Name: Weijie Xu
CNetID: weijiexu
Student ID: 12245277

Problem 1:

Theory I:

In theory I, there are two factors that would influence the outcome of the socially deviant aggressive behaviors: the rewards and the punishment that an individual would receive for his behavior. If we build a model based on this theory, the outcome of the socially deviant aggressive behaviors can be measured by its utility. That is, a higher utility of a certain behavior would yield to a higher commitment rate of such behavior. In this model, the dependent variable is the utility of the socially deviant aggressive behaviors, while the independent variables are the rewards and the punishment incurred by the individual's commitment to such deviant aggressive behaviors. Furthermore, in this model, the utility of socially deviant behaviors is positively correlated with the variable of rewards, whereas negatively correlated with the variable of punishment.

Given the model above, there are two ways for the policymaker to reduce the utility of socially deviant behaviors. Firstly, since higher rewards would lead to a higher outcome of the deviant behaviors, policymaker should look for ways to reduce the rewards that an individual would receive from such behaviors. Secondly, based on the negative correlation between punishment and socially deviant behaviors, making more punishment to increase the cost of such behavior would be another method.

Theory II:

Like what we did in the case of theory I, we can also build up a model based on theory II. In this theory, an individual has a certain amount of hostility to be expressed through different ways including socially deviant aggressive behaviors. Therefore, the dependent variable in the model we could build here is the amount of hostility to be expressed, and the independent variables are a series of outlets to express such hostility. The hostility expressed through these outlets is summed up to the amount of the hostility that a person needs to express.

Based on this model, the goal of the policymaker is to decrease the weight of the term of socially deviant aggressive behaviors. To achieve this, one solution is to add more independent variables to the model, so that, given a fixed total amount of hostility to be expressed, the contribution of socially deviant behaviors would be diluted by other variables. Therefore, the policymaker should seek to provide more emotional outlets, such as sports events, so that people could express their hostility in different ways.

Theory III:

In this theory, the factor that would influence the outcome of socially deviant aggressive behaviors is the discrimination of policy regarding such behaviors towards the individual. For example, some groups of people, if committing socially deviant

behaviors, are more likely to be punished to a more serious extent than other groups of people. Therefore, we can model the outcome of the socially deviant behaviors by measuring the variance of punishment of such behaviors across different demographic groups. The higher the variance is in a certain region, the higher the outcome of the socially deviant behaviors in that region would turn out to be.

Based on the positive correlation in this model, what the policymaker is supposed to do is to reduce the variance of punishment of socially deviant behaviors, which is to say that, to make the regulation of punishment fairer across different groups of people. Therefore, since people are less likely to be discriminated by the punishment, they are more willing to conform to policy and commit less socially deviant behaviors.

Theory IV:

In theory IV, the outcome of the socially deviant aggressive is influenced by the interaction with other people. Therefore, based on this theory, we would build a model whose independent variables would be the outcome of socially deviant behaviors of other people in the society. Then, the outcome of socially deviant behaviors of a specific individual is the weighted sum of the values of all the independent variables. And the weight of each term would reflect the intensity of interaction with the person represented by that variable.

A possible method to address the problem of socially deviant behaviors based on the model above is to reduce the weight of each term or variable in the contributed to the outcome of such behaviors of an individual. Since the weights, as explained above, refer to the interaction between people, what the policymaker is supposed to do is to intervene in the interaction between individuals committing deviant behaviors. For example, during the process of urban planning, the city should encourage community diversity and try to prevent demographic segregation, in order to reduce the interaction of people who have the same socially deviant behaviors.

Problem 2:

A possible explanation for the procrastination consists in the starting energy. That is, each task requires a certain amount of energy to start to work on it. Only when people's starting energy is over that threshold will they start to do things. Furthermore, the starting energy for an individual increases all the way up until the deadline, which means that the starting energy will be higher when it's closer to the deadline. Therefore, people always "wait until the last minute" because only at that time the starting energy is over the threshold for people to work on their tasks.

The model we would build based on the explanation above consists of two parts. Firstly, we would like to define a starting energy function in terms of the time left before the deadline. The less time left before the deadline, the higher the starting energy an individual would have. On the other hand, a threshold of starting energy is set for each task to be completed. Therefore, if we combine the starting energy function and the pre-set threshold together, we would get a point where the value of the starting energy function meets the threshold, and the time corresponding to this point is when an individual starts to work on his task.

This model would result in two predictions. Firstly, if the task is difficult, which means that the threshold of starting energy is relatively high, then people would be expected to start to work on their task later when the value of the starting energy gets high enough, and vice versa. Secondly, if there are two individuals with different instantiations of starting energy function, then the individual whose starting energy function is systematically higher is expected to start to work on his task earlier than the one whose function is systematically at a lower level.

An alternative possible explanation for the procrastination is that people tend to maximize the work efficiency, which is to say that, people always try to complete as many as possible tasks in the shortest time. On the other hand, the efficiency is always the highest right before the deadline. Therefore, in pursuit of high efficiency, people always choose to "wait until the last minute" to do things.

Based on the hypothesis above, the model we would build for this explanation has an efficiency function. The dependent variable is the work efficiency of an individual, and the independent variable, similar to the model of the first explanation above, is the time left before the deadline. Also, the value of the efficiency function is increasing all the way up until the deadline. Furthermore, according to this model, the workload of an individual is the integral of the efficiency function over a period of time. As a result, given a certain fixed amount of workload, if people want to complete their tasks using the shortest time, they will choose to work on it when the value of efficiency function is high. Therefore, since the value of the efficiency function is always the highest near the deadline, people always choose to "wait until the last minute" to start working.

This model also has two predictions. Firstly, if the slope of the efficiency function is relatively even, which means that the distribution of work efficiency is even across time, then the starting time to work would also be evenly distributed across the period of time before the deadline. In another word, the phenomenon that people always "wait until the last minute" would not appear. The reason for this prediction is that whenever

people start to work, the time they use to finish their work would be almost the same, therefore, people don't need to "wait until the last minute" in pursuit of high work efficiency. The second prediction is that, if the relationship between work efficiency and time left is in the opposite way, that is, the work efficiency always decreases all the way down until the deadline, then, instead of "wait until the last minute", people would choose to start to work as early as possible.

Problem 3:

a. The goal of selecting and fitting a model is to minimize the error function but also avoid over-fitting at the same time. In this case, a more flexible statistical learning method would be better than an inflexible one. Firstly, since the number of predictors is small, the information we have to simulate or predict the problem at hand is limited, which means that the model we build is supposed to be much more powerful and require more flexibility. On the other hand, since we have a large sample size, the impact of the problem of overfitting is limited. As a result, the cost of overfitting for a high flexibility is now affordable given the large amount of data.
b. In this case, a less flexible statistical learning method would be better. Firstly, since we have a large number of predictors, we have much more information at hand to model the problem of interest, which means that, compared to the situation where we only have a limited number of predictors, the model itself is already very complex and powerful. But the problem is that, the model may be too complex and powerful given the limited amount of data, which may lead to a serious problem of overfitting. Therefore, in order to avoid overfitting, a less flexible statistical learning method may be more appropriate.
c. In this case, a more flexible statistical learning method would be better than an inflexible one. This is because, given the highly non-linear relationship between the predictors and response, the problem itself of our interest to be modeled is very complex. Therefore, a simple inflexible linear model is no longer enough, and we thus need a more complex and flexible model in order to better simulate the complex problem.
d. In this case, a less flexible statistical learning method would be better. The high variance of the error term means that the data is full of noise. Therefore, if we tune our model to fit the data in a more flexible way, a large amount of noisy error would be added to the simulation, which would make our model more susceptible to the problem of overfitting and make it less powerful to predict unseen data. As a result, in order to avoid overfitting, a less flexible learning method is more appropriate.

Problem 4

a. The bias decreases when the complexity and flexibility is getting higher. This is because the bias is the error between the expected estimate value and the true target value given the training data. Therefore, if the model is more flexible, it would be easier to fit the training data, which would thus reduce the bias.

b. The variance increases when the complexity and flexibility is getting higher. The reason is that, if we imagine that we implement statistical learning many times by drawing multiple data sets, then variance refers to the extent to which the estimate value of a single data point fluctuates around the average estimate value of these learnings. Therefore, if the flexibility of the model increases, then the estimation would be more tuned to a specific data set instead of taking all the data sets into consideration comprehensively. This problem of overfitting would then lead to higher variance.

c. The training error decreases with higher complexity and flexibility. The reason is that when the model becomes more flexible, it could better capture the characteristics of the training data set, which would help the estimated value to be tuned closer to the true target value in the training data. Therefore, the training error would decrease when the model flexibility increases.

d. The testing error decreases at the beginning when the complexity and flexibility is getting higher. The reason is that, by capturing the characteristics of the training data, the model is also capturing the characteristics of the problem of our research interest. Therefore, at this stage, with the increasing flexibility, the model is more and more able to be generalized and to predict new data. However, in the later stage, the testing error increases when the model becomes more flexible. This is because high flexibility would render the model describe more attributes specific to the training data set, which means that the more randomness is taken into account and the problem of overfitting would occur. Therefore, the model is hard to be generalized to other data set and would then make poor predictions. As a result, the testing error would increase in the end when the flexibility of the model is getting further higher.

e. The irreducible error, as indicated by its name, cannot be reducible no matter how we build up the model and how we change the model complexity and flexibility. This is because of the variability and uncertainty of the sampling process itself. This property is inherent to the problem itself and cannot be reduced with any method.