Introduction: The SafeClear Program

In 2004, Bill White was elected mayor of Houston on the platform “Get Houston Moving,” a response to Houston’s #1 problem: traffic (Klineberg, 2014). In 2005, he established the SafeClear Program, an incident management program that separated the Houston highway system into 29 distinct segments. A specific tow company was made responsible for each segment. From 2005-2011, SafeClear tow operators cleared incidents free of charge. However, in 2011, a $50 charge was instituted for every tow to help pay for the program.

The justification for SafeClear was simple: making tow companies accountable for certain segments would decrease their response time and the time in which a vehicle was an obstruction in the roadway. This would thereby decrease the number of secondary collisions caused by the interfering vehicle.

Evaluating the SafeClear Program

The City of Houston charged a Houston Action Research Team (HART) - an interdisciplinary team composed of Rice University undergraduates - with determining whether SafeClear truly achieved the goal of reducing the number of “secondary” collisions and thus increasing public safety in the city. The justification for the program was that lingering disabled vehicles on the roadway inevitably caused disturbances like long queue lengths, rubbernecking (drivers getting distracted by obstructions in the roadway), debris, and drivers leaving their lane, which caused other collisions to occur.

Previous research has looked only at the effect of tow truck response times (time taken for tow operator to arrive at the incident) on secondary collisions. HART believes that the duration time (amount of time an incident stays on the highway) may also have an effect on the frequency of secondary collisions. Thus, to determine the efficacy of the program, HART evaluated the relationship between response time, duration time, and incidence of secondary collisions.

Primary incidents are defined as any obstruction to the highway that disrupts the flow of traffic. This includes abandoned cars, collisions, burns, flat tires, floods, mechanical failures, vehicles being driven by individuals with outstanding warrants, stalls, stolen vehicles, and medical emergencies. Studies have shown, to varying degrees, that primary incidents cause secondary collisions (Raub 1997; Karlaftis 1999; Latoski 1999; Moore 2004; Hirunyanitiwattana et al. 2006; Sun et al. 2006; Zhan 2008; Zhan 2009; Khattak et al. 2009; Pigman 2011). HART’s task was to determine whether this was also true for the City of Houston. Showing that reduced duration times lead to the decreased frequency of secondary collisions would show the efficacy of the SafeClear Program. This conclusion would support continuing and perhaps expanding Houston’s SafeClear Program.

Defining Secondary Collisions

In order to determine whether the SafeClear Program achieves the goal of increasing public safety, the term "secondary collision" must be clearly defined. Various U.S. traffic studies have been conducted to examine the relationship between primary incidents and secondary collisions. However the definition of secondary collisions differs significantly between the studies as shown in Table 1. Thus, after extensively reviewing the best available science, HART developed a definition for secondary collisions based on temporal and spatial relation of secondary collisions to primary incidents.

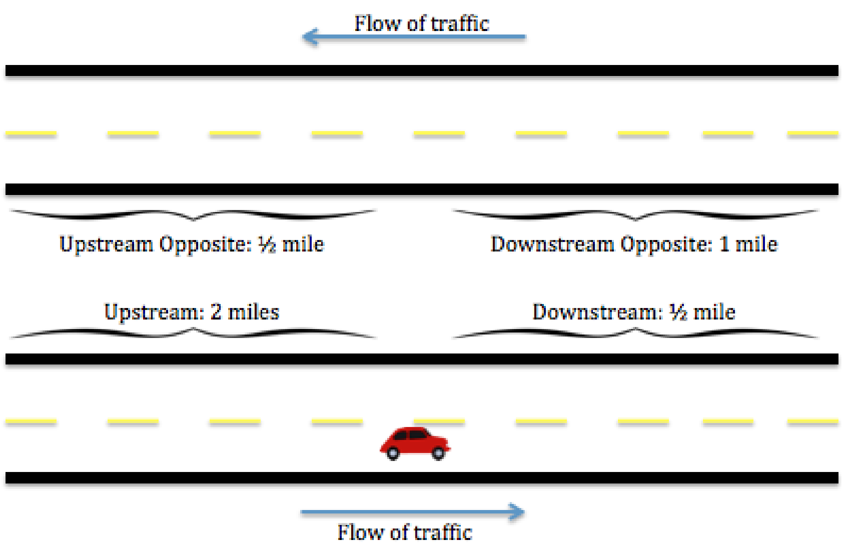
Table 1: Secondary Collision Spatial and Temporal Boundaries

Author	Spatial Boundary	Temporal Boundary
Raub 1997	1 mi radius from primary incident	Clearance time of primary incident plus 15 minutes
Karlaftis et al. 1998	½ mi upstream from primary incident (same side)	Clearance time of primary incident plus 15 minutes
Moore et al. 2004	2 mi upstream from primary incident (same side and opposite side)	2 hours from time of incident
Hirunyanitiwattana et al. 2006	2 mi upstream from primary incident (same side)	1 hour from time of incident
Zhan et al. 2008	2 mi upstream from primary incident (same side)	Clearance time of primary incident plus 15 minutes
Khattak et al. 2009	1 mi upstream from primary incident (same side)	Clearance time of primary incident plus 15 minutes
HART 2014	• ***2 mi upstream from primary incident (same side)*** • ***½ mi downstream from primary incident (same side)*** • ***½ mi upstream from primary incident (opposite side)*** • ***1 mi downstream from primary incident (opposite side)*** • ***½ mile from primary incident if primary incident and collision are on different roadways***	***Clearance time of primary incident plus 60 minutes (excluding collisions within 1 minute of the primary incident)***

First, a temporal filter for secondary collisions was created. When a primary incident occurs, any collision that occurs within one minute of the incident is considered concurrent and thus part of the primary incident. A secondary collision is then defined as any collision that occurs between one minute after an incident and 60 minutes after complete clearance of the incident.

The spatial proximity in which a secondary collision can occur depends on where the secondary collision is in relation to the primary incident (Figure 1). According to our definition, on the same side of the highway, a secondary collision occurs within two miles upstream or half a mile downstream from the primary incident. On the opposite side of the highway, a secondary collision occurs within half a mile upstream and one mile downstream from the primary incident.

Justifications for why HART chose certain temporal and spatial parameters can be found in



Appendix A.

Figure 1: Secondary Collision Definition: Spatial Constraints

Methodology

The following sections describe the method followed by the HART team to identify secondary collisions on Houston highways served by the SafeClear Program. HART first filtered the data set to remove observations with missing data and then dropped observations that included data that did not have a corresponding explanation in the code book. HART then created an algorithm to determine which collisions could be considered secondary according to its definition.

Filtering data

Since SafeClear's inception in 2005, detailed SafeClear tow information has been collected by Houston TranStar. Each incident has been recorded along with its spatial coordinates, notification time, tow arrival time, clearance time, type of incident, and segment location. In all, SafeClear recorded 482,704 incidents between 2005 and June 30, 2014. HART had to eliminate 199,047 recorded incidents based on missing data or data outside the range prescribed by the SafeClear data file codebook. An approximately equal number of incidents were dropped yearly except in 2008 where a significantly higher number of incidents had to be dropped because of incorrect geospatial data. More detailed information about how data was filtered can be found in Appendix B.

Determining Secondary Collisions

Secondary collisions were determined using an algorithm to compare incidents that occurred within prescribed spatial and temporal constraints. For every collision, HART considered previous incidents of all types which were proximate in time and space. If these fell within the previously discussed spatial and temporal constraints, the collision was marked as secondary. If multiple incidents were marked as causing one secondary collision, the closest incident in time and space was chosen as secondary to avoid overcounting. Finally, using a statistical data analysis software, HART was able to determine which factors most frequently contributed to the occurrence of secondary collisions. More information about how statistical regressions were done can be found in Appendix C.

Findings

The following paragraphs highlight the major findings of this study, including the physical characteristics of all recorded freeway incidents, the characteristics of secondary collisions, and the statistical regression analysis done on secondary collisions.

Descriptive Statistics: Incidents in Houston

Figure 2 shows the annual number of incidents towed by SafeClear during the study period (2007-2013). The number of incidents towed by SafeClear increased between 2007 and 2010 and decreased significantly between 2010 and 2013. Note that the low number of incidents in 2008 is due to the fact that many incidents in this year were logged with incorrect spatial codes and therefore had to be dropped from the data set during the filtering process discussed in the previous section.

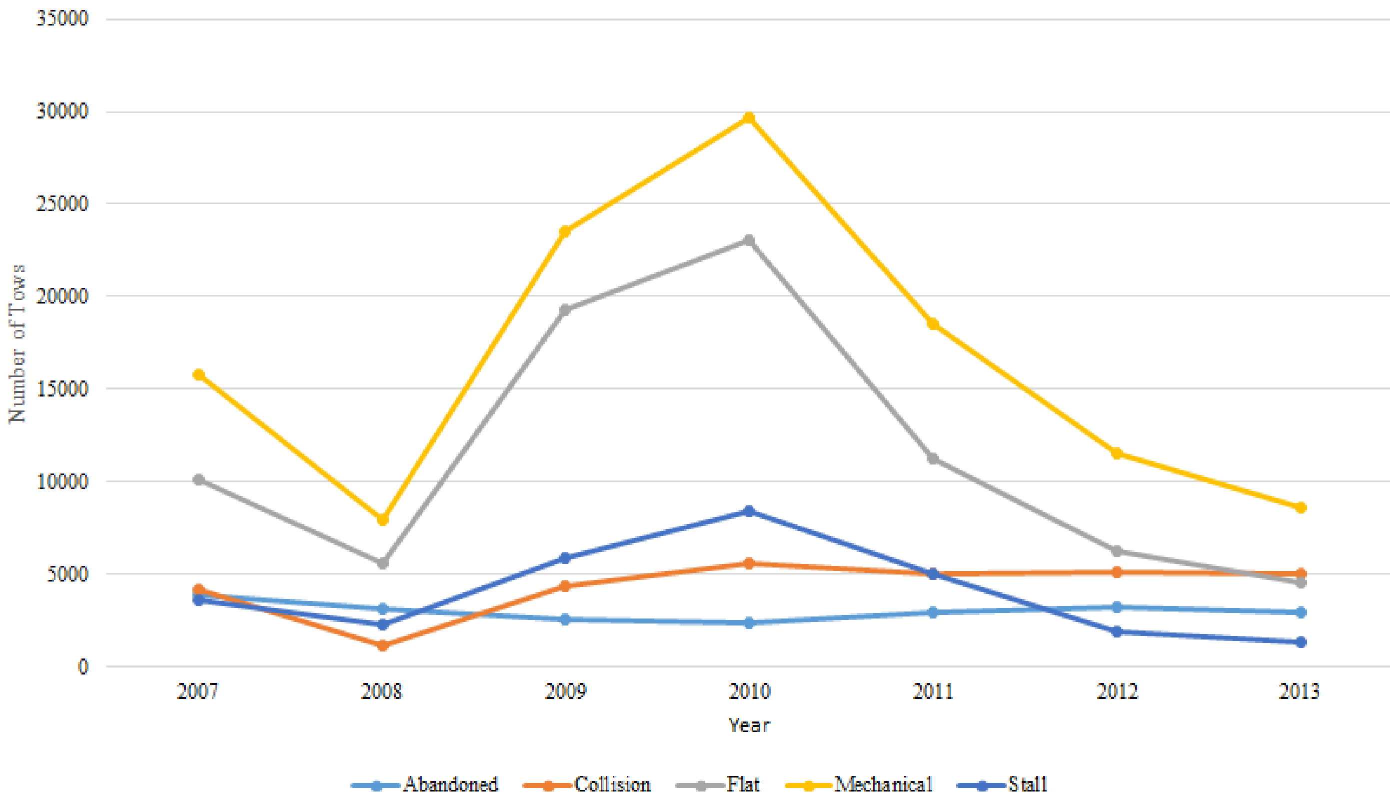


Figure 2: Type of Incident Towed by Year (2007-2013)

The number of tows decreased significantly after 2010. HART speculates two reasons for why this might have occurred: First, it is possible that the recession in 2008 caused less people to drive, and second, 2010 marked the year when citizens had to start paying for a SafeClear tow which had been a free service up until this point. Mechanical failures, flat tires, and stalls are the only types of incidents that decreased between 2010 and 2011. These are types of incidents that can potentially be cleared by the drivers themselves quickly before a tow operator arrives at the scene. Because these were the only types of incidents for which tows decreased after 2010, the assumption can be made that drivers were trying to avoid having to pay for a tow and were clearing incidents themselves. Additionally, since the addition of a fee coincided with the recession, citizens might have felt additional pressure to avoid the fee.

Figure 3 shows the trend in total number of incidents and collisions by hour. Both follow the same trend, however incidents peak before collisions do, leading to the supposition that incidents in the roadway may cause other collisions to occur.

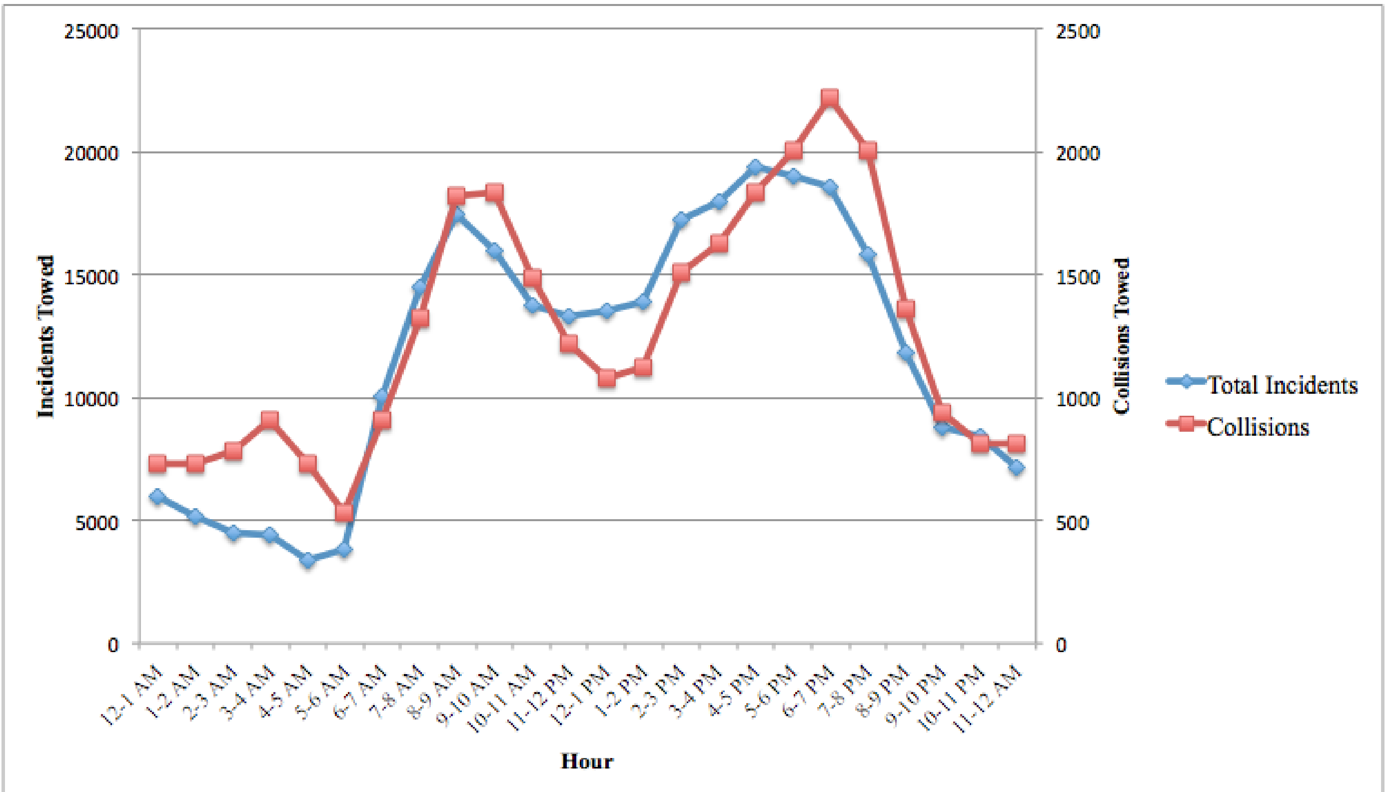


Figure 3: Incidents and Collisions by Hour (2007-2013)

Descriptive Statistics: Secondary Collisions

Secondary collisions were identified based on their temporal and spatial proximity to primary incidents. Figure 4 shows the locations of all 29 segments. Specific statistics by segment can be found in Appendix D.

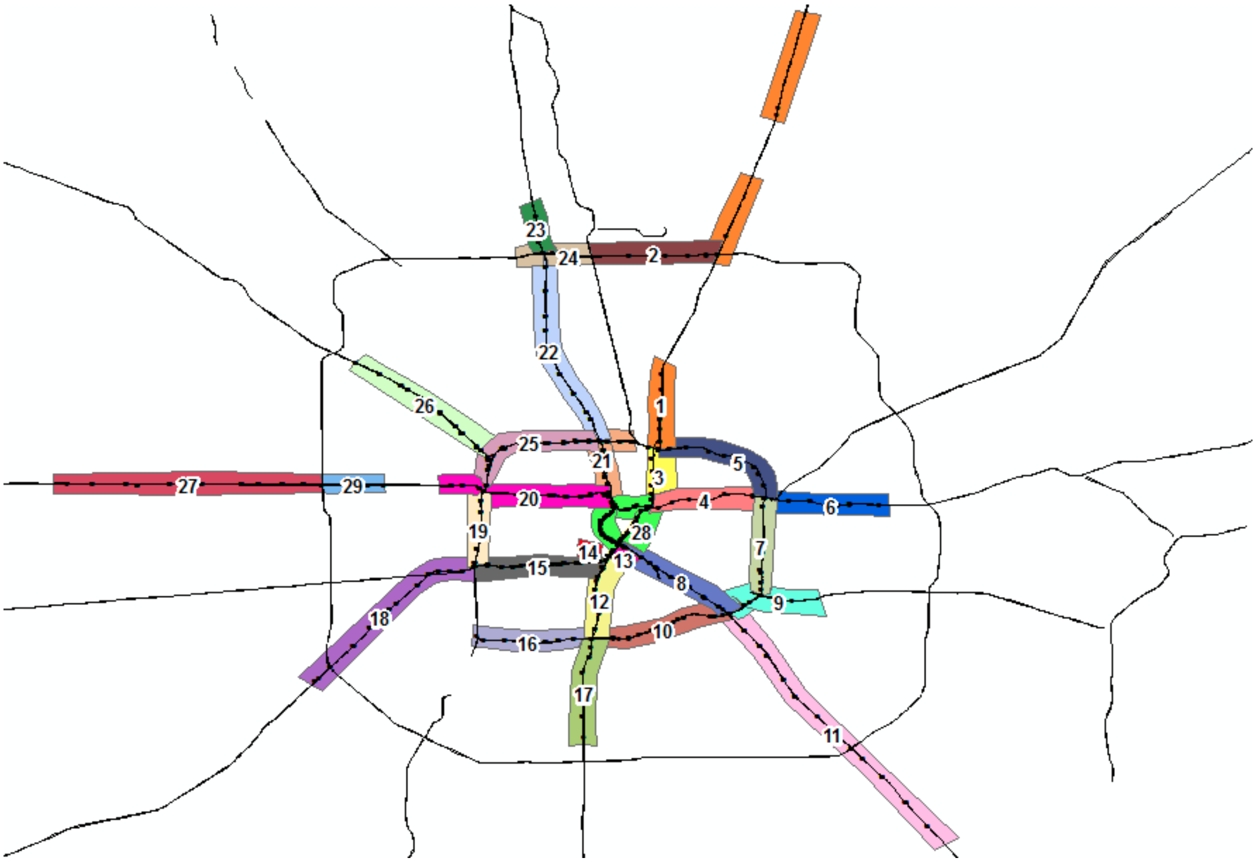


Figure 4: SafeClear Segments

Using ArcGIS, a geospatial information system, and the reported geospatial locations of incidents, a map of the hotspots of secondary collisions was created (Figure 5). This map shows where the highest density of secondary collisions occurs. Secondary collisions appear to occur more frequently at intersections between two or more highways. HART hypothesizes that secondary collisions were more likely to occur at the intersections of highways because drivers would be quickly changing lanes to exit or enter the highway. Thus, vehicles would be more likely to hit each other as they wove through traffic or tried to merge onto a highway. Additionally, if there was already a disturbance in the roadway, this would increase the likelihood of another collision as well.

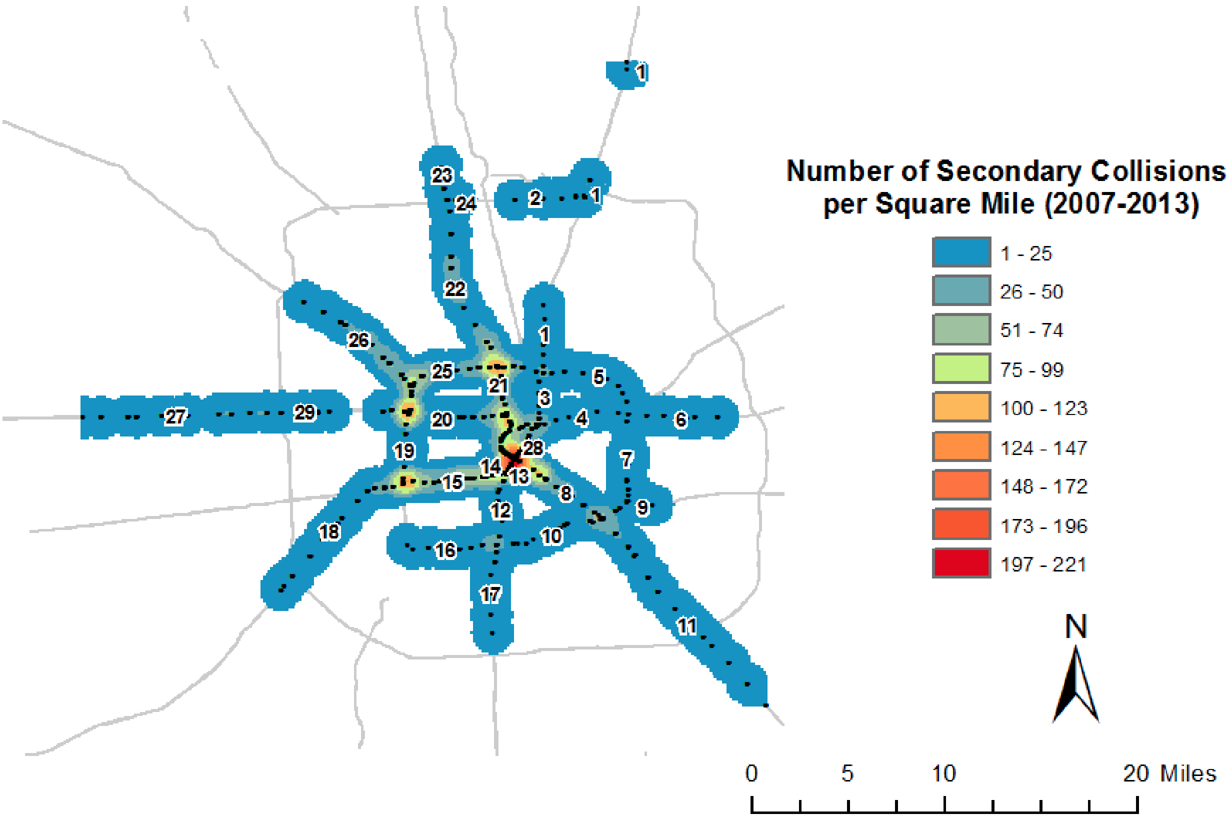


Figure 5: Hotspot analysis of Secondary Collisions (2007-2013)

Figure 6 shows a distribution of distances between primary incidents and secondary collisions. Distances of 0 can be accounted for by the fact that many incident locations were recorded at an intersection rather than at the specific incident site. The City of Houston does not collect data on traffic queue length. Thus, when constructing the definition of secondary collision, HART chose two miles as an upper bound based on previous studies of secondary collisions. However, based on the data below, most secondary collisions tend to occur within 1 mile of the primary incident leading to the conclusion that most queues do not build up greater than 1 mile in Houston.

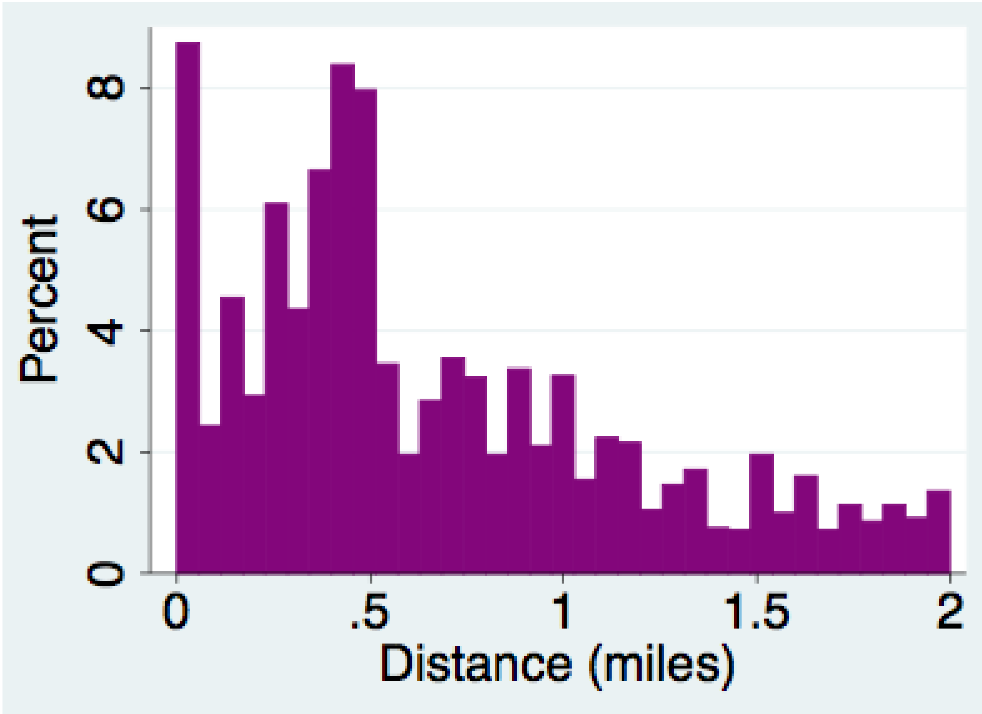


Figure 6: Distribution of Distances between Primary Incidents and Secondary Collisions (2007-2013)

Figure 7 shows the distribution of primary incident duration times. Most primary incidents that caused secondary collisions had a duration time of 30 minutes or less. Primary incident duration times above 30 minutes are rare and tend to have a lesser effect on the occurrence of secondary collisions. HART hypothesized that this could be true because, as the primary incident stays on the highway for a longer period of time, traffic slows and drivers are no longer surprised that it is there and thus drive more safely to avoid collisions.

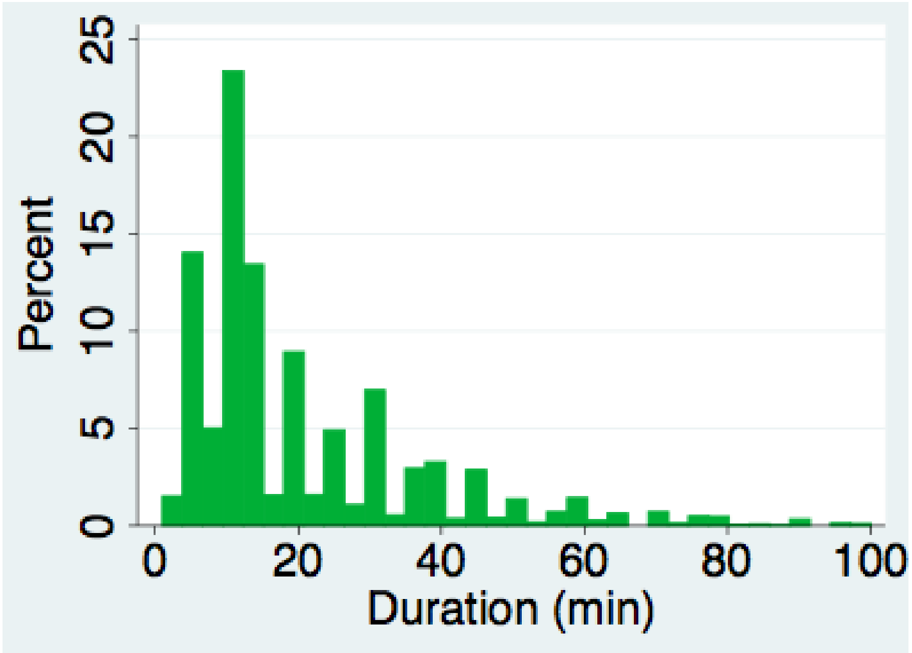


Figure 7: Distribution of Primary Incident Duration Times (2007-2013)

According to HART's definition of a secondary collision, there are four possible locations a secondary collision can occur in relation to a primary incident. Figures 8, 9, 10, and 11 show the distribution of primary incident duration times and distribution of distances between the primary incidents and secondary collisions based on the spatial relation of a primary incident and secondary collision.

Duration times of primary incidents all follow the same trend, peaking at around 10 minutes. Distances between primary incidents and secondary collisions differ based on where the primary incident is in relation to the secondary collision. The number of secondary collisions peak at around 0.5 miles from the primary incident when occurring downstream on the same side of the highway or upstream on the opposite side of the highway. When occurring upstream from the primary incident on the same side of the highway or downstream on the opposite side of the highway, the number of secondary collisions remains relatively constant no matter what distance they are from the primary incident.

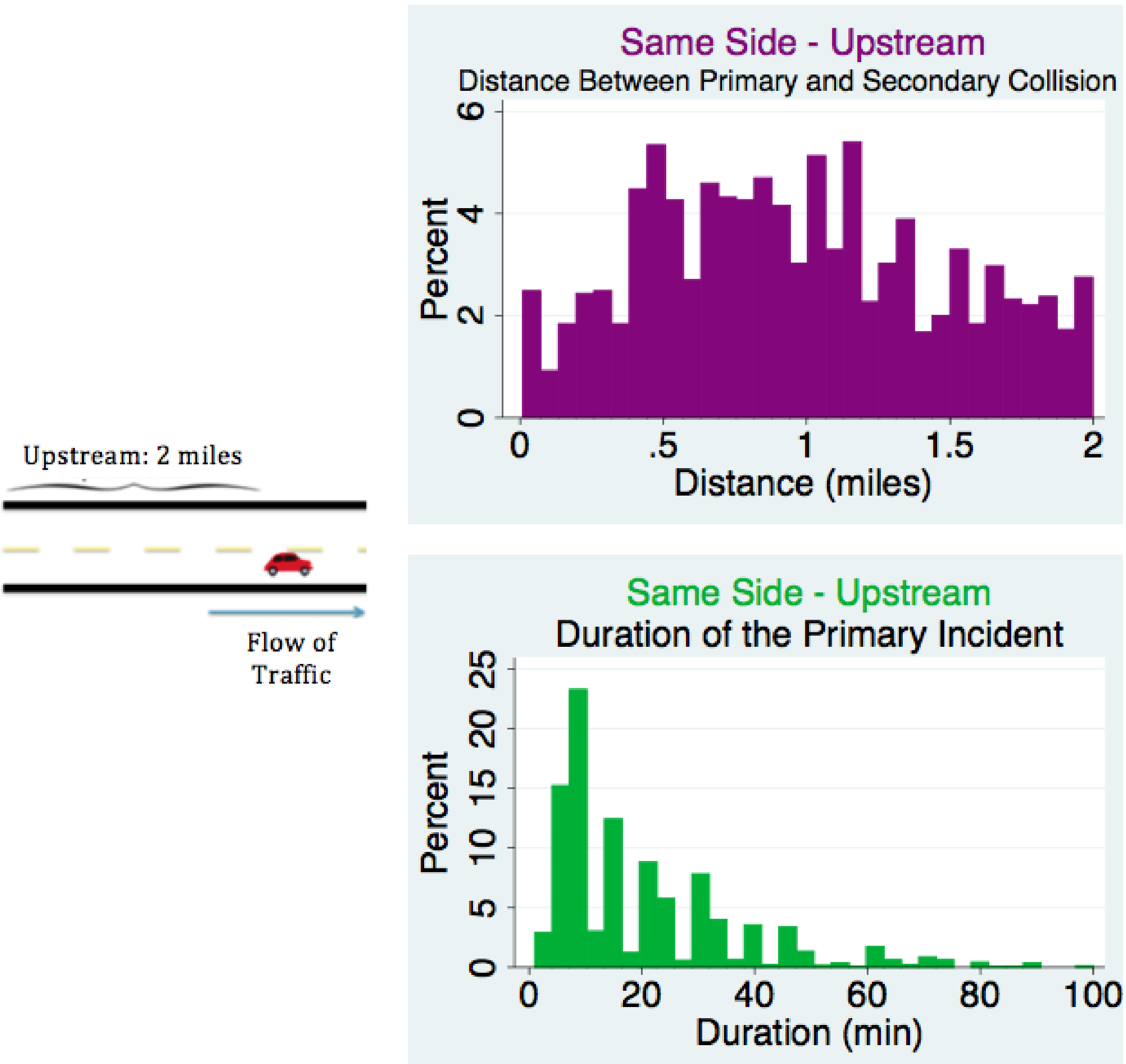


Figure 8: ***Distance between Primary Incident and Secondary Collision*** **and** ***Duration Time of Primary Incident*** **when secondary collision is** ***upstream*** **and on the** ***same side of the highway*** **as the primary incident**

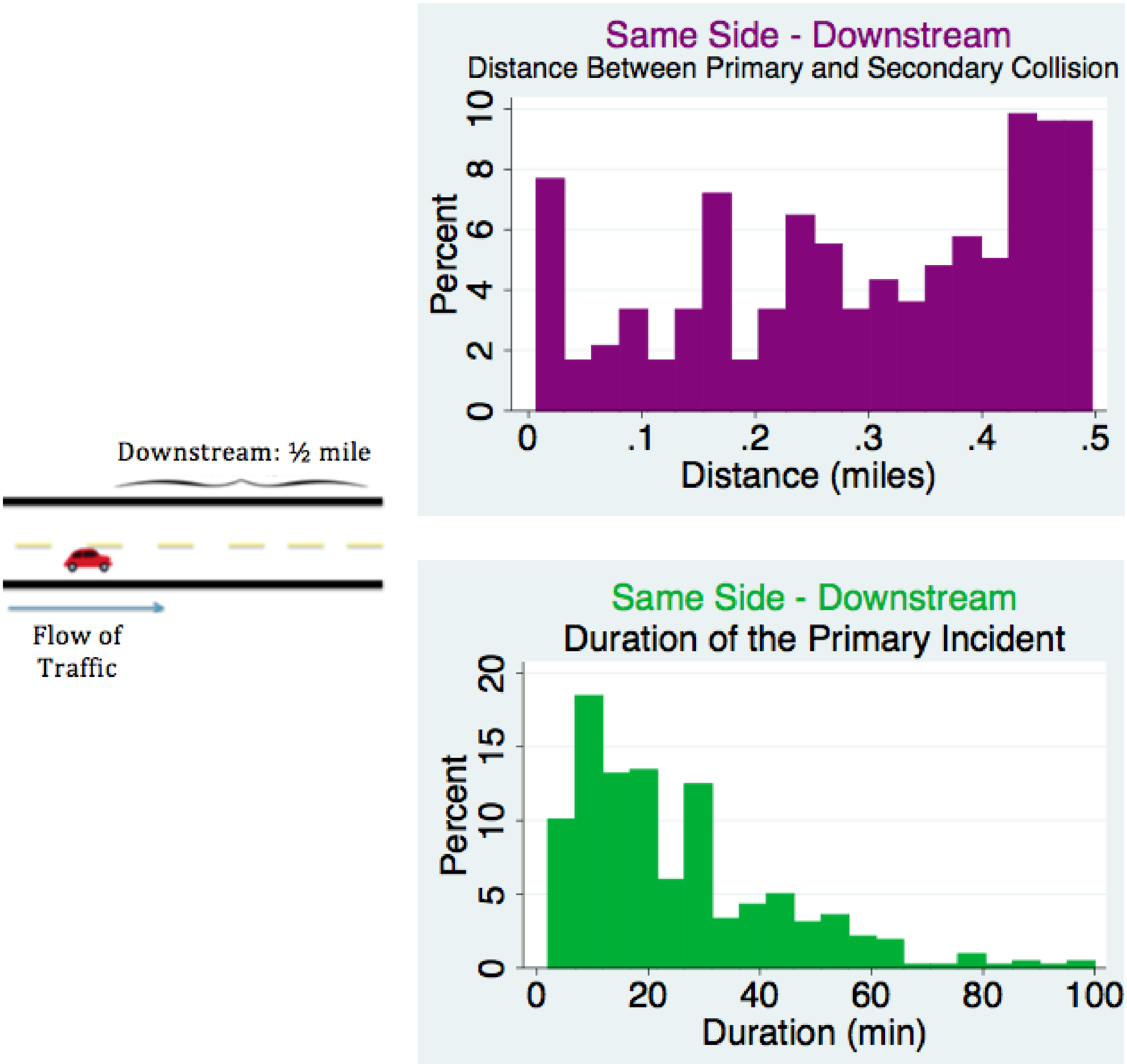


Figure 9: ***Distance between Primary Incident and Secondary Collision*** **and** ***Duration Time of Primary Incident*** **when secondary collision is** ***downstream*** **and on the** ***same side of the highway*** **as the primary incident**

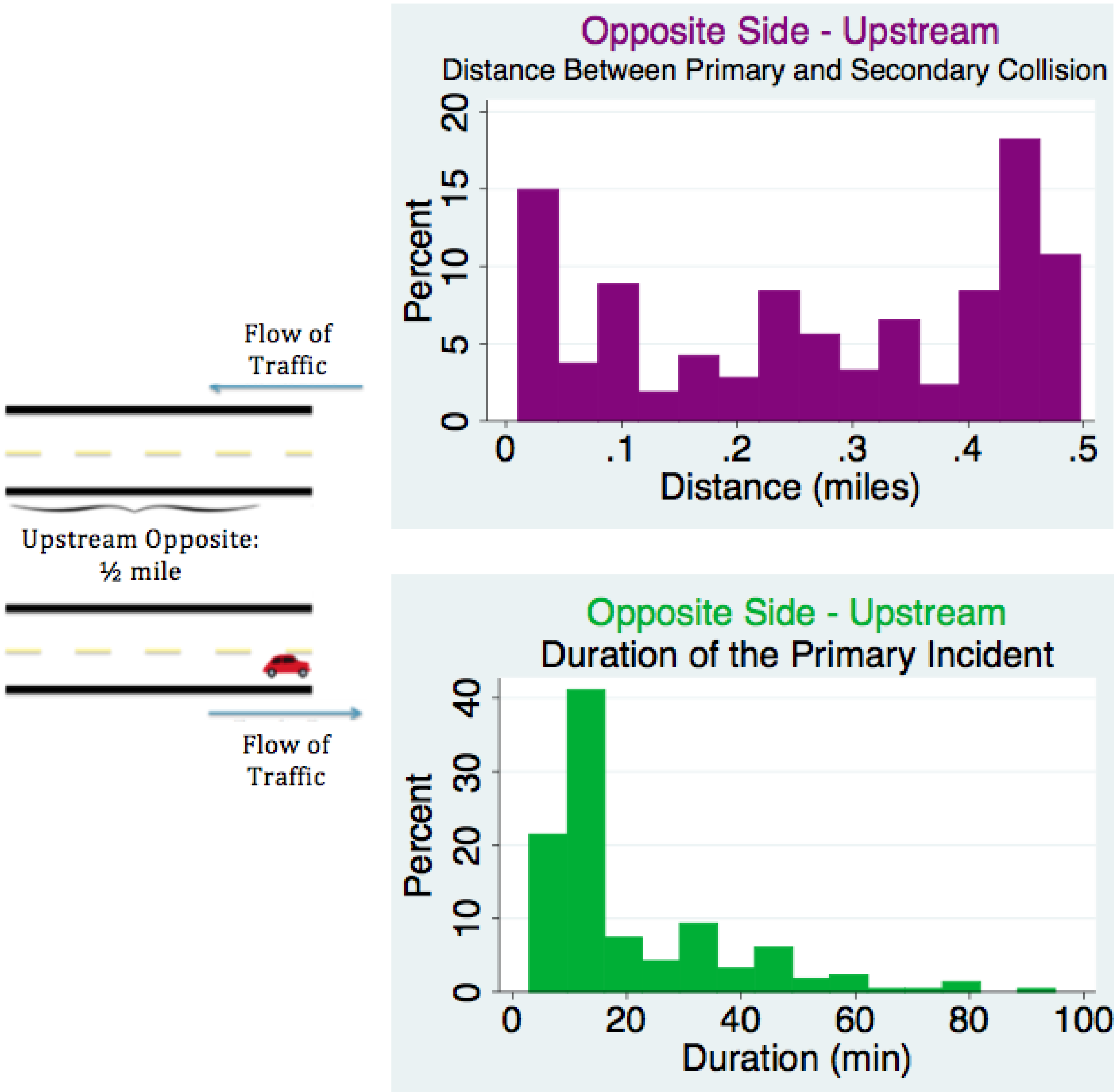


Figure 10: ***Distance between Primary Incident and Secondary Collision*** **and** ***Duration Time of Primary Incident*** **when secondary collision is <u>upstream</u> and on the <u>opposite side of the highway</u> as the primary incident**

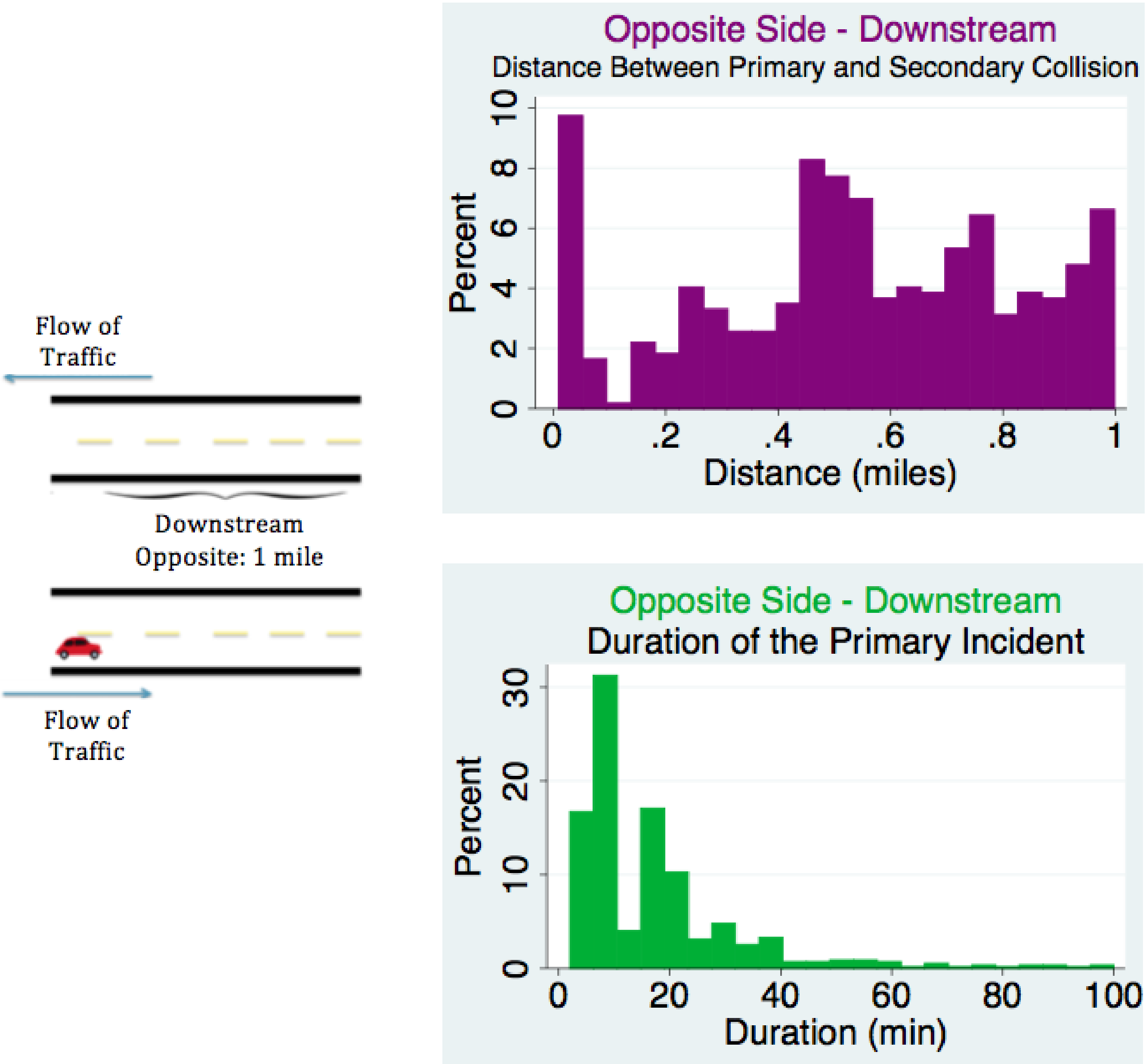
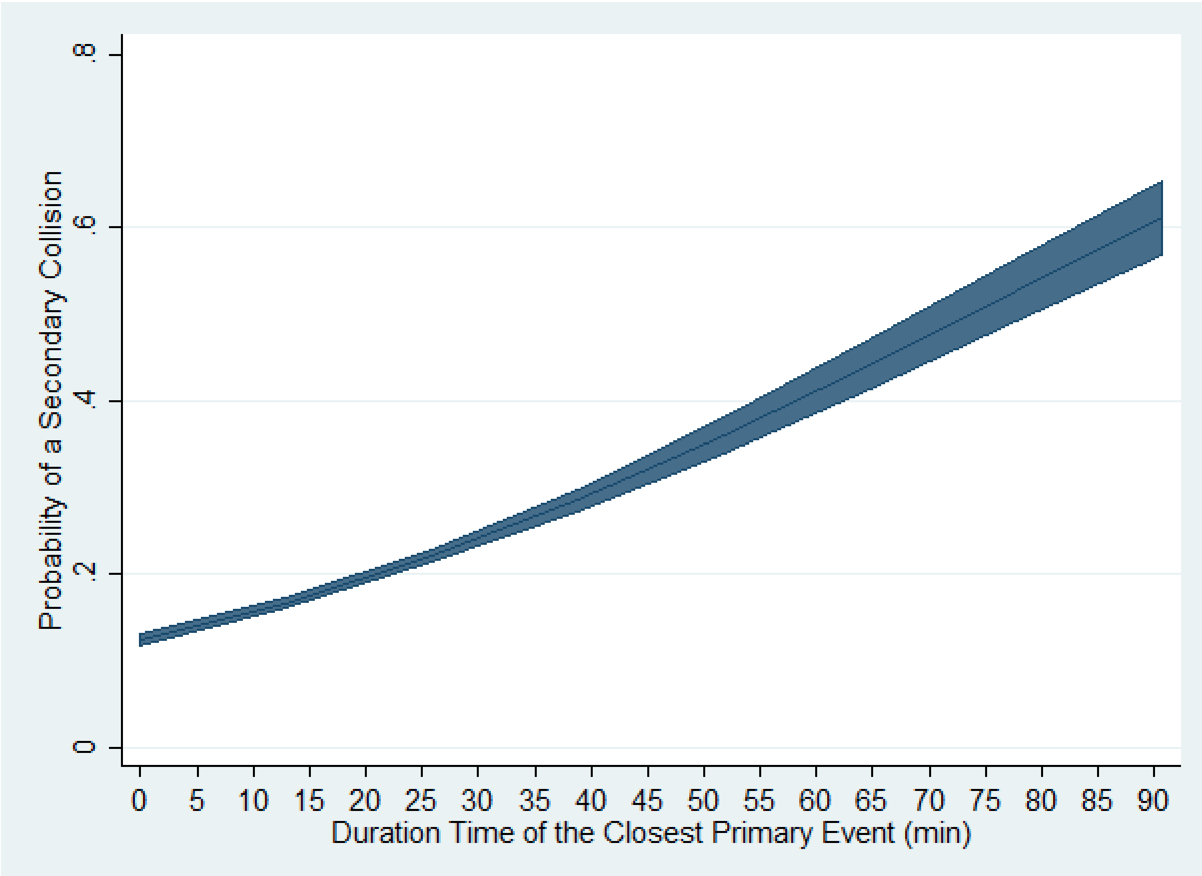


Figure 11: ***Distance between Primary Incident and Secondary Collision*** **and** ***Duration Time of Primary Incident*** **when secondary collision is <u>downstream</u> and on the <u>opposite side of the highway</u> as the primary incident**

Regression Analysis

Figure 12 supports HART's former hypothesis about a causal relationship between the duration time of an incident and the likelihood of a secondary collision. The relationship is highly significant (p-value = 0.000) and moderately strong, rising to a 60% probability at duration times of 100 minutes or greater. The slope of the curve is positive and increasing, meaning that the probability of a secondary collision increases more significantly with higher duration times. Moreover, the likelihood of a secondary collision is relatively small for the average incident



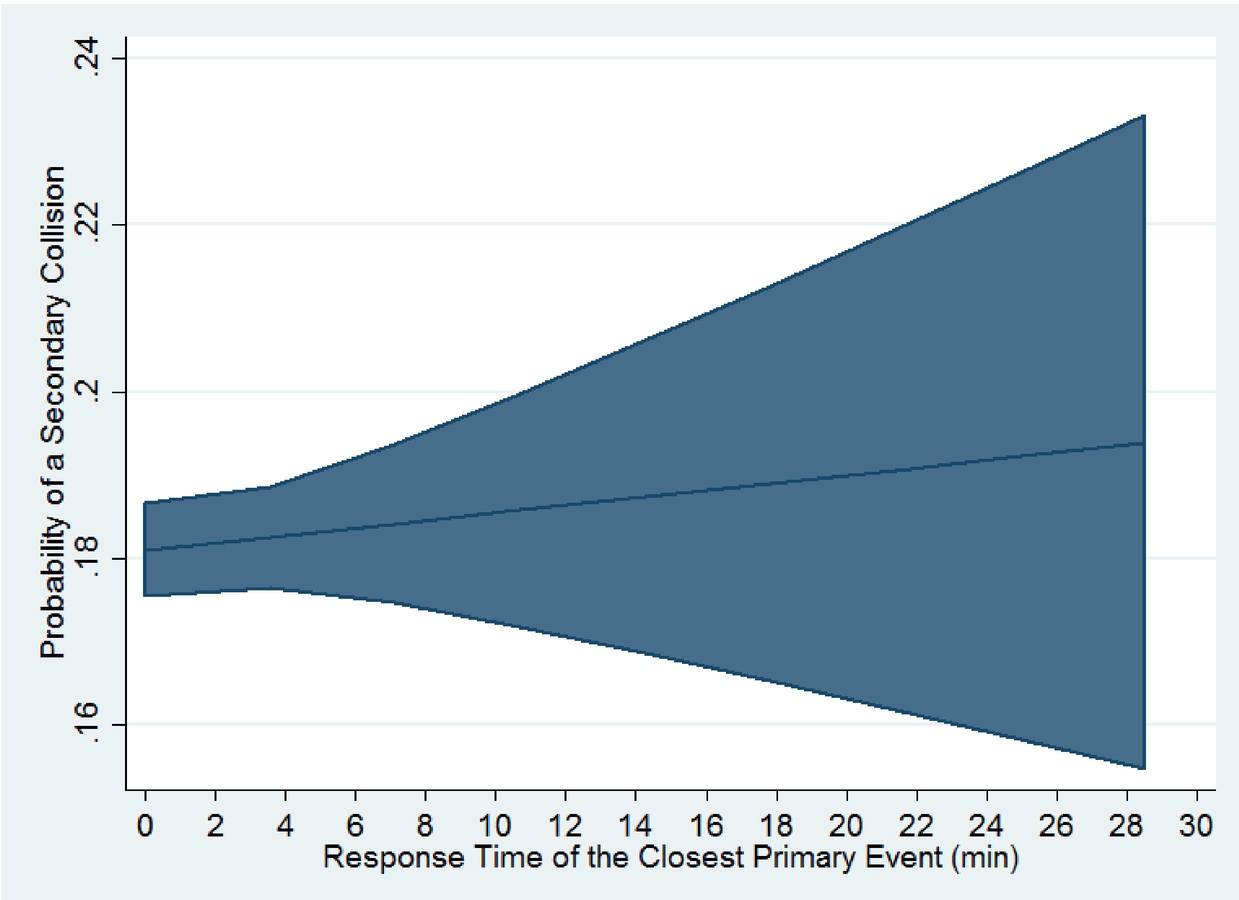
duration time (12.16 minutes) and does not exceed 20% unless duration exceeds 21 minutes.

Figure 12: Probability of a Secondary Collision vs Duration Time of the Closest Incident

To understand the implications of the the relationship between duration time and secondary collisions, it is important to examine the distribution of the duration times. The average duration time of 86% of all incidents (mechanical, flats, stall, abandoned) is a little over 12 minutes. For these incident types, the probability of a secondary collision based on duration time of the primary incident is 17%. However, the average duration time for collisions, which comprise 11% of all incidents, is 25.94 minutes. This increases the probability of a secondary collision to more than 20%. In other words, 1 out of 5 average-duration collisions causes a secondary collision.

HART did not find a significant relationship between the response time of the closest primary event and the likelihood of a secondary collision (p-value = 0.524). A possible explanation is that the vast majority of response times are shorter than 5 minutes. Because of missing data for longer response times and the fact that 84.38% of response times are 0, a regression analysis fails to accurately describe the relationship for times longer than 5 minutes. Another possible reason is that response time does not reflect total time an incident is on the road, whereas duration time does.

The extreme width of the curve in Figure 13 illustrates the uncertainty which we face due to an uneven distribution of the response times. It is impossible to draw a conclusion about whether



response time is related to the probability of a secondary collision with this much uncertainty.

Figure 13: Probability of a Secondary Collision vs Response Time of the Closest Incident

Conclusions:

HART concludes that there is a causal relationship between the duration time of a primary incident and the occurrence of a secondary collision. This is first shown through the fact that the total number of incidents that occur during a 24-hour period peaks before the number of collision. Additionally, the regression analysis shows a significant relationship between the duration time of the primary incident and the likelihood of a secondary collision by proving that 1 out of 5 average-duration collisions causes a secondary collision.

HART, however, did not find a significant relationship between the response time of a tow truck and the occurrence of a secondary collision leading to the conclusion that a change in response time will not decrease the number of secondary collisions.

Recommendations:

HART's recommendations fall into two categories: data collection methods and Safeclear policy implementation.

Data Collection

Future studies into the efficacy of the SafeClear Program can be improved if more accurate information regarding tows are collected. This was a more serious issue during the years 2005-2006 but still continues to be a problem.

First, the location of incidents must be more accurately logged. Many of the incidents in the 2007-2013 data set are located outside of the City of Houston, are not along a SafeClear highway segment, or are geolocated by intersection rather than by the actual location. Determining secondary collisions using spatial parameters can only be done if the location is correct.

Second, it is important that each tow is recorded on the correct SafeClear segment on which it is found. Logging an incorrect segment number skews any analysis by segment, and extra work has to be done to correctly place the observation.

Third, notification, arrival, and removal time must be logged correctly. When filtering data, many observations had to be dropped because negative response and duration times were calculated. Thus, the effect of removal and duration time on the probability of a secondary collision is less likely to be determined accurately.

Finally, a more effective analysis of the SafeClear Program can be made if more information is collected about the nature of the incidents. Severity, including damage and bodily harm, would allow conclusions to be made about whether the SafeClear Program decreases extreme collisions and increases public safety.

SafeClear Policy Recommendations

HART discovered that the main known factor contributing to secondary collisions was the duration time of primary incidents. Hence, measures to decrease these times should be taken. This can be accomplished by either shortening response or clearance times, the two parts which comprise duration time. Given that response times are already zero in nine out of ten cases, SafeClear should now focus on reducing clearance times - the amount of time between the tow operator's arrival and the vehicle's clearance.

As mentioned earlier, collisions take almost 26 minutes on average to clear, which is approximately twice as long as the clearance time of other incident types. Even though this might be due in part to injuries or debris on the road, HART believes that the fact that a police officer must be present at the site may inflate the clearance times.

A possible solution could be to employ more police officers that can be ready to respond to collision scenes. This, however, would likely be costly for the City. Therefore, other measures can be taken given a fixed number of officers. First, the the City of Houston could carefully consider which incidents require a police officer for clearance and which can go without. If an officer is not required at a scene, duration time of a collision would likely drop significantly and allow the officer to deal with another events. Second, the hotspot analysis in Figure 5 could be used to situate policemen closer to the areas with the highest collision concentration. This would decrease the officers' driving times and free them up for other incidents.

As for the segmentation, HART recommends redrawing some of the 29 segments. The main reason for such a change is that the number of incidents in various segments is highly asymmetrical. This leads to inefficiencies since there are likely segments with too few or too many tow operators. Working on a segment with few incidents incentivizes operators to leave those segments or help out operators on busier segments. HART found this might have been the case based on the fact that the segment ID's which towing operators were entering after each tow were wrong in thousands of cases.

Lastly, SafeClear's efficiency could improve if the City were to subsidize towing fees once again. Although this is a costly method, after implementing a $50 fee in 2011 for a SafeClear tow, the number of tows dropped dramatically and the duration times rose slightly. This reduction is most likely due to citizens trying to clear incidents themselves and arguing with the tow operators. Such responses by individual citizens surely increases duration times and thus increases the likelihood of secondary collisions. This contributes to the very effect that SafeClear targets to reduce.

HART concludes that the SafeClear Program does uphold the ideal of public safety by reducing duration time and thus reducing the number of secondary collisions. If these recommendations were to take effect, however, HART believes the policy would be even more efficacious.

References

Hirunyanitiwattana, W., and S. Mattingly. Identifying Secondary Crash Characteristics for California Highway System. Presented at 85th Annual Meeting of the Transportation Research Board, Washington, D.C., 2006.

Karlaftis, M., S. Latoski, N. Richards, and K. Sinha. ITS Impacts on Safety and Traffic Management: An Investigation of Secondary Crash Causes. ITS Journal, Vol. 5, No. 1, 1999.

Khattak, A. J., X. Wang, and H. Zhang. Are Incident Durations and Secondary Incidents Interdependent? In Transportation Research Record: Journal of the Transportation Research Board, No. 2099, Transportation Research Board of the National Academies, Washington, D.C., 2009.

Latoski, S. P., R. Pal, and K. C. Sinha. An Evaluation of the Cost Effectiveness of the Hoosier Helper Program and Framework for the Design of ITS Optimal System Configuration, Phase 1. Publication FHWA/IN/JTRP-97/09. Joint Transportation Research Program, Indiana Department of Transportation and Purdue University, West Lafayette, Indiana, 1997.

Moore, J. E., G. Giuliano, and S. Cho. Secondary Accident Rates on Los Angeles Freeways. 16 Journal of Transportation Engineering, Vol. 130, No. 3, 2004.

Pigman, Jerry G.; Green, Eric R.; and Walton, Jennifer R., "Identification of Secondary Crashes and Recommended Countermeasures" (2011). *Kentucky Transportation Center Research Reports*. Paper 25. http://uknowledge.uky.edu/ktc_researchreports/25

Raub, R. A. Occurrence of Secondary Crashes on Urban Arterial Roadways. In Transportation Research Record 1581, TRB, National Research Council, Washington, D.C., 1997.

Sun, C., and V. Chilukuri. Dynamic Incident Progression Curve for Classifying Secondary Traffic 40 Crashes. Journal of Transportation Engineering, Vol. 136, No. 12, 2010.

Zhan, C. J., L. Shen, M. A. Hadi, and A. Gan. Understanding the Characteristics of Secondary Crashes on Freeways. Presented at 87th Annual Meeting of the Transportation Research Board, Washington, D.C., 2008.

Zhan, C., A. Gan, and M. Hadi. Identifying Secondary Crashes and Their Contributing Factors. In Transportation Research Record: Journal of the Transportation Research Board, No. 2102, Transportation Research Board of the National Academies, Washington, D.C., 2009.

Appendix A - Justification for Secondary Collision Spatial and Temporal Filters

Temporal Filter

A buffer time from the moment of incident to one minute after the incident was chosen to account for chain collisions. A 60 minute buffer after clearance time was chosen to account for the time needed for normal traffic patterns to resume.

Spatial Filter

On the same side of the highway as the primary incident, more distance was chosen upstream as opposed to downstream because a queue will develop upstream from the primary incident. This creates more congestion behind the incident whereas any collision in front of the primary incident would likely be caused only by rubbernecking. The buffer distances chosen for secondary collisions occurring on the opposite side of the highway are shorter as compared to the same side because collisions occurring on the opposite side would likely be caused by rubbernecking and not by an obstruction in the roadway. Rubbernecking is less likely to cause secondary collisions because there is not a specific object impeding the movement of a car. The given distance for a secondary collision to occur upstream is shorter than the distance downstream because drivers upstream would have to turn their head completely to view the primary incident whereas drivers downstream would not.

Appendix B - Data Filter

Data from 2005-2006 and 2014 was disregarded (121,087 incidents) by HART because much of the data needed to measure secondary collisions was not reported in 2005 and 2006, and data for 2014 only covered the first six months of this year. Thus, HART chose 2007-2013 as the study period for the analysis. Observations missing highway segment numbers (1450 incidents) and any duplicate incidents (59 incidents) were also dropped.

The duration time of each incident was calculated by subtracting the time the tow truck operator for that segment was notified about a disabled vehicle from the time that the incident was cleared. Incidents with a duration time of zero minutes or less were dropped and assumed to be errors (8572 incidents). Next, a distribution of duration times was made and duration times longer than 100 minutes were calculated to be extreme outliers and were also dropped (1384 incidents).

Third, the response times of the tow truck operators were calculated by subtracting the notification time from the arrival time. Response times less than zero were dropped. which would have indicated a reporting error (6953 incidents) as well as response times above 30 minutes (2977 incidents). Response times of zero minutes were kept because, in these cases, tow operators driving on the highway noticed the incident before anyone notified them.

Finally, ArcGIS was used to determine which observations had latitude and longitude values outside the City of Houston or not along a segment of highway. These data points were also eliminated (56,565 incidents). Most of these incidents placed outside of possible segments occurred in 2008, although it is unknown what made 2008 such an erroneous year for reporting. Additionally, HART determined that 44,899 observations had latitudes and longitudes that placed them on a different segment than the one that was recorded. Using ArcGIS, these observations were placed on the correct segment.

After filtering, 283,657 incidents remained for HART to analyze.

Appendix C - Regression Analysis

Regression is a statistical process for estimating relationships among variables. The most widely used types of regressions are - Ordinary Least Squares Regression and Logit Analysis. The first method is used if the outcome variable takes on many values, whereas the second is used if the outcome takes exactly two values.

HART used logit analysis to determine the relationship between duration times, response times, and secondary collisions because the outcome (secondary or non-secondary collision) was binary. To control for unmeasured effects associated with each segment, such as road conditions, number of exit ramps, and traffic densities, each segment was included in the regression.

Secondly, HART tested the hypothesis that response times and duration times had a significant effect on secondary collisions. Since logit analyses are not easily understood from the regression coefficient, the marginal effects of response and duration times were examined using the statistical software Stata® functions “margins” and “marginsplot”. Figures 12 and 13, which were created using these functions, show the effect of duration time and response time of the primary event on the probability of an incidence of a secondary collision.