REU Project: Investigating Institutional Structure

Galen Harrison
Reed College

Mentor: Simon DeDeo
Omidyar Fellow, Santa Fe Institute
(Dated: August 20, 2013)

I. INTRODUCTION

Criminal justice systems are some of the most influential institutions in any society. They are the mechanism through which the state may legitimately exercise its power over its citizens. Furthermore, these systems are both ubiquitous, in that they must be possessed in one form or another by any state, and understudied, at least from a positive standpoint. As such, we would like to establish an understanding of how the decision making structure of these systems affects the functional behavior of the

The Proceedings of the Old Bailey provide an avenue into understanding the institutional decision making process of a criminal justice system. By providing a record of trials held in the Old Bailey courthouse, the Proceedings let us analyze the output of some sub-elements of the system. We are able to efficiently search and extract data from the Proceedings by the digitization effort lead by Tim Hitchcock et. al. [1]. We leverage this data to investigate the relationship between various factors at two scales of action, macro and micro.

Mutual information is a quantity that measures the strength of a relationship between two distributions, so it is well suited to identifying deterministic factors in decision making. By providing a quantifiable measurement of how much two quantities say about one another, mutual information is a useful metric in examining the decision-making structure of an organization.

In this paper, we will use mutual information to describe and theorize about phenomena from the Proceedings. We will look at the aggregate level interactions between the state and civil society, as well as the interactions within trials based on gender at the micro level.

II. METHODS

It's necessary to conceptualize the parties involved in the decision to convict and punish the defendant. While this goal may seem to be tautological with the previously stated one of determining the institutional structure of the system, we would like to determine the *functional* decision structure. In most systems we would like to describe, the decision makers or the group the decision makers belong to is known. However, this is not usually informative as to what factors in the decision making. We would like to know what these people or groups of people consider when coming to their decision.

In this case, we know that there are three parties that are involved, the judge, the jury and the prosecutors. We have no real insight into decisions made by the prosecutors since we only know about decisions being made one way (when they make the decision not to bring a case, we have no record of their not acting). We do have a record of the complete output space of both the jury and the judge however. The jury decides whether or not to convict, and the judge decides what the punishment should be.

Mutual information is a way to measure relationships between two sets of variables. Mutual information is informally, the average decrease in the number of yes/no questions one has to ask to determine y if x is known, versus if you don't know x. While there are exact methods for calculating mutual information, these rely upon the assumption that the distribution is discrete. Since we expect the quantities we would like to study (as will be explained later, conviction rate, and average severity) to be continuous, we must employ an estimator. For this purpose, we turn to the estimator described by Kraskov, et al. [2].

This estimator finds the k^{th} nearest neighbor to every point in the distribution and then for each point finds the number of points in X within some ϵ determined by the distance of that k^{th} point and the number of points in Y determined by that same ϵ. The mutual information is then estimated to be

$$\hat{MI}(X,Y) = \psi(k) - \langle \psi(n_x) + \psi(n_y) \rangle + \psi(N),$$

where N is the number of points in the distribution, and n_x and n_y are the number of points within ϵ_i (where ϵ_i is the ϵ set by the i^{th} point). The digamma functions of each n are averaged over the entire distribution. This will be used in examining distributions at both micro and macro levels of analysis.

III. CONVICTION RATE AND SEVERITY

At the aggregate level, we expect conviction rate to reflect the attitude of civil society towards criminal justice. The conviction rate is entirely determined by juries, which were samplings from middle class men. The severity is determined by the judge, who is an agent of the

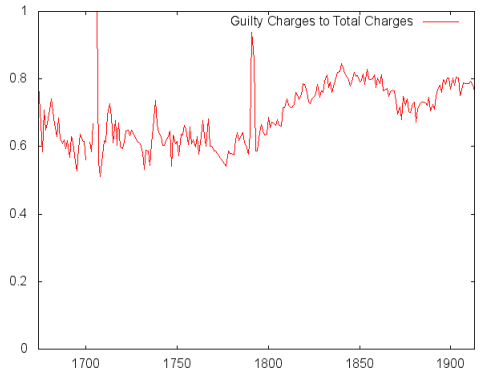


FIG. 1: The conviction rate is shown for 1674-1913. The spike in 1705 is due to there only being one record for the year. The spike in 1791 is due to the censorship of not guilty verdicts.

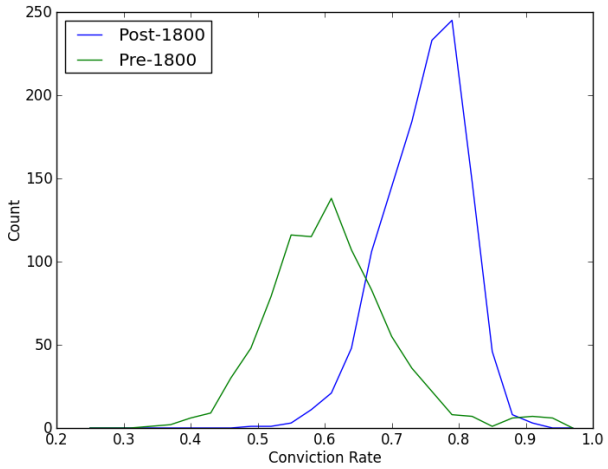


FIG. 2: The different distributions for conviction rate from pre-1800 to post 1800. Post-1800 average: 0.76 ± 0.06 Pre-1800 average: 0.6 ± 0.1

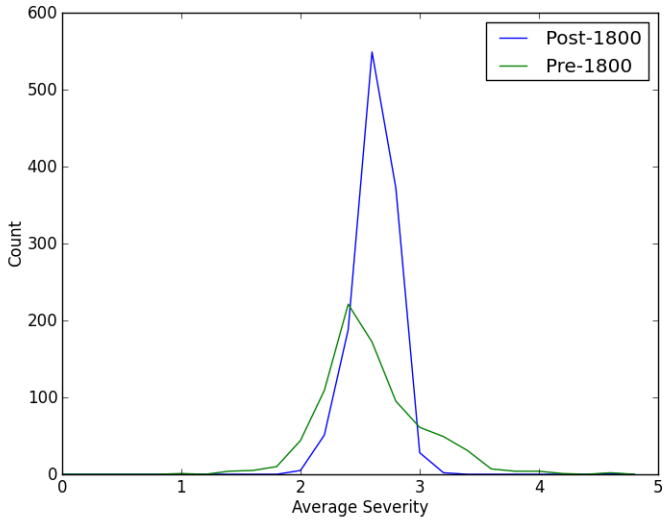


FIG. 3: The average severity by session of the punishments handed down post 1800 and pre 1800. Severity measured by no punishment 0, miscellaneous punishment 1, transport 2, imprisonment 3, corporal 4, and capital 5. Pre-1800 average: 2.7 ± 0.6, Post-1800 average: 2.8 ± 0.2

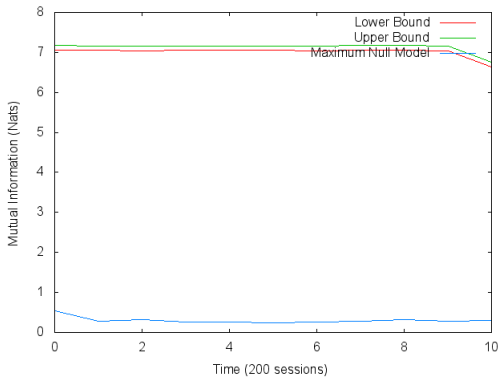


FIG. 4: Mutual information for groups of 200 sessions over time. Error bars are derived by bootstrapping 1000 times. The null model is also shown

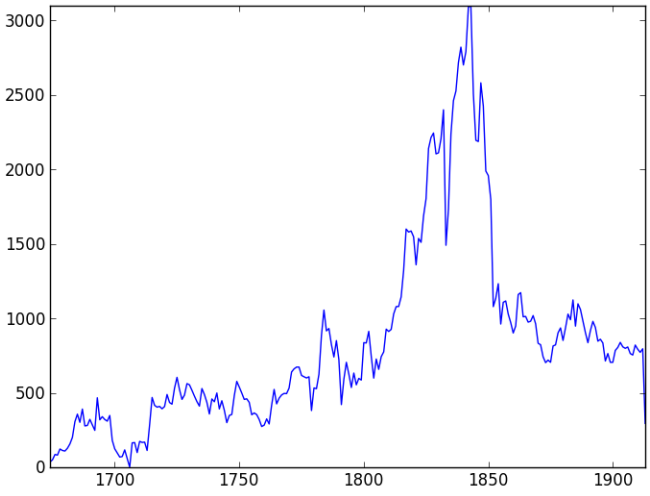
state. Therefore, we would like to see if there is a relationship between the actions of the state and the actions of civil society.

In figure 1, the conviction rate from 1674 till 1913 is shown. The conviction rate stays around 60% until 1800, when it climbs to around 80%. The different distributions are shown in 2. This is a shift in the actions taken by civil society. Looking at the actions of the state as measured by average severity in figure 3, we see that there is very little change in their actions. To see if there's any relationship between these two, we can employ mutual information.

If we form distributions of 200 sessions (sessions are the smallest subset of groups of trials), where each session is represented by its conviction rate and average severity, we can measure the mutual information. The average severity is measured on scale of 0-5, where 0 is no punishment and 5 is capital punishment. In figure 4, the graph of the mutual information through time is shown. The constant mutual information indicates that the strength of the relationship is constant across the entire time series.

This split is correlated with the growth of the criminal justice apparatus. The number of prosecutions grows post 1800 and the portion of punishments handed down that are imprisonment (as opposed to some punishment that requires less investment from the state) increases as well (fig 5). This suggests that there is a growth in the amount of resources put into the system, which suggests that the apparatus grows and concordant formalization increases.

While it is necessary to note the coinciding trends, it is



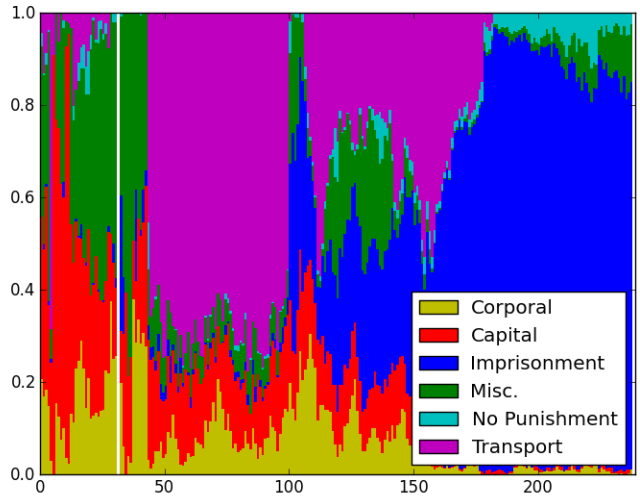


FIG. 5: The number of prosecutions by year (upper) and the portion of punishments by year (lower). The increase in both imprisonment, and the total number of prosecutions brought indicates an increase in the size of the of the criminal justice apparatus

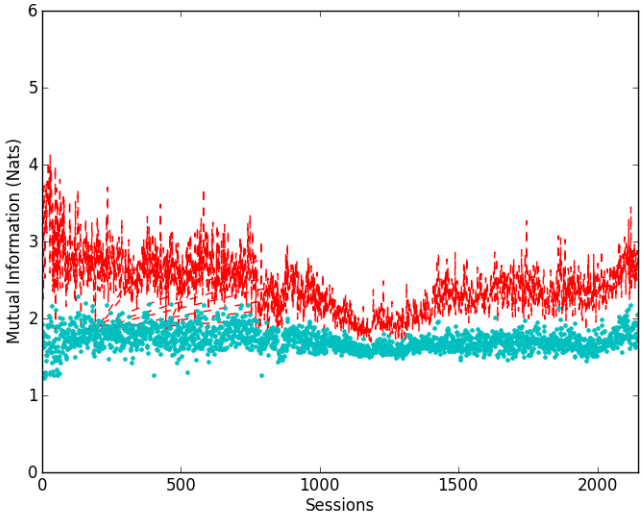


FIG. 6: The mutual information between the gender tuple and the conviction rate in a trial. Error bars produced via bootstrapping 1000 times, null model subtracted out.

impossible to say that there is a causal relationship one way or another. The mutual information suggests that there is some relationship between the actions of the state and the actions of civil society, and the split shows that the relationship is entirely static. It is our belief that there is a set of decision rules governing the interactions of the judge and the jury, and further investigation is necessary to discern them. Possibly looking at the charge will provide some insight into this.

IV. GENDER

We can apply a similar technique to gender. In each trial, there is a tuple defined by the number of male victims, the number of female victims, the number of male defendants and the number of female defendants. Looking at the distribution formed by each session, we can measure the mutual information of the gender tuple and the trial outcome, as measured by the conviction rate of the trial. This produces a constant mutual information, as shown in figure 6. While it might not matter in the same way, gender mattered just as much in 1674 as 1913.

The constancy shows that there is a relationship between the gender of the participants and the portion of the charges that the jury chose to affirm. What those rules are and whether or not they changed is not clear, and further investigation could try and determine what they are.

V. CONCLUSION

We have shown that there is some relationship between the way civil society acts and the way that the state acts in the criminal justice system. This relationship may change over time, whether this is true and what the exact relationship is an area for further investigation. This is also true for gender, we have evidence for some deterministic relationship, but do not know what the relationship is or whether it changes at any point. Preliminary investigations using decision trees have not shown any results, it is possible that it might be necessary to take the charge into account when trying to construct such a

tree.

Criminal justice systems are important and understudied institutions and therefore are important loci for research and investigation. By finding two strong relationships in the data, we show that the Proceedings provide a rich source of information for understanding the English criminal justice system.

[1] T. Hitchcock, R. Shoemaker, C. Emsley, S. Howard, and e. a. Jamie McLaughlin, "The Old Bailey Proceedings Online, 1674-1913," (????), URL (`www.oldbaileyonline.org,version7.0,24March2012`).

[2] A. Kraskov, H. Stögbauer, and P. Grassberger, "Estimating mutual information," Physical Review E **69**, 066138 (2004).