A nonparametric view of the civilizing process in London’s Old Bailey

UCDMS Research Report No. UCDMS2015/1, School of Mathematics and Statistics, University of Canterbury, Christchurch, NZ

Jasper Mackenzie ∗, Raazesh Sainudiin ∗ †, James Smithies ‡ and Heather Wolffram †

∗School of Mathematics and Statistics, and † School of Humanities and Creative Arts, University of Canterbury, Christchurch, New Zealand

An increasing number of large humanities data sets are becoming available, and new tools and methods are required to analyze them. There is a risk that statisticians and humanists will fail to recognize the historical contingency of such data, and make assumptions on the distribution. If appropriate methods do not appear, algorithmic analysis of large humanities dataset will only be able to be used in a heuristic sense, to augment current understanding and prompt new questions and angles of analysis but not to make strong empirical claims. This work develops Bayesian nonparametric models that allow researchers to ask longitudinal questions of large humanities data sets with confidence they have corrected for pre-existing bias derived from the received tradition.

The methods are tested on the Old Bailey Online (OBO), which draws its data from the notorious Old Bailey court of London England. The Court’s records span an unprecedented temporal landscape, from 1674 to 1913. The period witnessed significant legal and cultural change and is considered by historians to be a foundational period in the development of contemporary civil society. Coarse grain data regarding punishment and offence pairs are used to look for evidence of the so-called ’civilizing process’, a crucial marker of change associating increases in bureaucratic control with a decrease in social disorder. The chosen model, selected using the Akaike Information Criterion (AIC) for model selection, suggests the underlying process governing the Old Bailey data changes every three years. By partitioning offences it is found that over time the distribution of punishments for each offence becomes less haphazard, offering strong evidence of the civilizing process.

The methods used in this study are mathematically rigorous, based on open-source and widely available tools, and able to be used by a range of investigators. As such, this work contributes to the foundations being established for computational analysis of long-timescale historical trends using large humanities datasets.

Humanists have a long tradition of using statistical methods and empirical approaches in their research, but the appearance of major monographs exploring digital methods [1, 2, 3] and the proliferation of ‘Big Data’ experiments [4, 5, 6, 7, 8] reflects a need for the development of new methods, and the deployment of appropriate mathematical tools, to ensure conclusions drawn from the data are valid. The inter-disciplinary nature of the digital humanities requires that these methods be both mathematically rigorous and useful for domain experts. This study aims at providing a model for researchers wishing to explore digital archives of historical data to find longitudinal trends. It uses the Old Bailey Online (OBO) as a reference data source[9].

The OBO is a publicly accessible digital resource that stores digitized proceedings of trials held in the Old Bailey court in London between 1674 and 1913 [10, 9]. The Old Bailey tried crimes ranging from petty theft to violent offences that carried the death penalty. Because of this the Old Bailey data provides an excellent resource to investigate changing trends and attitudes towards violence. The OBO website offers a historical overview of the Old Bailey together with the digitized and marked up copies of all known surviving records of the Proceedings. This combination of easy to search data and rich historical and contextual information makes the OBO a flagship example of a digital humanities data set, a model anticipated to be reproduced as more funding makes becoming available to digitize archives. From this rich data source punishment frequencies per offence over the entire period of available data are extracted and stored as offence-punishment pairs for the purpose of modelling the underlying processes.

Despite its richness, few studies draw conclusions from quantitative analysis of the OBO data. This study is inspired by, and in many ways is a response to [11], in which the authors analyzed semantic changes in the data set to measure changing attitudes towards crime over time. Their central conclusion was that “changes in bureaucratic practice and in civil society as a whole” between 1760 and 1913 led to greater differentiation between violent and non-violent crimes. The authors view this as evidence of the so-called civilizing process. Other relevant studies include the multi-national effort conducted as part of a ’Digging into Data’ challenge [12], and a discussion of its implications for corpus linguistics [13]. An indication of the mathematical richness of the data is suggested in a study which developed a statistical bootstrap for the estimation of information theoretic quantities in OBO data [14]. The ”civilizing process” is a theory of social change developed by Norbert Elias, a German sociologist, in 1939[15]. It remains well regarded. Its central insight is that increasing levels of social control (professional policing and the regular

Significance

Appropriate mathematical tools and methods are needed by the digital humanities community to assist them with the analysis of large datasets. Due to the inter-disciplinary nature of the field, these methods would ideally be both mathematically rigorous and able to be used by researchers without advanced statistical training. It is also important that computational statisticians understand the complex nature of humanities data, and the constraints it imposes on big data research. This study contributes to both of these issues through a clear exposition of methods tailored specifically for the digital humanities community, a description of key analytical issues, and the publication of Python code in the public domain to encourage further collaboration and growth.

Author Contributions: JM, RS and JS initiated the project. JS provided digital humanities context. JS and HW provided historical interpretation. RS developed the mathematical models and methods. JM and RS analyzed the data. JM coded to obtain data and implement models. JM prepared figures and wrote the first draft and RS wrote the second draft with contributions from all authors.

The authors declare no conflict of interest.

This article contains supporting information online at https://github.com/spmp/OBO-APInection-and-Analysis
zation of judicial processes, for example) are tightly corre-
operatorized by reduced crime and violence, and the development of a middle-class[16, 17]. For Elias this increase in civility is a multifaceted and complex change that arises in many areas of society. The theory claims to account for multifaceted as-
pects of ‘civility’, ranging from long-term declines in homicide and state violence[18], and manners becoming more refined[19, 5-7]. The Old Bailey, being the court with responsibility for crimes deemed violent by both society and the law[20, 9], is an ideal place to look for long-timescale evidence of the civilizing process through the bias of offences heard and the punishments given for those offences.

Researchers working with humanities data sets must be mind-
ful of the complex nature of the data’s origin and the complex-
ity of knowledge required to bring meaning to results gener-
ated from it. The data in this study is of historic legal origin, firmly rooted in the history of crime and punishment in Eng-
land. Care must be taken when working with this kind of data not to influence the analysis with either limited historic underpinnings or invalid assumptions based on over identifi-
cation with qualitative (or previous quantitative) evidence. A critical awareness of the choices imposed on the data and mod-
els must be maintained. A non-parametric approach is key to
preserving the integrity of any robust signal in the data. In this approach, models of the data are created directly from the empirical frequencies and do not rely on fitting observations to any preconceived parametric model. In this way, the models themselves have intrinsic meaning, allowing the data to ‘speak for itself’ in historically interpretable ways. In our study, no presuppositions of what might be found in the data have been made other than to look for signs of the so-called civilizing process. Using non-parametric probabilistic models and the penalized likelihood principle called the Akaike Information Criterion (AIC) for model selection [21], it is found that the best model for the offence-punishment data is one where punishment depends on offence and where these proba-
bilities vary significantly across consecutive three year periods. This in itself is an intriguing finding, given it suggests a rel-
atively short procedural ‘rhythm’ for this key mechanism of State.

A more general model for partitioned offences is also created such that offences that are grouped within the same block of the partition share the same punishment distribution. This aspect of the study is motivated by exploratory data analysis, with the goal of finding specific evidence of the civilizing pro-
cess. The punishment outcomes for partitioned offences show
strong trends towards greater discrimination in both punishments per offence and in the group of offences most likely to receive death. Together with the decline in the death sen-
tence evident over time, these results provide strong evidence of the civilizing process in action: from the 1640s, where pun-
ishments were violent and haphazard, to 1913 where punish-
ments were well defined, commensurate with the offence and the likelihood of the death sentence was low.

The approach used provides visualizations of the OBO data that facilitates interpretation by historians exploring longitudi-
nal trends. In this way the work serves as a platform for addressing a rich array of historical questions.

The Python[22] programming language is chosen because of its
general suitability for data extraction and analysis, but also
because it is widely used in the digital humanities community. The Python library ‘Pandas’[23, 24, 25] provides a powerful framework for data storage and manipulation, avoiding the need to use additional mathematical software such as Matlab, GNU Octave, or R, and enabling the entire study to be com-
pleted within a single ecosystem. This has considerable benefit
for humanists who may be interested in the possibilities inher-
ent in computational analysis, but limited in the time they can
spend learning new packages. The study contends that these
kinds of ‘project design’ decisions are important for the future
of interdisciplinary studies in the digital humanities.

To help historians explore the civilizing process, and to ac-
count for the nature of the data set itself, the simplest account of the trials that the OBO has to offer (frequencies of offences and punishments) is used instead of a more complex analysis of coarse-grained spoken word testimony of trial transcripts as done in [11]. It is shown that non-parametric methods can extract statistically significant and historically meaningful results even from such a simple resolution of the data, provided the signals are strong enough.

Offence-punishment pairs are the fundamental units of this study and formed from frequencies of punishments per of-
ference as taken from the OBO web-API. In their markup of trials, the OBO include offence, verdict, and punishment catego-
ery per accused person, exposing the counts per categori-
cal bin and in temporal blocks. The counts are downloaded via the web-API. Verdicts are considered to be binary: guilty or not guilty, and without loss of gen-
erality, the NotGuilty verdict is combined with punishments such that all offences can map to a punishment-like outcome. Let \([n]\) be the set of natural numbers no larger than \(n\), i.e.,
\([n]\) = \{1, 2, ..., \(n\)\}. The nine offences are notated \(o_i\) where \(i\) ∈ [9], and the seven punishments are notated \(p_j\) where \(j\) ∈ [7] as defined in Figure 1.

Event counts are downloaded from the OBO in yearly time-
blocks and during the course of the analysis are re-sampled for
differing time-block lengths, \(\Delta t\), where typically:

\[\Delta t \in \{1, 2, 3, 4, 5, 10, 15, 20, 25, 30, 40, 50, 100, 240\}\]

and for a given \(\Delta t\) and a corresponding
deviation:

\[\tau = \tau(\Delta t) = \lfloor 240/\Delta t \rfloor,\]

the time-blocks are indexed by \(t\) ∈ \(\tau\). Thus, \(t\) is defined as done in [11]. It is shown that non-parametric methods can
extract statistically significant and historically meaningful
results even from such a simple resolution of the data, pro-
vided the signals are strong enough.

Methods

It is important that appropriate tools are selected for the study
of historical datasets, to encourage collaboration between ex-
erts in computational statistics and the humanities. A simple
and general reference methodology is chosen for this reason.

\[c_{i,j,t} = \sum_{r=1}^{n(t)} \mathbb{I}(o_r, p_j, o_{i(t)}),\]  

where

\[\mathbb{I}_A,B(x, y) = \begin{cases} 1 & \text{if } x \in A \text{ and } y \in B \\ 0 & \text{otherwise} \end{cases}\]

\[\Delta t \in \{1, 2, 3, 4, 5, 10, 15, 20, 25, 30, 40, 50, 100, 240\}\]

and for a given \(\Delta t\) and a corresponding

deviation:

\[\tau = \tau(\Delta t) = \lfloor 240/\Delta t \rfloor,\]

Thus, cases in a given time-block are heard at the Old Bailey as a sequence of \(n(t)\) many offence-punishment pairs:

\[(o_{i_1}, p_{j_1}), (o_{i_2}, p_{j_2}), ..., (o_{i_{n(t)}}, p_{j_{n(t)}})\].

This sequence is summed to obtain the counts for all 63 possible
\((o_i, p_j)\) pairs as:

\[c_{i,j,t} = \sum_{r=1}^{n(t)} \mathbb{I}(o_r, p_j, o_{i(t)}),\]  

where

\[\mathbb{I}_A,B(x, y) = \begin{cases} 1 & \text{if } x \in A \text{ and } y \in B \\ 0 & \text{otherwise} \end{cases}\]
Offence categories: $o_1 = $ BreakingPeace, $o_2 = $ Damage, $o_3 = $ Deception, $o_4 = $ Kill, $o_5 = $ Miscellaneous, $o_6 = $ RoyalOffences, $o_7 = $ Sexual, $o_8 = $ Theft, $o_9 = $ ViolentTheft

Punishment categories: $p_1 = $ NotGuilty, $p_2 = $ Corporal, $p_3 = $ Death, $p_4 = $ Imprison, $p_5 = $ MiscPunish, $p_6 = $ NoPunish, $p_7 = $ Transport

**Fig. 1:** The offence categories and punishment categories in use in this study, and as defined by markup in the OBO with the addition of ‘NotGuilty’ to the punishments.

Offence-punishment pairs are assumed to be identically distributed within each time-block and independent of all other offence-punishment pairs in any time-block. Observed empirical frequencies of offence-punishment pairs are therefore assumed to be realizations drawn from a true time-dependent probability distribution, with the tacit assumption that there is such a distribution, and that the events seen at the Old Bailey and represented in the OBO are driven by this as yet undiscovered process. In order to discover this true distribution, estimates are made of the process governing events at the Old Bailey from the data. To this end dependent models in which punishment depends on offence, and independent models in which punishment is independent of offence, are estimated and selected over each time-block, on the basis of their sparsefrequent fit to the observed data.

The probability of observing $(o_i, p_j, t)$, the event of offence $o_i$ resulting in punishment $p_j$ in a trial in time-block $t$, is given by the discrete probability distribution $(\theta_{i,j,t})$ as:

$$\Pr\{o_i, p_j, t\} = \begin{cases} \theta_{i,j,t} & \text{if } i \in [9], \text{ and } j \in [7] \\ 0 & \text{otherwise} \end{cases}$$

where the non-parametric model allows for $(\theta_{i,j,t})$ to be any discrete probability distribution:

$$\theta_{i,j,t} \geq 0 \text{ for each } i, j, t \text{ such that } \sum_i \sum_j \theta_{i,j,t} = 1,$$

for each $t$ in $[\tau]$. A model in this study is a time varying discrete probability distribution for offence-punishment pairs, with a separate estimated distribution for every time-block $t$ specified by a time-block length $\Delta t$. Since the offence-punishment pairs are assumed to be independently and identically distributed (i.i.d.) within each time-block, the joint probability of the sequence of offence-punishment pairs is:

$$\Pr\{(o_1, p_{j_1}, t), (o_2, p_{j_2}, t), \ldots, (o_{n(t)}, p_{j_{n(t)}}, t)\} = \prod_{r=1}^{n(t)} \Pr\{o_{r(t)}, p_{j_{r(t)}}, t\} = \prod_{r=1}^{n(t)} \prod_{t=1}^{\tau} (\theta_{i,j,t})^{c_{i,j,t}}.$$

Thus for the entire data $d$:

$$d = \{(o_1, p_{j_1}, t), (o_2, p_{j_2}, t), \ldots, (o_{n(t)}, p_{j_{n(t)}}, t)\}_{t=1}^{\tau},$$

its likelihood is given by:

$$L(\theta) = \Pr\{d|\theta\} = \prod_{t=1}^{\tau} \prod_{i=1}^{9} \prod_{j=1}^{7} (\theta_{i,j,t})^{c_{i,j,t}}.$$

where, the nonparametric model is specified by $\tau$ discrete probability distributions in $(9 \times 7) - 1 = 62$ dimensions given by $\theta$ as follows:

$$\theta = \{\theta_{i,j,t} : i \in [9], j \in [7], t \in [\tau]\}.$$
where the number of parameters for the time varying dependent and independent models are:

\[ K_{\text{step}} = \tau \times ((9 \times 7) - 1) = \tau \times 62 \]

\[ K_{\text{score}} = \tau \times ((9 \times 1) + (7 - 1)) = \tau \times 14 \]

We also applied the second-order AIC or AICc given by:

\[ \text{AICc} = \text{AIC} + \frac{2K(K + 1)}{n - K - 1} \]

for model selection and found the model ranking to be identical to that based on AIC. Here \( n = 218,288 \) is the sample size, i.e., the total number of trials that were analyzed from Old Bailey Online. This confirms that our model selection results are invariant to the penalty functions of AIC and AICc.

The confidence interval for the estimates \( \hat{\theta}_{i,t} \) of the best model is found from the inner 95 percentile of a series of estimates obtained from bootstrapped data. The bootstrapped data is obtained by randomly sampling (with replacement) the original data, where the number of samples must be the same as the number of events in each time-block of the original data. Parameters are estimated from each of the 10,000 bootstrapped realizations with the inner 95 percentile found for each \( \hat{\theta}_{i,t} \) from the bootstrapped estimates. Instead of obtaining Bayesian credible intervals by sampling from the Dirichlet posterior density over the probability simplex \( \theta \) for each \( t \), the uncertainty in our point estimates is obtained by bootstrapping the data. This is to ensure that the pragmatic Bayesian point estimate which is chosen to remain close to MLE (due to the small precision parameter \( \beta \)), has a classical frequentist interpretation. The Bayesian credible intervals are generally more concentrated (except for frequentist interpretation. The Bayesian credible intervals are generally more concentrated (except for \( t = 1 \)) than the bootstrapped intervals if we directly sampled from the Dirichlet posteriors for each time-block. Thus our confidence intervals are generally more conservative.

For visually exploring the estimates for a given time-block \( t \), \( \hat{\theta}_{i,t} \), the sum of the absolute differences between the average of a given pair of offences \( \varphi_i \) and \( \varphi_j \) is used as a distance measure given by:

\[ \delta(\varphi_i, \varphi_j; t) = \sum_{j=1}^{7} |\hat{\theta}_{i,j,t} - \hat{\theta}_{j,i,t}|. \]

Thus the smaller the value of \( \delta(\varphi_i, \varphi_j; t) \), the closer are the relative punishment outcomes for the pair of offences \( \varphi_i \) and \( \varphi_j \) in time-block \( t \). For a given time-block there are such \( \delta \) values from 36 pairs of offences and is used for visual explorations of the high dimensional estimates.

The dependent model can be made more general if the probability of receiving a punishment is allowed to depend on a block of offences. This is formalized by partitioning the set of offences into blocks such that the punishment probability depends only on the block, where all offences within the block have the same punishment distribution. We consider all 21,347 possible partitions of the 9 offences for each time-block \( t \). A partition \( \psi_m \) of the set of offences \( \{\varphi_1, \ldots, \varphi_9\} \) is denoted as:

\[ \psi_m = \{\psi_m^1, \psi_m^2, \ldots, \psi_m^{(m)}\} \]

where each offence-block \( \psi_{m}^{k} \) is non-empty, i.e., \( \psi_k^m \not= \emptyset \), and each pair of offence-blocks is disjoint, i.e., \( \psi_k^m \cap \psi_l^m = \emptyset \) for \( k \not= l \). Thus \( m \) enumerates all such partitions where \( m \in [21147] \) and \( \ell(m) \) is the number of offence-blocks within a partitions \( \psi_m \). Thus, a given offence, \( \varphi_i \), is in a particular offence-block of a given offence partition \( \psi_m \). Let the unique offence-block of \( \psi_m \) containing \( \varphi_i \) be \( \psi_i^{(m)} \), i.e.,

\[ \psi_i^{(m)} \cap \{\varphi_i\} \not= \emptyset \]

Finally, let the offence partition for the time-block \( t \) be \( \psi_{mt} \) with \( \ell(m) \) offence-blocks, where \( m \in [21147] \) for each time-block \( t \in [\tau] \).

Taking advantage of the temporal order of offence before punishment, the joint probability of observing \( \{\varphi_i, p, t\} \), the event of offence \( \varphi_i \) resulting in punishment \( p_j \) in a trial at time-block \( t \), is given by the discrete probability distribution \( \{\theta_{i,j,t}\} \) decomposed as follows:

\[ \text{Pr}(\varphi_i, p_j, t) = \text{Pr}(p_j|\varphi_i, t) \times \theta_{i,j,t}, \]

where, \( \theta_{i,j,t} = \text{Pr}(\varphi_i, t) \) is the probability of committing offence \( \varphi_i \) in time-block \( t \). Now the likelihood of \( \{\varphi_i, p_j, t\} \) under a given offence partition \( \psi_m \) with \( \ell(m) \) blocks is proportional to the following conditional probability:

\[ \text{Pr}(\varphi_i, p_j, t|\psi_m) = \text{Pr}(p_j|\varphi_i, t, \psi_m) \text{Pr}(\varphi_i, t|\psi_m) = \theta_{i,j,m} \times \theta_{i,t} \]

where,

\[ \theta_{i,j,m} \text{ and } \theta_{i,t} \]

are discrete probability distributions with free parameters in \( \ell(m) \times (7 - 1) = 6\ell(m) \) and \( 9 - 1 = 8 \) dimensions, respectively. Thus, the number of parameters for the time varying offence-partitioned model specified by

\[ \theta = \{\theta_{i,m,t} : t \in [\tau]\} \]

is:

\[ K_{\text{PART}} = \sum_{t=1}^{\tau} (8 + 6\ell(m)) = 8\tau + 6\sum_{t=1}^{\tau} \ell(m) \]

Therefore, for a given \( \Delta t \), the likelihood of a model specified by a sequence of \( \tau \) offence partitions is:

\[ L(\theta) = \text{Pr} \{d|\theta\} \]

\[ = \prod_{t=1}^{\tau} \left( \prod_{\ell=1}^{\ell(m)} \prod_{j=1}^{\ell(m)} \prod_{k=1}^{\psi_k^m} \theta_{i,j,m} \right) \prod_{i=1}^{9} \left( \theta_{i,t} \right)^{\Delta t} \]

and using the indicator function defined in Equation (2) the number of punishments \( p \), that are paired with any offence that belongs to the offence block \( \psi_{mt} \) in the time-block \( t \) is obtained as:

\[ c_{\psi_{mt}^k, i, \tau} = \sum_{t=1}^{\Delta t} \mathbf{1}(\psi_{mt}^k, i, \tau) \]

The Bayesian smoothed estimates are obtained analogously to Equation (5) with partitioned offences for \( \hat{\theta}_{mt} \) for each \( t \).

The AIC scores are found for all possible partitions of offences in a time-block and stored as a row in a table, such that the best performing partitions through time are determined as the minimum row scores per time-block. However, the difference between the best partitioning and the next best partitioning across all time-blocks is not significant, therefore we introduce a distribution over the partitioned models through time-blocks using a Markov Chain Monte Carlo method (MCMC).

A Bayesian perspective on AIC allows the use of the previously generated AIC scores to be re-used as transition probabilities of our MCMC. Consider a Markov chain whose states are the
21147 offence-partitioned models at each of the $\tau$ time blocks. Let $M_{m,t}$ denote a state, i.e., a model with offence-partition $m$ in time-block $t$. Let us choose the proper prior distribution over the models as

$$
Pr(M_{m,t}) = \frac{e^{-6\ell(m)} - 8}{\sum_{i=1}^{9} S(9, i) \cdot e^{-6i} - 8} = \frac{e^{-6\ell(m)}}{\sum_{i=1}^{9} S(9, i) \cdot e^{-6i}}
$$

where $S(9, i)$ is the Stirling number of the second kind, giving the number of partitions of the offences $\{o_1, \ldots, o_9\}$ with $i$ blocks. This Akaike proper prior distribution allows for a pragmatic Bayesian perspective of AIC as the energy function underlying the desired posterior distribution as follows:

$$
Pr(M_{m,t} | d) \propto Pr(d | M_{m,t}) Pr(M_{m,t})
\approx Pr(d | \hat{\theta}_{m,t}) Pr(M_{m,t})
\propto \mathcal{L}(\hat{\theta}_{m,t}) \cdot e^{-6\ell(m)}
= e^{-AIC(M_{m,t})/2}
$$

The chosen prior is well-behaved, penalizing a model by the number of parameters in it, as required. With our Akaike proper prior, the posterior probabilities are:

$$
P_{m,t} = Pr(M_{m,t} | d) = \frac{e^{-AIC(m,t)/2}}{\sum_{m} e^{-AIC(m,t)/2}}
$$

![Fig. 2: The transition diagram of an MCMC sampler, with edge weights given by Equation (9) for models $M_{m,t}$ where $t$ is the time-block, $m$ ranges from 1 to 21147 ($N$). Each vertex in time-block $t$ has an edge to every vertex in $t + 1$ with the edge weight $Pr_{x+1,m}$ from the connecting model in time-block $t + 1$. The colouring in each time represents the same edge weight, with the black path being an example of a random path through the state space of models that specifies a time-block varying partitioned offence model.](image)

Figure 2 illustrates the transition diagram where the vertices represent a partitioned model, $M_{m,t}$, for time-block $t$ and offence-partition $\psi_m$, which has edges to all possible models in time-block $t + 1$. Thus, we allow a transition from $M_{m',t}$ to $M_{m,t+1}$ for any $m'$ and $m$. The edge weights represent the transition probabilities of the Markov chain and is given by the posterior probabilities. In other words, the transition probability of MCMC algorithm visiting $M_{m,t+1}$ in time-block $t + 1$ from any state in time-block $t$ is given by the posterior probability $P_{m,t+1}$ in [10]. A stochastic MCMC path is obtained by starting at the top state in Figure 2 and iteratively choosing one of the 21,147 possible models according to its posterior probability for each time-block $t$ until terminating at time-block $\tau$. Many such independent sample paths are obtained from the MCMC algorithm to obtain a distribution over such model paths from the state space of 21147 $\times$ $\tau$ models.

The methods presented here, and the software tools created to achieve these methodologies, form the basis of a general and transferable framework that can be used to interpret time-varying humanities data. The intention is to find an optimal compromise between computational rigor and generalizable method, capable of supporting radically interdisciplinary research.

**Results and discussion**

**The Old Bailey and the Proceedings.** To make sense of the modelling results, a brief historical background of the Old Bailey and the Proceedings is needed. Further details are presented as appendices in the supplementary material.

![Fig. 3: Empirical frequencies per offence for the entire period of the Old Bailey on a log scale with frequency on the y-axis and date on the x-axis. Note the absence of data in 1701, the dramatic changes in Theft occurrences from 1840 to 1855 and increase in all other categories from 1840.](image)

Various acts of the British Parliament implemented during the nineteenth century had major effects on the Proceedings, particularly the removal of certain charges and the removal of the death penalty for others [20, p. 176]. Some changes had a greater effect on the observed frequencies of offences...
and punishments than others, as seen in the plots of empirical frequencies of offences and punishments, Figure 3 and Figure 4. Jurisdictional changes across all offence category frequencies represented the two strongest events influencing the Old Bailey proceedings. In 1834 the Old Bailey’s jurisdiction changed from just the square mile of London to include Middlesex, parts of Essex, Kent, and Surrey and maritime offences previously heard at the Admiralty sessions [28] [20, p. 147]. This resulted in a sudden increase in offences (as seen in Figure 3), especially those that were largely unrepresented before the change. In 1855, for example, the incidence of Theft offences plummeted when less serious crimes were relocated to magistrates courts.

Explosive population growth in London provides a backdrop to events recorded in the raw data. London grew from 400,000 in 1650 to 2.3 million in 1850, and 7.1 million in 1911 [29]. This is evident from the linearity of growth in Theft up to 1855, in the logarithmic plot Figure 3.

![Graph](image)

**Fig. 4:** Empirical frequencies per offence for the entire period of the Old Bailey on a log scale with frequency on the y-axis. Note the absence of data in 1701, the beginning of Transport in 1718 with a sudden break in 1780, ceasing in 1857, and the dramatic increase in Imprison and decrease in Death from 1840.

Changes in law are also seen in individual offence and punishment categories. Transport to America was introduced in 1718 as an alternative punishment for death [9], halting briefly during Britain’s war with America in 1776, resuming in 1787 with transport to Australia, and finally ended in 1857 [30][31]. The signal of these key dates is strongly imprinted in the raw data (Figure 4), providing a compelling validation of existing historiography. The rise in imprisonments to compensate for the halting of transport to America and Australia is also evident in Figure 4.

The removal of the death penalty for most forms of theft in 1832 sees incidence of death and corporal punishments reduce suddenly and dramatically in the raw data, from a peak of 181 sentenced to death in 1827 to 5 in 1838 [9]. This reduces the empirical frequency of punishments by death by an order of magnitude despite the population growth, and thereby moving the judicial system towards retributive justice, where punishment is commensurate with crime. The impact of legal and administrative changes were not always straight-forward, however. The Prisoners Counsel Act of 1836 [32] gave defendants the right to make their “full answer and defence by counsel of learned law” [20][33, p. 176], and resulted in counsel having an increasing role in the trial process, did not lead to a marked increase in NotGuilty verdicts. Instead Figure 4 shows a decrease in NotGuilty from 1836, with dramatic declines in 1855 coinciding with the change in jurisdiction.

Jurisdictional change is not the only driving factor behind observed frequencies at the Old Bailey. During the period reported in the Proceedings policing in London underwent radical change, from the 1670s when the onus of apprehending culprits was the duty of the public, to the early 20th century where a modern police force with special investigative units was in place [34, 8-15]. Interestingly, unlike changes in the court and the Law, developments in policing method did not produce obvious signals in the data.

**Table 1:** AIC scores sorted in ascending order for dependent and independent models for various ∆t.

<table>
<thead>
<tr>
<th>No.</th>
<th>Model</th>
<th>K</th>
<th>log(C)</th>
<th>AIC</th>
</tr>
</thead>
<tbody>
<tr>
<td>4</td>
<td>Dep 3yr</td>
<td>4960</td>
<td>-480030.60</td>
<td>969981.219</td>
</tr>
<tr>
<td>6</td>
<td>Dep 4yr</td>
<td>3720</td>
<td>-481425.54</td>
<td>970291.088</td>
</tr>
<tr>
<td>2</td>
<td>Dep 2yr</td>
<td>7440</td>
<td>-447773.84</td>
<td>970427.684</td>
</tr>
<tr>
<td>8</td>
<td>Dep 5yr</td>
<td>2976</td>
<td>-483284.22</td>
<td>972520.448</td>
</tr>
<tr>
<td>10</td>
<td>Dep 10yr</td>
<td>1488</td>
<td>-487993.85</td>
<td>977163.715</td>
</tr>
<tr>
<td>0</td>
<td>Dep 1yr</td>
<td>14880</td>
<td>-474199.77</td>
<td>978159.557</td>
</tr>
<tr>
<td>12</td>
<td>Dep 15yr</td>
<td>992</td>
<td>-490532.03</td>
<td>983048.072</td>
</tr>
<tr>
<td>14</td>
<td>Dep 20yr</td>
<td>744</td>
<td>-491854.43</td>
<td>985196.869</td>
</tr>
<tr>
<td>1</td>
<td>Indep 1yr</td>
<td>3360</td>
<td>-491684.85</td>
<td>990089.712</td>
</tr>
<tr>
<td>3</td>
<td>Indep 2yr</td>
<td>1680</td>
<td>-493694.84</td>
<td>990749.699</td>
</tr>
<tr>
<td>5</td>
<td>Indep 3yr</td>
<td>1120</td>
<td>-495180.65</td>
<td>992601.306</td>
</tr>
<tr>
<td>7</td>
<td>Indep 4yr</td>
<td>840</td>
<td>-496195.71</td>
<td>994071.436</td>
</tr>
<tr>
<td>16</td>
<td>Dep 25yr</td>
<td>620</td>
<td>-496804.44</td>
<td>994848.883</td>
</tr>
<tr>
<td>9</td>
<td>Indep 5yr</td>
<td>672</td>
<td>-497728.45</td>
<td>996800.919</td>
</tr>
<tr>
<td>18</td>
<td>Dep 30yr</td>
<td>496</td>
<td>-498682.93</td>
<td>998357.866</td>
</tr>
<tr>
<td>11</td>
<td>Indep 10yr</td>
<td>336</td>
<td>-500741.72</td>
<td>1001255.45</td>
</tr>
<tr>
<td>13</td>
<td>Indep 15yr</td>
<td>224</td>
<td>-503764.69</td>
<td>1007977.38</td>
</tr>
<tr>
<td>20</td>
<td>Dep 40yr</td>
<td>372</td>
<td>-504682.61</td>
<td>1010109.23</td>
</tr>
<tr>
<td>15</td>
<td>Indep 20yr</td>
<td>168</td>
<td>-504943.42</td>
<td>1010206.84</td>
</tr>
<tr>
<td>17</td>
<td>Indep 25yr</td>
<td>140</td>
<td>-506243.00</td>
<td>1018766.01</td>
</tr>
<tr>
<td>19</td>
<td>Indep 30yr</td>
<td>112</td>
<td>-510927.84</td>
<td>1022079.78</td>
</tr>
<tr>
<td>21</td>
<td>Indep 40yr</td>
<td>84</td>
<td>-517644.45</td>
<td>1034546.91</td>
</tr>
<tr>
<td>23</td>
<td>Indep 50yr</td>
<td>70</td>
<td>-522660.80</td>
<td>1045461.60</td>
</tr>
<tr>
<td>24</td>
<td>Dep 100yr</td>
<td>186</td>
<td>-528491.45</td>
<td>1057361.03</td>
</tr>
<tr>
<td>25</td>
<td>Indep 100yr</td>
<td>42</td>
<td>-540626.11</td>
<td>1081336.23</td>
</tr>
<tr>
<td>26</td>
<td>Dep 240yr</td>
<td>62</td>
<td>-573432.44</td>
<td>1146988.89</td>
</tr>
<tr>
<td>27</td>
<td>Indep 240yr</td>
<td>14</td>
<td>-587599.35</td>
<td>1175226.71</td>
</tr>
</tbody>
</table>
subset of all the $\Delta t$ values considered in our model selection procedure are reported in Table 1.

Fig. 5: Plot of offence probability estimates per offence for the best model (Dep 3yr)

Fig. 6: Plot of punishment probability estimates per offence for the best model (Dep 3yr)

Referring to Figure 5 and Figure 6 it can be seen that trends in the model follow those of the empirical data. Theft continues to dominate offences, dropping sharply at the 1855 jurisdiction change. From this date a consistent increase in many other offences (notably Miscellaneous, Deception, Sexual and BreakingPeace) can be seen. Punishments show a continual downward trend in Death and NotGuilty and an upward trend in Imprison. The 1834 jurisdictional change created similarly significant effects in punishment estimates, with Death and Corporal all but disappearing and Imprison rising sharply. These suggest a move towards a more forgiving correctional system from the highly punitive system of the first several decades — another nonparametric signal of the civilizing process.

The offence-punishment pair estimates in Figure 7 show that Theft-NotGuilty, Theft-Transport, and Theft-Miscpunish dominated the period before 1834, with Theft-Imprison dominating the remainder of the period. This is also consistent with Elias theory of the civilizing process, and in line with related historical understanding.

We also applied the Bayesian information criterion (BIC) to rank the models. The penalty term is $K \log(n) > 12K$ under BIC, due to the sample size $n = 218,288$. Thus, it is over six times more penalizing than the AIC criterion with a penalty of $2K$. We found that the dependent model with $\Delta t = 11$ had the best BIC score. Thus, the penalty term from an information-theoretic point of view or the prior choice from a Bayesian point of view can affect the ranking of the models. A natural hold-out procedure from a longitudinally aware partitioning of the data into training, validation and test set can be done to address this smoothing problem. We can introduce a smoothing parameter $\lambda$ into a parametric penalty function $\lambda K$ to find an optimal $\lambda \in (1, 15)$ that maximizes say the likelihood of the held-out validation data, for instance. However, such a smoothing problem based on hold-out or cross-validation methods is beyond the current scope of this project.

Nonetheless, the fundamental processes driving events at the Old Bailey change significantly every three to eleven years if we conservatively allow the penalty criterion in model selection to vary over a wide range that contains the AIC, AICc and BIC criteria. We focus on the AIC criterion for the rest of this paper after noting that the estimates from the best model under BIC is smoother but in qualitative agreement with that based on the AIC.

Referring to Figure 7 it can be seen that Theft-NotGuilty, Theft-Transport, and Theft-Miscpunish dominated the period before 1834, with Theft-Imprison dominating the remainder of the period. This is also consistent with Elias theory of the civilizing process, and in line with related historical understanding.

It is accepted that the easily interpretable blanket partition of time into multi-year blocks of uniform size may appear too unrealistic. To infer greater causality than the uniform partition of time, further models where time-blocks are dynamically defined based on peaks in results from taking the absolute difference between $\hat{\theta}_{i,j,t}$ and $\hat{\theta}_{i,j,t+1}$ for each $t$ were explored. The peaks show where two subsequent years differ significantly and as such these points are candidates for time-block breakpoints.
Partitioned offence model. Exploratory data analysis, specifically the $\delta(o, o', t)$ plots in Figure 8, motivates the search for evidence of the civilizing process through partitioning of offences. The finest offence partition was used implicitly in the model selection done in Table 2. The best model (Dep 3yr) based on the AIC criterion with the finest offence partition, where each offence was allowed to have its own possibly distinct probability distribution over punishments, varied every three years. Here, we build on this best model by allowing the punishment distributions to possibly be identical for one or more offence(s), as specified by the offences within a block of a given offence-partition. An AIC based model selection process is first used to select the best model from the class of time-block varying partitioned offence models whose offence partitions are allowed to vary arbitrarily among all possible offence-partitions in each of the $[240/3]=80$ time-blocks. This model selection process using AIC did not produce a single best time varying partitioned offence model that is significantly better than the next best model.

We thus drew independent and identically distributed samples from the posterior distribution over time-block varying partitioned offence models by producing sample paths from an MCMC algorithm with the transition diagram depicted in Figure 2. The partitions are ordered from best to worst per time-block such that the partition number representing the best partition is 0 and 21,146 for the worst. This partition number mapping is convenient for visualization of the MCMC sample paths where the partition number is presented on the $y$ axis of Figure 9. The partitioned offence models over the 80 time-blocks visited by 1000 sample paths of the MCMC algorithm are shown in Figure 9 together with the mean number of blocks within the partitions and the number of offences grouped in the block most likely to receive Death as a punishment. Evidence of the civilizing process is immediately evident in the increasing average number of blocks and the decreasing number of offences in the block most likely to receive Death as a punishment. The posterior mode corresponds to the model made up of the partitioned offence models for each time-block with the lowest AIC score. This is tabulated for closer inspection of the estimates from the best time-block varying partitioned offence partitions from the perspectives of being punished by Death, Imprisonment or Transport in Tables 2 to 4, respectively. The second and third best models from the MCMC sampler are depicted in Tables 5 and 6 to show the agreement between the posterior samples around the best model. The general increase through time in the average number of blocks evident in Figure 9 relates to the punishments associated with offences becoming more differentiated. Similarly, the decrease in the average number of offences grouped in the block most likely to receive Death as a punishment is probably related to the judiciary using more discrimination with this most severe of punishments.

Table 2 shows that from 1839 onward the block most likely to receive Death contains Kill as the offence exclusively in all but a few cases. Figure 10 shows the proportion of the thousand transverses in the top ten and remaining partitions. The first three partitions account for the majority visited during the MCMC, except short time periods with large spreads in partitions such as between 1752 and 1758. Referring to Tables 5 and 6 containing the second and third best unique models, it is clear that they are very similar to the best partitioning especially in the 1839 demarcation of Kill being grouped alone and leading the probability of death. Figure 11 shows the probability of death for offences grouped by the best partition. It indicates a disturbing trend, where all offences except theft have a higher probability of death than killing. The probability of death for killing is consistent when compared to the dra-

Exploratory data analysis. The best model found by AIC is thoroughly investigated graphically. Many different perspectives on $\hat{\theta}_{i,o,t}$ are generated: from the general plots of estimated probabilities of offences, punishments, and offence-punishment pairs as presented in Figures 5 to 7, to per offence punishment distributions and animations (available from the authors) of the offence-punishment pairs through time. The most information in a single graphic is found in plots of $\delta(o,o',t)$, showing the absolute difference between estimated probabilities of punishments for pairs of offence slices. Figure 8 shows such a plot for the best model (Dep 3yr) where the colouring is chosen to highlight the strong grouping of punishment distributions for offences, especially before 1855. The grouping by color codes for a given offence (or group of offences) being fixed in each one of a set of offence-pairs defining the corresponding set of $\delta(o,o',t)$ plots. Some groups of offences have very distinct punishment distributions while others are similar. For example, Theft and Sexual offences have the most different punishment profiles as reflected by the red colored $\delta(o,o',t)$ plot in Figure 8 taking large values. The sharp transition in 1855 is an important aspect of this visualization, where evidence of some change in behaviour is seen for all offences. It is clear from the plot that some offences are more strongly grouped than others, and that how strongly they are grouped changes in time, especially after 1855. This supports a hypothesis that evidence of the civilizing process may be found by allowing the offences to be grouped together, to assess how the grouping of offences changes in time and whether punishment distributions become more offence-specific as time progresses. This is investigated in the next Section.

![Fig. 8: Plot of $\delta(o,o',t)$ between punishment slices for all possible pairs of offences from dependent estimates with $\Delta t = 3$. The orange lines are Theft compared to all other offences, the green lines are Sexual offences compared with all other offences, the red line is Theft compared with Sexual, the violet lines for the offences of BreakingPeace and Deception with all remaining offences, and the blue lines contain the $\delta$ values for the remainder of the offence pairs.](image-url)
matic order of change in the probability of death for all other offences circa 1839. The transition in 1839 from high probabilities of death for a wide range of offences to a relatively low probability of death only for killing is a crucial moment in the civilizing of the Old Bailey, where violence by the state begins to be applied with significantly more discrimination.

These plots, together with the best time-block varying partitioned offence models by AIC score, and the relative frequency of ‘Death’ and ‘Corporal’ punishments in Figure 6, clearly show that not only is the violence meted out by the state reducing, but that this coincides with an increase in discrimination in punishments for offences. The civilizing process is thus multi-faceted. The data presented here offers a picture of court culture striving towards civility.

Conclusion
This study aspires to function as an example for longitudinal studies on large humanities data sets. Using The Old Bailey Online (OBO) as the reference data set, it introduces non-parametric modelling techniques supported by publicly available Python code, to expose long-timescale trends and to seek evidence of the civilizing process. The OBO, the digitized records of trial events from London’s longest running and most notorious court, is an ideal resource for the study of long timescale trends. It covers an extremely long period, from 1674 to 1913, which witnessed vast social and political change.

Offence and punishment frequencies are downloaded from the OBO in one year blocks covering the full period of available data. Using non-parametric modelling techniques on this course-grained data, models on offence-punishment pairs are generated for various time period resamplings and to test whether punishment depends on offence. By the Akaike Information Criterion (AIC) the chosen model is one with a three year time period in which punishment depends on offence. This is a sensible model selection for a court of law, indicating that the fundamental process driving the Old Bailey changes every three years. This model of the Old Bailey, together with the exploratory visualizations, exposes the data to a wide audience of historians and digital humanists for further in-depth inquiry. It also provides a straightforward but robust approach to statistical modelling that can provide a useful methodological baseline for further efforts with this and other data sets. By exhaustively partitioning offence categories, it is found that the per offence punishment distributions become more distinct with time and coincide with jurisdictional changes. The punishment distributions for violent offences become strongly grouped and have a much higher probability of receiving ‘Death’ as a punishment. These factors, together with the general reduction in relative frequency of violent punishments such as ‘Corporal’ and ‘Death’, provide strong evidence of the civilizing process in the court of the Old Bailey, supporting the previous study conducted by [36].

This study, along with its methods and presented graphics, will contribute to the foundations being established for future researchers eager to ask big questions of big humanities data.
ACKNOWLEDGMENTS. R.S. was partly supported by a visiting scholarship at Department of Mathematics, Cornell University, Ithaca, NY, USA, a sabbatical grant from College of Engineering, University of Canterbury, and consulting revenues from Wynyard Group, Christchurch, NZ. We thank Elena Moltchonova and Amandine Veber for their valuable comments on an earlier version of this manuscript, and Tim Hitchcock and the team at Old Bailey Online for maintaining and providing the dataset.
<table>
<thead>
<tr>
<th>Date</th>
<th>Partition with blocks sorted by probability of death</th>
</tr>
</thead>
<tbody>
<tr>
<td>1758</td>
<td>[27.27, RoyalOffences, Sexual], [13.13, Deception, Miscellaneous, ViolentTheft], [3.67, Theft]</td>
</tr>
<tr>
<td>1761</td>
<td>[31.88, BreakingPeace, Damage, Deception, Miscellaneous, RoyalOffences], [20.97, Kill, RoyalOffences, Sexual], [4.11, Theft]</td>
</tr>
<tr>
<td>1767</td>
<td>[37.50, Damage, Miscellaneous, RoyalOffences, ViolentTheft], [19.70, BreakingPeace, Deception, Kill, Sexual], [5.55, Theft]</td>
</tr>
<tr>
<td>1770</td>
<td>[48.28, BreakingPeace, Damage, Miscellaneous], [41.21, Damage, ViolentTheft, Deception, RoyalOffences], [18.46, Deception, RoyalOffences, Sexual], [15.15, Kill, Sexual], [5.64, Theft]</td>
</tr>
<tr>
<td>1776</td>
<td>[37.44, BreakingPeace, RoyalOffences, Sexual, ViolentTheft], [17.11, Deception, Miscellaneous], [14.29, Damage, Kill], [14.20, Damage, Miscellaneous], [7.76, Theft]</td>
</tr>
<tr>
<td>1782</td>
<td>[48.84, Miscellaneous, RoyalOffences, ViolentTheft], [41.21, Damage, Deception, RoyalOffences], [14.29, Kill, RoyalOffences, Sexual], [7.76, Theft]</td>
</tr>
<tr>
<td>1791</td>
<td>[48.20, ViolentTheft], [30.65, Damage, Deception, Miscellaneous], [14.29, Kill, Sexual], [5.48, BreakingPeace, Theft, ViolentTheft], [0.00, RoyalOffences]</td>
</tr>
<tr>
<td>1794</td>
<td>[43.36, BreakingPeace, Damage, Deception, Miscellaneous, ViolentTheft], [16.33, Kill, Sexual], [4.88, Theft], [2.59, RoyalOffences]</td>
</tr>
<tr>
<td>1803</td>
<td>[48.84, Miscellaneous, RoyalOffences, ViolentTheft], [24.64, Damage, ViolentTheft], [18.80, Deception, Miscellaneous, RoyalOffences], [14.74, Kill, RoyalOffences, Sexual], [4.53, Theft]</td>
</tr>
<tr>
<td>1809</td>
<td>[29.03, Deception, ViolentTheft], [26.92, BreakingPeace, Damage, Miscellaneous, RoyalOffences], [18.18, RoyalOffences, Sexual], [12.04, Kill, Miscellaneous], [5.99, Theft]</td>
</tr>
<tr>
<td>1812</td>
<td>[29.19, Damage, RoyalOffences, Sexual, Deception, Miscellaneous], [20.27, BreakingPeace, Deception, RoyalOffences], [15.22, Kill, RoyalOffences, Sexual], [7.52, Theft]</td>
</tr>
<tr>
<td>1815</td>
<td>[36.56, ViolentTheft], [18.06, Damage, Miscellaneous], [12.07, BreakingPeace, Deception, RoyalOffences], [9.26, Kill, RoyalOffences, Sexual], [6.85, Theft]</td>
</tr>
<tr>
<td>1818</td>
<td>[43.35, ViolentTheft], [28.38, BreakingPeace, Damage, Deception, Miscellaneous, ViolentTheft], [18.18, RoyalOffences, Sexual], [7.23, Theft]</td>
</tr>
<tr>
<td>1821</td>
<td>[35.51, Damage, ViolentTheft], [33.33, RoyalOffences], [17.14, BreakingPeace, Deception, Kill, Sexual], [13.33, Miscellaneous], [5.27, Theft]</td>
</tr>
<tr>
<td>1824</td>
<td>[33.58, ViolentTheft], [20.83, BreakingPeace, Miscellaneous], [14.93, BreakingPeace, Deception, RoyalOffences], [14.29, Kill, Sexual], [3.73, Theft]</td>
</tr>
<tr>
<td>1842</td>
<td>[32.15, Deception, RoyalOffences, Sexual, ViolentTheft], [16.85, BreakingPeace, Damage, Deception, RoyalOffences], [5.57, Theft]</td>
</tr>
<tr>
<td>1845</td>
<td>[10.08, Damage, Kill], [1.04, BreakingPeace, Sexual], [0.00, Deception, Miscellaneous, RoyalOffences], [0.00, Theft]</td>
</tr>
<tr>
<td>1848</td>
<td>[48.84, Miscellaneous, RoyalOffences, ViolentTheft], [41.21, Damage, Deception, RoyalOffences], [14.29, Kill, RoyalOffences, Sexual], [7.76, Theft]</td>
</tr>
<tr>
<td>1851</td>
<td>[36.56, ViolentTheft], [18.06, Damage, Miscellaneous], [12.07, BreakingPeace, Deception, RoyalOffences], [9.26, Kill, RoyalOffences, Sexual], [6.85, Theft]</td>
</tr>
<tr>
<td>1854</td>
<td>[48.20, ViolentTheft], [30.65, Damage, Deception, Miscellaneous], [14.29, Kill, Sexual], [5.48, BreakingPeace, Theft, ViolentTheft], [0.00, RoyalOffences]</td>
</tr>
<tr>
<td>1857</td>
<td>[48.84, Miscellaneous, RoyalOffences, ViolentTheft], [24.64, Damage, ViolentTheft], [18.80, Deception, Miscellaneous, RoyalOffences], [14.74, Kill, RoyalOffences, Sexual], [4.53, Theft]</td>
</tr>
<tr>
<td>1860</td>
<td>[10.83, Kill, Sexual], [5.56, Damage, Miscellaneous], [0.00, Deception, ViolentTheft, RoyalOffences], [0.00, RoyalOffences], [0.00, Theft]</td>
</tr>
<tr>
<td>1866</td>
<td>[5.93, Kill, Sexual], [0.00, BreakingPeace, Damage, [0.00, Deception, Miscellaneous, RoyalOffences], [0.00, RoyalOffences], [0.00, Theft]</td>
</tr>
<tr>
<td>1870</td>
<td>[5.93, Kill, Sexual], [0.00, BreakingPeace, Damage, [0.00, Deception, Miscellaneous, RoyalOffences], [0.00, RoyalOffences], [0.00, Theft]</td>
</tr>
<tr>
<td>1875</td>
<td>[48.84, Miscellaneous, RoyalOffences, ViolentTheft], [41.21, Damage, Deception, RoyalOffences], [14.29, Kill, RoyalOffences, Sexual], [7.76, Theft]</td>
</tr>
<tr>
<td>1880</td>
<td>[48.84, Miscellaneous, RoyalOffences, ViolentTheft], [41.21, Damage, Deception, RoyalOffences], [14.29, Kill, RoyalOffences, Sexual], [7.76, Theft]</td>
</tr>
<tr>
<td>1884</td>
<td>[48.84, Miscellaneous, RoyalOffences, ViolentTheft], [41.21, Damage, Deception, RoyalOffences], [14.29, Kill, RoyalOffences, Sexual], [7.76, Theft]</td>
</tr>
<tr>
<td>1889</td>
<td>[48.84, Miscellaneous, RoyalOffences, ViolentTheft], [41.21, Damage, Deception, RoyalOffences], [14.29, Kill, RoyalOffences, Sexual], [7.76, Theft]</td>
</tr>
<tr>
<td>1894</td>
<td>[48.84, Miscellaneous, RoyalOffences, ViolentTheft], [41.21, Damage, Deception, RoyalOffences], [14.29, Kill, RoyalOffences, Sexual], [7.76, Theft]</td>
</tr>
<tr>
<td>1900</td>
<td>[48.84, Miscellaneous, RoyalOffences, ViolentTheft], [41.21, Damage, Deception, RoyalOffences], [14.29, Kill, RoyalOffences, Sexual], [7.76, Theft]</td>
</tr>
<tr>
<td>1906</td>
<td>[48.84, Miscellaneous, RoyalOffences, ViolentTheft], [41.21, Damage, Deception, RoyalOffences], [14.29, Kill, RoyalOffences, Sexual], [7.76, Theft]</td>
</tr>
<tr>
<td>1911</td>
<td>[48.84, Miscellaneous, RoyalOffences, ViolentTheft], [41.21, Damage, Deception, RoyalOffences], [14.29, Kill, RoyalOffences, Sexual], [7.76, Theft]</td>
</tr>
<tr>
<td>Date</td>
<td>Partition with blocks sorted by probability of imprisonment</td>
</tr>
<tr>
<td>----------</td>
<td>---------------------------------------------------------------------------------------------------------------------------</td>
</tr>
<tr>
<td>1878</td>
<td>{ [15.43, Deception, RoyalOffences, Sexual, [4.55, Damage, Miscellaneous, ViolentTheft]], [0.82, Theft], [0.20, Kill] }</td>
</tr>
<tr>
<td>1875</td>
<td>{ [69.52, RoyalOffences, Theft], [64.34, BreakingPeace, Deception, Sexual], [61.21, Damage, ViolentTheft], [40.71, Kill], [36.82, Miscellaneous] }</td>
</tr>
<tr>
<td>1869</td>
<td>{ [79.98, RoyalOffences], [71.57, Deception, Theft], [57.21, Damage, Sexual], [56.86, ViolentTheft], [56.03, BreakingPeace], [34.36, Kill], [24.84, Miscellaneous] }</td>
</tr>
<tr>
<td>1854</td>
<td>{ [87.39, RoyalOffences], [73.80, Theft], [61.64, BreakingPeace, Sexual], [58.16, Deception, ViolentTheft], [29.56, Miscellaneous], [27.89, Damage, Kill] }</td>
</tr>
<tr>
<td>1845</td>
<td>{ [73.92, RoyalOffences], [59.45, Theft], [53.07, BreakingPeace, Sexual], [46.35, Deception], [32.47, Miscellaneous], [28.22, ViolentTheft], [21.04, Damage, Kill] }</td>
</tr>
<tr>
<td>1836</td>
<td>{ [72.75, RoyalOffences], [60.10, Sexual], [47.23, BreakingPeace], [45.28, Theft], [38.94, Deception], [32.37, Miscellaneous], [24.87, Kill], [9.39, Damage, ViolentTheft] }</td>
</tr>
<tr>
<td>1830</td>
<td>{ [28.44, Theft], [25.45, Miscellaneous, RoyalOffences], [23.28, Kill, Sexual], [14.56, Damage, Deception], [2.40, BreakingPeace, ViolentTheft], [0.95, Theft] }</td>
</tr>
<tr>
<td>1812</td>
<td>{ [22.75, Theft], [17.25, Damage, Miscellaneous], [15.53, Kill], [15.05, RoyalOffences, Sexual], [3.49, BreakingPeace, Deception, ViolentTheft], [0.93, Theft] }</td>
</tr>
<tr>
<td>1809</td>
<td>{ [26.98, Kill, Miscellaneous], [24.94, Theft], [15.58, RoyalOffences, Sexual], [4.40, BreakingPeace, Damage], [3.32, Deception, ViolentTheft], [1.50, RoyalOffences, BreakingPeace], [0.21, Miscellaneous] }</td>
</tr>
<tr>
<td>1806</td>
<td>{ [23.99, Kill, Miscellaneous, RoyalOffences, Sexual], [23.24, Theft], [7.63, BreakingPeace, Damage, Deception, ViolentTheft], [0.91, Theft] }</td>
</tr>
<tr>
<td>1803</td>
<td>{ [30.06, RoyalOffences], [24.22, Theft], [16.76, BreakingPeace, Deception, Kill], [11.96, Miscellaneous], [1.66, Damage, ViolentTheft], [0.95, Theft] }</td>
</tr>
<tr>
<td>1794</td>
<td>{ [33.74, RoyalOffences], [20.70, Kill, Sexual], [18.60, Theft], [5.20, BreakingPeace, Damage, Deception, ViolentTheft], [0.05, RoyalOffences, Sexual], [0.02, Theft] }</td>
</tr>
<tr>
<td>1782</td>
<td>{ [34.92, RoyalOffences], [19.89, Theft], [6.41, BreakingPeace, Kill, Sexual], [6.27, Miscellaneous], [3.90, Damage, Deception, ViolentTheft], [0.91, Theft] }</td>
</tr>
<tr>
<td>1776</td>
<td>{ [26.50, Deception, Miscellaneous], [26.23, Theft], [14.58, Kill], [5.69, BreakingPeace, RoyalOffences, Sexual], [0.97, ViolentTheft], [0.95, Theft] }</td>
</tr>
<tr>
<td>1773</td>
<td>{ [14.46, Deception, RoyalOffences, Sexual], [13.57, RoyalOffences], [1.02, BreakingPeace, Damage, Miscellaneous, ViolentTheft], [0.95, Theft] }</td>
</tr>
<tr>
<td>1761</td>
<td>{ [13.13, Kill, Sexual, RoyalOffences, Sexual], [1.55, BreakingPeace, Damage, Deception, Miscellaneous, ViolentTheft], [0.34, Theft], [0.24, Death], [0.14, Damage, RoyalOffences, Sexual], [0.08, Theft] }</td>
</tr>
<tr>
<td>1752</td>
<td>{ [11.90, RoyalOffences, Damage, Kill, RoyalOffences, Sexual], [1.54, Deception, Miscellaneous, ViolentTheft], [0.24, Theft], [0.20, RoyalOffences, Sexual], [0.14, Damage, Miscellaneous, RoyalOffences], [0.09, Theft] }</td>
</tr>
<tr>
<td>1748</td>
<td>{ [10.71, BreakingPeace, Deception, Kill], [6.14, Damage, Miscellaneous, RoyalOffences, Sexual], [0.98, Theft], [0.95, Theft] }</td>
</tr>
<tr>
<td>1745</td>
<td>{ [7.91, Deception, RoyalOffences], [7.39, BreakingPeace, Miscellaneous], [6.28, Kill, Sexual], [0.42, Theft], [0.09, Damage, ViolentTheft], [0.08, Theft] }</td>
</tr>
<tr>
<td>1730</td>
<td>{ [16.69, BreakingPeace, Deception, Kill], [12.27, Theft], [10.48, Deception, Miscellaneous, Sexual], [9.02, Damage, ViolentTheft], [0.91, Theft] }</td>
</tr>
<tr>
<td>1728</td>
<td>{ [15.64, BreakingPeace, Deception, Kill], [12.57, Theft], [9.61, BreakingPeace, Miscellaneous, Sexual], [7.41, RoyalOffences, Sexual], [6.27, RoyalOffences, Deception], [4.54, RoyalOffences, Sexual], [3.29, RoyalOffences, Sexual], [1.99, Miscellaneous, RoyalOffences], [0.98, Theft] }</td>
</tr>
<tr>
<td>1725</td>
<td>{ [15.13, RoyalOffences, Sexual], [13.37, RoyalOffences, Sexual], [11.53, RoyalOffences, Sexual], [9.47, RoyalOffences, Sexual], [7.58, RoyalOffences, Sexual], [5.89, RoyalOffences, Sexual], [3.90, RoyalOffences, Sexual], [1.99, Miscellaneous, RoyalOffences], [0.98, Theft] }</td>
</tr>
<tr>
<td>1722</td>
<td>{ [14.86, RoyalOffences, Sexual], [13.57, Deception], [1.02, BreakingPeace, Damage, Miscellaneous, ViolentTheft], [0.95, Theft] }</td>
</tr>
<tr>
<td>1719</td>
<td>{ [14.96, RoyalOffences, Sexual], [13.57, Deception], [1.02, BreakingPeace, Damage, Miscellaneous, ViolentTheft], [0.95, Theft] }</td>
</tr>
<tr>
<td>1710</td>
<td>{ [12.57, RoyalOffences, Sexual], [12.57, Deception], [1.02, BreakingPeace, Damage, Miscellaneous, ViolentTheft], [0.95, Theft] }</td>
</tr>
<tr>
<td>1708</td>
<td>{ [11.50, RoyalOffences, Damag, Kill, RoyalOffences, Sexual], [1.54, Deception, Miscellaneous, ViolentTheft], [0.24, Theft], [0.20, RoyalOffences, Sexual], [0.14, Damage, Miscellaneous, RoyalOffences], [0.09, Theft] }</td>
</tr>
<tr>
<td>1707</td>
<td>{ [10.71, BreakingPeace, Deception, Kill], [6.14, Damage, Miscellaneous, RoyalOffences, Sexual], [0.98, Theft], [0.95, Theft] }</td>
</tr>
<tr>
<td>1704</td>
<td>{ [9.56, Deception, RoyalOffences, Sexual], [9.56, Deception, RoyalOffences, Sexual], [6.26, Kill, Sexual], [0.42, Theft], [0.09, Damage, ViolentTheft], [0.08, Theft] }</td>
</tr>
<tr>
<td>1701</td>
<td>{ [8.31, Deception, RoyalOffences, Sexual], [8.31, Deception, RoyalOffences, Sexual], [6.26, Kill, Sexual], [0.42, Theft], [0.09, Damage, ViolentTheft], [0.08, Theft] }</td>
</tr>
<tr>
<td>Date</td>
<td>Partition with blocks sorted by probability of transportation</td>
</tr>
<tr>
<td>---------</td>
<td>--------------------------------------------------------------</td>
</tr>
<tr>
<td>1800</td>
<td>[20.70, Theft, Deception, Miscellaneous, RoyalOffences, Sexual, Damaged, Kill]</td>
</tr>
<tr>
<td>1806</td>
<td>[18.90, Theft, BreakingPeace, Damage, RoyalOffences, Sexual, Kill, RoyalOffences]</td>
</tr>
<tr>
<td>1809</td>
<td>[17.88, Theft, Deception, ViolentTheft, RoyalOffences, Sexual, Kill, Miscellaneous]</td>
</tr>
<tr>
<td>1816</td>
<td>[16.65, Deception, Miscellaneous, ViolentTheft, [0.24, BreakingPeace, Damage, Kill, RoyalOffences, Sexual]]</td>
</tr>
<tr>
<td>1827</td>
<td>[34.18, Theft, Deception, Miscellaneous, RoyalOffences, Sexual, Kill, Sexual]</td>
</tr>
<tr>
<td>1830</td>
<td>[38.13, Theft, Deception, Miscellaneous, RoyalOffences, Sexual, Kill, Sexual]</td>
</tr>
<tr>
<td>1836</td>
<td>[31.46, Theft, Deception, Miscellaneous, RoyalOffences, Sexual, Kill, Sexual]</td>
</tr>
<tr>
<td>1842</td>
<td>[37.11, ViolentTheft, Deception, Miscellaneous, RoyalOffences, Sexual, Kill, Sexual]</td>
</tr>
<tr>
<td>1845</td>
<td>[45.57, ViolentTheft, Deception, Miscellaneous, RoyalOffences, Sexual, Kill, Sexual]</td>
</tr>
<tr>
<td>1851</td>
<td>[37.28, ViolentTheft, Deception, Miscellaneous, RoyalOffences, Sexual, Kill, Sexual]</td>
</tr>
<tr>
<td>1857</td>
<td>[31.46, Theft, Deception, Miscellaneous, RoyalOffences, Sexual, Kill, Sexual]</td>
</tr>
<tr>
<td>1869</td>
<td>[21.10, Deception, Miscellaneous, RoyalOffences, Sexual, Kill, Sexual]</td>
</tr>
<tr>
<td>1872</td>
<td>[20.44, Deception, Miscellaneous, RoyalOffences, Sexual, Kill, Sexual]</td>
</tr>
<tr>
<td>1881</td>
<td>[10.99, Deception, Miscellaneous, RoyalOffences, Sexual, Kill, Sexual]</td>
</tr>
<tr>
<td>1890</td>
<td>[14.00, Deception, Miscellaneous, RoyalOffences, Sexual, Kill, Sexual]</td>
</tr>
<tr>
<td>1893</td>
<td>[17.10, Deception, Miscellaneous, RoyalOffences, Sexual, Kill, Sexual]</td>
</tr>
<tr>
<td>1905</td>
<td>[19.02, Deception, Miscellaneous, RoyalOffences, Sexual, Kill, Sexual]</td>
</tr>
<tr>
<td>1911</td>
<td>[21.10, Deception, Miscellaneous, RoyalOffences, Sexual, Kill, Sexual]</td>
</tr>
</tbody>
</table>
Table 5: The partitions for the second best offence-partitioned model that can vary every three years, with each block preceded by the probability of death within that block as a percentage.

<table>
<thead>
<tr>
<th>Year</th>
<th>Partition with blocks sorted by probability of death</th>
</tr>
</thead>
<tbody>
<tr>
<td>1674</td>
<td>38.52, Damage, Kill, Miscellaneous, Sexual, Theft, ViolentTheft, [14.29, BreakingPeace, Deception, RoyalOffences]</td>
</tr>
<tr>
<td>1675</td>
<td>45.58, RoyalOffences, Violence, [28.86, Deception, Kill, Miscellaneous, Sexual, [20.37, BreakingPeace, Damage, Theft]]</td>
</tr>
<tr>
<td>1676</td>
<td>43.77, Damage, Deception, ViolentTheft, Kill, Miscellaneous, Sexual, Theft, [20.37, BreakingPeace, Deception, RoyalOffences]</td>
</tr>
<tr>
<td>1677</td>
<td>65.93, ViolentTheft, [22.08, Kill, Miscellaneous, Sexual, [18.88, Deception, RoyalOffences, [11.82, Theft, [10.48, BreakingPeace, Deception, RoyalOffences]]]</td>
</tr>
<tr>
<td>1678</td>
<td>46.39, Sexual, ViolentTheft, [23.20, Kill, Miscellaneous, [10.87, BreakingPeace, Theft, [5.75, Damage, Deception, Miscellaneous, Theft]]]</td>
</tr>
<tr>
<td>1679</td>
<td>38.03, BreakingPeace, RoyalOffences, Violence, [17.93, Kill, Sexual, [10.65, Damage, Deception, Miscellaneous, Theft]]</td>
</tr>
<tr>
<td>1681</td>
<td>45.39, Theft, [6.72, BreakingPeace, Damage, Deception, Kill, Miscellaneous, RoyalOffences, Sexual, Theft, [6.60, Deception, Theft]]</td>
</tr>
<tr>
<td>1682</td>
<td>18.67, Damage, Deception, RoyalOffences, [6.51, BreakingPeace, Miscellaneous, Theft, [6.51, Damange, Miscellaneous, Theft]]</td>
</tr>
<tr>
<td>1684</td>
<td>45.58, RoyalOffences, Violence, [28.86, Deception, Kill, Miscellaneous, Sexual, [20.37, BreakingPeace, Theft, [15.49, Deception, Miscellaneous, Theft]]]</td>
</tr>
<tr>
<td>1685</td>
<td>46.39, Sexual, ViolentTheft, [23.20, Kill, Miscellaneous, [10.87, BreakingPeace, Theft, [5.75, Damage, Deception, Miscellaneous, Theft]]]</td>
</tr>
<tr>
<td>1686</td>
<td>38.03, BreakingPeace, RoyalOffences, Violence, [17.93, Kill, Sexual, [10.65, Damage, Deception, Miscellaneous, Theft]]</td>
</tr>
<tr>
<td>1688</td>
<td>45.39, Theft, [6.72, BreakingPeace, Damage, Deception, Kill, Miscellaneous, RoyalOffences, Sexual, Theft, [6.60, Deception, Theft]]</td>
</tr>
<tr>
<td>1689</td>
<td>18.67, Damage, Deception, RoyalOffences, [6.51, BreakingPeace, Miscellaneous, Theft, [6.51, Damange, Miscellaneous, Theft]]</td>
</tr>
<tr>
<td>1691</td>
<td>45.58, RoyalOffences, Violence, [28.86, Deception, Kill, Miscellaneous, Sexual, [20.37, BreakingPeace, Theft, [15.49, Deception, Miscellaneous, Theft]]]</td>
</tr>
<tr>
<td>1692</td>
<td>46.39, Sexual, ViolentTheft, [23.20, Kill, Miscellaneous, [10.87, BreakingPeace, Theft, [5.75, Damage, Deception, Miscellaneous, Theft]]]</td>
</tr>
<tr>
<td>1693</td>
<td>38.03, BreakingPeace, RoyalOffences, Violence, [17.93, Kill, Sexual, [10.65, Damage, Deception, Miscellaneous, Theft]]</td>
</tr>
<tr>
<td>1695</td>
<td>45.39, Theft, [6.72, BreakingPeace, Damage, Deception, Kill, Miscellaneous, RoyalOffences, Sexual, Theft, [6.60, Deception, Theft]]</td>
</tr>
<tr>
<td>1696</td>
<td>18.67, Damage, Deception, RoyalOffences, [6.51, BreakingPeace, Miscellaneous, Theft, [6.51, Damange, Miscellaneous, Theft]]</td>
</tr>
<tr>
<td>1698</td>
<td>45.58, RoyalOffences, Violence, [28.86, Deception, Kill, Miscellaneous, Sexual, [20.37, BreakingPeace, Theft, [15.49, Deception, Miscellaneous, Theft]]]</td>
</tr>
<tr>
<td>1699</td>
<td>46.39, Sexual, ViolentTheft, [23.20, Kill, Miscellaneous, [10.87, BreakingPeace, Theft, [5.75, Damage, Deception, Miscellaneous, Theft]]]</td>
</tr>
<tr>
<td>1700</td>
<td>38.03, BreakingPeace, RoyalOffences, Violence, [17.93, Kill, Sexual, [10.65, Damage, Deception, Miscellaneous, Theft]]</td>
</tr>
<tr>
<td>1702</td>
<td>45.39, Theft, [6.72, BreakingPeace, Damage, Deception, Kill, Miscellaneous, RoyalOffences, Sexual, Theft, [6.60, Deception, Theft]]</td>
</tr>
</tbody>
</table>

As a percentage.
Table 6: The partitions for the third best offence-partitioned model that can vary every three years, with each block precede by the probability of death within that block as a percentage.
32. (1836) Prisoners’ counsel. A bill for enabling prisoners to make their defence by counsel or attorney. 19th Century House of Commons Sessional Papers IV:595.

16

Mackenzie, Sainudin, Smithies and Wolffram
Appendix: The Proceedings and the Old Bailey

The historic context of the Old Bailey and the Proceedings is critical to making sense of the data, and is in itself a fascinating story of the progress of society. This supplement presents a short historic background of the Proceedings and policing in London from the fifteenth to the twentieth century.

History of the Proceedings. The Proceedings of the Old Bailey, or simply the Proceedings, is the publication of transcripts from court hearings of the Old Bailey, produced between 1674 and 1913. It represents a unique data set. The Proceedings consistently record political, social and linguistic change spanning a vast temporal landscape, from the Enlightenment, the industrial revolutions, mass urbanisation, and the rise of the global British Empire, to the beginning of the First World War [11]. The content and quality of records within the Proceedings ranges from cases with only a few words in the 1670s to cases of particular public interest requiring volumes of their own—often with vast verbatim testimonies [9]. When using it as a historical source it is important to consider issues of bias and reliability induced by social and political change over time.

Audience. The Proceedings were a commercial venture published by private publishing houses for profit and licensed by the City. Until 1778 it was designed for a middle and upper class audience, the ‘property owning middle and upper classes’ [37], but was widely read and critiqued by Londoners of all classes. There is strong evidence it was carefully read by the criminal underbelly as well [37]. In this early period, and for most publishers, the Proceedings was another mainstream publication within their portfolio, much like our gossip magazines today. The social and political importance of the Proceedings is indicated by the fact that it was licensed annually by the City (exactly who was responsible changed over time). The names of at least 25 different printers appear on the title pages between 1720 and 1778. Londoners were by no means a passive audience, however. Readers were aware of the tendency toward selective reporting in the Proceedings, and many read them sceptically. Many hearings had a large audience, and there was public outcry surrounding numerous trials when what was published in the Proceedings was not representative of the trial, bringing its reliability into question [37]. After 1778 the target audience of the Proceedings changed from a general readership to lawyers, making the content significantly longer and drier [38, p. 468].

Reliability. The Proceedings therefore has questionable reliability as a truly representative sample of what crimes were tried and the accounts of those crimes. These issues were known to Londoners at the time and have been discussed in contemporary setting by the likes of Shoemaker [37]. Due to the “abstracted” nature of the shorthand even if the publishers could afford to publish every word from the scribes, the Proceedings would still be incomplete, because much of what was said in court was not recorded. Much of the material left out of the Proceedings was repetitive or insubstantial. As one scribe put it “It is not to be expected I should write every unintelligible word that is said by the evidence.” [39]. Much information of interest to modern researchers does not appear in the Proceedings because of this, including the judges summation at the conclusions of trials. At the conclusion of his term of office as Lord Mayor, John Wilkes complained of the inadequacy of the Proceedings, a sentiment that was by no means isolated. According to the Middlesex Journal, “the learned and humane judge who tried the Perreaus had complained to him of the inexplicit manner in which those trials had been taken down, and that many parts of them were contrary to his own notes.” [37]. There was also criticism about the use of censorship to decide which trials were included [9]. It was claimed that the under-representation of cases resulting in not-guilty charges made the Old Bailey appear more effective than it actually was. Whether rightly or wrongly, it was believed that the inclusion of not-guilty charges might aid the would-be criminal. In 1778 these and other issues led to a condition being imposed on the printer that the Proceedings should, “contain a true, fair, and perfect narrative of the whole condition being imposed on the printer that the Proceedings had been taken down, and that many parts of them were contrary to his own notes.” [37]. There was also criticism about the use of censorship to decide which trials were included [9].

Crime and Policing. During the period reported in the Proceedings, policing in London underwent radical change: from the 1670s when the onus of apprehending culprits was the duty of the public, to the early 20th century when an established modern police force with special investigative units was in place. In the context of the Proceedings these changes (amongst many others) need to be considered, with an expectation that the increase in organised policing will have a noticeable effect on trials brought before the courts of the Old Bailey.

As an example of early policing, from 1674 to 1829 the victims of many crimes had to identify and apprehend the accused before contacting a constable or a justice of the peace to secure their arrest. A constable could require that members of the public assist in the pursuit of suspected criminals. Every male homeowner was required to assist in the role of constable part-time for a year [9]. An increase in crime in the late 17th century led to bounties being offered for offenders of serious crimes, giving rise to those known as thief-takers, i.e. professional bounty hunters. This represented the beginnings of modern policing; in the 1730s thief-takers were hired on retainer and sent by the magistrates to apprehend suspected criminals. The thief-taker would receive an extra reward on apprehension of the culprit. This coincided with magistrates becoming available in an office during fixed hours, in the same manner as a modern police station. This was followed by the introduction of the “Watch Acts” where the rotation of appointed householders was replaced by full time watchmen paid by the state.

Many of these innovations came from the famous Bow Street office where they were concerned with investigating and disseminating patterns of crime. Armed with this knowledge they organised patrols in areas of high risk as a preventative measure. In this way the Bow Street office become the equivalent of an intelligence organisation. In 1792 seven new police stations were opened within London, with three magistrates and six constables each. In 1800, the Thames police station was opened with three magistrates and one hundred officers [9]. Continuing this trend towards professional policing, in 1829 the Metropolitan Police Act established a professional police force under the control of the home secretary, who had responsibility for policing the area surrounding but not including the city of London. This was an attempt to reduce crime in the area outside the jurisdiction of the Old Bailey. It would be natural to see the effect of this change in the Proceedings five years later, when the jurisdiction of the Old Bailey was increased to the whole of England [28][29][17]. In 1839 a second Metropolitan Police Act increased the policing jurisdiction by a third, and the number of officers to 4300[9].
At the same time another Act created a similar police organisation for the square mile of the City of London[34, 24-32]. In 1842 when authority was given to create a distinct detective force within the Metropolitan Police, the modernisation of the police force was complete. It remained this way until the Proceedings ceased publication on the eve of the First World War in 1913 [9].

Death and Transportation. For many years, most offences heard at the Old Bailey were punishable by death. The predominant method of execution was hanging. Convicted persons were hung at Tyburn from the 1670 until 1783, after which they were hung at Newgate prison [10, 9]. Despite many convicted felons receiving death as a sentence, however, few were actually executed because their sentences could be deferred through various mechanisms. Execution was normally only carried out on those convicted of murder. By the 1840s the death sentence was no longer applicable to most offences. This was formalised in the ‘Offences Against the Persons Act’ of 1861 [41], which reserved death only for murder and high treason[9]

Transportation as a punishment has an interesting history of its own. Introduced in the 1718 Transportation Act as an alternative to death, offences whose punishment could include benefit of the clergy (a mechanism of grace of historic origin whereby for certain offences guilty parties could avoid the death penalty) were eligible to receive seven years transport to America instead of execution. The outbreak of Britain’s war with America in 1778 halted transportation, which resumed again in 1787 with transport to Australia. By the early 1800s transportation was the maximum sentence of many crimes. Due to mounting doubts about transportation’s effectiveness as a deterrent, and the appalling conditions in the penal colonies, the punishment was abolished in 1857 and replaced in most cases by hard labour[9].

Key dates for the Old Bailey and the Proceedings. A subset of the key dates likely to have a visible affect on the recorded data:
- 1674 First publication of the Proceedings
- 1678 Regular accounts of trials begin.
- 1679 Publication of the Proceedings require official approval.
- 1718 Transportation introduced.
- 1776 War with America, no transportation.
- 1778 The Proceedings is subsidised and must be a “True, fair, and perfect narrative”
- 1787 Transportation resumes.
- 1808 No death for pick-pocketing.
- 1832 Removal of the death penalty for most forms of theft.
- 1834 The Old Bailey given jurisdiction over all of England, method of recording changes.
- 1836 Counsel Act, increases the role of council.
- 1855 Less serious crimes no longer seen at the Old Bailey.
- 1857 Transportation abolished.
- 1861 Offences Against the Persons Act, revocation of death penalty for most offences.
- 1913 Publication of the Proceedings ends.

Appendix: The journey of data
This study presents one stage in the journey from historic events to historical understanding. It is tempting to take the available data as fact; when realised as tables of data on screen it appears concrete. The truth, should it exist, is far more complicated and requires us to be circumspect with the claims made and the certainty ascribed to the data. This section attends to that issue by discussing the data in the context of the data creation process. Fundamentally, of course, it is important to remember that the data originates with trials entering the OBO and is therefore entirely dependent on the Proceedings for the record of those events. As discussed in ??, the Proceedings is fraught with issues of reliability, to the extent that the quality and relative meaning of the published trial accounts change depending on the year in question. It is unlikely, moreover, whether the extent to which trials are unreported or have incorrect or false outcomes printed can ever be known.

Physical records of the Proceedings have survived as either printed editions or microfilm copies, both of which were scanned to 400ppi TIFF image files to begin the digital stage of the data’s journey. For the 1864-1834 period of the Proceedings, the OBO team had two typists rekey the same TIFF images to text, then compared them to reduce error. A similar process was employed for the period 1834 to 1913, this time with a single typist and computer optical character recognition (OCR).

Markup such as defendant and victim name and gender, offence type, verdict, and punishment were added in a similar manner as the rekeying: manually for records from 1864 to 1834 and in an automated manner thereafter. As The Old Bailey Online team note, it is impossible to avoid some issues caused by changes in printed language, especially relating to the old English long s ( keeper k) which looks like an f to both humans and computers. Despite the 1% error it is still advisable to search for alternative spellings.

The second greatest source of error in terms of this study (second to biases and misreporting in the Proceedings) result from decisions made about text markup, including the choice of markup categories. Accused and victim names have a certainty of form, for example, that the categorisation of offence, verdict and punishment do not. Similarly, changing interpretations of law aside, the grouping of offences and punishment into major and minor categories despite no such a-priori list existing represents a large bias - albeit a well intentioned and utterly necessary step.

For the purpose of this project, the next step in the journey of the data was from information stored in the OBO SQL server into our Pandas tables via the web-API.

Appendix: Retrieving data from the OBO
The OBO web-API. A web-based Application Programming Interface or web-API is generally characterised as a documented method for querying a web site with the query being generated as a URL. The web content returned is of a known format, to allow the information to be reliably parsed. Throughout this section Listing 1 will be used to illustrate a typical OBO web-API query.

Listing 1: OBO web-API example query.
http://www.oldbaileyonline.org/obapi/ob? term0=fromdate_1774 &term1=todate_1824 &term2=offcat_theft&terms0= defgen_female &breakdown=puncat&count=1

The first term is the web-API URL:
http://www.oldbaileyonline.org/obapi/ob?
with subsequent query terms being built up beginning with a term", where n is incremented with every term and k separating terms. In our example the terms result in a search for the occurrence of offences in the theft category (offcat_theft) committed by women (defgen_female) between 1771 and 1824 (fromdate_1774 &term=todate_1824), breaking down the results by punishment category (breakdown=puncat). The
Listing 2: OBO web-API example of returned data.
```json
{
  "total": 10234,
  "breakdown": [
    {"term": "imprison", "total": 3479},
    {"term": "miscPunish", "total": 2239},
    {"term": "transport", "total": 533},
    {"term": "corporate", "total": 867},
    {"term": "noPunish", "total": 128}
  ],
  "hits":
  [117740112-2]
}
```

Referring to Listing 2, the query returns JavaScript Object Notation (JSON), where “total” : 10234 shows the total number of events matching our criteria, and the terms following “breakdown” are the breakdown of that total into punishment categories. For example “term” : “transport”, “total” : 2038 shows that for the period of the query 2038 women accused of Theft were Transport as a punishment.

Categories and Subcategories. An important distinction in the OBO (and in particular in the web-API) is that between Categories and Subcategories. Offences, verdicts and punishments are first divided into Categories, and then further partitioned into more fine grained Subcategories. In the web-API terms can be searched over either Categories, subcategories, or a mix of both. From Listing 1, the `offcat_theft` term is used to search for offences in the Category of Theft. An example of a similar Subcategory query term is `offsubcat_highwayRobbery` for the offence Subcategory of Highway robbery. Only valid terms as defined in the API doc can follow the underscore. Categories and Subcategories have been kept separate, essentially resulting in two distinct versions of the downloaded data. The study has been conducted on the Categories data.

Downloading data from the OBO. The data is programmatically downloaded from the OBO via the web-API using Python, and relies heavily on advanced features of `urllib3`[42] to keep the connection open for multiple queries and to back off (dynamically reduce the time between requests) when the server delivering the content gives errors. The servers hosting the OBO provide no information as to how to politely issue requests, responding inconsistently to large volumes of requests. Hence the need to dynamically adjust the time between queries. Listing 3 shows the code for generating URL’s.

Listing 3: URL generating code
```python
def generateURLCategoryRange(StartDate, EndDate, OffenceCategory, Gender='None'):
    """Generate an OBOAPI url for insertion to an http 'GET' request.
    """
    if Gender is 'None':
        return 'http://www.oldbaileyonline.org/obapi/ob?term0=fromdate&term1=startdate&term2=offcat0&term3=' + OffenceCategory + '&Gender=' + Gender + '&breakdown=1'
    else:
        return 'http://www.oldbaileyonline.org/obapi/ob?term0=fromdata&term1=startdate&term2=offcat0&term3=' + OffenceCategory + '&Gender=' + Gender + '&breakdown=1'

def generateURLSubCategoryNotGuiltyRange(StartDate, EndDate, OffenceCategory, Gender='None'):
    """Generate an OBOAPI url for insertion to an http 'GET' request.
    """
    if Gender is 'None':
        return 'http://www.oldbaileyonline.org/obapi/ob?term0=fromdate&term1=startdate&term2=offcat0&term3=NotGuilty&term4=offsubcat&Gender=' + Gender + '&breakdown=1'
    else:
        return 'http://www.oldbaileyonline.org/obapi/ob?term0=fromdata&term1=startdate&term2=offcat0&term3=NotGuilty&term4=offsubcat&Gender=' + Gender + '&breakdown=1'
```

Listing 4 is a code snippet for opening the socket, getting the URL and returning JSON. Note that the `http` object is initialised outside the loops for getting each date range and gender etc.

Listing 4: Socket opening and URL getting code
```python
import urllib3
retry = urllib3.util Retry(total=100, read=100, connect=100, backoff_factor=1)
timeout = urllib3.util.Timeout(connect=4.0, read=8.0)
http = urllib3.PoolManager(retry, timeout=timeout, maxsize=5)
def URLtoJSON(URL, httpPool):
    import json
    return json.loads(httpPool.request('GET', URL).data.decode('utf8'))
```

Listing 5 shows how the JSON returned by the OBO web-API is processed into arrays, then concatenated together to become a row of our DataFrame. `c2n.puncat` is an array containing the punishment categories (sim. `c2n.offcat`).

Listing 5: JSON processing
```python
import CategoryToNumberAssignment as c2n
TempCategories = [0]

for Totals in json['breakdown']:
    TempCategories[c2n.puncat.index(Totals['term']) + 1] = Totals['total']

# Append Temps to the Rows
Row = Row + TempCategories
```

Data representation. Raw data obtained from the OBO are frequencies of punishments for a given offence in a given time period. As such the retrieved data points are offence-punishment pairs. These offence-punishment pairs can be represented as a matrix as in Figure 12 where the i,jth element is the frequency or probability estimate of p_{ij} given α. With ∆t = 1 there are 240 such matrices (one for every year) for the period of the Old Bailey. It is more convenient to store the matrix for a single year as a row in a table indexed by the date as in Figure 13 such that the data for all years is represented in a single table. In the matrix representation the punishments associated with an offence is simply a column, and the offences conditional on a punishment is a row. In the table view the punishments associated with an offence, α_{i}, is found as a slice of the columns from o_{pi} to o_{pi}, and the offences associated with a punishment, p_{i}, is the selection of columns {o_{pi}, i ∈ {1...9}}. Note that when working with the tabular data in Python Pandas the columns and rows can be accessed via name and index. The indexes start at 0 and the columns are named as ‘DeceptionNotGuilty’ where the offence and punishment names are joined without spaces in camel-case.
as used by Python scripts for manipulating this data.

\[
\begin{bmatrix}
    x_{1,1} & x_{1,2} & x_{1,3} & x_{1,4} & x_{1,5} & x_{1,6} & x_{1,7} & x_{1,8} & x_{1,9} \\
    x_{2,1} & x_{2,2} & x_{2,3} & x_{2,4} & x_{2,5} & x_{2,6} & x_{2,7} & x_{2,8} & x_{2,9} \\
    x_{3,1} & x_{3,2} & x_{3,3} & x_{3,4} & x_{3,5} & x_{3,6} & x_{3,7} & x_{3,8} & x_{3,9} \\
    x_{4,1} & x_{4,2} & x_{4,3} & x_{4,4} & x_{4,5} & x_{4,6} & x_{4,7} & x_{4,8} & x_{4,9} \\
    x_{5,1} & x_{5,2} & x_{5,3} & x_{5,4} & x_{5,5} & x_{5,6} & x_{5,7} & x_{5,8} & x_{5,9} \\
    x_{6,1} & x_{6,2} & x_{6,3} & x_{6,4} & x_{6,5} & x_{6,6} & x_{6,7} & x_{6,8} & x_{6,9} \\
    x_{7,1} & x_{7,2} & x_{7,3} & x_{7,4} & x_{7,5} & x_{7,6} & x_{7,7} & x_{7,8} & x_{7,9} \\
\end{bmatrix}
\]

\[
\delta_{(p_1)} = \frac{1}{\sum_{i=1}^{7} x_{i,j}} \delta_{(i,j)}
\]

\[
\Delta t \text{ represents the empirical data is done by the changePeriod(EmpFrame, Delta) function on the one year raw data, where the empirical frequencies are summed down the rows in } \Delta t \text{ blocks resulting in a raw data table with } \frac{\Delta t}{\Delta t} \text{ rows. For } \Delta t > 1 \text{ these reduced raw data tables are used to generate the probability estimate tables.}
\]

Listing 6: Raw data resampling code snippet from the function: changePeriod.

Start date = 1674

Check the number of rows in the EmpFrame
EmpRows = EmpFrame.shape[0]

Create an empty list the same length as the number of rows
\[\text{indexGroup} \equiv \left(\text{EmpRows} \div \Delta\text{Delta}\right)\text{i}
\]

Create a list with EmpRows/Delta(ish) unique entries the
same size as EmpRows for row in range(0, math.ceil(EmpRows/Delta)): \text{indexGroup[\text{row}+\Delta\text{Delta}*(\text{row}+1)+\Delta\text{Delta}] = \left[\text{row}\right]+\Delta\text{Delta}+}

Remove extra entries
\[\text{del } \text{indexGroup[EmpRows]} \]

Resample the EmpFrame:
\[\text{return EmpFrame.groupby(indexGroup).sum()}\]

Bayesian smoothed estimates. Referring to Equation (5), the code to produce a Bayesian smoothed estimate of a DataFrame of empirical data is presented in Listing 7, and the resulting table in Table 8.

Listing 7: Laplace smoothing code snippet from the function: generateBetaDependentModelLaplace.

DFlater = EmpFrame.shape[1]

For the first row set alpha = 1
ModelFrame.loc[\text{row} : \text{loc}[\text{row}] + \text{div}(\text{EmpFrame.loc[\text{row}] + 1})]

\[\text{ModelFrame.loc[1] = \text{EmpFrame.loc[1] + ModelFrame.loc[1-\Delta\text{Delta}].div(EmpFrame.loc[\text{row}]+1)}\]

Table 8: Laplace smoothed probability estimate subset as percentage. Note that the 0 values are due to truncation.

**Conditional estimate.** The probability estimates of punishments conditional on offence, \(\theta(p_i | \alpha_i)\), and offences conditional on punishment, \(\theta(\alpha | p_j)\), are found from Laplace

20 Mackenzie, Sainudiin, Smithies and Wolffram
smoothed probability estimates on the entire offence-punishment pairs data with the following code:

**Listing 8:** Probability estimate conditional on offence \((\theta(p_j | o_i))\) code snippet from the function: `conditionPunishmentEstimatesOnOffence`

\[
p = 7 \\
o = 9 \\
\text{for } c \in \text{range}(0, o) \\
\text{CatModCond} . iloc[:, \text{exp}((c)\cdot p)] = \text{CatModCond} . iloc[:, \text{exp}((c)\cdot p)] . \text{div}((\text{CatModCond} . iloc[:, \text{exp}((c)\cdot p)] . \text{sum}(\text{axis}=1), \text{axis}=0)
\]

**Listing 9:** Offence estimate conditional on punishment \((p_j | o_i))\) code snippet from the function: `conditionOffenceEstimatesOnPunishment`

\[
\text{ModLen} = \text{CatMod} . \text{shape}[1] \\
p = 7 \\
o = 9 \\
\text{for } c \in \text{range}(0, p) \\
\text{CatModCond} . iloc[:, \text{list} \text{range}(c, \text{ModLen}, p)] = \text{CatModCond} . iloc[:, \text{list} \text{range}(c, \text{ModLen}, p)] . \text{sum}(\text{axis}=1), \text{axis}=0
\]

Combine offence-punishment pairs on offence or punishment.

From the offence-punishment pairs estimates or raw datasets, subsidiary datasets on offences or punishments only can be achieved by slicing and summing the original data as in Listing 10.

**Listing 10:** Combining offence-punishment pairs on offence or punishment.

\[
\text{def combineOnOffence(DataFrame)}: \\
\text{offenceGroup} = [\text{offence for offence in c2n.offsetUp for punishment in c2n.punUp}] \\
\text{OffenceFrame} = \text{DataFrame.groupby(offenceGroup, axis=1, sort=False).sum()} \\
\text{return OffenceFrame}
\]

\[
\text{def combineOnPunishment(DataFrame)}: \\
\text{punishmentGroup} = [\text{punishment for offence in c2n.offsetUp for punishment in c2n.punUp}] \\
\text{PunishmentFrame} = \text{DataFrame.groupby(punishmentGroup, axis=1, sort=False).sum()} \\
\text{return PunishmentFrame}
\]

**Log likelihood.** Referring to the In term of Equation (6), the log-likelihood between the empirical frequencies and the probability estimate table are found table wide assuming that both tables are of the same size as in Listing 11.

**Listing 11:** The log-likelihood code snippet from the function: `logLikelihood`

\[
(\text{EmpiricalFrame} . \text{np} . \text{log}(\text{ModelFrame})) . \text{sum}(\text{axis}=1) . \text{sum}()
\]

**AIC.** Referring to Equation (6), the code to produce an AIC score from a estimated probability model DataFrame frame and the raw data is in Listing 12.

**Listing 12:** The AIC code snippet from the function: `generateAllStandardModels`

\[
\text{if 'Dependent' in filename:} \\
t = (8-1)\cdot(7-1)\cdot\text{CatEmp} . \text{shape}[0] \text{\number of rows of data} \\
\text{elif 'Independent' in filename:} \\
nt = (9-1)\cdot(8-2)\cdot\text{CatEmp} . \text{shape}[0] \\
\text{AIC} = 2nt - 2nt
\]

**Bootstrap.** Generating the bootstrap data \(N\) (large) times and finding the upper and lower percentile is non trivial. This is achieved through the use of the three dimension version of the DataFrame, the Pandas Panel, which is essentially a series of DataFrame's. The code to generate the confidence intervals is in Listing 13.

**Listing 13:** Bootstrap for confidence interval code: `generateBootStrapEstimateConfidenceIntervals`

\[
\text{BootstrapPanel} = \text{pd.Panel(items=list(range(N)), major_axis=EmpFrame.index, minor_axis=EmpFrame.columns)} \\
\text{for row in range(DFlon)}: \\
\text{empCountList} = [n for n in enumerate(EmpFrame)] \\
\text{EmpCountLen = len(EmpCountList)} \\
\text{RandomSampleFrame} = \text{pd.DataFrame( EmpCountLen, EmpCountLen, N)} . \text{columns = list(range(N), index = list(range(EmpCountLen)))} \\
\text{RandomSampleFrame} = \text{RandomSampleFrame.applymap(lambda x : x.value_counts( bins=[-2, -1, 1, range(N)], sort=False, shift(1)[2]}} \\
\text{RandomSampleEvents} = \text{RandomSampleFrame.apply(map(lambda x : x.value_counts( bins=[-2, -1, 1, range(DFlon+1)], sort=False).shift(1)[2])} \\
\text{PairWiseTV = BootstrapPanel.iloc[:, PairwiseIndex]} \\
\text{ModelFrame} = \text{ModelFrame.apply(lambda x : x.unique(ConfLowFrame) .quantile(ConfLowFrame))} \\
\text{ModelFrame} = \text{ModelFrame.apply(lambda x : x .quantile(ConfHighFrame, axis=0))} \\
\text{return ModelFrame}
\]

**Total variation distance.** Three different TVD's are used: Consecutive TVD between each adjacent \(t\), TVD between a fixed \(t\) and all \(l\in\{1, 2, \ldots, [240/\Delta]\}\), and the TVD between punishment distributions\(^1\) of all pairs of offences in the same time-block. The code to perform these operations is in Listing 14 and Listing 15.

**Listing 14:** The TVD code snippet from the function: `generateTVD`

\[
\text{\#Calculate consecutive total variation distance and TV per year} \\
\text{Variation[ 'CumulativeTVD' ] = 0.5 \cdot \text{ModelFrame} . \text{diff()} . \text{abs()} . \text{sum()} . \text{axis}=2} \\
\text{\#For each row in ModFrame calculate the TVD between it and every other} \\
\text{\#In ModFrame index} \\
\text{\#\text{Variation[row] = 0.5 * (\text{ModFrame - ModelFrame.iloc[row]}).abs().sum()} . \text{axis}=2} \\
\text{\#Generate the TVSum, and auto scale it for plotting} \\
\text{\#\text{TVSum} = \text{Variation.iloc[::1].sum() .values} . \text{values}} \\
\text{\#\text{TVSumScaled} = \text{TVSum} / \text{TVSum} . \text{max()}}
\]

**Listing 15:** The pairwise TVD code snippet from the function: `generatePairwiseTV`

\[
\text{PairWiseIndex} = 0 \\
\text{for i in range(0, OffenceArrayLen):} \\
\text{\text{for j in range(1, OffenceArrayLen-1):} \\
\text{\#This captures all \(N\) choose \(2\) possibilities} \\
\text{\#\text{PairWiseTV.iloc[:, PairWiseIndex] = 0.5 * (ModelFrame.iloc[:, i] + \text{ModelFrame.iloc[:, j]} + \text{ModelFrame.iloc[:, i]} + \text{ModelFrame.iloc[:, j]}).abs().sum()} . \text{axis}=2} \\
\text{\#PairWiseIndex += 1}
\]

\(^1\text{Note that a pairwise TVD between non conditional probability estimates for the punishments of a pair of offences can be taken. This is not a true TVD as neither pair of probabilities is a distribution, rather subsets of the offence-punishment distribution}\)

Mackenzie, Sainauldin, Smithies and Wolffram 21
Partitioning. The partitions are found using *partition.Partitions* from the 'PartitionSets' *Python* library. This has the advantage of giving an ordering to the partitions, such that a given partition number can be associated with the actual partition, with the bonus that partitions are only calculated when requested. For every possible partitioning, the AIC score for each time block is calculated and stored as a column of a DataFrame by the `generateAICTable` function which finds the AIC score per partitioning by calling `partitionAIC`. These functions also return a DataFrame showing the partitioning together with the probability of death per block.

The DataFrame of AIC scores is then sorted by row from smallest to largest at the same time as the partitions associated with these scores are recorded to give a new DataFrame of sorted partition numbers. The best partitioning by AIC is then in the first column of these sorted DataFrame’s. The same ordering is applied to the DataFrame of partitionings with probability of death. The sorting is achieved by calling `sortAICTable`. The partitioning and sorting is very slow in pure *Python*. To improve the performance of these functions are copied to a *.pyx* file for compilation by *Cython*, allowing the modules to be imported into a Python session, yet with the actual work having been converted to C code and compiled to machine language as a library. All this is transparent to the user. Details of the compilation process are in `OBOpartitioning.pxd`. The speed increases, with almost no optimisation, are impressive at over ten times.

Code validation. The pivotal step of the data mangle is that of validating whether the code is functioning correctly. The programming workflow of *Python* in an interactive environment, where each step can be atomically checked, helps ensure that functions will perform as expected. In order to test the operation of a whole programme flow, simulated data is generated which will give a known result upon correct operation of the functions. To test the model selection process a selection of probability distributions on offence-punishment pairs are generated, repeating the data row wise for a given number of rows to simulate $\Delta t$. The punishment distribution is not dependent on offence. The dummy data is produced with the functions `generateDummyDataFrame` whose resulting data frame is fed into `validateAICModelSelection` which returns the best model. This works as expected, returning an independent model with the correct $\Delta t$.

This procedure is similarly repeated on the partition selection process for simulated data with three sets of punishment distributions distributed amongst the offences. The partition selection process validation function, `validatePartitioning`, successfully finds which offences had the same punishment distributions.

Data generation workflow. This section is an explicit guide to the functions used to generate the various results. The code is available as a *git* repository from .

During the course of this study the *iPython* interactive *Python* shell has been used exclusively, it is recommended that you do the same. The one exception is in the downloading of the data from the OBO, where the module can be run from a terminal. This was done so that it could be called in a never ending *while* loop, to ensure that it would restart if the OBO server started issuing errors.

Some of the functions have hard coded relative paths to the raw data and various models. It is assumed that the following relative paths exist. (By relative it is meant that the directories are within the current directory which contains the *Python* code), and that *iPython* is launched from the directory containing the following OBO *Python* modules: (i) `Data/Raw` is the directory for empirical (raw) data from the OBO, (ii) `Data/AICModels` is the directory for generated models, (iii) `Data/AIC` is the directory for AIC results on non-partitioned models, (iv) `Data/Bootstrap` is the directory for bootstrap confidence values, and (v) `Data/Partition` is the directory for partition model selection data.

Upon entering the *iPython* shell it is useful to load all the required libraries and modules. It is more convenient to do this using a script loaded at startup such as Listing 16 running *iPython* from a terminal with the command: 

```
iPython3 -i iPython3-startup.py
```

Listing 16: *iPython* startup script

```python
import argparse, pathlib, re, glob
import pandas as pd
import numpy as np
from partitionSets import partition
get_python() magic('load_autoreload
')
get_python() magic('autoreload 2
')
import CategoryToNumberAssignment as c2n
import OBapiExtraction as oboA
import OBOModelling as oboM
import OBOPartitioning as oboP
import OBOValidation as oboV
```

```python
import argparse, pathlib, re, glob
import pandas as pd
import numpy as np
from partitionSets import partition
get_python() magic('load_autoreload
')
get_python() magic('autoreload 2
')
import CategoryToNumberAssignment as c2n
import OBapiExtraction as oboA
import OBOModelling as oboM
import OBOPartitioning as oboP
import OBOValidation as oboV
```

Download the raw empirical frequencies for ‘NoGender’, ‘Female’, ‘Indeterminate’ and ‘Male’ defendants is done by issuing the following command from a terminal: 

```
python3 OBapiExtraction.py
```

Data can be selectively downloaded from within the *iPython* shell using the `generateDataFramesRangeSaving` function as: to download data with no gender specification on defendants, between the years 1674 and 1913 yearly, saving every iteration (year). This downloads data on categories and subcategories from the OBO concatenating them together in the Raw directory as `OBDataExtractNoGenderXyr.csv`

To load the empirical data on categories use: Now that the data is loaded models can be generated from it. This can be done individually returning a new DataFrame, for example to generate a dependent model resampling the period to two years: For model selection, as many models as reasonable need to be generated. This is achieved en masse by the `generateModelFiles` as: Once all the models have been generated the AIC scores can be calculated. The function `generateAICstandardModels` assumes that all the models have been generated and placed in the appropriate directory, and generates the AIC score table by either writing it to file or returning it to the *iPython* shell. To return the AIC table to the shell:

```
which shows the best model as the top line, being the dependent model with $\Delta t = 2$.
```

The bootstrap data for $\Delta t = 2$ is generated as: These confidence frames can be easily saved as: and To view the offence-punishment pair data in terms of offence or punishment alone, two functions perform this action on either empirical data or model data as: and The partitioning is best performed using the *Cython* version of `OBOPartitioning.py`. `OBOpartitioning.pxd` which is loaded and run as: `import OBOpartitioning as oboP; AICFrame, DeathProbFrame = oboP.generateAICTable(CatEmp)` — Once the partitioning frames have been generated they need to be sorted before carrying out the MCMC run: From these sorted values a whole series of MCMC runs can be generated for $N = [1000, 10000, 100000]$ saving the results in the Data/Partitions directory as: These steps will allow the complete reproduction of this study.

22

Mackenzie, Sainudin, Smithies and Wolffram