

University of Montana

ScholarWorks at University of Montana

Undergraduate Theses and Professional Papers

2019

Investigating a Macro Level Trial Tax: An Analysis of the Relationship Between a Circuit's Plea Proportion and Median Sentence Length in U.S. Federal Courts

Denise LaFontaine

University of Montana, Missoula, denise.lafontaine@umontana.edu

Follow this and additional works at: <https://scholarworks.umt.edu/utpp>



Part of the [Criminal Law Commons](#), and the [Criminology Commons](#)

Let us know how access to this document benefits you.

Recommended Citation

LaFontaine, Denise, "Investigating a Macro Level Trial Tax: An Analysis of the Relationship Between a Circuit's Plea Proportion and Median Sentence Length in U.S. Federal Courts" (2019). *Undergraduate Theses and Professional Papers*. 254.

<https://scholarworks.umt.edu/utpp/254>

This Thesis is brought to you for free and open access by ScholarWorks at University of Montana. It has been accepted for inclusion in Undergraduate Theses and Professional Papers by an authorized administrator of ScholarWorks at University of Montana. For more information, please contact scholarworks@mso.umt.edu.

Investigating a Macro Level Trial Tax:
An Analysis of the Relationship Between a Circuit's Plea Proportion and Median Sentence Length in U.S.
Federal Courts

Denise LaFontaine

Abstract: This study investigates the relationship between the proportion of cases resolved by plea in the federal court circuits and the median sentence length of those circuits. Through OLS regression, this study models the relationship, controlling for variation in circuits, year, and violent crime proportion. With an understanding of two common theoretical frameworks surrounding the trial tax discussion, this study theorized under a focal concerns perspective that as the proportion of cases resolved by plea in a circuit increases, the median sentence length of that circuit would decrease. The results of this study are consistent with results of prior research showing the existence of a trial tax. Additionally, this study investigates the relationship between the caseload of a circuit and the proportion of cases resolved by plea in a circuit to understand the extent to which caseload influences plea bargaining practices in a circuit. The results of this analysis suggest that caseload differences have little influence on plea proportions when the caseloads are below one hundred. However, caseloads larger than this do tend to be correlated with high plea proportions, supporting the focal concerns theory prediction of increased plea pressure in high caseload courts. Overall, this study provides evidence of a macro level trial tax that is consistent with individual level findings.

I. Introduction

The Sixth Amendment of the Constitution reads “In all criminal prosecutions, the accused shall enjoy the right to a speedy and public trial.” (United States Constitution, 1789). The founding fathers deemed this right imperative for due process. However, in today’s justice system, less than 10% of criminal defendants assert this right (Kim, 2015; Ulmer & Bradley, 2006). The other 90-95% of criminal defendants resolve their cases through plea bargains (Johnson, 2019). Plea bargaining is an ongoing process that generally begins when the defendant is formally charged and continues up until trial. In this process, the prosecutor offers to lessen the charge or present specific sentence recommendations to the judge in return for the defendant accepting a guilty plea or *nolo contendere* plea (Guidorizzi, 1998; Johnson, 2019; Ulmer, Eisenstein, & Johnson, 2010). This process is heavily relied on in the American criminal justice system because it helps balance the Crime Control and Due Process models of justice.

The differences in these two models lies in their primary goals. The Crime Control model of justice prioritizes crime suppression and efficiency in case disposition (Packer, 1963). A justice system built solely on this interest would process cases quickly with little regard for defendant’s due process rights, resulting in few guilty people going unpunished but a greater likelihood of innocent defendants being convicted (Packer, 1963). Conversely, the Due Process model prioritizes the rights of the defendants (Packer, 1963). A justice system based solely on this interest would process cases slowly with great scrutiny, minimizing the probability of an innocent defendant being wrongly adjudicated and punished but increasing the probability that a guilty defendant avoids conviction (Packer, 1963). Therefore, a justice system that emphasizes fair treatment of defendants and protecting the community from crime must balance these two models and their components (Packer, 1963). Plea bargaining is one successful mechanism for doing so.

Plea bargaining allows cases to be dealt with efficiently, requiring less time and fewer resources than trials (Guidorizzi, 1998; Johnson, 2019). Thus, it increases the efficiency of the justice system. Additionally, plea bargaining is optional. A defendant must knowingly and willingly relinquish their due process rights to trial and to confront the witnesses against them (Guidorizzi, 1998). Hence, theoretically, plea bargaining still values and respects the due process rights of criminal defendants. Due to large caseloads, limited resources, and plea bargaining’s ability to improve efficiency without violating individual’s rights, there is a heavy reliance on plea bargaining at all levels of the American criminal justice system (Johnson, 2019). This heavy reliance warrants a critical examination of the plea-bargaining process and its effects to make sure that it is in fact being used fairly and not infringing on defendant’s rights.

Prior research suggests that empirically, plea bargaining may not be striking the balance it theoretically should. Some scholars suggest that plea bargaining is used disproportionately against the poor, who are less able to devote time and resources to their case processing (Hall, 1999). Other studies suggest that defendants who opt for trial rather than resolve their cases through plea receive harsher punishments (Bradley-Engen, Engen, Shields, Damphousse & Smith, 2012; Bushway, Redlich, & Norris, 2014; Kim, 2015; Johnson, 2003; Johnson, 2006; Spohn, 1992; Ulmer & Bradley, 2006; Ulmer et. al, 2010, Ulmer & Johnson, 2004). This phenomenon, often called the “trial tax” or “trial penalty”, is controversial in the legal and criminal justice fields (Guidorizzi, 1998; Johnson, 2019; Ulmer et. al, 2010). Some argue that a trial tax is necessary. In order to discourage individuals from taking their case to trial and using valuable system resources, there must be a threat of greater punishment for doing so (Steffensmeier, Kramer & Ulmer, 1998; Ulmer et. al, 2010). Otherwise, everyone would assert their right to trial, and the issues of backlog and scarce resources already faced by the criminal justice system would be exacerbated. Meanwhile, others argue that it is less of a “trial tax” than a “plea discount” (Bushway et al., 2014; Johnson, 2019). However, either way one chooses to look at it, evidence suggests that those who opt to go to trial are often receiving harsher punishments than those who agree to plead guilty

(Bradley-Engen et al., 2012; Bushway et al. 2014; Kim, 2015; Johnson, 2003; Johnson, 2006; Spohn, 1992; Ulmer & Bradley, 2006; Ulmer et. al, 2010, Ulmer & Johnson, 2004). Therefore, the existence of a “trial tax/plea discount” means that defendants are penalized for asserting their Constitutional right to trial.

This study investigates the existence of a trial tax on a macro level. Using federal sentencing data from the United States Sentencing Commission, this study analyzes the relationship between the median sentence length of the federal courts in the U.S. Circuits (excluding the District of Columbia Circuit) and the proportion of cases resolved by plea in those circuits. If a trial tax does exist, circuit courts that have a lower proportion of cases resolved by plea should have a higher median sentence length. Conversely, circuit courts with a higher proportion of cases resolved by plea should have a lower median sentence length if a trial tax exists. This study identifies a nonlinear model that explains the relationship between plea bargaining and sentencing length at a macro level while controlling for other variables. Evidence of a trial tax is found.

II. *Conditional or Unconditional*

In addition to the trial tax’s potential impact on a Constitutional right, it has been controversial because there is conflicting evidence on its existence. Many studies have shown that trial convictions tend to result in harsher punishments—either in the sanction itself (Johnson, 2006; Spohn, 1992; Ulmer & Bradley, 2006) or in its duration (Bushway et al., 2014; Engen et al., 2012; Johnson, 2006; Kim, 2015; Spohn, 1992; Ulmer & Bradley, 2006)—than convictions by guilty or *nolo contendere* pleas. However, a couple prominent studies resulted in no evidence of a trial tax and even some evidence in support of a trial discount (Abrams, 2013; Breen, 2011; Breen & Johnson, 2018). While these differences in findings have created much controversy, a careful examination of the methods used in these studies reveals that these differences can be explained in terms of conditionality.

Both types of studies essentially compare the expected sentence length for a defendant convicted by plea to a similar defendant convicted by trial to see if differences exist. However, one type of study uses the conditional expected value while one uses the unconditional expected value (Abrams, 2013). While these may seem the same on the surface, in reality they are not. For example, think about the expected number of heads when a coin is flipped ten times. Most people would say that the expected number of heads is five. However, if the coin being flipped was a coin with heads on both sides, the expected number of heads would increase to ten. This is the conditional expected value. The additional knowledge that the coin has heads on both sides changes the outcomes of the event. This same nuance is seen in trial tax research.

Whether a researcher looks at the expected sentence length of those who have been convicted versus the whole population of defendants matters greatly. This is because a portion of the defendants who assert their right to trial will be acquitted (Abrams, 2013). Whether or not these defendants are included in the calculations turns out to be very important. In fact, most studies that exclude these defendants from their data sets find evidence of a trial tax, while those who include them tend to find no evidence of a trial tax (Abrams, 2013; Kim, 2015). The conditional and unconditional expected values for those who plead guilty are the same whether one looks at those convicted or the whole population of criminal defendants. This is because a guilty plea always results in a conviction. However, this is not the case for trial. The unconditional expected sentence length of a trial is the expected sentence length if convicted multiplied by the probability of conviction plus the expected sentence length if acquitted (0) multiplied by the probability of acquittal (Abrams, 2013)¹. Since this average includes many zero values,

¹ The following equation calculates the unconditional expected sentence length where S represents the sentence length, p represents probability, and C represents convicted:

the resulting number will be much less than just the average sentence length of those convicted at trial. Hence, the conditional expected sentence length will always be greater than the unconditional expected sentence length for those who go to trial. Therefore, studies that use the unconditional expected sentence length find smaller differences between the expected sentence length of those who go to trial and those who plead guilty (Abrams, 2013; Breen, 2011). This results in less evidence of a trial tax in these studies.

Controversy exists over which of these methods is appropriate due to their different conclusions on the trial tax. Overall, the appropriate method to use depends on the question of the researcher. For example, if the researcher is interested in whether going to trial or taking a plea deal results in a more favorable outcome, weighing the possibility of acquittal is important. Similarly, if they want to know if the average sentence length is different after accounting for the possibility of acquittal, they would want to include this information. However, when determining the existence of a trial tax, one should use the conditional expected values. This is what most researchers and legal scholars refer to when speaking of a trial tax (Kim, 2015; Ulmer & Bradley, 2006; Ulmer et al., 2010). Further, by definition, the trial tax is the excess of punishment received by those who go to trial compared to those who plead guilty (Johnson, 2006; Kim, 2015; Ulmer & Bradley, 2006; Ulmer & Johnson, 2004). Including acquittals into this calculation would be like including the fact that any other citizen not involved with the criminal justice system has a sentence length of zero punishment. The presumption of innocence requires that no judgements be made about an individual and no actions taken against them until the State proves beyond a reasonable doubt that they are guilty (Anderson, Baradaran, Lopez, Perry, & Center, 2011; Kim, 2015). Including acquittal values of no punishment would violate this foundational principle (Kim, 2015). In the eyes of the law, someone acquitted of their charges is no guiltier than someone who has never faced charges. Thus, in the eyes of the law, including their sentence length values of zero would be no different than including a sentence length value of zero for every other person not involved in the criminal justice system and then averaging. This would not result in any sort of meaningful value. Thus, the trial tax should use conditional expected sentence lengths rather than unconditional (Kim, 2015). Accordingly, this study uses the median sentence length in a circuit such that the median is calculated only upon the non-zero sentences handed out in that circuit.

III. *Theoretical Background*

Due to the controversy over the existence of the trial tax, two prominent theories have been used to explain the relationship between the sentences of those convicted at trial and those who pled guilty. One of these theories, the focal concerns perspective, focuses on how criminal judges make sentencing decisions (Steffensmeier et al., 1998), and this in turn, has been applied to explain the existence of a trial tax (Bradley-Engen, et al., 2012; Bushway et al., 2014; Ulmer & Bradley, 2006). It argues that as a result of going to trial, a defendant is viewed more negatively by the sentencing judge (Ulmer, 1997; Ulmer & Bradley, 2006). Hence, they receive a harsher punishment. On the contrary, the shadow of the law perspective explains why theoretically, a trial tax should not exist or in the least, should be a modest compensation for the odds of acquittal (Abrams, 2013; Bibas, 2004; Bushway et. al, 2014). Based upon an economic/rational actor theory, this perspective argues that ideally, both sides will only agree to a

$$E(S) = E(S|C) * P(C) + 0 * P(C^c).$$

Compare this to the equation for the conditional expected sentence length where S_k represents the sentence length of a convicted defendant and N represents the number of convicted defendants:

$$E(S|C) = \frac{1}{N} \sum_{k=1}^N S_k.$$

plea deal with a sentence equal to what they would expect if convicted at trial (Abrams, 2013; Bibas, 2004; Kim, 2015) Therefore, the disparity in sentences for those convicted by trial and those who plea should be minimal (Abrams, 2013; Bibas, 2004). Thus, a significant trial tax should not be present. Understanding the different components of the theories will provide insight into their different conclusions on the existence of a trial tax.

The focal concerns theory has three components: blameworthiness, practical considerations, and protection of the community (Steffensmeier et al., 1998). It argues that judges primarily consider these three components to predict future behavior and analyze the 'just desert' of the offender when sentencing an individual (Steffensmeier et al., 1998). Defendants who go to trial compared to those who plead guilty are likely perceived more negatively in respects to these components (Ulmer & Bradley, 2006). Defendants convicted at trial may be perceived more negatively for a variety of reasons. One of these reasons is that defendants convicted at trial are viewed as less remorseful and less accepting of responsibility than those convicted by plea (Breen, 2011; Ulmer & Bradley, 2006). If the defendant is convicted for a violent offense against a person, their blameworthiness further intensifies (Ulmer & Bradley, 2006). This is because rather than saving the victim from the trauma of a trial and admitting to their crimes, the defendant forced them to relive it. Further, throughout the trial process the prosecution paints the defendant in the most negative light possible in order to secure a conviction (Ulmer & Bradley, 2006). "Bad facts" pertaining to the crime itself or to the defendant's character may come out during trial. These "bad facts" increase the culpability of the defendant, resulting in the judges imposing harsher sentences (Ulmer & Bradley, 2006; Ulmer et al., 2010). The same is not true for convictions by plea deal. In a majority of the plea deals, the sentencing judge typically imposes the sentence the prosecutor recommends in the plea deal (Ulmer et al., 2010). Most often, the judge knows much less information about the case or the defendant if they are only presented with a plea deal than they would following a trial (Bradley-Engen et al., 2012; Ulmer & Bradley, 2006; Ulmer et al., 2010). While some may argue that this means the judge would not know any mitigating circumstances and thus result in greater perceived blameworthiness, the structure of a trial creates a situation in which aggravating circumstances are more highly emphasized and dramatized than mitigating circumstances (Bradley-Engen et al., 2012; Ulmer & Bradley, 2006; Ulmer et al., 2010). Some evidence has been found supporting this component of the theory in a study using qualitative survey data from judges and attorneys (Ulmer et al., 2010).

The practical constraint component of the focal concerns theory is closely related to the issue with the Due Process Model of justice—organizational efficiency. As mentioned earlier, the American criminal justice system has an impossible caseload. Thus, it helps to process cases efficiently and prevent a backlog if judges either impose harsher penalties for those who go to trial or impose less severe penalties to those who plead guilty (Dixon, 1995; Ulmer & Bradley, 2006). By doing this judges encourage the use of plea deals to resolve cases and save the time and resources for more legitimate trials (Ulmer & Bradley, 2006). Accordingly, the more illegitimate a trial is seen to be, the harsher the punishment that will be imposed for it (Ulmer et al., 2010). Another practical constraint that affects how the judge sentences is the fact that defense attorneys, judges, and prosecutors are all a part of the same workgroup. Since trials put the members of the workgroup against each other and force the judge to mediate, they are viewed as unpleasant and disruptive (Bradley-Engen et al., 2012; Ulmer & Bradley, 2006). As a result of this, penalties are given to those who opt for trial in order to discourage such experiences (Bradley-Engen et al., 2012). Findings that caseload and opting for a trial over a guilty plea interact to increase sentence length in the federal district court system suggest that this theory has merit (Ulmer et al., 2010).

The last component of the focal concerns perspective is protection of the community (Steffensmeier et al., 1998). Similar to the blameworthiness component, a trial will unearth aggravating factors like a defendant's criminal history that a plea deal would not unearth (Bradley-Engen et al., 2012;

Ulmer & Bradley, 2006; Ulmer et al., 2010). This factor may negatively affect the way the judge perceives the defendant and the threat they pose to society. Thus, this may result in harsher sentences. Additionally, trials generate more publicity than a plea deal (Brereton & Casper, 1981). Due to this, there is increased pressure on the judge to emphasize their interest in community protection (Ulmer & Bradley, 2006). In order to do so, the judge may give a sentence longer than the one they would otherwise impose (Ulmer & Bradley, 2006).

While applying the focal concerns theory to the concept of a trial tax argues that going to trial changes the judge's perception of the defendant and this altered perception results in more severe punishments, the shadow of the law theory² argues that theoretically the difference in the sentence lengths of those who go to trial and those who plead guilty should be slight or nonexistent (Abrams, 2013). This theory has its roots in economic and rational actor theory. It argues that reasonably, the defense will only accept plea offers if they are less than or equal to the sentence the defendant would expect from trial (Abrams, 2013; Bushway et al., 2014). On the other side, the prosecutor will only offer pleas that carry a sentence greater than or equal to what they expect as the outcome of a trial after accounting for the probability of acquittal (Abrams, 2013; Bibas, 2004; Bushway et al., 2014). Thus, the expected plea sentence would be about equal to the expected trial sentence since that is the small range the defense and prosecutors agree upon (Abrams, 2013). Theoretically, this theory suggests that the trial tax should be minimal if even existent. In reality, there is great variation in this. This is because not all cases are the same. Some cases lack convincing physical evidence or strong witnesses. Thus, the prosecutor may be more willing to decrease the plea sentence length in order to guarantee a conviction they otherwise could not secure (Bibas, 2004). Further, not all court actors act solely based upon logic. For example, a risk averse defendant and risk averse court actors will favor the certainty that a plea deal offers compared to a trial (Bibas, 2004; Bushway et al., 2014). Due to their inclination towards certainty, the defendant may be more willing to accept a plea sentence longer than what they would expect from trial in order to have assurance in their sentence and a quick resolution of their case (Bushway et al., 2014).³ This could potentially result in a negative trial tax.⁴ Further, this theory is based on human perception—something that can vary greatly depending on other factors. For example, a young attorney will have less of an idea of the expected outcome of a given case if it goes to trial than an experienced attorney who has had many similar cases in the past. Additionally, if the presiding judge has been on the bench for many years, both sides will be better able to estimate the outcome of the trial and the resulting sentence than if the judge has presided over few cases (Abrams, 2013). Factors like these may affect one or both sides perceived expected value, affecting the sentence lengths they would agree to in plea deals.

These two prominent theories provide different views on a trial tax. The focal concerns application argues that going to trial may carry negative consequences for the defendant's image and for the court's efficiency, resulting in harsher punishments (Bradley-Engen et al., 2012; Breen, 2011; Ulmer & Bradley, 2006; Ulmer et al., 2010). The shadow of the law perspective argues that plea deals only happen when the two sides agree on the expected outcome of trial and that outcome is what is

² This theory is also commonly referred to as the bargaining in the shadow of the trial theory.

³ This factor presents another potential issue with plea bargaining that should be explored. Studies show that criminal offenders tend to be much less risk averse than their non-offending counterparts, with some studies even showing a tendency for risk seeking among this population (Kim, 2015). Thus, guilty defendants would tend to be much more likely to risk trial than innocent defendants. Due to innocent defendant's tendency to avoid risk, they will feel increased pressure to plead guilty despite their innocence. Therefore, this factor of risk aversion and its effect on plea bargaining decisions could create a due process issue in which innocent people are wrongly convicted (Bibas, 2004; Kim, 2015). Further research should investigate this potential relationship.

⁴ This variation on the shadow of the law theory often is referred to as a distinct theory called the uncertainty avoidance theory. For further elaboration see (Breen, 2011; Albonetti, 1986; Albonetti, 1987).

stipulated to in the plea deal (Abrams, 2013; Bibas, 2004; Bushway et al., 2012). The former explains the existence of a trial tax while the latter explains why a trial tax should be minimal, especially if using unconditional expected values for comparison. As stated above, there is potential variation in the latter theory and its predictions regarding a trial tax based upon whether risk aversion is taken into account and the characteristics of the court actors (Breen, 2011; Bushway et al., 2014). The findings of this study show the existence of a significant negative nonlinear relationship between the proportion of cases resolved by plea in a circuit and the median sentence length for that circuit. Overall, the findings support the existence of a significant trial tax on a macro level in the federal U.S. circuit Courts. Thus, these results provide support for the focal concerns theory over the conventional shadow of a trial theory. However, alterations of the shadow of a trial theory including potential differences in courtroom actor experience between circuits could be supported by this findings as well.

IV. Trial Tax Research

An abundance of literature exists suggesting that defendants who assert their rights to trial face harsher punishments than those who plead guilty (Bradley-Engen et al., 2012; Bushway et al., 2014; Johnson, 2006; Kim, 2015; Spohn, 1992; Steffensmeier et al., 1998; Ulmer & Bradley, 2010; Ulmer & Johnson, 2004; Ulmer et. al, 2006). One common measure of a trial tax is the difference in likelihood of incarceration between those convicted at trial and those who pled guilty. Studies estimate that those convicted by trial are two to three times more likely to be imprisoned than those convicted by plea (Johnson, 2006; Steffensmeier et al., 1998). The other measure of trial tax commonly used is the difference between the lengths of the sentences between the two modes of conviction. Estimates for this trial tax vary greatly with some studies suggesting differences of a few months between the two modes of conviction while other studies suggest that a defendant asserting their right to trial could increase their incarceration sentence by years (Bradley-Engen et al., 2012; Johnson, 2006; Spohn, 1992; Ulmer & Bradley, 2006). Overall, a review of the research suggests the existence of a trial tax.

The size of the trial tax observed is not consistent. Research suggests that the trial tax may vary from court to court (Breen, 2011; Ulmer & Bradley, 2006; Ulmer et al., 2010) and from offender to offender (Spohn, 1992; Ulmer & Bradley, 2006; Ulmer et al., 2010). Depending on the characteristics of the offender (Spohn, 1992), the nature of the convicted offense, and the characteristics of the courts studied, the trial tax may fluctuate (Ulmer & Bradley, 2006; Ulmer et al., 2010). As mentioned earlier, the size and direction of the trial tax also varies by the method of the study depending on the inclusion or exclusion of acquittal data. While studies excluding acquittal data provide the most accurate measure of a trial tax, the studies including acquittal data still contribute to the discussion of the trial tax in their identification of control variables and investigation into other previously unstudied characteristics. Thus, the following review of the trial tax literature will include both conditional and unconditional studies.

Throughout the criminal justice system, defendants vary greatly in their characteristics. From race to gender to criminal history, defendants have a variety of attributes that differentiate them from other defendants. Differences in these attributes may affect the trial tax imposed. While many studies suggest that race does not significantly impact the trial tax of a given court or set of courts (Breen, 2011; Spohn, 1992; Ulmer et. al, 20016), others suggest that African American defendants face a greater trial tax than their white counterparts (Johnson, 2003; Ulmer 1997). Similarly, conflicting evidence exists as to whether being male increases the trial tax imposed upon a defendant (Ulmer & Bradley, 2006; Breen, 2011). While the literature generally suggests that there is a significant difference in the punishment imposed based upon these characteristics, it provides no evidence that these characteristics further interact with the mode of conviction variable (Spohn, 1992; Ulmer & Bradley, 2006; Ulmer et al., 2010). One variable that the research suggests may interact with the mode of conviction variable is the criminal

history of a defendant. However, the direction of this interaction is not clear. Ulmer's early study suggested a positive interaction effect between the criminal history of a defendant and the trial tax imposed (1997), while the subsequent study done by Ulmer & Bradley suggested the opposite—a negative interaction effect (2006). This characteristic along with gender and race appear to be important in their direct effects on sentence length, but with conflicting evidence from different studies, the interaction effects of individual defendant characteristics are not well understood. As findings vary from study to study, more individual level research is needed to understand these variables. While this study does not expand on these variables because it is a macro level study rather than individual level study, future research could include aggregate macro level measures of these variables for analysis as well.

Just as differences in offender characteristics have been investigated in the context of the trial tax, differences in the nature of the convicted offense have also been analyzed. This analysis has taken two different forms. While one looks broadly at offense severity (Bradley-Engen et al., 2012; Ulmer & Bradley, 2006; Ulmer et al., 2010), the other looks specifically at offense type (Breen, 2011; Kim, 2015; Ulmer & Bradley, 2006). The measures of the variables themselves also take on many different form. For example, offense severity could be measured by the harm caused to the victim (Bradley-Engen et al., 2012; Ulmer & Bradley, 2006) or the guideline recommended sentence length for that offense (Ulmer et al., 2010). Likewise, offense type can be measured by general person versus property offense categories (Breen, 2011) or more specific offense categories (Kim, 2015; Ulmer & Bradley, 2006). In their 2006 study, Ulmer et al. looked specifically at serious violent offenses defined as “third degree murder, aggravated assault, rape, involuntary deviant sexual intercourse, and robbery”, finding that the trial tax still existed within this set of offenses. Regarding interaction effects for the nature of the offense, studies have shown that the trial penalty varies by offense type. Studies suggest more severe offenses result in harsher punishments (incarceration versus no incarceration) for those who opt for jury trial versus those who plead guilty (Breen, 2011; Ulmer & Bradley, 2006). However, when the term trial tax is used to refer to differences in lengths of incarceration between the two modes of conviction, there are mixed results for the interaction effect between offense severity and the trial tax (Kim, 2015; Ulmer & Bradley, 2006; Ulmer et al., 2010). Based upon this prior research, this study will control for the nature of the offenses in a given circuit by including a proportion of violent crime variable. However, it will not investigate an interaction effect.

Trial tax research has examined a variety of courts from state to federal to military courts. Although the criminal justice system attempts to be as uniform and consistent as possible, large variations exist from court to court and circuit to circuit. Courts vary in individual characteristics such as caseload, location, and the primary types of crime dealt with in that court (see United States Sentencing Commission Federal Sentencing Data). Additionally, there are different court systems. To name a few, there are county courts, municipal courts, federal courts, and military courts. The county and municipal courts belong to the state court system while the federal and military courts each belong to their own court system (Breen, 2011; Glick, 1983). While these court systems share many characteristics, there are distinctions between them in their jurisdiction and structure. Prior research suggests that trial tax varies depending upon these types of variations in court characteristics.

Regarding individual characteristics, different courts vary greatly. For example, in 2017 the federal district courts in the first circuit handled 1993 cases, while the federal courts in the ninth circuit handled 13,122 cases (United States Sentencing Commission “2017 Federal Sentencing Statistics” [USSC]). This is because the ninth circuit consists of a much greater population than the first circuit. These differences in the number of cases per year are somewhat remedied by the increased number of judges in the ninth federal circuit. In fact, a standardized measure of caseload per judge has been created by taking the number of cases in a given court district divided by the number of judgeships in that district. Calculating this for the given example, there are still vast differences between the two circuits. The ninth circuit federal judges manage a caseload of 117 cases per year while the first circuit federal judges handle a

caseload of only 69 cases per year. These differences carry importance when it comes to analyzing the relationship between mode of conviction and sentence length. Some studies show a positive interaction effect between the trial tax and the caseload of a given court (Ulmer & Bradley, 2006; Ulmer et al., 2010). In other words, as a court's caseload increases, the difference between the sentence lengths of those convicted by trial and those convicted by plea tends to increase. These findings lend support to the focal concerns theory, specifically to its practical constraints component. Additionally, some courts have greater trial rates than others. Ulmer et. al found this characteristic conversely affects the trial tax (2010). In fact, one extra trial per judge was associated with a .004 decrease on the trial tax coefficient ($p < .05$) (Ulmer 2010). Thus, these individual characteristics can have a large effect on the observed trial tax in that court.

In addition to differences in caseload and trial rates, some courts are distinct in nature from other courts. For example, while most research on trial tax focuses on civilian courts, some research has extended into military courts (Breen, 2011; Breen & Johnson, 2018). The military court system differs from the civilian court system in structure, punishments, caseloads, trial rates, and many other factors (Breen, 2011). One of these main factors is that the military court system allows for sentencing done by a jury. Few states allow this, and even fewer allow for jury sentencing following a guilty plea (Breen, 2011). Mixed research surrounds this type of sentencing with some evidence suggesting jurors to be more lenient than judges and other evidence suggesting the contrary (Breen, 2011). However, concerning this type of sentencing's effect on trial tax, Breen (2011) found no difference in the sentences for those convicted at trial versus those convicted by guilty plea whether they were sentenced by a judge or a jury (Breen & Johnson, 2018).⁵

While most of the research suggesting evidence of a trial tax is done on state courts (Bushway et al., 2014; Johnson, 2003; Johnson, 2006; Ulmer & Bradley, 2006), a few researchers have looked specifically at federal courts to understand how the trial tax there differs from their more commonly studied state counterparts (Kim, 2015; Ulmer et al., 2010). While the differences between federal and states courts are much smaller than those between civilian and military courts, there are still some important distinctions that have potential to impact the trial tax. The selection of judges differs between these two courts with federal judges being appointed by the President while state court judges can either be elected or appointed ("Comparing Federal & State Courts"). Further, federal courts tend to have fewer cases than state courts ("Comparing Federal & State Courts"). Most importantly in the context of the trial tax, federal courts have the federal sentencing guidelines (Ulmer et al., 2010).

The federal sentencing guidelines are important in terms of investigating the trial tax for a few reasons. The federal sentencing guidelines are the product of the Sentencing Reform Act of 1984 ("2018 Guidelines Manual Annotated"). In order to establish a more consistent and rational sentencing process throughout the federal justice system, Congress established the United States Federal Sentencing Commission and tasked them to create just sentencing guidelines ("2018 Guidelines Manual Annotated"). As a result, once a defendant was convicted of an offense, a small range of sentences was given based upon the offense, the offense characteristics, and the defendant's backgrounds ("2018 Guidelines Manual Annotated"). From their creation until the Supreme Court decision in *United States v. Booker* in 2005, these "guidelines" were mandatory for judges (Ulmer et al., 2010). Hence, they had to choose a sentence within the range prescribed. As detailed later, *Booker* made these guidelines advisory (*United States v. Booker*, 543 U.S. 220 (2005)). As a result, the guidelines have been amended, and judges must only consider their recommendations rather than impose them.

Important in the context of trial tax, the federal sentencing guidelines include specific exceptions for which a judge may consider lowering the imposed sentence. This includes if the defendant accepts

⁵ In the 2011 study, Breen includes acquittal data, but the findings referenced are based only on the subsample of convicted defendants sentenced to confinement making them accurate measures of trial tax.

responsibility for their actions or provides assistance to the government in apprehending and prosecuting others (“2018 Guidelines Manual Annotated”). Since these accepted departures more likely result after a plea deal than a trial, they may explain some of the trial tax. In fact, Ulmer et al. found that these accepted departures explain a significant amount of the trial tax observed in federal courts (2010). Despite this, the findings still suggested the existence of a 15% trial tax after accounting for these legally accepted departures, implying the disparity comes from more than plea defendants providing substantial assistance or accepting responsibility (Ulmer et al. 2010).

Despite differences in caseloads, judicial selection, and judicial discretion, evidence of a trial tax persists in these courts in the literature (Ulmer et al., 2010; Kim, 2015). Similarly, evidence of a trial tax is present in this study, which focuses on federal rather than state courts. This study hopes to contribute to the discussion on how court characteristics may alter the trial tax by focusing on a less studied court type with federal courts. Further, unlike other studies, this study examines an extensive time period both before and after the *Booker* decision to understand how the trial tax may have changed as a result of this shift in federal judicial practices. Finally, this study departs from the typical individual level analysis of the trial tax, where here we examine whether a trial tax can be observed on a macro level for the federal court system. That is, we examine whether circuits with a higher proportion of their cases resolved by plea have a lower median sentence length.

V. *Data*

This macro level analysis uses aggregate data from the United States Sentencing Commission’s Federal Sentencing Statistics. Information including number of total cases, number of violent crime cases⁶, proportion of cases resolved by plea, and median sentence length was recorded for each circuit (excluding Washington D.C.) for the years 1995-2017. The Washington D.C. circuit was excluded in this study because it is believed to differ in important ways from the others. While other circuits are made up several different states and/or U.S. territories, the D.C. circuit includes only Washington D.C. Due to its smaller size, the D.C. circuit handles much fewer cases per year than any other circuit. For example, in 2017, the D.C. circuit had 232 sentencing cases. The circuit with the next lowest number of cases was Circuit 1, with 1993 cases—over 8 times the amount observed in the D.C. Circuit. These vast differences in size between the circuits means that the median sentence length observed in D.C. is based upon a much smaller sample size than that of the other circuits. In addition to this, due to D.C. being the nation’s capital, it is likely that the breakdown of crime there is much different than that of other circuits. More white-collar crime would be expected in this circuit than others, which may affect the median sentence length observed. Due to these differences, this study chose to exclude the D.C. circuit, focusing only on the circuits that are made up of multiple states and/or U.S. territories.

This data provides accurate information on federal sentencing. By law, every federal sentencing court is required to submit a detailed document to the USSC within thirty days of making a judgement on a case (“Research Notes”).⁷ Thus, the resulting USSC datafiles represent a census of all federal sentences handed down in the U.S. and its territories from 1995 to 2017. This means that this data is representative of the population of inference (federal circuit courts), and thus, the results from analysis of this data will be reliable in answering the question of existence of a macro level trial tax. In addition to the USSC data, information on the number of authorized judgeships in each district from 1995 to

⁶ For this study, the offenses categorized as violent offenses includes murder, manslaughter, sexual abuse, robbery, and assault.

⁷ For more information on the USSC data and the Commission’s collection processes, visit https://www.uscourts.gov/sites/default/files/pdf/research-and-publications/research-notes/20190719_Research-Notes-Issue1.pdf.

2017 was collected from the “Additional Authorized Judgeships-Since 1960” PDF file (uscourts.gov) in order to calculate the caseload in each circuit.

i. Response Variable

The response variable in this study is *Median Sentence Length*. This value represents the median number of months of incarceration that defendants were sentenced to in a given circuit for a given year. For analysis, we considered taking the log of the median sentence length obtained from USSC for each circuit-year combination, as this is something commonly done in the literature and recommended by many to remedy the skew in the sentence length distributions. However, unlike the distribution of individual sentence lengths, the median sentence length distribution does not display a great skew. Further, transforming these data with the log function creates heterogeneity in the residuals. Larger log median sentence lengths tend to have more variation than smaller log median sentence lengths. This violation of the ordinary least squares (OLS) assumptions would unnecessarily complicate inference. The residuals resulting from an OLS regression on the raw median sentence lengths do not violate the assumption of homoscedasticity. Thus, this raw values method will be used.

ii. Explanatory Variables

The primary variable of interest is the *Plea Proportion*. This variable is the proportion of cases in a circuit for a given year that were resolved by plea. In any circuit in any given year, a vast majority of the cases are resolved by plea. Thus, this variable takes on values between .86 and .99 with a majority of the observations having between .93 and .97 of their cases resolved through plea. Based on prior research suggesting the existence of a trial tax, this study hypothesizes that the proportion of cases resolved by plea in a circuit will have a negative relationship with the median sentence length of that circuit. Hence, circuits with a higher proportion of their cases resolved by plea will tend to have a lower median sentence length while circuits with a lower plea proportion will tend to have a higher median sentence length.

Hypothesis 1: Circuits with a lower proportion of their cases resolved by plea will tend to have a higher median sentence length than circuits with higher plea proportion resolution tend to after controlling for other variables.

Since violent offenses tend to result in harsher punishments, a control variable called *Violent Offenses* will be used. This is a quantitative variable representing the proportion of convicted offenses in a circuit for a given year that classify as violent offenses. For the purposes of this study, offenses comprising the violent offenses category are murder, manslaughter, sexual abuse, assault, and robbery. This definition is adapted from the Uniform Crime Report’s definition of violent crime, which consists of murder, nonnegligent manslaughter, rape, robbery, and aggravated assault. Since the USSC data does not differentiate between manslaughter and nonnegligent manslaughter, rape and sexual abuse, or aggravated assault and simple assault, the broader categories were included in the violent crime definition of this study. Although this measurement overestimates the true proportion of violent crime in a circuit, this measurement will still help control for differences in circuit-years with more serious offenses like those included under this variable and with less serious offenses like drug trafficking, arson, and fraud. This variable will be used only as a control variable, rather than as a potential interaction variable.

Due to differences in geography and the population make-up of each circuit, the offense breakdown varies from circuit to circuit. For example, 55.5% of the offenses in the 5th Circuit were immigration offenses.⁸ Meanwhile, only 7.7% of the offenses in the 3rd Circuit were immigration offenses.⁹ These large differences in offense type will likely influence the median sentence length of a given circuit. In order to measure these differences, ten indicator variables called “*Circuit i*” where *i* ranges from 2 to 11 will be used. These variables are defined below. Any coefficient other than zero for these variables will suggest that that circuit is different from the first circuit in its median sentencing length, and a cross coefficient comparison can be made to see if other circuits are significantly different from each other. We expect that there will be difference between circuits. More specifically, we predict that Circuit 5 will have longer median sentence lengths and Circuit 10 will have shorter median sentence lengths. Circuit 5 is predicted to have a higher median sentence length because it includes the state of Texas, which is notorious for being tough on crime. Circuit 10 is predicted to have shorter median sentence lengths than the other circuits because of the border it shares with Mexico and the resulting higher proportion of immigration crime in this circuit.

$$Circuit\ i = \begin{cases} 0 & \text{if not Circuit } i \\ 1 & \text{if Circuit } i \end{cases}$$

Hypothesis 2: There are differences in median sentence length by circuit after controlling for other variables. More specifically, Circuit 10 will tend to have a lower median sentence length and Circuit 5 will tend to have a higher median sentence length.

As mentioned earlier, an important Supreme Court decision was handed down in the *U.S. v. Booker* case in 2005 that changed the structure of sentencing in the federal court system. Prior to this decision, the federal sentencing guidelines were mandatory to follow (United States v. Booker, 543 U.S. 220 (2005)). A judge had to select a sentence within the narrow range provided based upon the offense and its characteristics as well as the defendant’s background (“2018 Guidelines Manual Annotated”). Importantly, the federal sentencing guidelines allowed for judges to weigh their own findings of fact against the defendant (United States v. Booker, 543 U.S. 220 (2005)). These facts did not have to be facts found by a jury and proven beyond a reasonable doubt. Rather, the judge’s finding of an aggravating factor like the defendant committing a crime with excessive cruelty was enough for the judge to impose a harsher sentence than that dictated by the guidelines (United States v. Booker, 543 U.S. 220 (2005)). In *Booker*, the Supreme Court questioned the constitutionality of such practices. They ruled that allowing a judge to enhance sentences beyond the maximum of the mandatory guideline range based upon facts not submitted to and found by the jury or admitted by the defendant themselves violated the defendant’s Sixth Amendment right to trial by jury (United States v. Booker, 543 U.S. 220 (2005)). This is because such finding was based solely on the preponderance of the evidence rather than the burden of proof required in criminal law of beyond a reasonable doubt (United States v. Booker, 543 U.S. 220 (2005)). As a result of this ruling, the federal sentencing guidelines were invalidated and ruled unconstitutional (United States v. Booker, 543 U.S. 220 (2005)). To remedy this, the Supreme Court ruled that from that point forward, the guidelines would serve as advisory to judges, giving them a basis for their sentencing decisions (United States v. Booker, 543 U.S. 220 (2005); “2018 Guidelines Manual Annotated”).

This study contains data from before and after this major shift in federal sentencing practices. Thus, an indicator variable *Year* as defined below will be controlled for in the model. Using the focal concerns

⁸ The 5th Circuit consists of Louisiana, Mississippi, and Texas.

⁹ The 3rd Circuit consists of Delaware, New Jersey, Pennsylvania, and the Virgin Islands.

theory as a framework, this study hypothesizes that the greater discretion allotted to federal judges following *Booker* allows for greater disparity in sentence length depending on the perceived blameworthiness of the defendant. Thus, this study hypothesizes that the median sentence lengths for the 11 circuits will be significantly different before 2005 and after 2005.

$$Year = \begin{cases} 1 & \text{if after 2005} \\ 0 & \text{Otherwise} \end{cases}$$

Hypothesis 3: *Booker* has had a significant impact on the median sentence length of the circuits.

Prior research suggests that higher caseloads are correlated with a greater trial tax (Ulmer & Bradley, 2006; Ulmer et al., 2010). Focal concerns theory explains that this relationship exists because higher caseloads create greater pressure to process cases quickly to keep the system efficient (Ulmer & Bradley, 2006; Ulmer et al., 2010). Thus, there is a heavier reliance on plea deals to resolve cases. Additionally, the court actors in these areas will place a greater value on plea deals (Bradley-Engen et al., 2012; Dixon, 1995; Ulmer et al., 2010). Hence, prosecutors may be more willing to offer a plea discount, or judges may be more likely to impose a sentence penalty for those who assert their right to trial. In order for this theory to be supported, there should thus be a relationship between a circuit's caseload and plea proportion. More specifically, it would be expected that circuits with greater caseloads should also have a longer median sentence lengths. In order to test this theoretical and research-based prediction, a separate investigation was done on the relationship between cases resolved by plea in a circuit and the caseload of that circuit. The *Caseload* variable will be plotted against the *Plea Proportion* variable and a loess smoother will be put through this scatterplot to examine its trends. The *Caseload* variable will be a quantitative measure of the number of cases per judge in a given circuit for a given year.¹⁰

Hypothesis 4: Among the circuits, those with a larger caseload will have a higher proportion of cases resolved by plea.

VI. Analysis and Findings

To begin analysis, a descriptive overview of the data was created. The average median sentence length of the federal district courts in the eleven circuits from 1996 to 2017 was about 39 months with a standard deviation just over 12 months. The highest median sentence length of 60 months was observed during several years in both the fourth and seventh circuit. The lowest median sentence length of 3 months was observed in the tenth circuit in 2016. For the independent variable, on average 95.18% of cases were resolved by plea in the eleven circuits over the range of years. This variable took on a limited range of values with a standard deviation of .0211. Concerning the control variables, each circuit had 24 years of observations¹¹ for a total of 264 circuit-years. Further, 121 of these circuit years

¹⁰ $Caseload = \left(\frac{\text{Number of cases in a circuit}}{\text{Number of approved judgeships in that circuit for that year}} \right)$

¹¹ The data set included information from 1995-2017 for a total of 23 years. However, the year 2004 was split into two sets based on the Washington Supreme Court decision in *Blakely v. Washington* decided this year. 2005 was similarly divided into Pre-*Booker* data and Post-*Booker* data. Thus, there are 25 total observations per circuit. Running a model on these 25 observations showed serial correlation in the residuals. As a result, an additional variable for the prior year's median sentence length for that circuit was included in the regression to reduce the autocorrelation. Since median sentence length predictions for a circuit relied on the previous years' response value, the final model could only include years from 1996-2017. Therefore, there were a total of 24 observations per circuit.

preceded the *Booker* decision while 143 followed *Booker*. The average caseload was 89.31 cases per judge per year. However, caseload varied greatly from circuit to circuit and year to year with a standard deviation of 47.17 cases. Lastly, on average, 4% of the convicted offenses in the eleven circuits from 1996-2017 were violent offenses. This control variable had a standard deviation of about .02. Table 1 below summarizes the descriptive statistics for the data set.

Descriptive Statistics for the Data					
	Coding	N	%	Mean	SD
Dependent Variables					
Median Sentence Length	Quantitative Value Range: (3,60)			39.30	12.65
Independent Variables					
Plea Proportion	Quantitative Value Range: (0.86,0.99)			0.952	0.021
Control Variables					
Circuit 2	1 if Circuit 2 0 otherwise	24	9.09		
Circuit 3	1 if Circuit 3 0 otherwise	24	9.09		
Circuit 4	1 if Circuit 4 0 otherwise	24	9.09		
Circuit 5	1 if Circuit 5 0 otherwise	24	9.09		
Circuit 6	1 if Circuit 6 0 otherwise	24	9.09		
Circuit 7	1 if Circuit 7 0 otherwise	24	9.09		
Circuit 8	1 if Circuit 8 0 otherwise	24	9.09		
Circuit 9	1 if Circuit 9 0 otherwise	24	9.09		
Circuit 10	1 if Circuit 10 0 otherwise	24	9.09		
Circuit 11	1 if Circuit 11 0 otherwise	24	9.09		
Year	1 if after Booker 0 if before Booker	143 121	54.167 45.83		
Caseload	Quantitative Value Range: (7,249)			89.31	47.17
Violent Offenses	Quantitative Value Range: (0,.15)			0.040	0.019
Regress	Quantitative Value Range: (3,60)			39.34	12.28

Table 1: Descriptive statistics for the variables of analysis.

Table 3 in Appendix A presents the correlation matrix of all independent and control variables. This table provides no evidence of collinearity among the explanatory and control variables. The strongest correlation between the explanatory and a control variables is only 0.577 between caseload

and plea proportion. This is only a moderate correlation. Additionally, caseload will not be used in a model with plea proportion. Caseload will be regressed on plea proportion separately. The strongest correlation between the remaining independent and control variables that will be used together in the model is even weaker at only -0.475. With no major issues of collinearity, analysis was carried out through OLS modelling.

Looking at Figure 1 of median sentence length by plea proportion, a nonlinear relationship is apparent. Thus, nonlinear least squares regression was used to identify a model in R. Originally, the median sentence length was regressed on the plea proportion, year, violent offenses and all eleven circuit variables. However, the resulting residuals violated the independence assumption, displaying positive autocorrelation from year to year within the circuits. Thus, for subsequent analysis a variable called Regress that represents the prior year's median sentence length for that circuit was added. Adding the regress variable greatly decreased this autocorrelation, and little evidence remained to suggest a lack of independence of the residuals of the other models created from OLS regression. Error normality was also assessed along with error homogeneity. Figure 3 in the Appendix B shows no evidence of a violation of the homoscedasticity assumption. Further, Figure 4 in the Appendix B shows some slight departure from normality. However, t-procedures which are robust to nonnormality, will be used for inference. Since there is no evidence of extreme skew or outliers in the model residuals, there is no major concern in performing inference based on this model.

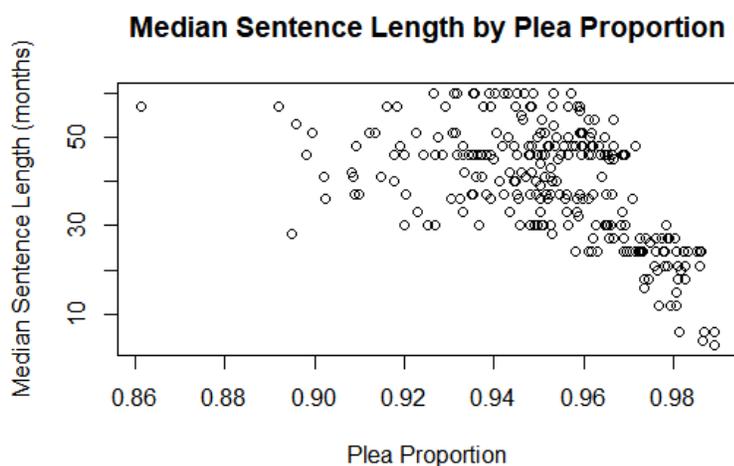


Figure 1: This scatterplot displays the nonlinear relationship between median sentence length in months for a circuit and the proportion of cases resolved by plea in that circuit.

In order to test hypotheses 1, 2, and 3 simultaneously, the nonlinear least squares function *nls* in R was used to create two models. The results of the regression are shown below in Table 2. In Model 1, median sentence length was regressed on the plea proportion, violent offenses, year, regress, and all circuit variables. In Model 2, median sentence length was regressed on the plea proportion, violent offenses, regress, and all circuit variables (so that year was omitted from Model 2).

Concerning Hypothesis 2, this model provides evidence that the median sentence length does differ from circuit to circuit. Circuits 2, 5, 9, and 10 tend to have median sentence lengths below that of Circuit 1 after controlling for other variables, while Circuits 4 and 7 tend to have median sentence lengths above that of Circuit 1 after controlling for other variables. Model 2 shows similar results regarding the circuit variables. Figure 2 gives a side by side comparison of the parameters estimates for each circuit.

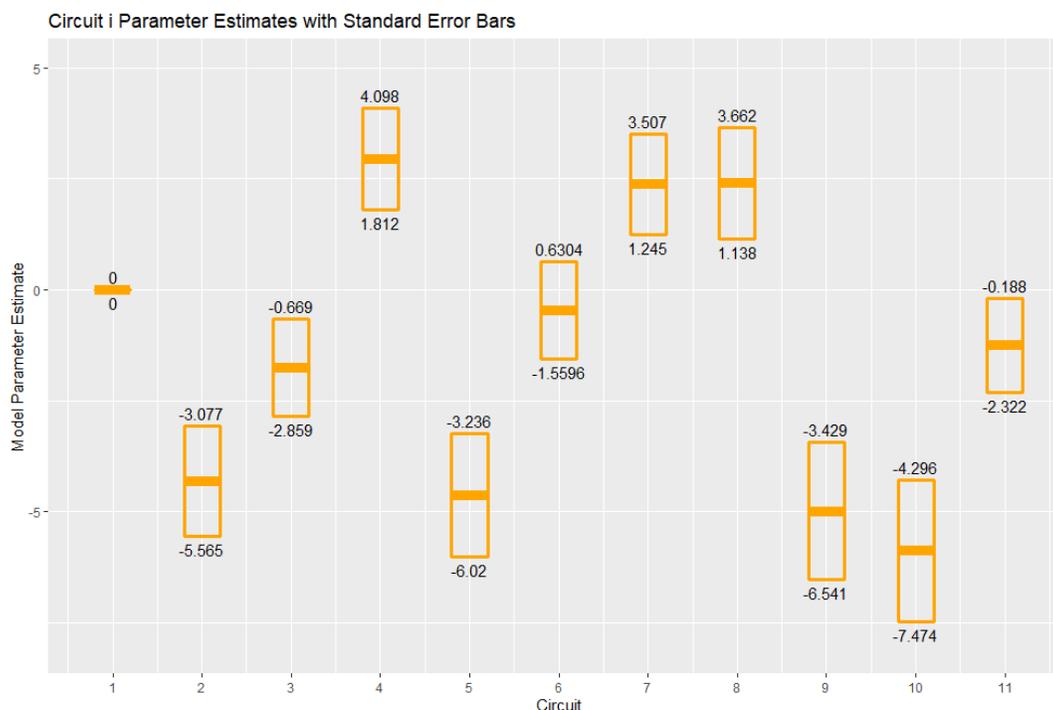


Figure 2: A side by side comparison of parameter estimates by circuits with corresponding standard error bars.

Concerning hypothesis 3, Model 1 shows no evidence that the median sentence length changed following *Booker*. With 95% confidence, this study concludes that the median sentence length of the eleven circuits following the 2005 *Booker* decision is between .98 months shorter and 1.22 months longer than the median sentence length prior to the *Booker* decision after accounting for the other model variables. Thus, this variable does not significantly add to the existing model. Therefore, this variable was excluded from Model 2.

The exclusion of the *Year* variable in Model 2, increased the significance of the *Violent Offenses* variable. In Model 1, there was only weak evidence to suggest that the proportion of convicted offenses in a circuit that were violent impacted the median sentence length of that circuit. However, by excluding the *Year* variable, this study argues with 95% confidence that an increase of .01 in the proportion of violent offenses in a circuit decreases the median sentence length of this circuit between 1.81 days and 24.71 days. This result is surprising, as it was argued that the proportion of violent offenses should be controlled for because the median sentence length would be expected to increase as the proportion of violent offenses increased. There is no obvious theoretical explanation as to why this would occur.

Concerning hypothesis 1, the results suggest that there is a significant relationship between plea proportion and median sentence length even after controlling for other variables. Based on the large amount of research suggesting the existence of an individual level trial tax, Hypothesis 1 predicted that circuits with higher plea proportions would tend to have lower median sentence lengths than circuits with lower plea proportions. Model 1 and Model 2's significant exponential factors support this prediction even after accounting for other variables ($p < .01$). The effect a .01 increase in plea proportion has on median sentence length decreases as the size of the plea proportion decreases. In fact, this study finds that holding all other variables constant, an increase from 98% of cases resolved by plea to 99% of cases resolved by plea in a given circuit would decrease the median sentence length by over 4 months. Meanwhile, an increase from 88% of cases resolved by plea to 89% of cases resolved by plea in a given circuit holding all other variables constant would decrease the median sentence length by less than a

day. Thus, changes in plea proportion of a circuit have more drastic consequences on the median sentence length when the plea proportion is relatively high to begin with.

	<i>Model 1</i>			<i>Model 2</i>		
	B	SE	95% Confidence Interval	B	SE	95% Confidence Intervals
Intercept	18.278	2.514	(13.326, 23.230)	18.346	2.489	(13.443, 23.249)
Exponential Constant	-55.061	18.092	(-90.694, -19.428)	-55.031	18.230	(-90.936, -19.125)
Variables						
Plea Proportion	57.932	18.320	(21.850, 94.014)	57.891	18.460	(21.534, 94.248)
Circuit 2	-4.309	1.248	(-6.766, -1.852)	-4.322	1.244	(-6.772, -1.872)
Circuit 3	-1.793	1.106	(-3.971, 0.3842)	-1.764	1.095	(-3.922, 0.3927)
Circuit 4	2.954	1.145	(0.6992, 5.210)	2.955	1.143	(0.7042, 5.206)
Circuit 5	-4.579	1.413	(-7.361, -1.797)	-4.628	1.392	(-7.369, -1.889)
Circuit 6	-0.4873	1.102	(-2.658, 1.683)	-0.4646	1.095	(-2.621, 1.692)
Circuit 7	2.350	1.140	(0.1042, 4.595)	2.376	1.131	(0.1478, 4.604)
Circuit 8	2.345	1.290	(-0.1954, 4.885)	2.400	1.262	(-0.0851, 4.885)
Circuit 9	-4.998	1.560	(-8.070, -1.926)	-4.985	1.556	(-1.921, -0.0851)
Circuit 10	-5.911	1.597	(-9.055, -2.766)	-5.885	1.589	(-9.014, -2.755)
Circuit 11	-1.255	1.069	(-3.361, 0.8516)	-1.255	1.067	(-3.357, 0.8476)
Year	0.1206	0.5602	(-0.9828, 1.224)			
Violent Offenses	-42.328	21.280	(-84.241, -0.4163)	-44.200	19.376	(-82.361, -6.038)
Regress	0.6616	0.0465	(0.570, 0.753)	0.6628	0.0460	(0.5722, 0.7535)

Table 2: Nonlinear Model output for Median Sentence Length Regressed on Plea Proportion and control variables.

In order to test Hypothesis 4, Figure 3 was constructed to illustrate the relationship between the proportion of cases resolved by plea in a circuit and the caseload of that circuit. The scatterplot displays a lack of association between the two variables for caseloads under 50 cases. In fact, when the caseload is below 50, there is little to no association between caseload and plea proportion. This lack of association is shown by the horizontal loess line across this portion of the graph. A moderate to strong positive relationship appears for caseloads greater than 50. This trend does support a focal concerns theory. Although low caseloads tend to have little affect on the plea proportion, high caseloads in a circuit are associated with a high proportion of cases resolved by plea. This association is not consistent, but it does suggest that high caseloads result in increased reliance on plea bargaining and increased pressure to plea.

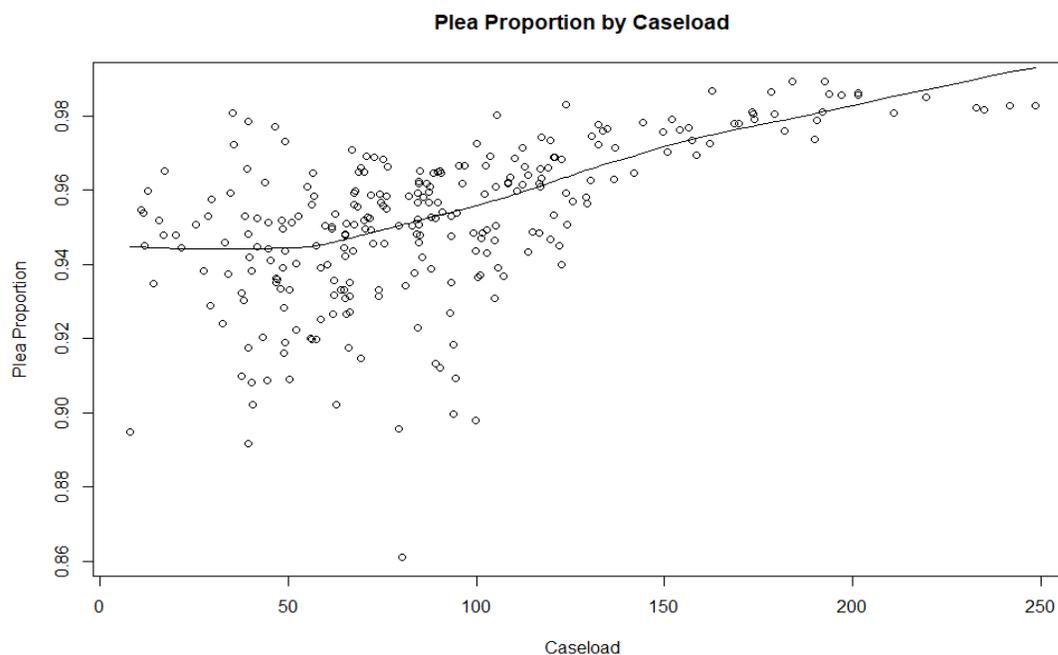


Figure 3: This graph illustrates the relationship between the caseload of a given circuit and the proportion of cases resolved by plea in that circuit.

VII. Discussion

As with all research, this study had some important limitations. First, this study only controlled for a few possible variables. The data used did not provide a breakdown of the characteristics of the offenders in each circuit. Prior research suggests African Americans, Hispanics, and males tend to face longer sentences (Bradley-Engen et al., 2012; Ulmer & Bradley, 2006). Additionally, some evidence suggests that older convicted offenders tend to receive more lenient sentences (Ulmer & Bradley, 2006), and offenders with a longer criminal history tend to receive harsher sentences (Bradley-Engen et al., 2012; Ulmer et al., 2010). Thus, relationships between high plea proportions and lower median sentence lengths may be mediated by the proportion of older offenders or proportion of female offenders in that circuit. This study was unable to control for these specific variations that could have created differences in the median sentence length. However, the indicator variables for each circuit are designed to help control for differences between the eleven circuits. Thus, these indicator variables should be able to capture and control for a majority of this variation as long as the characteristics of offenders in a given circuit do not vary much from year to year. Second, this research is based on data solely from federal courts. Due to important distinctions between different court systems, the results of this study cannot be generalized to military courts or state courts such as county or municipal court. Therefore, this study is limited in the contributions it can make to the trial tax discussion. Future research should consider examining state level courts to see if evidence of a trial tax similarly exists on the macro level. Additionally, this would be interesting to investigate in military courts where no evidence of a trial tax has been found (Breen, 2011; Breen & Johnson, 2018). Third, the nonlinear nature of this relationship made it difficult to test interaction effects among the variables. The focal concerns theory, specifically the practical constraints component, suggest that as a court's caseload increases, the size of the trial tax in that court also increases (Ulmer & Bradley, 2006). Much of the prior research has found evidence supporting this theory (Ulmer & Bradley, 2006; Ulmer et al., 2010). This study was unable to test for

such an interaction effect on the macro level. Future research is encouraged to find ways to investigate this interaction effect.

Despite the limitations of this study, it is able to contribute to the many areas of sentencing research. This research provides strong evidence of differences in sentencing between the different federal circuits. With 95% confidence this study concludes that Circuit 2 has a median sentence length between 1.79 and 6.72 months shorter after accounting for the proportion of cases resolved by plea, the proportion of convicted offenses that are violent offenses, and the prior year's median sentence length for that circuit. Similarly, we are 95% confident that Circuit 5 has median sentence lengths between 1.92 and 7.48 months shorter. The confidence interval for Circuit 9 is very similar to that of Circuit 5. For Circuit 10, we are 95% confident that the median sentence length is between 2.69 and 8.95 months lower. These differences are likely due to the different types of crime observed in each circuit. In fact, Circuits 5, 9, and 10 all share borders with Mexico resulting in higher proportions of immigration crime within these circuits compared to others (USSC Federal Sentencing Statistics). The sentences imposed for such crimes tend to be less severe than for other crimes. Thus, the higher proportion of these crimes in these circuits likely bring down their median sentence lengths. Conversely, Circuits 4 and 7 tend to have higher median sentence lengths. In fact, we are 95% confident that Circuit 4 has median sentence lengths between 1.16 and 5.76 months longer while Circuit 7 has median sentence lengths between .57 and 5.08 months longer. There is no apparent explanation as to why these circuits would tend to have higher median sentence lengths or Circuit 2 would tend to have a lower median sentence length. Overall, this study shows strong evidence that the median sentence length varies from circuit to circuit. This variation suggests the necessity of either controlling for circuit or controlling for specific circuit characteristics in future research.

This study also contributes an important finding on the relationship between the violent offenses sentence length. Individual level research on this relationship suggests that the more serious/violent the convicted offense is, the longer the sentence length (Bradley-Engen et al., 2012; Ulmer et al., 2010). Thus, on a macro level it was theorized that a higher proportion of violent crime in a circuit would be associated with greater median sentence lengths. However, the findings of this study do not support this hypothesis. In fact, the findings of this study suggest the opposite; circuits with a higher proportion of violent crime tend to have lower median sentence lengths than circuits with a lower proportion of violent crime. While some may dismiss this finding due to the relatively small difference a .01 increase in violent crime proportion is expected to make in sentence length, it is important to note that some circuits in our data set differ by more than .1 in their violent crime proportion. We are 95% confident that differences this large would create differences in average median sentence lengths between 9.31 days and 8.05 months. Thus, this variable does have important implications for circuits that differ greatly in their violent crime proportion. Further, it is important to note that even a ten-day difference in median sentence length of a circuit would result in great differences in financial requirements for that circuit and great differences in the lives of those incarcerated. Thus, this finding should not be considered negligible. In fact, the contradiction between the macro level finding and the individual level findings on this relationship is particularly interesting and future research should investigate this further.

In addition to these contributions, this study adds to the conversation on the effect of *Booker*. Prior research has mixed findings on the impact the Supreme Court decision had on federal sentencing practices. Some studies have suggested an increase in sentencing disparity following *Booker* and increased sentencing guidelines departures (Scott, 2009) while others have observed little difference in sentencing disparity by regions (Bowman, 2006). More related to this study, prior research suggests no differences in the average sentence lengths before and after the *Booker* decision (Hofer, 2007; Starr & Rehavi, 2013). The findings of non-significance in the year variable support these previous findings of no change in the median/average sentence lengths. While this study did not test sentencing disparity before and after *Booker*, this relationship could be investigated with a year-circuit interaction variable.

Further research should consider this, specifically mimicking Ulmer et al.'s (2010) approach of analyzing disparity within the context of designated guideline departures.

Most importantly, in line with the focus of this study, these results contribute to the research on trial tax. A significant negative exponential relationship is apparent between the proportion of cases resolved by plea in a circuit and the median sentence length of that circuit. As a circuit's plea proportion increases, its median sentence length tends to decrease. The rate at which it decreases depends on the plea proportion. An increase in the proportion of cases resolved by plea in a circuit that is at the lower range of this variable, would tend to result in a smaller decrease in median sentence length than a similar increase in plea proportion for a circuit that already has a higher plea proportion. This may be due to the relationship observed with caseloads as those with the highest plea proportions were the circuits with the highest caseloads. Through a focal concerns framework, these higher caseload circuits may impose greater trial penalties to maintain efficiency. Caseload disparity may help mediate the exponential nature observed in the trial tax.

This study provides evidence of the existence of a trial tax at the macro level. Circuits with lower plea proportions tend to have harsher punishments than circuits with higher plea proportions. Thus, circuits with a higher proportion of defendants who assert their constitutional right to trial impose harsher punishments on their convicted offenders on average. The existence of this relationship despite controlling for differences in circuit, year, and proportion of violent crime supports the claim that defendants who opt for trial and are convicted are punished more severely than those convicted by plea.

VIII. Conclusion

Despite its limitations, this research provides insight into many areas of sentencing research. Building upon prior researchers' investigations into the individual level effect of mode of conviction on sentence length, this study finds evidence that circuits with higher plea proportions tend to have higher median sentence lengths. This finding is consistent with previous findings that defendants who assert their right to trial and are found guilty suffer harsher consequences than those that resolve their cases through a guilty or *nolo contendere* plea. These findings coupled with prior research suggest that the American legal system tends to punish those who assert their due process rights. Despite arguments that this is necessary in order to preserve efficiency and functionality of the system, this phenomenon cannot be condoned. Punishing individuals who assert their Constitutional rights creates pressure on the individuals not to do so. Thus, this practice creates a system in which individuals are not truly free and willing in their choice to waive trial. Rather, they do so and are advised to do so by their attorneys out of fear of receiving harsher sentences. The rights of the Constitution are meant to protect American citizens and their liberties. These rights should be respected and safeguarded. If the assertion of such rights carries drastic consequences for future liberty, these rights are not being safeguarded as the Constitution meant them to be. Further research into this potential violation of Constitutional due process rights should be carried out. Importantly, legal scholars should begin brainstorming how to abolish the trial tax while maintaining a functional system, whether this means allocating more resources to the judicial system or higher selectivity in prosecutor's charging decisions.

References

- Abrams, D. S. (2013). Putting the trial penalty on trial. *Duquesne Law Review*, 51(3), 777-786.
- Albonetti, C. A. (1986). Criminality, prosecutorial screening, and uncertainty: Toward a theory of discretionary decision making in felony case processings. *Criminology*, 24(4), 623-644.
- Albonetti, C. A. (1987). Prosecutorial discretion: The effects of uncertainty. *Law and society review*, 291-313.
- Anderson, D. R., Baradaran, S., Lopez, A. B., Perry, R., & Center, K. B. (2011). Restoring the Presumption of Innocence. *Ohio State Law Journal*, 72(4).
- Bibas, S. (2004). Plea Bargaining outside the Shadow of Trial. *Harvard Law Review*, 117(8), 2463-2547
- Bowman III, F. O. (2006). The Year of Jubilee... or Maybe Not: Some Preliminary Observations about the Operation of the Federal Sentencing System after Booker. *Hous. L. Rev.*, 43, 279.
- Bradley-Engen, M. S., Engen, R. L., Shields, C., Damphousse, K. R., & Smith, B. L. (2012). The time penalty: Examining the relationship between time to conviction and trial vs. plea disparities in sentencing. *Justice Quarterly*, 29(6), 829-857.
- Breen, P. D. (2011). The Trial Penalty and Jury Sentencing: A Study of Air Force Courts-Martial. *Journal of Empirical Legal Studies*, 8(1), 206-235.
- Breen, P. D., & Johnson, B. D. (2018). Military Justice: Case Processing and Sentencing Decisions in America's "Other" Criminal Courts. *Justice Quarterly*, 35(4), 639-669.
- Brereton, D.; Casper, J. D. (1981). Does it pay to plead guilty differential sentencing and the functioning of criminal courts. *Law Society Review*, 16(1), 45-70.
- Bushway, S. D., Redlich, A. D., & Norris, R. J. (2014). An explicit test of plea bargaining in the "shadow of the trial". *Criminology*, 52(4), 723-754.
- Comparing Federal & State Courts. (n.d.). Retrieved November 5, 2019 from <https://www.uscourts.gov/about-federal-courts/court-role-and-structure/comparing-federal-state-courts>
- Dixon, J. (1995). The Organizational Context of Criminal Sentencing. *American Journal of Sociology*, 100(5), 1157-1198.
- Guidelines Manual Annotated. (2018). Retrieved November 2, 2019 from <https://www.ussc.gov/guidelines/2018-guidelines-manual-annotated>
- Guidorizzi, D. D. (1998). Should we really ban plea bargaining: The core concerns of plea bargaining critics. *Emory Lj*, 47, 753.

- Hall, J. S. (1999). Guided to Injustice: The Effect of the Sentencing Guidelines on Indigent Defendants and Public Defense. *Am. Crim. L. Rev.*, 36, 1331.
- Hofer, P. J. (2007). United States v. Booker as a natural experiment: Using empirical research to inform the federal sentencing policy debate. *Criminology & Public Policy*, 6(3), 433-460.
- Johnson, B. D. (2003). Racial and ethnic disparities in sentencing departures across modes of conviction. *Criminology*, 41(2), 449-490.
- Johnson, B. D. (2006). The multilevel context of criminal sentencing: Integrating judge-and county-level influences. *Criminology*, 44(2), 259-298.
- Johnson, B. D. (2019). Trials and tribulations: The trial tax and the process of punishment. *Crime and Justice*, 48(1), 313-363.
- Kim, A. (2015). Underestimating the trial penalty: An empirical analysis of the federal trial penalty and critique of the abrams study. *Mississippi Law Journal*, 84(5), 1195-1256.
- Kitchens, Christine. USSC Research Notes Issue 1. (n.d.). Retrieved October 29, 2019 from https://www.usc.gov/sites/default/files/pdf/research-and-publications/research-notes/20190719_Research-Notes-Issue1.pdf
- Military Courts. (n.d.). Retrieved November 4, 2019 from <https://www.fjc.gov/history/courts/military-courts>
- Packer, H. L. (1964). Two models of the criminal process. *U. Pa. L. Rev.*, 113, 1.
- Scott, R. (2009). The Effects of Booker on Inter-Judge Sentencing Disparity. *Federal Sentencing Reporter*, 22(2), 104-108.
- Spohn, C. (1992). An analysis of the "jury trial penalty" and its effect on black and white offenders. *Justice professional*, 7(1), 93-112.
- Starr, S. B. & Rehavi, M. (2013). Mandatory sentencing and racial disparity: Assessing the role of prosecutors and the effects of booker. *Yale Law Journal*, 123(1), 2-81.
- Steffensmeier, D., Ulmer, J., & Kramer, J. (1998). The interaction of race, gender, and age in criminal sentencing: The punishment cost of being young, black, and male. *Criminology*, 36(4), 763-798.
- Ulmer, J. T. (1997). *Social worlds of sentencing: Court communities under sentencing guidelines*. Albany, NY: State University of New York Press.
- Ulmer, J. T., & Bradley, M. S. (2006). Variation in trial penalties among serious violent offenses. *Criminology*, 44(3), 631-670.
- Ulmer, J. T., Eisenstein, J., & Johnson, B. D. (2010). Trial penalties in federal sentencing: extra-guidelines factors and district variation. *Justice Quarterly*, 27(4), 560-592.

Ulmer, J. T., & Johnson, B. (2004). Sentencing in context: A multilevel analysis. *Criminology*, 42(1), 137-178.

United States v. Booker, 543 U.S. 220 (2005)

Appendix A

	Median Sentence Length	Plea Proportion	Booker	Violent Offenses	Caseload	Regress
Median Sentence Length	1.000	-0.538	0.008	0.185	-0.477	0.949
Plea Proportion	X	1.000	0.304	-0.454	0.577	-0.508
Year	X	X	1.000	-0.372	0.381	0.027
Violent Offenses	X	X	X	1.000	-0.311	0.173
Caseload	X	X	X	X	1.000	-0.465
Regress	X	X	X	X	X	1.000

Table 3: This correlation matrix for all variables of interest shows no evidence of a collinearity issue.

Appendix B

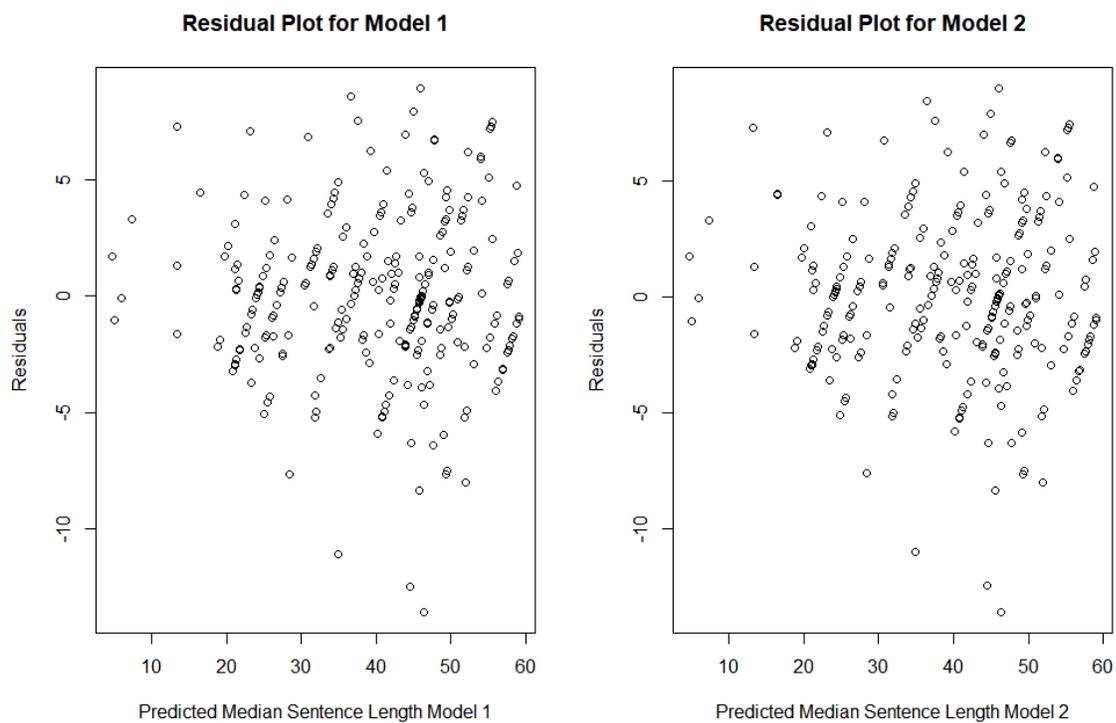


Figure 4: Residual Plots for Model 1 and Model 2.

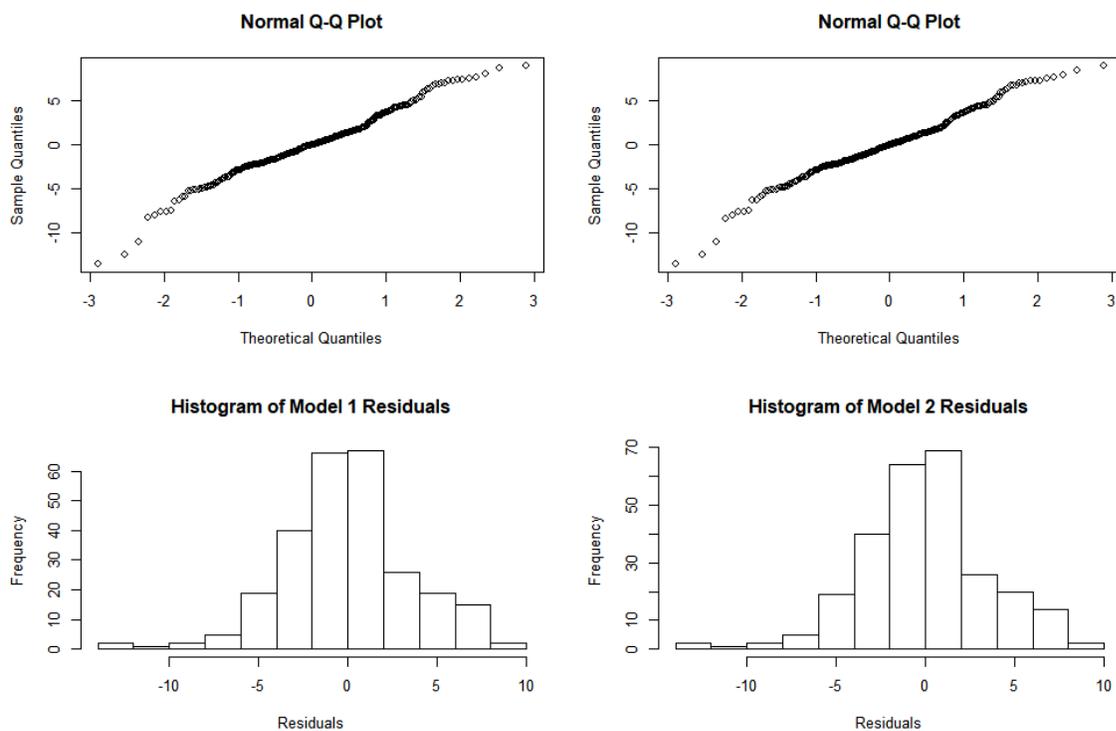


Figure 5: Normal quantile plots and residual histograms for Model 1 and Model 2, respectively.