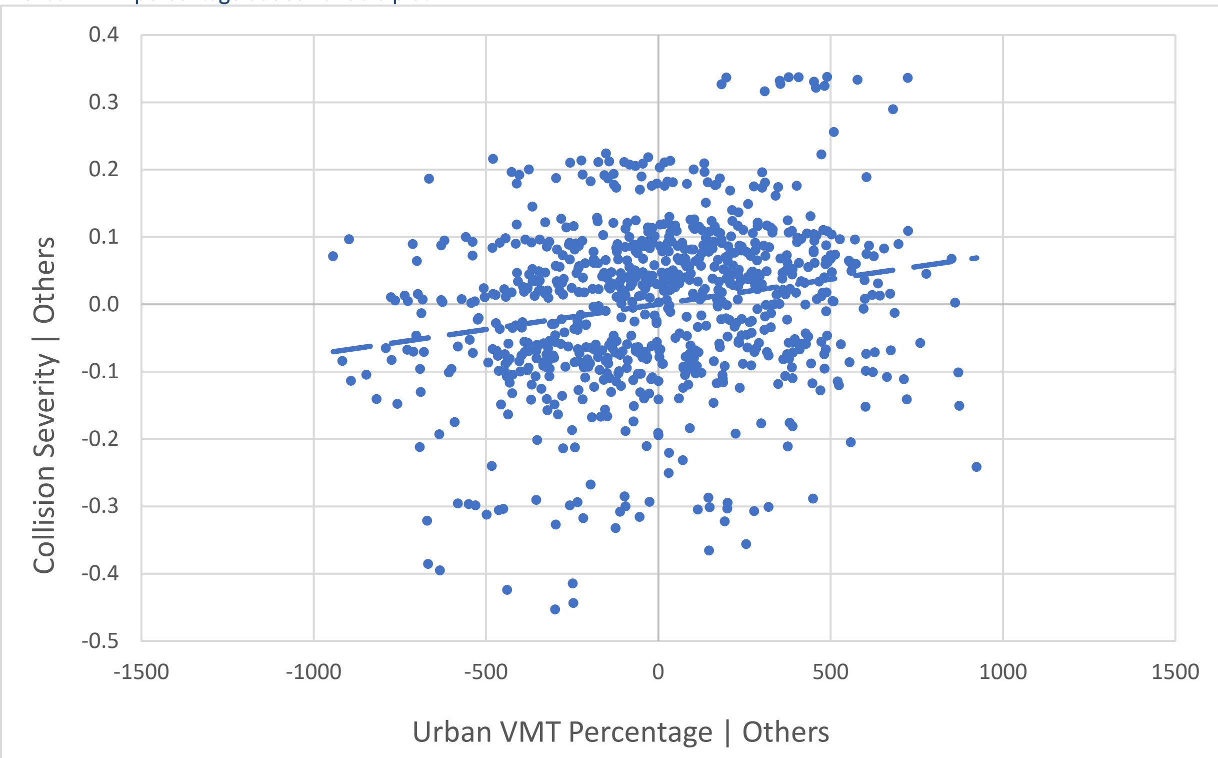
Average Miles per Driver

Average miles per driver is positively correlated with collision claim severity (see Figure 10). This trend may have a similar explanation to rural vehicle miles traveled per lane. If cars can drive more miles, they are probably in rural areas, going much faster with less traffic. If the cars are going faster, they may be also having more severe accidents when they occur. There are a number of states (Alaska, Wyoming, Oklahoma, and Mississippi) outside the domain of the rest of the states. Even when excluding these states, there is still a positive effect of average miles per driver on collision claim severity.

Percentage of Vehicle Miles Traveled that are Urban (vs. Rural)

The higher the percentage of vehicle miles traveled that are urban in a state, the higher the collision claim severity will likely be. While this seems counterintuitive given the results in the previous sections, it is largely driven by Nevada. It accounts for all of the points in the top right corner in Figure 11. The majority of the driving in the state occurs in and around Las Vegas. Nevada has the 4th highest percentage of accidents in urban zones.

Figure 11
Urban VMT percentage added variable plot



An Additional Seasonal Trend

A clear trend (shown by Figures 12 and 13) in states where weather varies dramatically by the season, collision claim severity fluctuates with each quarter. For example, California is much more stable in each quarter than states like Montana or Alaska, where the claim severity increases dramatically during the last and first quarters of every year.

Conclusion

In summary, the economic indicators are the most influential in predicting the severity of collision claims. The seasonality of the states also has a large impact on the quarterly data, as shown in Figure 13. Some important states to consider include Nevada due to its urban collision percentage, Alaska because of the low average miles per driver, and New York with a low percentage of commuters with private vehicles.

Figure 12
Collision claim frequency (claims per car year) in western states

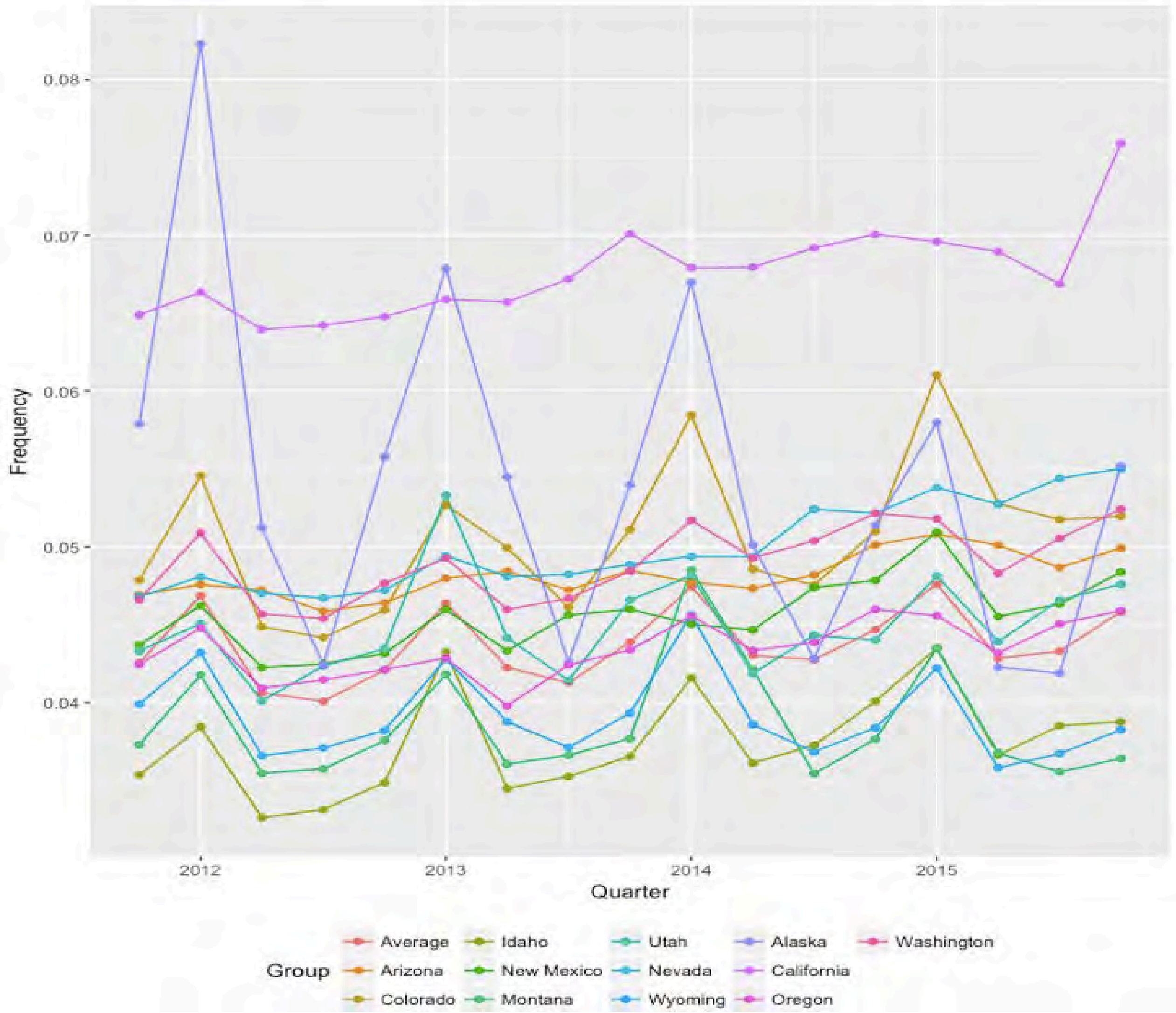
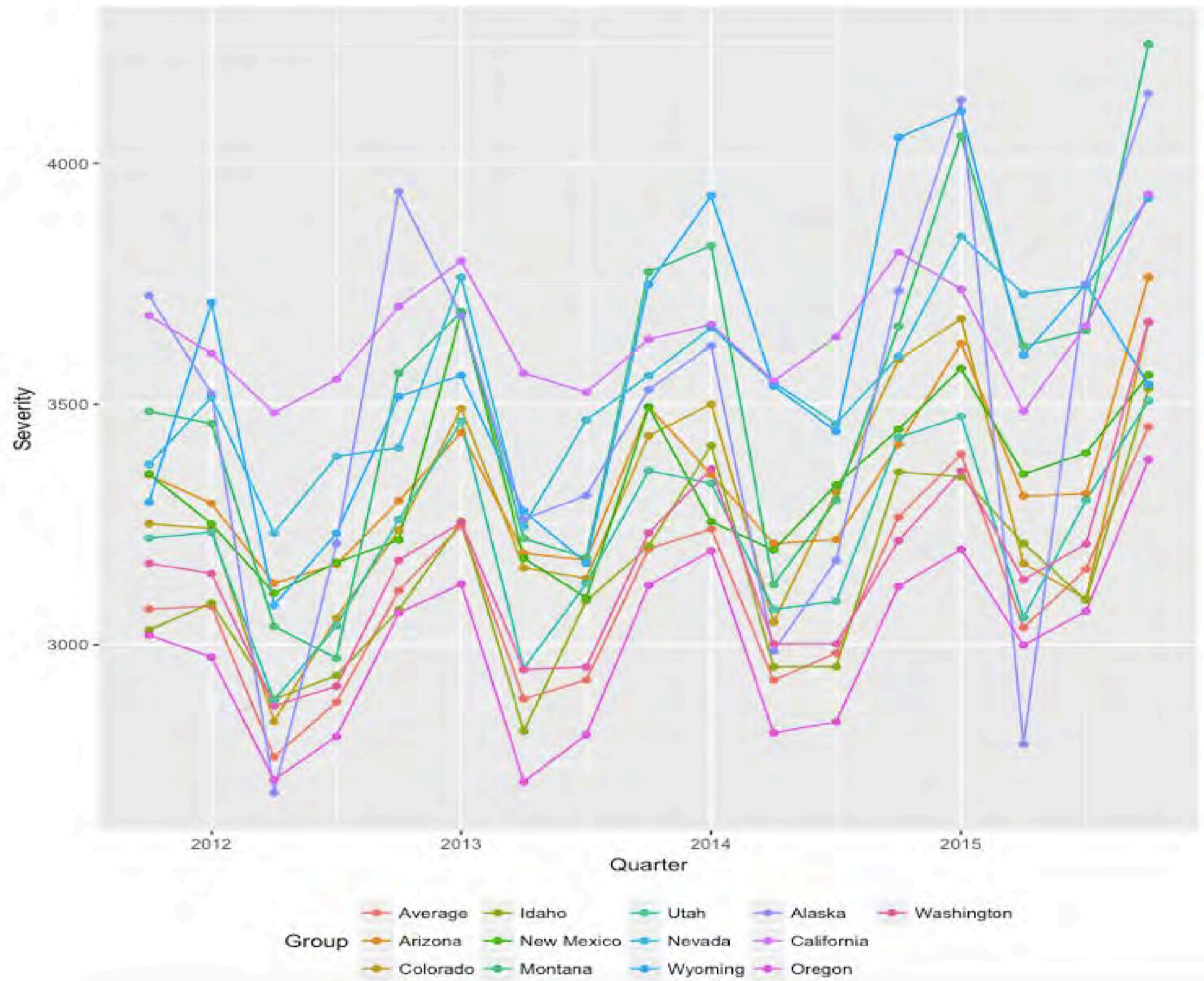


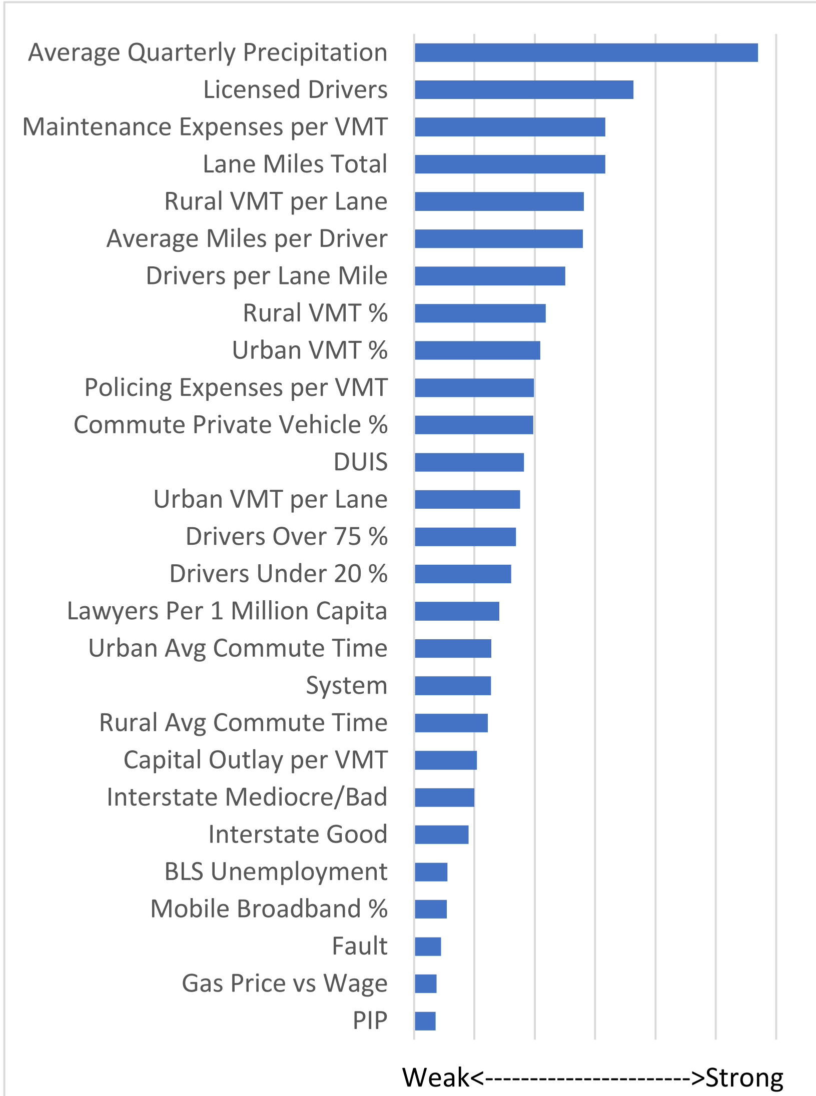
Figure 13
Collision claim severity (dollars per claim) in western states



Comprehensive Frequency

When modeling comprehensive frequency, the variable importance chart (Figure 14) shows that the average quarterly precipitation is the most important explanatory variable.

Figure 14
Variable importance plot



Precipitation

It makes sense that the precipitation would be related to comprehensive frequency. More precipitation means more chance for damaging precipitation (hail, ice, etc.) But, we observe from the quintile plots (Figure 15) that average quarterly precipitation is negatively related to comprehensive frequency. We see a similar negative trend when looking at the quintile plots for licensed drivers (Figure 16).

Both negative trends seem to be due to spatial variation rather than temporal trends. States in the mid-west and southwest have lower average precipitation and higher comprehensive frequency. Likewise, there are proportionally fewer licensed drivers in the mid-west (a higher comprehensive frequency area) than on either coast. Figure 17 shows the average comprehensive frequency and Figure 18 shows the average precipitation by state.

Figure 15
Quarterly precipitation quintile plot

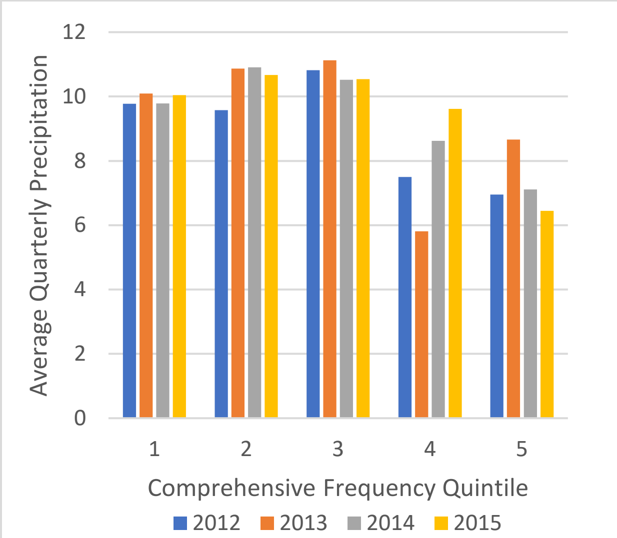


Figure 16
Licensed drivers quintile plot

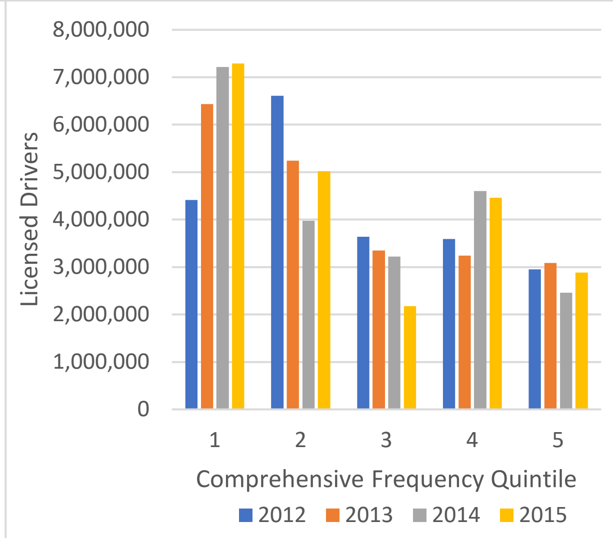


Figure 17
Comprehensive frequency (claims per car year)

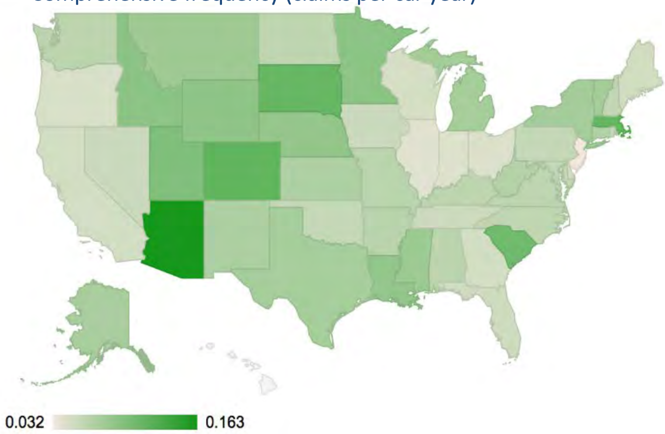
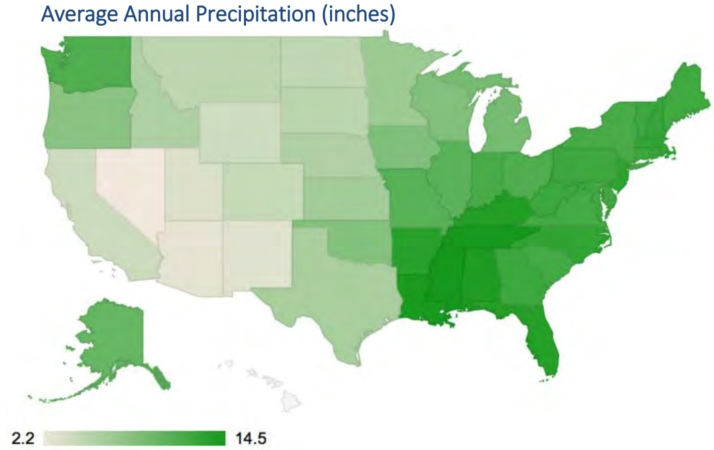


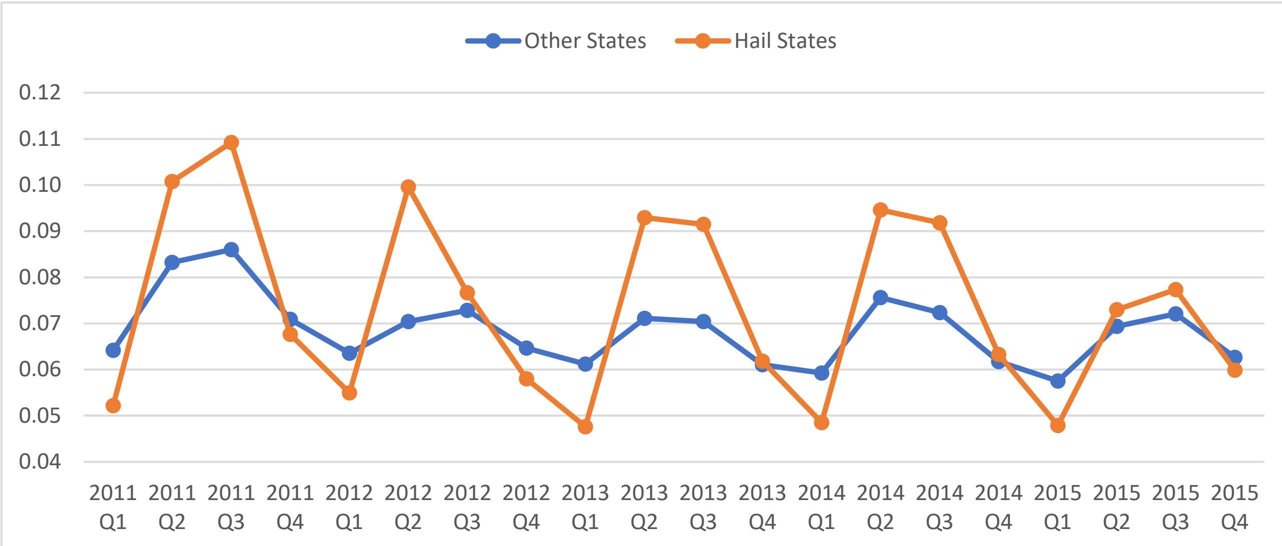
Figure 18
Average Annual Precipitation (inches)



Hail

Hail storms and comprehensive frequency follow similar seasonal trends. Using data on hail from NOAA, the ten states with the most hail are Texas, Kansas, Nebraska, Oklahoma, South Dakota, Missouri, Iowa, North Carolina, Colorado, and Illinois. The comprehensive frequency in those ten states has a stronger seasonal trend than the other states (Figure 19). Most severe hailstorms occur in the second and third quarters. The comprehensive frequency trend in those ten states is also significantly higher in the second and third quarters than in the first or fourth.

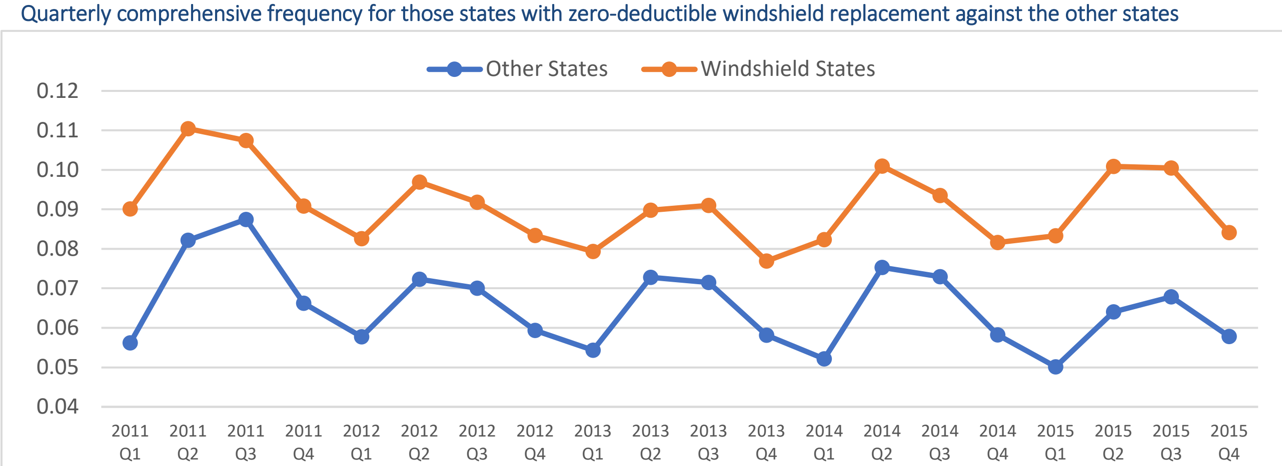
Figure 19
Quarterly comprehensive frequency for the ten highest hail states against the other states



Zero-Deductible Windshield Replacement

Eight states mandate zero-deductible windshield replacement (Arizona, Connecticut, Florida, Kentucky, Massachusetts, Minnesota, New York, and South Carolina) with conditions varying between each state. These states, on average, have higher comprehensive frequency than states without the requirement (Figure 20). Much of the difference can be attributed to Arizona (the state with the highest comprehensive frequency), but even after removing Arizona the zero-deductible states still have significantly higher comprehensive frequency.

Figure 20
Quarterly comprehensive frequency for those states with zero-deductible windshield replacement against the other states



Conclusion

The seasonal pattern in comprehensive frequency seems to be related to a similar seasonal pattern in the prevalence of hailstorms. States with mandated zero-deductible windshield replacement also seem to have higher comprehensive frequency. Many of the variables we use to analyze the other coverages do not appear to be significantly related to comprehensive frequency.

Comprehensive Severity

In extreme cases, the severity of comprehensive claims will be driven by natural disasters. One disaster within the range of this data is Hurricane Sandy, which hit New York and New Jersey in the fourth quarter of 2012. Comprehensive severity in those quarters is far above any other observations. Therefore, our analysis will proceed without including these observations, as they give unnecessary leverage to specific explanatory variables.

The variable importance plot (Figure 21) shows us that the number of drivers per lane mile and the average miles per driver are the most important explanatory variables. The following quintile plots (Figures 22 and 23) show that drivers per lane mile demonstrates a negative correlation with comprehensive severity, and the average miles per driver displays a positive correlation.

Figure 21
Comprehensive severity variable importance plot

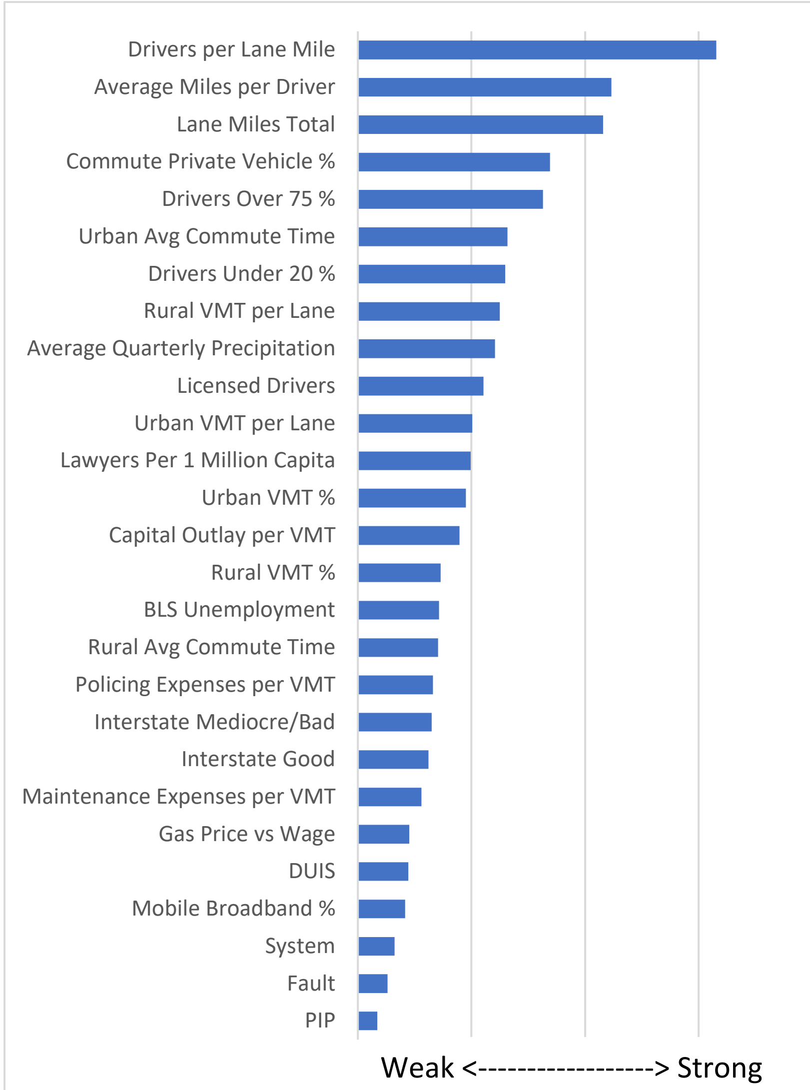


Figure 22
Drivers per lane mile quintile plot

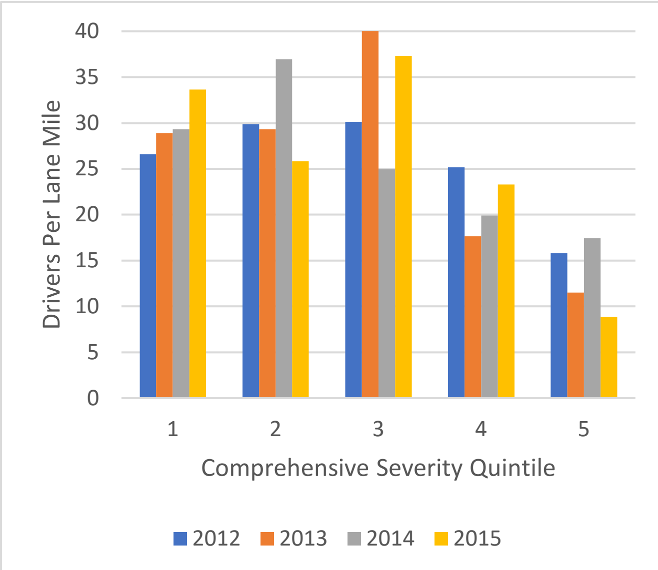
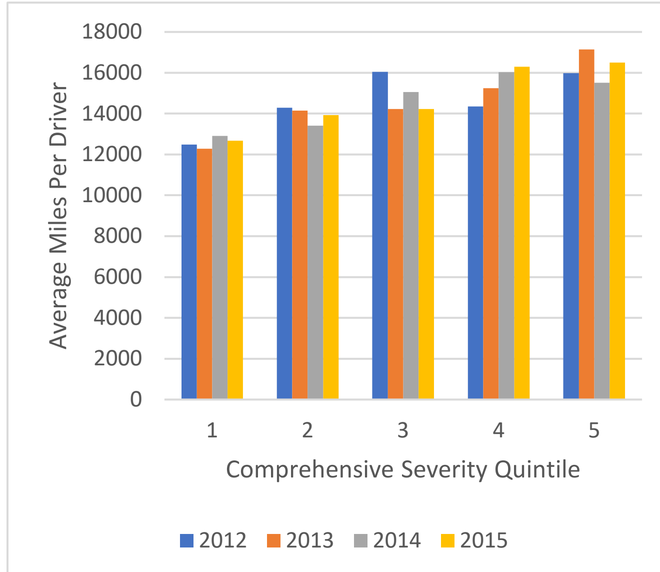


Figure 23
Average miles per driver quintile plot

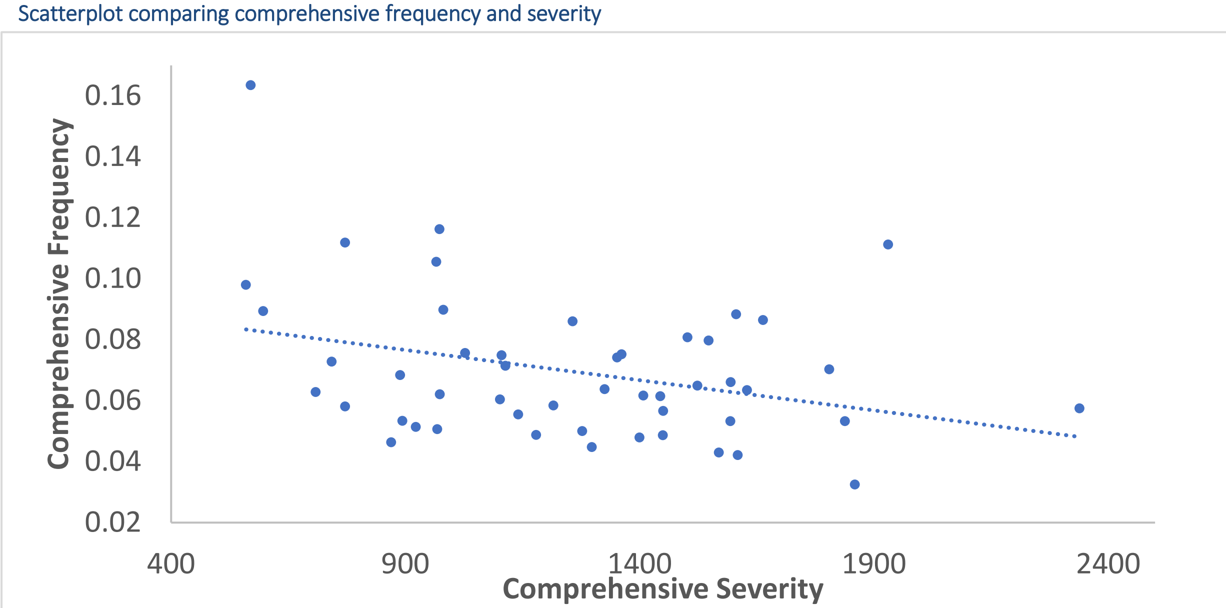


Relationship with Comprehensive Frequency

Figure 19 demonstrates how on an average state-by-state basis, comprehensive frequency and severity are negatively correlated. Arizona is an extreme example of this – it has the highest average comprehensive frequency, and the second lowest comprehensive severity. This also makes sense with more (relatively low cost) windshield claims in those states because of the zero-deductible requirement. It is likely the case that factors such as this drive the negative correlation between frequency and severity.

Additionally, frequency tends to decrease over this time and severity tends to increase. Severity increases largely due to inflation while the decrease in frequency could be due to improved safety features preventing animal collisions and theft.

Figure 24
Scatterplot comparing comprehensive frequency and severity

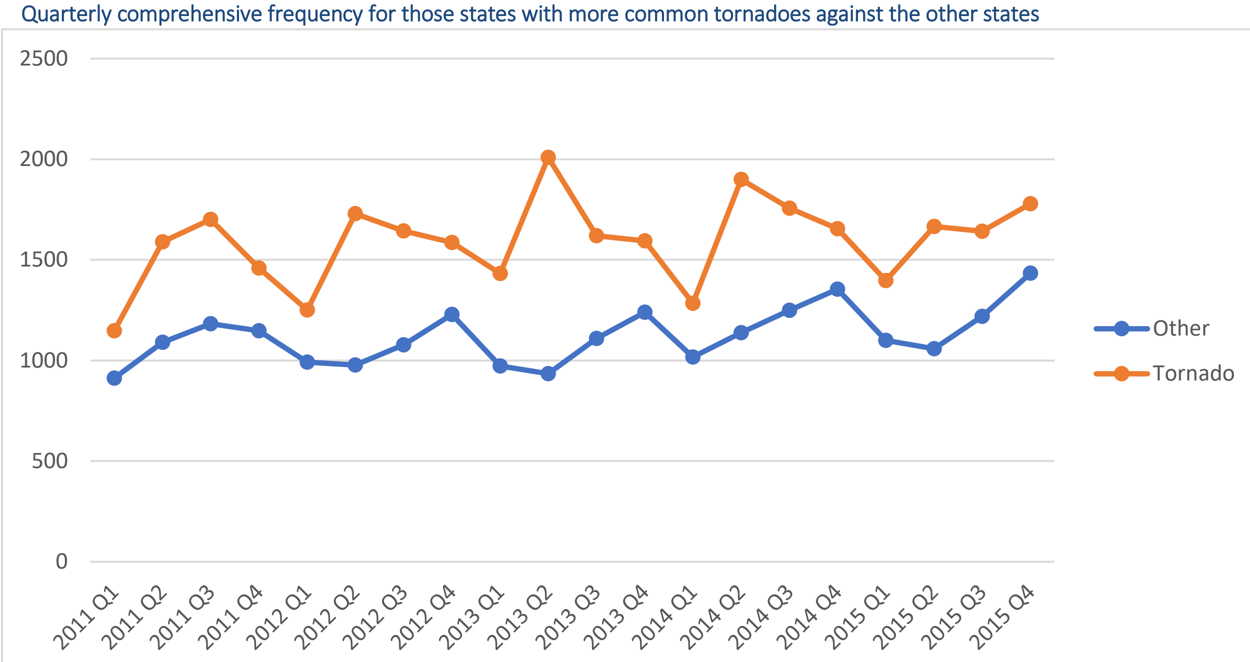


Tornadoes

Of the possible causes of severe comprehensive claims, tornadoes could have a large influence. States that rank high in tornado occurrences include Delaware, Florida, Illinois, Indiana, Iowa, Kansas, Louisiana, Mississippi, Missouri, Nebraska, Oklahoma, South Dakota, and Texas.

Figure 25 shows the difference between those 13 tornado states and states with fewer tornadoes. Oklahoma has the highest comprehensive severity and is also one of the highest tornado states. Even after removing Oklahoma, the tornado states have higher average severity than the other states.
Like the trend of hail, tornadoes follow a seasonal trend. Tornadoes seem to spike during the second quarter. This follows the severity pattern in the tornado states in Figure 20. Notice that the quarterly pattern is not as

Figure 25
Quarterly comprehensive frequency for those states with more common tornadoes against the other states



pronounced in the non-tornado states. To compare the spatial distribution, Figure 26 plots the comprehensive severity while Figure 27 shows the tornado count. The pattern is similar to the one between hail and comprehensive frequency, though not as obvious.

Figure 26
Average comprehensive severity (dollars per claim)

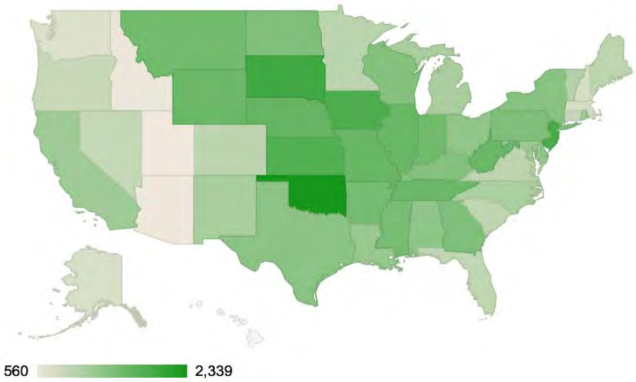
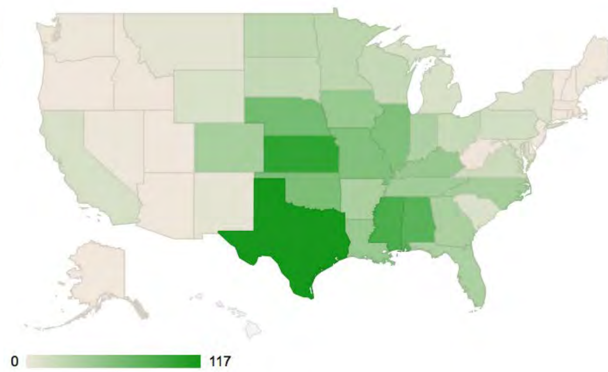


Figure 27
Tornado count (2011-2015)



Conclusion

Comprehensive severity is driven to extremes by natural disasters. Drivers per lane mile and average miles per driver also appear to be related to comprehensive severity.