ESE 420/520 Midterm 3
Zhiyang Chen, Bari Gordon, Meiqing Li
Prof. Barry Silverman
December 12th, 2017

Autonomous Vehicles and their Effects on Safety and Destination Utility

I. Introduction

This semester our ESE 420/520 team built a model that analyzes the effects of autonomous vehicles on traffic conditions, overall safety, and satisfaction of values being met by different agent types. Our model consists of three agent types: Autonomous Vehicles (AVs), Human-operated Vehicles (HVs), and Pedestrians moving around and interacting at a busy intersection. Our goal was to model a future scenario where autonomous vehicles will interact in traffic with human-operated vehicles and pedestrians. In Midterm 1, our team built the main functionality of the model by creating the agent types, the environment, and the different rules of the model. Midterm 2 involved adding decision making for our agents, mainly in choosing their destination and deciding if they should go or stop at an intersection. Additionally, in Midterm 2 we added an emotion component to our agents and their decision making process. In this final Midterm we will be randomizing our model by designing a stochastic model and simulation to analyze the robustness and emergence of system equilibria as a function of sensitivity to parameter shifts and policy changes.

II. Input Analysis

Our literature research and survey questionnaires from Midterm 2 were extensive and produced reliable evidence tables. We have therefore used the same results for this Midterm and have reproduced our results from Midterm 2 below for reference.

Literature Review

From Midterm 2 we have found from various literature sources that the two main factors that influence risk level of agents are age and gender. We therefore created a distribution of risk factor based on our findings:

- Female drivers showed a slightly higher percentage of compliance than male drivers at feature sites (69% vs. 64%);
- Mid-age drivers (25–59) showed the highest percentage of compliance at feature sites (83%), followed by older drivers (60+, 69%) and younger drivers (16–24, 61%).

Table 1. Four groups of agents with different risk-factors

Demographic group	**Risk-factor (risk-averse vs. risk-seeking)**	**Percentage of population**
Female (25-59)	1	1/6
Male (25-59) and Female (>59)	2	1/3
Male (>59) and Female (<25)	3	1/3
Male (<25)	4	1/6

*The higher the risk-factor, the more risk-seeking.

Knowledge Engineering

Below in Table 2, we have created a knowledge engineering table to prove and disprove some of our initial hypotheses. We used the evidence from our survey in order to do so. Below are the five hypotheses that we used in our analysis:

1. Drivers will run red lights under high emotions
2. Pedestrians will jaywalk under high emotions
3. Agents will prioritize destination over safety in high emotional states.
4. Drivers are more likely to avoid Pedestrians
5. Drivers are more likely to avoid other cars

The confidence index is the tool to actually determine if a hypothesis has been proven or disproven and was calculated with the following formula:

$$CI_{Avg}(Hj) = 1/n * \Sigma K * Cij * Rj \quad \text{where } K = \{1 \text{ when } Cij >= 0, \text{ and } 5 \text{ when } Cij < 0\}$$

Where CI is the confirmation index of all competing hypotheses (Hj) that weighs disconfirming evidence (5) higher than confirming evidence (1). Basically K is used to assign weights of disconfirming and confirming evidences. Based on our findings, H2 and H4 had the highest confidence indexes, while H1 and H3 had the lowest. H1 and H3 however still were within a range of acceptability and therefore aren't disproved, but rather will be weighted lower than H2 and H4 (refer to the above hypotheses and the table below for more information on H1, H2, H3, and H4).

Table 2: Knowledge Engineering

Source	**Evidence (Ei)**	Reliability (Ri)	H1: Drivers will run red lights under high emotions	H2: Pedestrians will jaywalk under strong emotions	H3: Agents will prioritize destination over safety when there is an urgency to get to location and emotions are high	H4: Drivers are more likely to avoid Pedestrians	H5: Drivers are more likely to avoid other cars
Q 1 and 2	Over 50% of Respondents reported not making the best decisions while feeling strong negative and/or strong positive emotions	0.6	0.6	0.6	0.5	0	0
Q3	Over 50% of Respondednts reported that they would be at least slightly likely to run a red light while driving and emotions are high	0.8	0.5	0	0.1	0	0
Q4	Over 75% of Respondents reported that they would be at least slightly likely to jaywalk when emotions are high; over 50% were even more likely to jaywalk	0.8	0	0.8	0.2	0	0
Q 5 and 6	Over 75% of Respondents reported that they would be at least slightly likely to jaywalk when destination was urgent and emotions were high	0.6	0	0.5	0.6	0	0
Q9	Over 85% of Resondents were very likely to slow down or stop if they saw a pedestrian	0.7	-0.2	0	-0.2	0.8	0
Q10	Over 25% of Respondents were very likely to stop when they saw other cars nearby	0.6	-0.1	0	-0.1	0	0.2
	Confidence Index		-0.04	0.22	-0.17	0.093	0.02

Weight Estimation for Agents

Below we have created weighted estimation tables for how each of our agents will make their safety decision within our MODM functionality. Our HV and Pedestrian agents will have to decide whether they want to prioritize safety or reaching their destination. AVs will be rule based and will always choose to prioritize safety. However, all agents will use MODM to determine which destination they should go to next. We will use the weights that we find below to enhance our utility function for HV and Pedestrian safety decision at intersections.

Table 3: Pairwise Comparison Score Description

Level	Descriptor	Score
1	Equally important or preferred	1
2	Slightly more important or preferred	3
3	Strongly important or preferred	5
4	Very strongly important or preferred	7
5	Extremely important or preferred	9

However, the total weight given to an attribute in the given node can have a maximum weight of +1 and a minimum weight of 0. Therefore, normalized values of the row averages are used as the weight inputs to the trees.

Table 4: Weighted Estimates for Pedestrians

	Avoiding Cars in intersection	Stopping at Red-light ahead	Increasing Money	Increasing Education	Increasing family time	Increasing materialism	Geometric Average	Weight
Avoiding Cars in intersection	1	3	1/3	3	3	5	1.89	0.23
Stopping at Red-light ahead	1/3	1	1/7	1/5	1/5	1/3	0.29	0.03
Increasing Money	3	7	1	3	3	7	3.31	0.40
Increasing Education	1/3	5	1/3	1	1/3	5	0.99	0.12
Increasing family time	1/3	5	1/3	3	1	7	1.51	0.18
Increasing materialism	1/5	3	1/7	1/5	1/7	1	0.37	0.04

Table 5: Weighted Estimates for HVs

	Avoiding Pedestrians in intersection	Avoiding Cars in intersection	Stopping at Red-light ahead	Increasing money	Increasing education	Increasing family time	Increasing materialism	Geometric Average	Weight
Avoiding Pedestrians in intersection	1	7	5	3	5	5	7	5.14	0.46
Avoiding Cars in intersection	1/7	1	3	1/3	3	3	5	1.36	0.12
Stopping at Red-light ahead	1/5	1/3	1	1/3	1	1	3	0.64	0.06
Increasing money	1/3	3	3	1	3	3	7	2.40	0.21
Increasing education	1/5	1/3	1	1/3	1	1/3	5	0.58	0.05
Increasing family time	1/5	1/3	1	1/3	3	1	7	0.88	0.08
Increasing materialism	1/7	1/5	1/3	1/7	1/5	1/7	1	0.18	0.02

Parameters for PDF Analysis

The two parameters chosen for our PDF analysis were the weight of perceived risk when seeing a red light ahead for Pedestrians and HVs separately. These two parameters were highly researched in our literature review and evidence tables above. From our evidence tables we determined a weight of 0.03 for stopping at a red light for Pedestrians and a weight of 0.06 for stopping at a red light for HVs. This is based on our survey results which concluded that pedestrians are less sensitive to a red light than HVs and that both pedestrians and HVs are less sensitive to red lights than other risk factors in the simulation.

In order to determine the PDF for each parameter, we took our initial survey results and categorized the questions in order to extract information from our findings by using R. A detailed report of our findings and results can be found below:

Empirical PDF

Categorize questions into two different factors

```
data2 <- mutate(data, redLight_HV = E3 + E7 + E8)
data3 <- mutate(data2, redLight_Ped = E4 + E5 + E6)
```

Scale the variable so that it has range from 0.03 to 1.03

```
range01 <- function(x){(x-min(x))/(max(x)-min(x)) + 0.03}

data3$redLight_HV_01 <- range01(data3$redLight_HV)
data3$redLight_Ped_01 <- range01(data3$redLight_Ped)
```

Red Light Parameter for HV

Draw PDF for hv

```
hist(data3$redLight_HV_01,probability=TRUE)
```

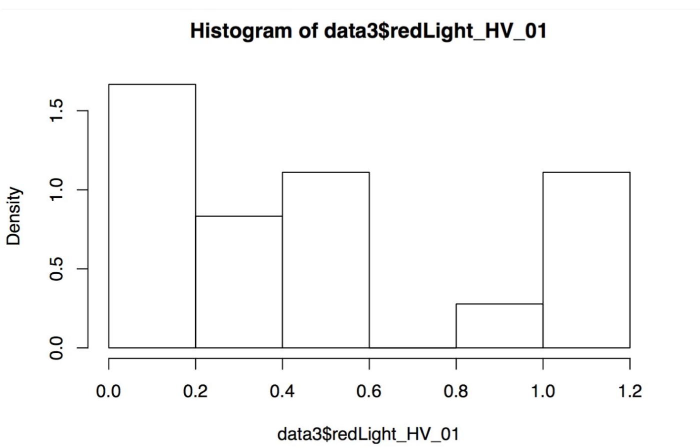


Figure 1: Histogram for red light parameter for HVs

Figure 1 shows the PDF for the red light parameter for HVs, this is a scaled histogram from 0.03 to 1.03. We then tried to fit a few distributions to our PDF and using AIC, we found that the exponential distribution would be the best fit.

AIC of exponential fit is lower, we will use exponential

```
c(fit.weibull$aic, fit.gamma$aic, fit.norm$aic, fit.exp$aic)
```

```
## [1]  9.525360  9.302110 20.966006  7.848888
```

```
fit.exp <- fitdist(data3$redLight_HV_01, "exp")
plot(fit.exp)
```

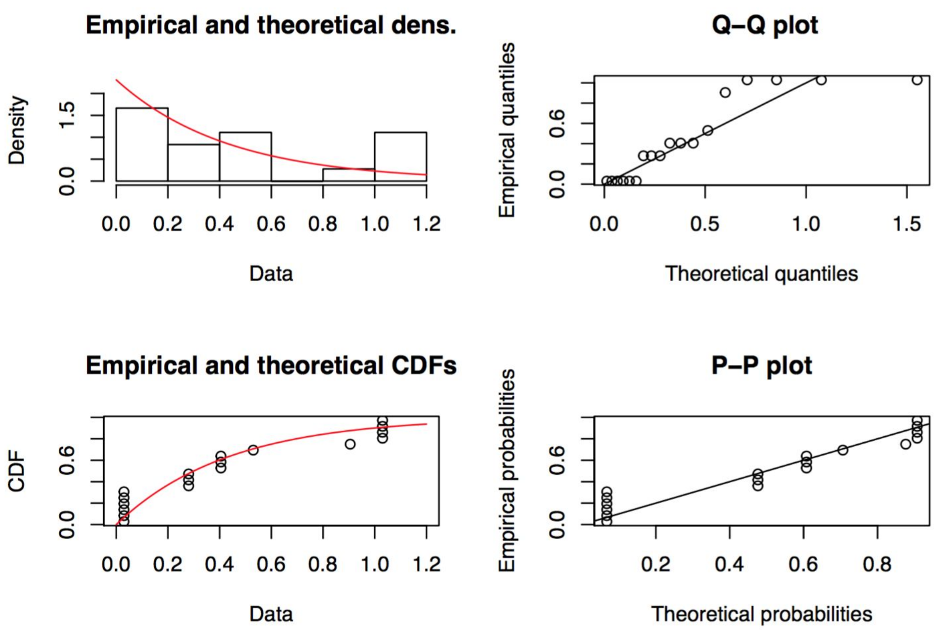


Figure 2: Exponential distribution fit for red light parameter for HVs

For the red light parameter for HVs, we found the distribution to be exponential with lambda = 2.310655.

Red Light Parameter for Pedestrians

Draw PDF for ped

```
hist(data3$redLight_Ped_01,probability=TRUE)
```

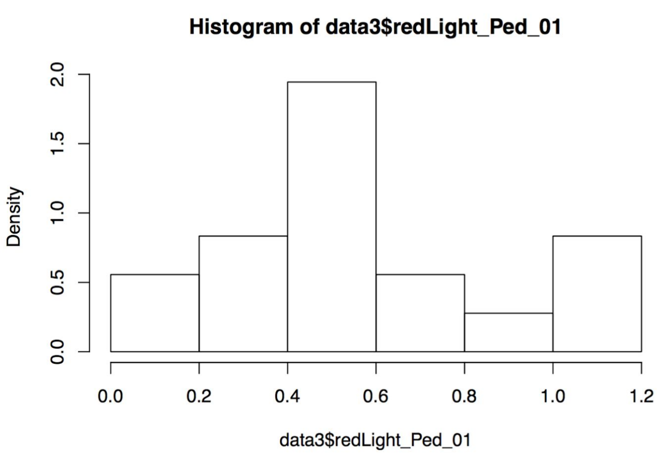


Figure 3: Histogram for red light parameter for Pedestrians

Figure 3 shows the PDF for the red light parameter for Pedestrians, this is a scaled histogram from 0.03 to 1.03. We then tried to fit a few distributions to our PDF and using AIC, we found that the normal distribution would be the best fit.

AIC of normal fit is lower, we will use normal.

```
c(fit.1.weibull$aic, fit.1.gamma$aic, fit.1.norm$aic, fit.1.exp$aic)
```

```
## [1] 13.65881 16.06532 12.14552 17.14601
```

```
fit.1.norm <- fitdist(data3$redLight_Ped_01, "norm")
plot(fit.1.norm)
```

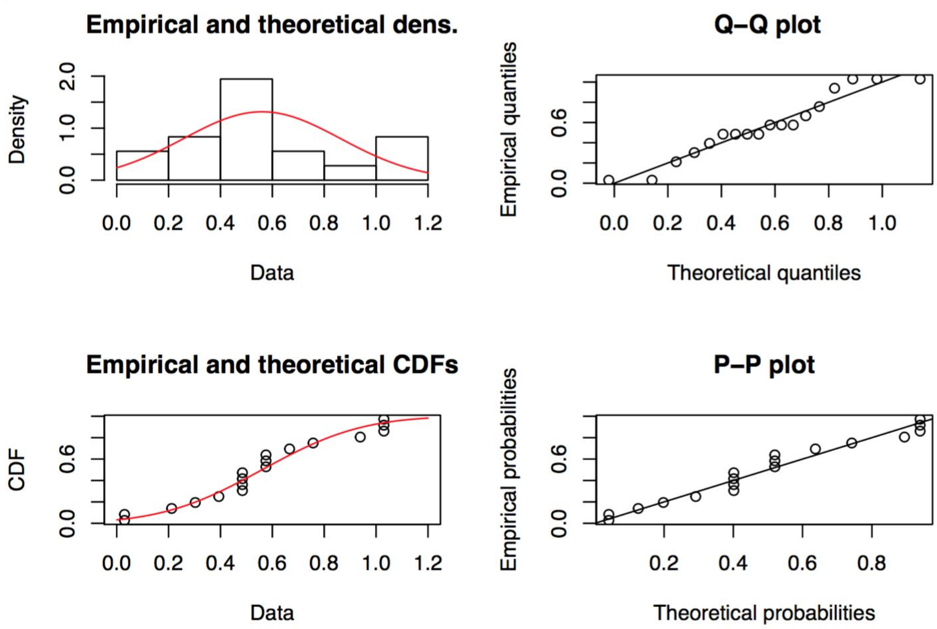


Figure 4: Normal distribution fit for red light parameter for Pedestrians

For the red light parameter for Pedestrians, we found the distribution to be normal with mean = 0.56, sd = 0.30.

III. Monte Carlo Model Control

Based on the above mentioned empirical PDFs, we created a parameter for both pedestrian and HV weight of perceived risk for red light ahead called red-light_Ped and red-light_HV respectively. While the parameter for HV follows an exponential distribution with a factor of 2.310655, the parameter for pedestrians follows normal distribution N(0.56, 0.30). We ran combinatorially the two input parameters for 100 iterations, each run goes 1000 ticks in NetLogo. The outcomes recorded were number of traffic offenses committed throughout the duration of the simulation. The code to run this simulation and the plots obtained can be found below (see attachment for complete code):

```
# we ran cominatorically the two PDFs
d_HV <- rexp(100,2.310655)
d_Ped <- rnorm(100,0.56,0.30)
traffic_offenses <- list()
HV <- list()
Ped <- list()

for(i in 1:100) {
  print (d_HV[i])
  print (d_Ped[i])
  NLCommand("setup")
  NLCommand("set red-light_Ped", d_Ped[i])
  NLCommand("set red-light_HV", d_HV[i])
  a <- NLReport("red-light_HV")
  print(a)
  b <- NLReport("red-light_Ped")
  print(b)
  NLDoCommand(1000,"go")
  traffic_offenses[[i]] <- NLReport(c("number_traffic_offenses"))
  HV[[i]] <- d_HV[i]
  Ped[[i]] <- d_Ped[i]}
```

Our NetLogo model is non-deterministic in that it already incorporates multiple randomization processes. For example, we assign risk-factor to pedestrian from a probability distribution based on studies and our assumption of the distribution of agents' age and gender.

```
    ;; assigns the risk-factor to pedestrian according to probability distribution (takes into account gender and age)
    ;; the probability comes from studies and is presented in Table 1 of midterm report.
    ;; we generate a random number range from 1 to 100. This number falls in to one of the four ranges below, and we
    ;; assign turtles' risk factors based on this.

;; Comment Below out when doing What If Scenario
    let rd random 100
    ;; risk-factor = 1
    if rd >= 0 and rd < 16 [
      set risk-factor 1
    ]
    ;; risk-factor = 2
    if rd >= 16 and rd < 50 [
      set risk-factor 2
    ]
    ;; risk-factor = 3
    if rd >= 50 and rd < 80 [
      set risk-factor 3
    ]
    ;; risk-factor = 4
    if rd >= 80 and rd < 100 [
      set risk-factor 4
    ]
```

Also the initial values of each agent's value factors in terms of money, family, education and materialism are randomly generated.

```
  set emotion_level 0  ; initialize emotion-level to be 0
  move-to one-of side-walks with [ not any? turtles-on self ]   ; move pedestrian onto the empty side walks
  set money random max-money              ;initialize the money that a pedestrian has
  set family random max-family            ;initialize the family time that a pedestrian has
  set education random max-education      ;initialize the education level that a pedestrian has
  set materialism random max-materialism  ;initialize the materialism that a pedestrian has
  set list-foes []                        ;initialize the foe list
  set from one-of goal-candidates         ; choose at random a location to start
  set goto destination-choice-people from ; choose the destination according to its decision making
  choose-best-direction-people           ; set up the best direction that pedestrians are facing
  set jaywalk false                      ; pedestrian is not jaywalking initially
  set time-to-dest-people 0              ; initialize the time to reach the destination to be 0
  set time-to-dest-incr-people 0         ; initialize the time spent on the way to the destination to be 0
]
```

The outputs from 100 runs are illustrated in the following graphs. As we can see from the plot, the perceived risk factor for HV, which is exponentially distributed, demonstrates a more systematic impact on the number of traffic offences (negatively correlated to the perceived risk factor of HV) (Figure 5).

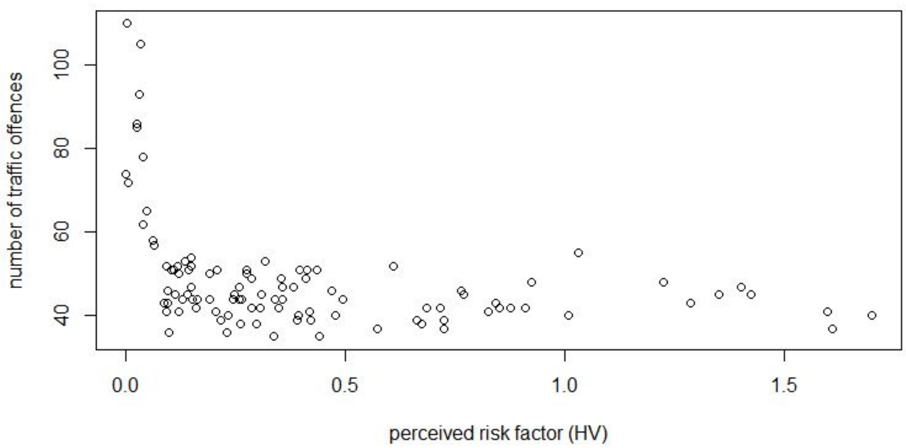


Figure 5: Monte Carlo analysis of red light parameter for HVs versus number of traffic offences

On the other hand, the normally distributed perceived risk factor for pedestrian does not show as much impact on the total number of traffic offences as the perceived risk factor for HV does.

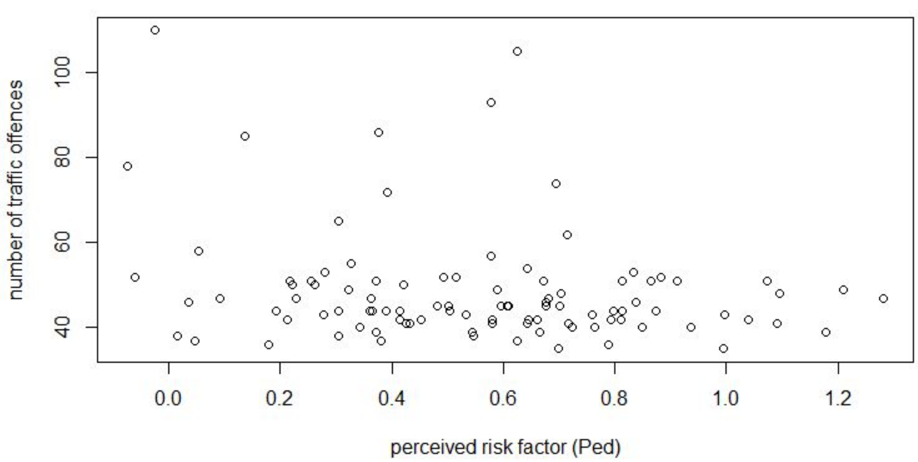


Figure 6: Monte Carlo analysis of red light parameter for Pedestrians versus number of traffic offence

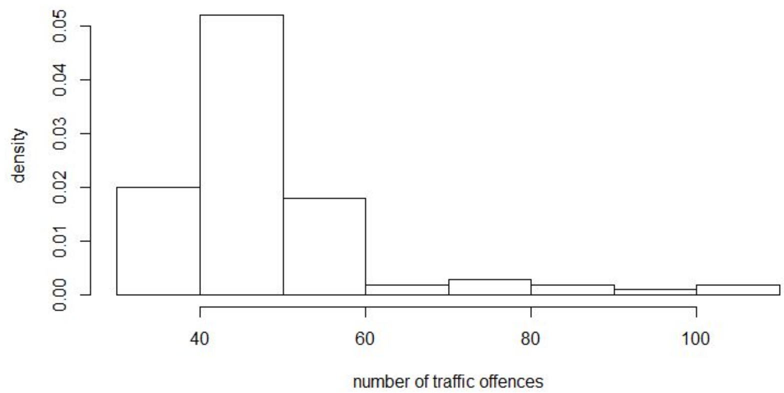


Figure 7: Histogram of number of traffic offences from combinatorial effect

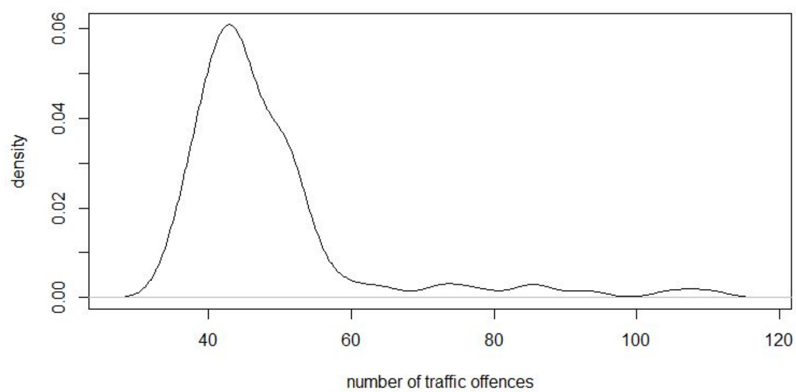


Figure 8: Density of number of traffic offences from combinatorial effect

The combinatorial impact of the two input parameters are illustrated in the above graphs. Based on simulation output, the combinatorial effect from the two input parameters makes the number of traffic offences most likely be around 40, controlling all other factors (Figure 7 & 8).

IV. Model Analysis Design

<u>Parametric Sensitivity MEA</u>

Using the 2^k factorial design we have designed a sensitivity analysis intended to find significant factors affecting our model. Specifically, we will be looking at the effects on average emotion levels, population counts, and percentage of lawbreakers.

Input Factors:

- Parking ratio
- Percentage of AVs
- Destination Utility for Work
- Destination Utility for School
- Destination Utility for Home
- Destination Utility for Shop

Using these six factors, we will have $2^6 = 64$ combinations and therefore 64 possible responses. We will simplify this to 8 runs by applying the L8 orthogonal array. In general when doing a sensitivity analysis in MEA, you often do not choose the very extreme points for the + and - ends. Something close to there is more suitable.

Table 6: Apply the L8 orthogonal array

A	B	C	D	E	F
-1	-1	-1	-1	-1	-1
-1	-1	-1	1	1	1
-1	1	1	-1	-1	1
-1	1	1	1	1	-1
1	-1	1	-1	1	-1
1	-1	1	1	-1	1
1	1	-1	-1	1	1
1	1	-1	1	-1	-1

Parking Ratio:

Parking ratio is an input that indicates the ratio of parking spots available for the number of HVs on the road i.e. if there are 50 HVs and the ratio is 100% then there will be 50 spots distributed amongst the 4 locations. If there are any HVs in the model, then the minimum number of parking spots would be 4 since we automatically place a parking spot in each location. Parking ratio therefore ranges from 0 to 100 and number of parking spots range from 4 to the number of HVs if the number of HVs is greater than 0.

Factor	Low	High
Parking Ratio	10	90

Percentage of AVs:

Similar to parking ratios, percentage of AVs can range from 0 to 100. This value indicates the percentage of car users that use AVs versus HVs.

Factor	Low	High
Percentage of AVs	10	90

Destination Utility:

Below are the possible destination utilities for both HVs and Pedestrians for work, school, home, and shop. For each location and for each agent type these values range from high, medium, and low and this is based on their level of need to reach their destination. In our sensitivity analysis we will be looking at the values for low, medium, and high levels of need. By analyzing this range of input floats from 0 to 2 we can see the affected destination utility output. Once again as mentioned above, in MEA it is not common practice to use the highest and lowest value, so we used 0.2 and 1.8 as our range of inputs.

Destination Utilities HV			*range: 0 - 0.42*
level of need	multiplier	value	
Work			
high - 2	0.21	0.42	
med - 1	0.21	0.21	
low - 0	0.21	0	
School			
high - 2	0.05	0.1	
med - 1	0.05	0.05	
low - 0	0.05	0	
Shop			
high - 2	0.02	0.04	
med - 1	0.02	0.02	
low - 0	0.02	0	
Home			
high - 2	0.08	0.16	
med - 1	0.08	0.08	
low - 0	0.08	0	

Destination Utilities Pedestrians			*range: 0 - 0.8*
level of need	multiplier	value	
Work			
high - 2	0.4	0.8	
med - 1	0.4	0.4	
low - 0	0.4	0	
School			
high - 2	0.12	0.24	
med - 1	0.12	0.12	
low - 0	0.12	0	
Shop			
high - 2	0.04	0.08	
med - 1	0.04	0.04	
low - 0	0.04	0	
Home			
high - 2	0.18	0.36	
med - 1	0.18	0.18	
low - 0	0.18	0	

Factors	Input Low	Input High	Output Low	Output High
Destination Utility for Work	0.2	1.8	0.042	0.720
Destination Utility for School	0.2	1.8	0.01	0.216
Destination Utility for Home	0.2	1.8	0.016	0.324
Destination Utility for Shop	0.2	1.8	0.004	0.072

Below is our problem statement in which Y is our outcome (MOP values of number of offences and number of deaths) and X are our factors for our factorial design:

H: Y = f(Parking ratio, Percentage of AVs, Destination Work, Destination School, Destination Home, Destination Shop)

We hypothesize that parking ratio will affect the average emotion level for HVs. The more parking spots there are available, the more stable HVs emotions will be. Stable emotions may have a slight effect on the number of traffic offenses and therefore the total population count. Destination utility will most likely have the highest effect on the number of traffic offenses and the population. This is because the higher the destination utility, the more likely an agent is to de-prioritize laws and safety. Percentage of AVs will affect average emotions in that AVs don't experience emotions and it will also affect the number of traffic offenses and the population count. AVs follow the laws and therefore less accidents are likely to occur.

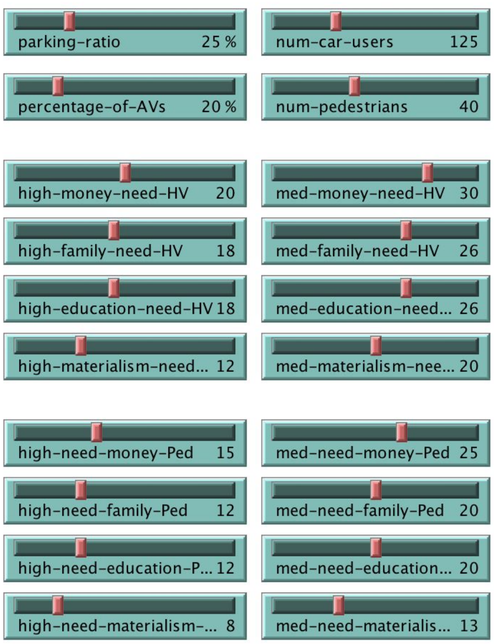
Using Main Effect Analysis (MEA) later in this report, we will be able to identify which factors are significant and which have little effect. This is done by testing the effect of independent variables (parking ratio, percentage of AVs, and destination utility weights) on a dependent variable (emotion level, number of traffic offenses, and population count).

<u>"What-if-Scenario"</u>

In our "What-if-Scenario" we will be analyzing what happens to our model results when all agents have the same risk factor. We will be looking at when all agents are risk averse with a risk factor of 1 and when all agents are risk seeking with a risk factor of 4.

Inputs:

- Base Case Scenario:
- Number of car users: 125
- Number of pedestrians: 40
- Percentage of AVs: 20%
- Level of need for reaching each destination
- Number of ticks: ~1000



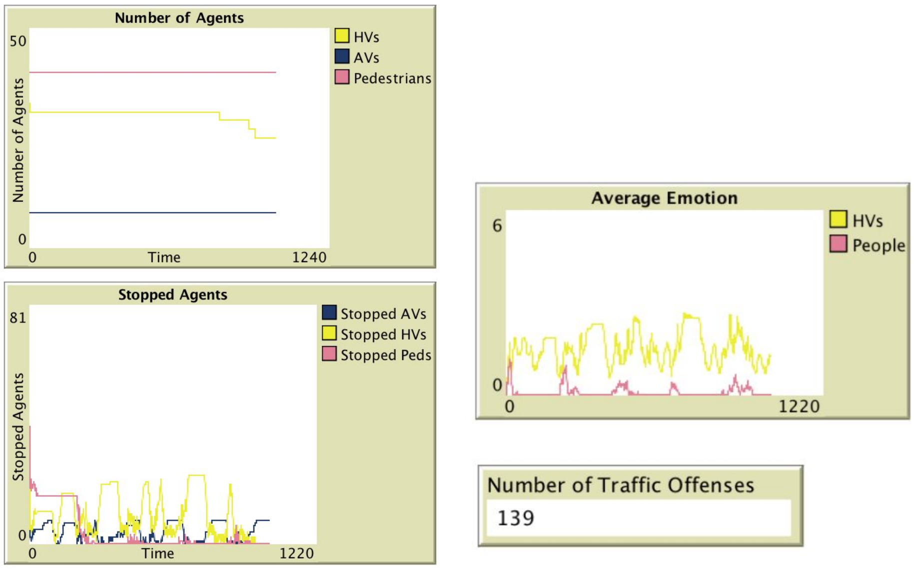
Outputs:

- Final population count for each agent
- Average number of stopped agents of each type
- Number of Traffic Offenses (instances of jaywalking)
- Average emotion level of HVs and Pedestrians

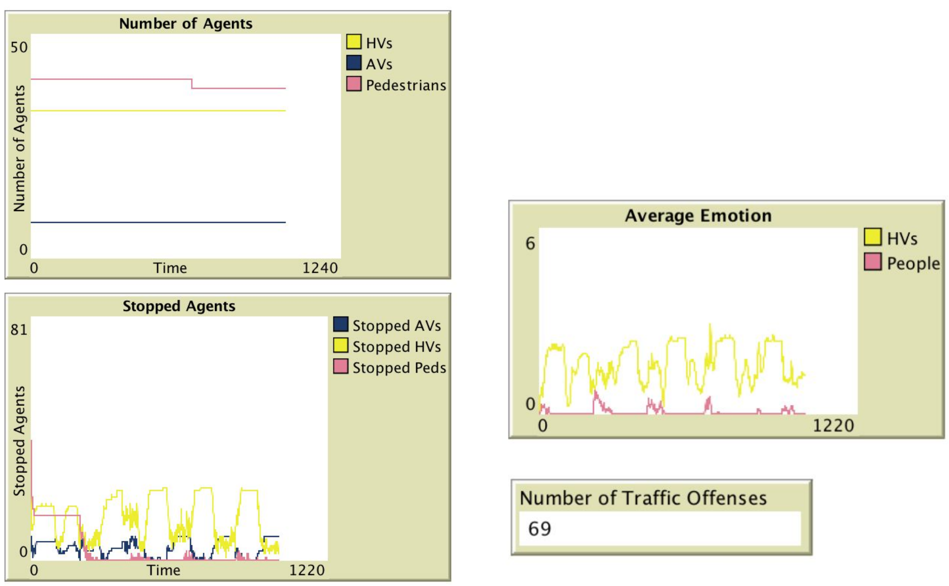
We are looking for both a numerical output and behavioral pattern with our what-if-scenario. Our outputs are looking for numeric values of population change and number of traffic offenses as well as behavioral components that check for changes in average emotion levels. We used a run length of 1,000 time ticks in the NetLogo model. In previous analysis, trends seem to stabilize around this time. We used 3 replications in a trial.

Our hypothesis for this what-if-scenario is that all HVs and pedestrians having the same risk factor will slightly alter the population number and number of traffic offenses, but will most likely not alter the emotion levels. In order to test a few different scenarios, we decided to see what would happen if everyone had a risk factor of 1 (most risk averse) and 4 (most risk seeking). Since our model already has randomization, it is a non-deterministic model. We therefore were unable to determine our base case with just a single run but rather by averaging three runs. I found that all three runs had similar results and have therefore chosen to display the results from one of the runs.

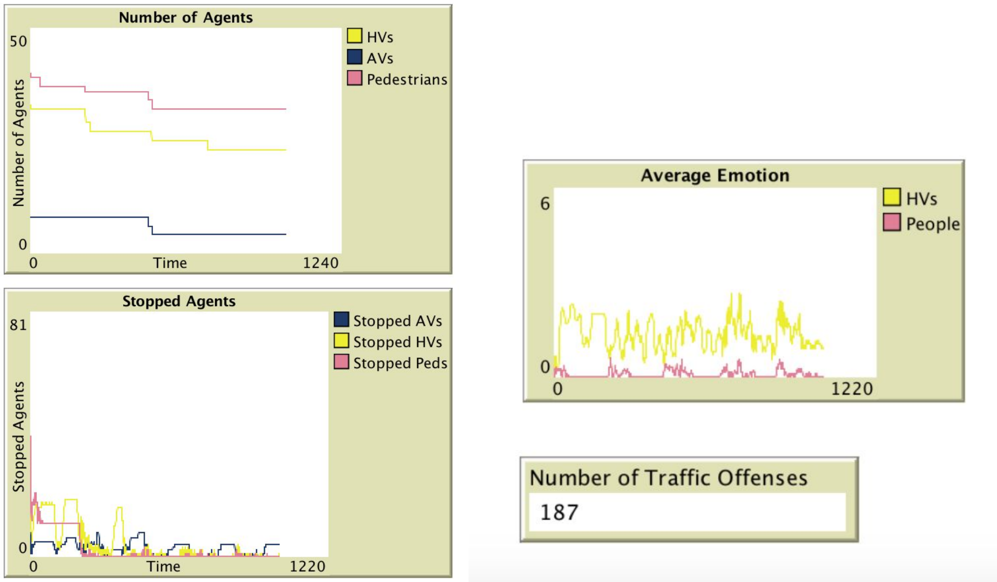
Base Case Outputs with original randomized risk levels



Base Case Outputs with equal risk level of 1



Base Case Outputs with equal risk level of 4



From this analysis the main outcomes that stick out are the number of remaining agents and the number of traffic offenses. It is clear that when all agents have a risk factor of 1 (most risk averse) the number of traffic offenses is lowest at 69 offenses, while when all agents have a risk factor of 4 (most risk seeking) the number of traffic offenses is highest at 187 offenses. The randomized risk factor model shows an outcome of 139 offenses. Additionally, when all agents have a risk factor of 4, more agents of all types seem to be dying whereas when all agents have a risk factor of 1 it is very rare that any accidents occur.

V. Output Analysis

Sensitivity Analysis

After running our parametric sensitivity analysis with respect to the six input parameters that we used in our MEA design, we analyzed the results across three replications and two outputs (number of traffic offences and number of deaths) and found that the most significant factor affecting the number of traffic offenses was Destination Utility for Home and the most significant factor affecting the number of deaths was Destination Utility for Work.

	Number of Offences					Number of Deaths				
system	X1	X2	X3	mean	sd	X1	X2	X3	mean	sd
1	117	110	530	252.33	240.49	0	0	6	2	3.46
2	908	836	441	728.33	251.43	2	2	8	4	3.46
3	411	42	41	164.67	213.33	14	4	4	7.33	5.77
4	47	48	39	44.67	4.93	4	2	0	2	2
5	37	225	172	144.67	96.93	4	20	20	14.67	9.24
6	265	744	175	394.67	305.86	12	6	20	12.67	7.02
7	846	48	43	312.33	462.18	2	0	0	0.67	1.15
8	49	41	51	47	5.29	0	2	0	0.67	1.15

Output from three replications

Our MEA plots and corresponding sensitivity analysis can be found below:

Number of Offences

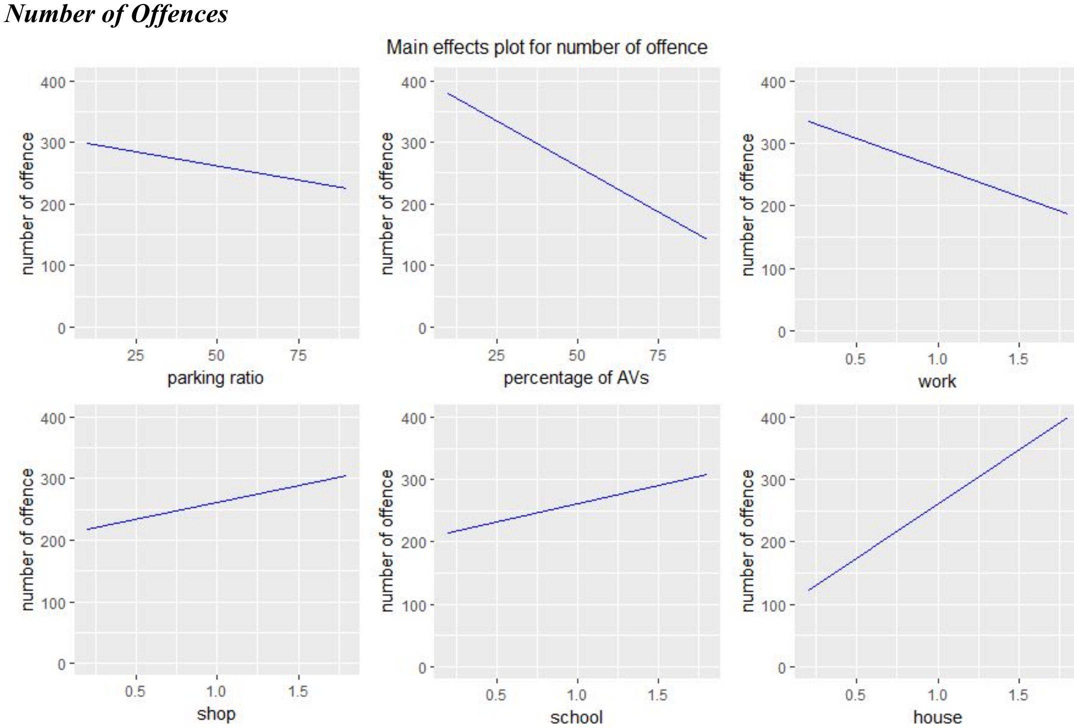


Figure 9: MEA plots for number of offences

It is clear from these plots that the most significant factors in affecting the number of traffic offences are the percentage of AVs and Destination Utility for Home. This makes sense with our initial hypotheses in that as the percentage of AVs increase, the number of traffic offences decrease and as the need to reach home increase, the number of traffic offenses increase.

```
> fit1 <- lm(mean ~ park_ratio+percent_av+work_u+shop_u+school_u+house_u, data = output1)
> summary(fit1)

Call:
lm(formula = mean ~ park_ratio + percent_av + work_u + shop_u +
    school_u + house_u, data = output1)

Residuals:
     1      2      3      4      5      6      7      8
-10.08  10.08  10.08 -10.08  10.08 -10.08 -10.08  10.08

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept) 262.7462    32.4800   8.089   0.0783 .
park_ratio   -0.9104     0.2521  -3.611   0.1720
percent_av   -2.9729     0.2521 -11.793   0.0539 .
work_u      -92.3922    12.6047  -7.330   0.0863 .
shop_u       53.2297    12.6047   4.223   0.1480
school_u     58.0203    12.6047   4.603   0.1362
house_u     173.6453    12.6047  13.776   0.0461 *
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 28.52 on 1 degrees of freedom
Multiple R-squared:  0.9977,    Adjusted R-squared:  0.9839
F-statistic: 72.44 on 6 and 1 DF,  p-value: 0.08969
```

From the magnitude and of coefficient estimates above, we can see that Destination Utility for Home is the most significant factor for number of traffic offences. Due to the different units of analysis for the two ratio parameters and four value parameters, the significance should be compared by groups.

Number of Deaths

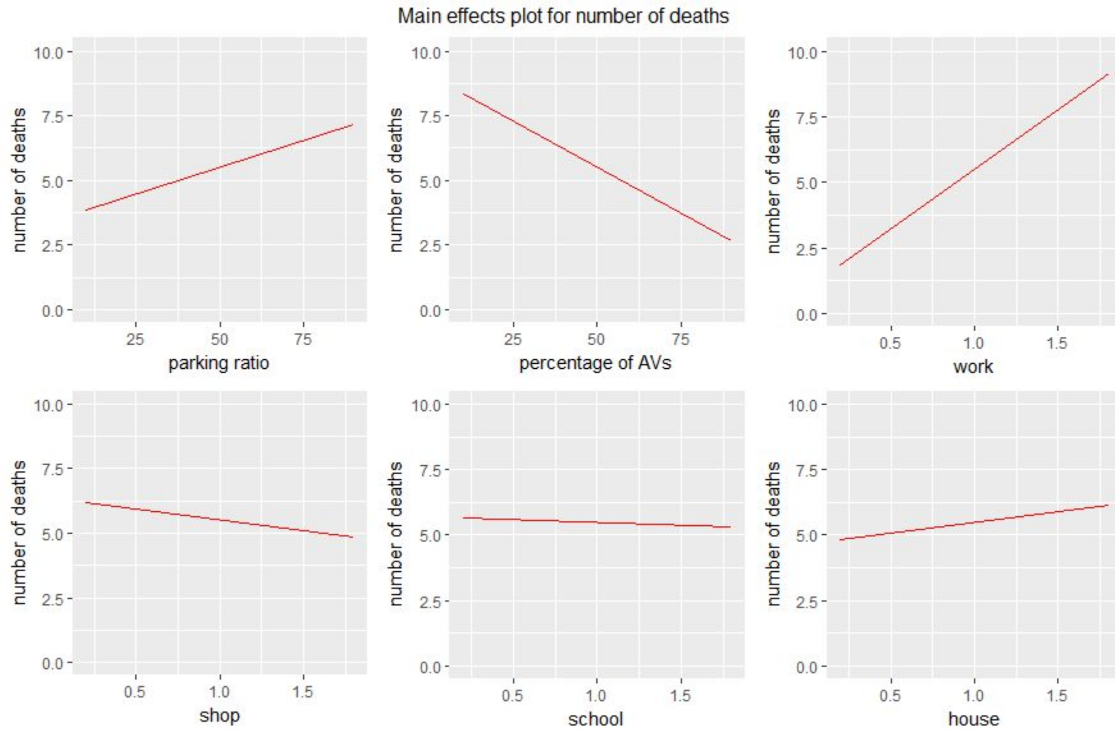


Figure 10: MEA plots for number of deaths

Likewise, it is clear from these plots that the most significant factors in affecting the number of deaths are the percentage of AVs and Destination Utility for Work. This makes sense with our initial hypotheses in that as the percentage of AVs increase, the number of deaths decrease and as the need to reach work increase, the number of deaths increase.

```
> fit2 <- lm(mean ~ park_ratio+percent_av+work_u+shop_u+school_u+house_u, data = output2)
> summary(fit2)

Call:
lm(formula = mean ~ park_ratio + percent_av + work_u + shop_u +
    school_u + house_u, data = output2)

Residuals:
     1      2      3      4      5      6      7      8
-1.166  1.166  1.166 -1.166  1.166 -1.166 -1.166  1.166

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)  2.58250    3.75652   0.687    0.617
park_ratio   0.04172    0.02916   1.431    0.388
percent_av  -0.07084    0.02916  -2.430    0.249
work_u       4.58281    1.45781   3.144    0.196
shop_u      -0.83281    1.45781  -0.571    0.670
school_u    -0.20781    1.45781  -0.143    0.910
house_u      0.83281    1.45781   0.571    0.670

Residual standard error: 3.299 on 1 degrees of freedom
Multiple R-squared:  0.9487,    Adjusted R-squared:  0.6411
F-statistic: 3.084 on 6 and 1 DF,  p-value: 0.4102
```

What-if-Scenario of Most Significant Factor

From our MEA analysis above, we found that **destination utility for work** was the most significant factor affecting our outcome of **number of deaths**, while **destination utility for home** was the most significant factor for **number of offence**. We then ran our "What-if-Scenario" (i.e. all agents' risk factor is 4) while analyzing the changes in destination utility for work and its effects on the number of total deaths of agents.

Below is our problem statement in which Y is our outcome (MOP value number of deaths) and X is our most significant factor from our sensitivity analysis done above:

H: Y = f(Destination Utility Work)

In this "What-if-Scenario", we ran three iterations for a sequence from 0.2 (low value) to 1.8 (high value) with an increment of 0.01. Within each iteration, there are 161 runs with 1,000 steps (ticks).

```
dat <- seq(0.2,1.8,0.01)

deaths <- list()
work_utility <- list()

for (i in 1:161){
  NLCommand("setup")
  NLCommand("set MEA_work", dat[i])
  a <- dat[i]
  print(a)
  NLDoCommand(1000,"go")
  work_utility[[i]] <- dat[i]
  deaths[[i]] <- NLReport(c("number_deaths"))}

whatif_output_work <- data.frame(t(sapply(work_utility,c)))
whatif_output_deaths <- data.frame(t(sapply(deaths,c)))
whatif_output <- rbind(whatif_output_work,whatif_output_deaths)
```

We then computed the mean and variance of outputs from three iterations of the whatif scenarios with risk-factor of 4, each running 100 ticks. We analyzed the results from the three iterations. The following graph plots the mean value of output (number of deaths) against different values of destination utility of work. We fitted this relationship by both linear and nonlinear functions. Since the results are a little dispersed in a small range, it is not definite whether such relationship is nonlinear.

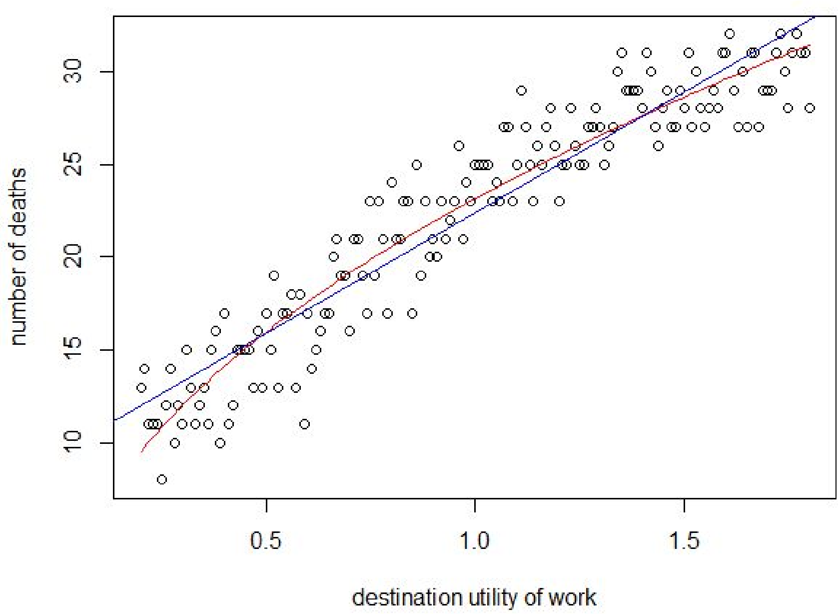


Figure 11: Outcome of most significant factor using our "What-if-Scenario"

According to the *figure 11* above and the statistical test below, both fits are statistically significant with an adjusted R-squared of 0.8851. We can see that as destination utility of work increase, number of deaths increase as well. This seems sound with our model since as agents begin to place a higher need and importance on reaching work they will be more likely to make risky and unsafe decisions at intersections and more deaths will result.

```
Call:
lm(formula = whatif$deaths ~ whatif$work)

Residuals:
    Min      1Q  Median      3Q     Max 
-6.0312 -1.4804 -0.0523  1.6039  5.2179 

Coefficients:
            Estimate Std. Error t value Pr(>|t|)    
(Intercept)   9.3715     0.4076   22.99   <2e-16 ***
whatif$work  12.9825     0.3696   35.12   <2e-16 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 2.18 on 159 degrees of freedom
Multiple R-squared:  0.8858,	Adjusted R-squared:  0.8851 
F-statistic:  1234 on 1 and 159 DF,  p-value: < 2.2e-16
```

```
Formula: deaths ~ a + b * work^(-c)

Parameters:
  Estimate Std. Error t value Pr(>|t|)    
a -2.18415    4.48738  -0.487    0.627    
b 25.33801    4.63307   5.469 1.74e-07 ***
c -0.48103    0.09682  -4.968 1.74e-06 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 2.008 on 158 degrees of freedom

Number of iterations to convergence: 12 
Achieved convergence tolerance: 3.887e-06
```

VI. Conclusion

In Midterm 1 we focused on the effects of autonomous vehicles on traffic congestion and parking. We found that AVs are able to meet their destination goals more quickly and accommodate over 10x the amount of people as HVs. Therefore, larger percentages of AVs lead to less cars on the road and therefore less congestion. We are also able to see from running tests on our simulation from Midterm 1 that the initial introduction of AVs could actually worsen traffic and parking conditions; however, once the AV percentage gets to 80% and then finally full adoption, improvements are clearly seen.

Moving into Midterm 2, we added MODM and emotions for our agents. We decided to focus on the decision of which destination to go to based on a value system incorporated in our model as well as safety decisions made at intersections. Due to the addition of values for destinations, we were also able to see that congestion on the work and house roads were higher than that of the shop road since agents value going to work and home more than going to the shop.In terms of safety, we were able to see from our final results that pedestrians tend to decrease overall safety at intersections. Pedestrians are emotional and often don't follow traffic rules; this leads to higher emotions amongst HVs and more accidents. We also saw that when we increased the percentage of AVs, there was less of a chance of traffic gridlock, and overall safety conditions were much higher. Overall, decreasing the percentage of AVs and increasing the number of pedestrians in our model led to more jaywalkers and more agent deaths.

Midterm 3 wrapped up our model analysis by incorporating randomness to our model. We were able to use sensitivity analysis techniques learned in class to determine the most important factors in our model for analyzing certain output values. Our main results from our MEA indicated that percentage of AVs and destination utility for work and home were the most sensitive factors affecting number of deaths and number of traffic offences in our model. Additionally, we were able to see the value of our risk factor variable, by using our What-if-Scenario in which we set all agents' risk factor to be equal. It was quite clear that when all agents had a risk factor of 1 (most risk averse), number of traffic offences was lower; while when all agents' had a risk factor of 4 (most risk seeking), the number of traffic offences was higher. Tying back to our initial goal from Midterm 1 in which we aimed to analyze the traffic conditions and safety effects of increasing the number of AVs on the road, we can see that increasing the percentage of AVs does in fact seem to improve traffic and safety conditions by decreasing the number of traffic offenses, deaths, and overall congestion on the roads.

VII. References

Florida Department of Transportation, *Understanding Interactions between Drivers and Pedestrian Features at Signalized Intersections*, USF Center for Urban Transportation Research, 2015.

World Health Organization, *World Report on Road Traffic Injury Prevention*, 2004.

Resnik, Michael D. *Choices: An introduction to Decision Theory*. University of Minnesota Press, 1987.

VIII. Appendix:

Refer to R-code and NetLogo submitted for additional reference.