Auto Loss Costs: Bodily Injury

January 2020

APCIA apci.org

CAS

SOCIETY OF ACTUARIES

Auto Loss Costs: Bodily Injury

January 2020

In both no-fault and tort states, rural and urban congestion measures had strong, positive relationships with loss cost, but only in tort states did the percent of licensed drivers that are male and over the age of 80 and percent of licensed drivers that are male and under 24 have a significant negative relationship with loss costs. Similarly, in both no-fault and tort states we see a strong positive relationship between rural congestion and frequency, but we also see that in both no-fault and tort states the percent of licensed drivers that are male and over 80 has a negative relationship with bodily injury frequency.

Introduction

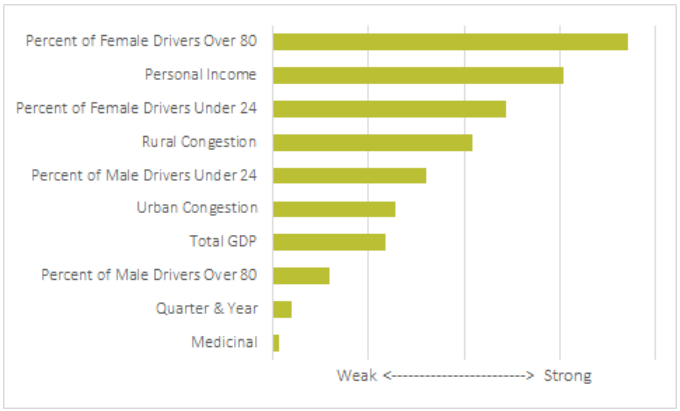
Bodily injury coverage pays for injuries and lost work time caused to another person while the insured person is operating a vehicle whether or not the injured person(s) are in a vehicle. In our analysis we look for the most important variables in predicting bodily injury claims using bodily injury frequency, severity, and loss cost data reported quarterly from 2010 through 2018.

One of the main factors that influences bodily injury claims is whether the state is a no-fault or tort state. In no-fault states, the injured person's insurance pays for damages regardless of who is at fault, while in a tort state the insurance of the person who is determined to be at fault pays for the claim. Because of the large influence these two different state laws has on bodily injury, we started our analysis by separating our data into no-fault states and tort states. After initial examination of the data, we selected a subset of the variables to examine based on whether they had consistent data or were overly similar to another variable.

Variables in each analysis:

Percent Licensed Male Drivers Under 24 (Male Drivers Under 24 / Total Male Drivers)
Percent Licensed Female Drivers Under 24
Percent Licensed Male Drivers Over 80
Percent Licensed Female Drivers Over 80
Urban Congestion (Urban Vehicle Miles Traveled in Millions / Urban Road Miles)
Rural Congestion (Rural Vehicle Miles Traveled in Millions / Rural Road Miles)
Total GDP
Personal Income
Medicinal Marijuana Legal Status

Figure 1



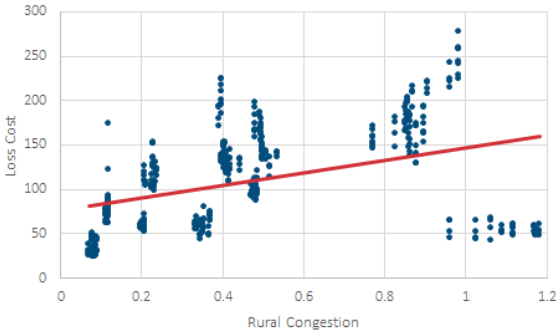
No Fault States

No fault states include Florida, Hawaii, Kansas, Kentucky, Massachusetts, Michigan, Minnesota, New Jersey, New York, North Dakota, Pennsylvania, and Utah.

Since loss cost is a combination of frequency of claims and severity of claims, we found it more important to emphasize the significant factors in bodily injury loss cost. In our random forest model, which we used to get a preliminary idea of the most influential factors on bodily injury claims data, we found that the percent of female drivers over the age of 80, percent of female drivers under the age of 24, personal income, and rural congestion were the top factors (Figure 1).

We then performed variable selection using two different metrics of penalized model fit: Akaike Information Criterion (AIC) and Bayesian Information Criterion (BIC). Both AIC and BIC showed that rural congestion was significant in determining bodily injury loss cost, having a strong, positive relationship. The scatter plot of loss cost by rural congestion (Figure 2) shows this trend. It is interesting to note that the data points in the bottom right hand corner are all from Hawaii. This could be because the number of road miles is so low compared to other states, the congestion measure is much higher. Surprisingly, urban congestion was also extremely significant, but with a strong, negative effect on loss cost.

Figure 2
Loss Cost by Rural Congestion



Frequency was mostly driven by urban congestion, which had a negative effect. Again, there was a positive correlation between frequency and rural congestion, although not as significant. The second most important factor in modeling bodily injury frequency was the percent of drivers that are female and over the age of 80 or under the age of 24, which surprisingly had a negative effect. Added variable plots show the effect percent of female drivers over 80 (Figure 3) and under 24 (Figure 4) have on bodily injury claim frequency given all other variables are provided. We see that it has a negative effect, which is curious because male drivers over 80 and under 24 both had significant and positive effects. Both of these same observations about male and female drivers were made in the loss cost models.

Figure 3
Percent of Female Drivers Over 80 Added-Variable Plot

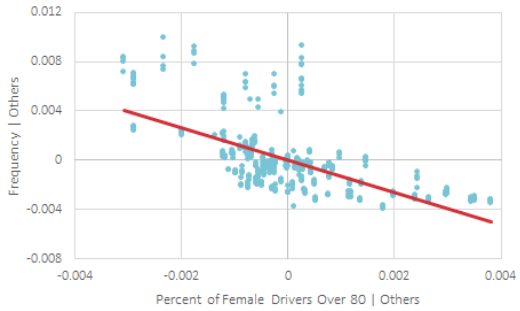
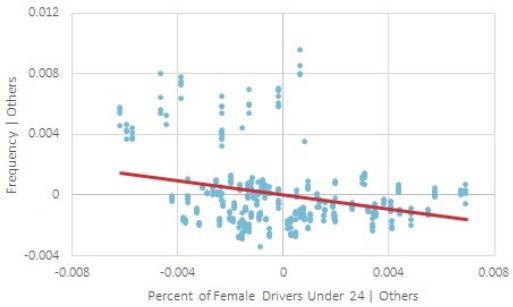


Figure 4
Percent of Female Drivers Under 24 Added-Variable Plot



The random forest model that focused on severity showed that the most important factors included both rural and urban congestion and percent of female drivers above 80. The AIC and BIC variable selections agreed, choosing urban congestion as the most important. This had a highly significant positive relationship with bodily injury severity, while rural congestion had a significantly negative effect on severity. Two graphs are provided to visualize this point, one showing the scatter plot of rural congestion plotted against severity (figure 5), and one showing the added variable plot displaying the negative relationship rural congestion has on severity (figure 6) when all the other variables are considered.

Figure 5
Severity by Rural Congestion

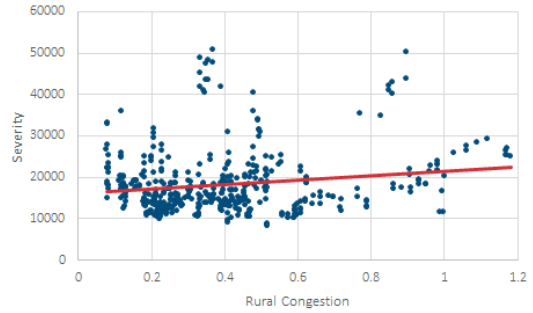


Figure 6
Rural Congestion Added-Variable Plot

In contrast to the random forest model, which highlighted the percent of female drivers over the age of 80, the variable selection techniques included percent of male drivers over the age of 80 as being highly significant and having a positive effect.

Tort States

Tort states include all the states that are not no-fault states. In finding the most important variables for the tort states, we followed a similar pattern to the no-fault states by emphasizing the loss cost factors in our variable selection models. In our random forest model, we found that urban and rural congestion measures were once again important, and while the model showed importance significantly dropped off after the urban congestion measure, total GDP was the next most important factor (Figure 7).

In the AIC- and BIC-based models we saw a strong, positive correlation between urban congestion and loss cost (Figure 8). Also, similar to the previous models, total GDP has a positive relationship with bodily injury loss cost in tort states (Figure 9). Rural congestion was not significant on loss cost in tort states.

Figure 7
Tort States Variable Importance

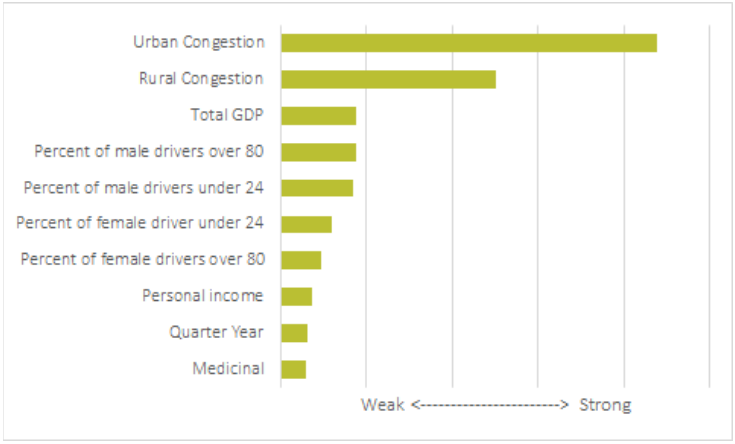


Figure 8
Loss Cost by Urban Congestion

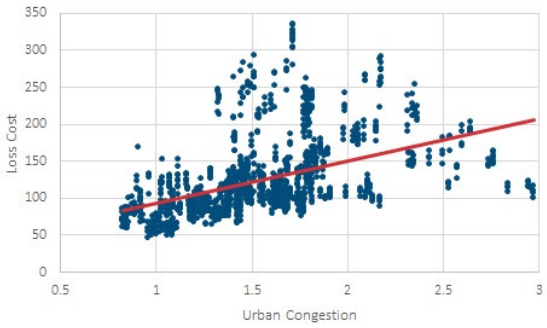
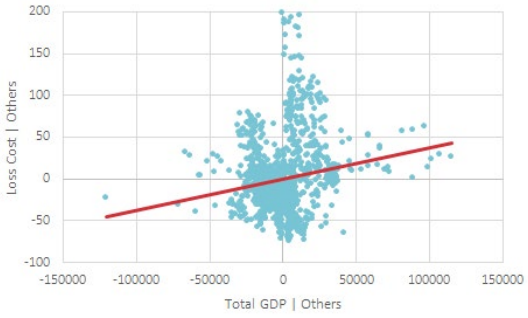


Figure 9
Total GDP Added-Variable Plot



Besides urban and rural congestion, the percent of licensed male drivers under 24 came out to be highly significant and had negative effects on bodily injury loss cost.

Bodily injury frequency had very similar results in tort states as was shown before in no fault states. The percent of male drivers over 80 had a significantly negative effect on frequency. Urban and rural congestion have highly positive relationships with bodily injury frequency.

In tort states, bodily injury severity has a large negative relationship with urban congestion. The percent of licensed drivers that are male and over the age of 80 also came out to be positively related with severity.

Conclusion

In both no-fault and tort states, rural and urban congestion measures had strong, positive relationships with loss cost, but only in tort states did the percent of licensed drivers that are male and over the age of 80 and percent of licensed drivers that are male and under 24 have a significant negative relationship with loss costs.

Similarly, in both no-fault and tort states we see a strong positive relationship between rural congestion and frequency, but we also see that in both no-fault and tort states the percent of licensed drivers that are male and over 80 has a negative relationship with bodily injury frequency.

Additional Tables

Parameter	Loss Cost Linear Model			
	Estimate	Std. Error	t value	Pr(>|t|)
(Intercept)	325.5	17.59	18.505	< 2e-16
Rural Congestion	54.87	11.91	4.607	0.0000054
Urban Congestion	-82.04	7.132	-11.504	< 2e-16
Percent of Female Drivers Over 80	-10220	677	-15.090	< 2e-16
Percent of Male Drivers Over 80	7892	948.3	8.323	1.2E-15
Decriminalized Reduced	-22.65	6.023	-3.76	0.000194
Decriminalized Yes	19.67	17.32	1.135	0.256847
Personal Income	3.357E-07	3.786E-08	8.867	< 2e-16
All Industry Total GDP	-1.913e-04	0.00003041	-6.291	7.88E-10

Parameter	Frequency Linear Model			
	Estimate	Std. Error	t value	Pr(>|t|)
(Intercept)	0.0245643	0.0028454	8.633	< 2e-16
Rural Congestion	0.0053122	0.0010114	5.252	0.000000239
Urban Congestion	-0.0086716	0.0006129	-14.148	< 2e-16
Percent of Female Drivers Under 24	-0.2315962	0.037928	-6.106	2.32E-09
Percent of Female Drivers Over 80	-1.3229699	0.0834389	-15.856	< 2e-16
Percent of Male Drivers Under 24	0.236457	0.0411278	5.749	1.72E-08
Percent of Male Drivers Over 80	1.2822106	0.0843173	15.207	< 2e-16
Medicinal Yes	-0.0021062	0.0003549	-5.934	6.17E-09
Decriminalized Reduced	-0.0040623	0.0005834	-6.963	1.28E-11
Decriminalized Yes	0.0045237	0.0016264	2.781	0.00566

Parameter	Severity Linear Model			
	Estimate	Std. Error	t value	Pr(>|t|)
(Intercept)	-46410	9157	-5.068	0.000000602
Rural Congestion	-17350	3205	-5.414	0.000000104
Urban Congestion	24860	1907	13.031	< 2e-16
Percent of Female Drivers Under 24	96900	29360	3.300	0.00105
Percent of Female Drivers Over 80	3082000	222600	13.847	< 2e-16
Percent of Male Drivers Over 80	-3283000	290100	-11.318	< 2e-16
Medicinal Yes	8966	1027	8.730	< 2e-16
Decriminalized Reduced	10890	2067	5.268	0.000000221
Decriminalized Yes	-10350	5214	-1.985	0.04781
Personal Income	0.00002376	0.000002267	10.481	< 2e-16