

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

2nd revision 20151129.

Jasper Mackenzie^{*}, Raazesh Sainudiin^{*†}, James Smithies[‡] and Heather Wolfram[‡]

^{*}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. If appropriate methods don't appear, algorithmic analysis of large humanities datasets 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. Frequencies of 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 selected model suggests that the underlying process governing the Old Bailey data changes every three years. By partitioning offences, we find that the distribution of sentences (punishments) for each offence becomes less haphazard over time, offering strong evidence of the civilizing process.

Old Bailey Online | social evolution | nonparametrics | civilizing process | digital humanities | open-source

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 (<http://www.oldbaileyonline.org>) 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 antic-

ipated to be reproduced as more funding becomes 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 regularization of judicial processes, for example) are tightly correlated with the development of a more civil society characterized by reduced crime and violence, and the development of

[†]To whom correspondence should be addressed. E-mail: r.sainudiin@math.canterbury.ac.nz

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-APIextraction-and-Analysis>

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 aspects 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 mindful of the complex nature of the data’s origin and the complexity of knowledge required to bring meaning to results generated from it. The data in this study is of historic legal origin, firmly rooted in the history of crime and punishment in England. Care must be taken when working with this kind of data not to influence the analysis with either limited historic understanding or invalid assumptions based on over identification with qualitative (or previous quantitative) evidence. A critical awareness of the *choices* imposed on the data and models 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, estimated 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 probabilities vary significantly across consecutive three year periods. This in itself is an intriguing finding, given it suggests a relatively 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 process. 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 sentence evident over time, these results provide strong evidence of the civilizing process in action: from the 1640s, where punishments were violent and haphazard, to 1913 where punishments were well defined, commensurate with the offence and the likelihood of the death sentence was low. The data we work with are actually offence-sentence pairs, whereby a sentence may not necessarily result in its execution due to the possibility of being pardoned. We use the term punishment with this possibility in mind as the sentence still captures the essence of the intended punishment.

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

Methods

It is important that appropriate tools are selected for the study of historical datasets, to encourage collaboration between ex-

perts in computational statistics and the humanities. A simple and general reference methodology is chosen for this reason. 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 completed within a single ecosystem. This has considerable benefit for humanists who may be interested in the possibilities inherent 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. All supporting code for this project is available via <https://github.com/spmp/OBO-APIextraction-and-Analysis>.

To help historians explore the civilizing process, and to account 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 offence as taken from the OBO web-API. In their markup of trials, the OBO include offence, verdict, and punishment category per accused person, exposing the counts per categorical event over trials via the web-API. Verdicts are considered to be binary: guilty or not guilty, and without loss of generality, 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., $[n] = \{1, 2, \dots, n\}$. The nine offences are notated \mathbf{o}_i where $i \in [9]$, and the seven punishments are notated \mathbf{p}_j where $j \in [7]$ 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, Δt , where typically:

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

and for a given Δt and a corresponding

$$\tau = \tau(\Delta t) = \lceil 240/\Delta t \rceil,$$

the time-blocks are indexed by $t \in [\tau]$. Thus, cases in a given time-block t are heard at the Old Bailey as a sequence of $n(t)$ many offence-punishment pairs:

$$(\mathbf{o}_{i_1}, \mathbf{p}_{j_1}), (\mathbf{o}_{i_2}, \mathbf{p}_{j_2}), \dots, (\mathbf{o}_{i_{n(t)}}, \mathbf{p}_{j_{n(t)}}).$$

This sequence is summed to obtain the counts for all 63 possible $(\mathbf{o}_i, \mathbf{p}_j)$ pairs as:

$$c_{i,j,t} = \sum_{r=1}^{n(t)} \mathbb{1}_{\{\mathbf{o}_i\}, \{\mathbf{p}_j\}} (\mathbf{o}_{i_r}, \mathbf{p}_{j_r}) \quad [1]$$

where

$$\mathbb{1}_{A,B}(x, y) = \begin{cases} 1 & \text{if } x \in A \text{ and } y \in B \\ 0 & \text{otherwise} \end{cases} \quad [2]$$

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 Δt on the basis of their parsimonious fit to the observed data.

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

$$\Pr\{(\mathbf{o}_i, \mathbf{p}_j, t)\} = \begin{cases} \theta_{i,j,t} & \text{if } i \in [9], \text{ and } j \in [7] \\ 0 & \text{otherwise} \end{cases}$$

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 Δt . 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:

$$\begin{aligned} & \Pr\{(\mathbf{o}_{i_1}, \mathbf{p}_{j_1}, t), (\mathbf{o}_{i_2}, \mathbf{p}_{j_2}, t), \dots, (\mathbf{o}_{i_{n(t)}}, \mathbf{p}_{j_{n(t)}}, t)\} \\ &= \prod_{r=1}^{n(t)} \Pr\{(\mathbf{o}_{i_r}, \mathbf{p}_{j_r}, t)\} = \prod_{i=1}^9 \prod_{j=1}^7 (\theta_{i,j,t})^{c_{i,j,t}}. \end{aligned}$$

Thus for the entire data \mathbf{d} :

$$\mathbf{d} = \left((\mathbf{o}_{i_1}, \mathbf{p}_{j_1}, t), (\mathbf{o}_{i_2}, \mathbf{p}_{j_2}, t), \dots, (\mathbf{o}_{i_{n(t)}}, \mathbf{p}_{j_{n(t)}}, t) \right)_{t=1}^{\tau}, \quad [3]$$

its likelihood is given by:

$$\mathcal{L}(\theta) = \Pr\{\mathbf{d}|\theta\} = \prod_{t=1}^{\tau} \prod_{i=1}^9 \prod_{j=1}^7 (\theta_{i,j,t})^{c_{i,j,t}}, \quad [4]$$

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

$$\theta = \{\theta_{i,j,t} \mid i \in [9], j \in [7], t \in [\tau]\}.$$

For a given time-block length Δt and the corresponding number of time-blocks τ , the point estimates, $\hat{\theta}_{i,j,t}$ for each time block t , are obtained as the posterior mean under the Dirichlet prior with parameters:

$$\begin{aligned} & \beta \alpha_{i,j,t}, \text{ where } i \in [9], j \in [7] \text{ such that} \\ & \beta > 0, \alpha_{i,j,t} > 0, \text{ and } \sum_i \sum_j \alpha_{i,j,t} = 1 \text{ for each } t \in [\tau]. \end{aligned}$$

Such a Bayesian estimation or Laplace smoothing is required to take into account years of zero data, such as 1701, and to avoid unsmoothed frequentist maximum likelihood estimates for offence-punishment pairs that have no empirical realization in a given time-block. According to our modelling assumption that all offence-punishment pairs are realizable due to the positive prior parameters, every offence-punishment pair

must have a non-zero estimate. The point estimates are given by the mean aposteriori estimates [26, Chapter 12] with an initially uninformed Dirichlet prior on offence-punishment pairs as:

$$\hat{\theta}_{i,j,t} = \frac{c_{i,j,t} + \beta \alpha_{i,j,t}}{\sum_i \sum_j (c_{i,j,t} + \beta \alpha_{i,j,t})} \quad [5]$$

where

$$\alpha_{i,j,t} = \begin{cases} 1/63 & \text{if } t = 1 \\ \hat{\theta}_{i,j,t-1} & \text{if } 1 < t \leq \tau. \end{cases}$$

Choosing $\alpha_{i,j,t} = \hat{\theta}_{i,j,t-1}$ when $t > 1$ takes advantage of the immediate past information. Choosing the precision parameter $\beta = 1$ weakens the prior distribution in favour of the observed data, makes the point estimates close to the frequentist maximum likelihood estimates, and allows the point estimates to vary, possibly sharply, between time-blocks. These choices can account for sudden variations in data that reflect actual historical changes in law or jurisdiction. Thus, these point estimates with the pragmatically chosen prior and precision parameters circumvent the usual problems with the frequentist maximum likelihood estimate (MLE) while remaining close to the MLE. Our choice for the prior and precision parameters are thus justified by wanting to keep point estimates close to the frequentist MLE. We defer the problem of Bayesian prior selection to a future study. A truly Bayesian approach integrating over all possible partitions of the yearly time-blocks, as opposed to several regularly partitioned time-blocks, with priors over time-block partitions and probability simplices over such partitions is beyond the scope of this study.

To test for independence, models are generated where the estimate of an offence-punishment pair is the product of the estimate of the punishment and that of the offence as:

$$\hat{\theta}_{i \times j, t} = \hat{\theta}_{i, t} \cdot \hat{\theta}_{j, t},$$

where likelihood for the dependent model is obtained by substituting $\theta_{i,t} \cdot \theta_{j,t}$ for $\theta_{i,j,t}$ in Equation (4). The model which best describes the data is selected by the penalized likelihood principle called the Akaike Information Criterion (AIC)[21]. The model with the lowest AIC score is chosen as the most parsimonious explanation of the data through time if the difference between the lowest and the next lowest score is more than two [27, 446]. The AIC score of a model is found by penalizing its log likelihood by K (the number of parameters in it) as:

$$\text{AIC} = 2K - 2 \log(\mathcal{L}) \quad [6]$$

where the number of parameters for the time varying dependent and independent models are:

$$\begin{aligned} K_{\text{DEP}} &= \tau \times ((9 \times 7) - 1) = \tau \times 62 \\ K_{\text{INDEP}} &= \tau \times ((9 - 1) + (7 - 1)) = \tau \times 14 \end{aligned}$$

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 $\widehat{\theta}_{i,j,t}$ 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 $\widehat{\theta}_{i,j,t}$ from the bootstrapped estimates. Instead of obtaining Bayesian credible intervals by sampling from the Dirichlet posterior density over the probability simplex $\theta_{\cdot,\cdot,t}$ 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 β), has a classical 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 , $\widehat{\theta}_{\cdot,\cdot,t}$, the sum of the absolute differences between a given pair of offences \mathbf{o}_i and \mathbf{o}_j is used as a distance measure given by:

$$\delta(\mathbf{o}_i, \mathbf{o}_j, t) = \sum_{j=1}^7 \left| \widehat{\theta}_{i,j,t} - \widehat{\theta}_{j,i,t} \right|.$$

Thus the smaller the value of $\delta(\mathbf{o}_i, \mathbf{o}_j, t)$, the closer are the relative punishment outcomes for the pair of offences \mathbf{o}_i and \mathbf{o}_j in time-block t . For a given time-block there are such δ 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 21,147 possible partitions of the 9 offences for each time-block t . A partition ψ_m of the set of offences $\{\mathbf{o}_1, \dots, \mathbf{o}_9\}$ is denoted as:

$$\psi_m = \{\psi_m^1, \psi_m^2, \dots, \psi_m^{\ell(m)}\}.$$

where each offence-block ψ_m^k is non-empty, i.e., $\psi_m^k \neq \emptyset$, and each pair of offence-blocks is disjoint, i.e., $\psi_m^k \cap \psi_m^{k'} = \emptyset$ for $k \neq k'$. Thus m enumerates all such partitions where $m \in [21147]$ and $\ell(m)$ is the number of offence-blocks within a partition ψ_m . Thus, a given offence, \mathbf{o}_i , is in a particular offence-block of a given offence partition ψ_m . Let the unique offence-block of ψ_m containing \mathbf{o}_i be $\psi_m^{k(i)}$, i.e.,

$$\psi_m^{k(i)} \cap \{\mathbf{o}_i\} = \{\mathbf{o}_i\}, \text{ for a particular } k(i) \in [\ell(m)].$$

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

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

$$\Pr\{\mathbf{o}_i, \mathbf{p}_j, t\} = \Pr\{\mathbf{p}_j | \mathbf{o}_i, t\} \times \theta_{i,t},$$

where, $\theta_{i,t} = \Pr\{\mathbf{o}_i, t\}$ is the probability of committing offence \mathbf{o}_i in time-block t . Now the likelihood of $\{\mathbf{o}_i, \mathbf{p}_j, t\}$ under a given offence partition ψ_{m_t} with $\ell(m_t)$ blocks is proportional

to the following conditional probability:

$$\begin{aligned} \Pr\{\mathbf{o}_i, \mathbf{p}_j, t | \psi_{m_t}\} &= \Pr\{\mathbf{p}_j | \mathbf{o}_i, t, \psi_{m_t}\} \Pr\{\mathbf{o}_i, t | \psi_{m_t}\} \\ &= \theta_{\psi_{m_t}^{k(i)}, j} \times \theta_{i,t}, \end{aligned}$$

where,

$$\theta_{\psi_{m_t}^{k(i)}, j} \text{ and } \theta_{i,t}$$

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

$$\theta = \{\theta_{m_t} : t \in [\tau]\}$$

where,

$$\theta_{m_t} = \left(\left\{ \theta_{\psi_{m_t}^k, j} : k \in [\ell(m_t)], j \in [7] \right\}, \left\{ \theta_{i,t} : i \in [9] \right\} \right)$$

is:

$$K_{\text{PART}} = \sum_{t=1}^{\tau} (8 + 6\ell(m_t)) = 8\tau + 6 \sum_{t=1}^{\tau} \ell(m_t)$$

Therefore, for a given Δt , the likelihood of a model specified by a sequence of τ offence partitions is:

$$\begin{aligned} \mathcal{L}(\theta) &= \Pr\{\mathbf{d} | \theta\} \\ &= \prod_{t=1}^{\tau} \left(\left(\prod_{k=1}^{\ell(m_t)} \prod_{j=1}^7 (\theta_{\psi_{m_t}^k, j})^{c_{\psi_{m_t}^k, j, t}} \right) \prod_{i=1}^9 (\theta_{i,t})^{c_{i,t}} \right) \quad [7] \end{aligned}$$

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

$$c_{\psi_{m_t}^k, j, t} = \sum_{r=1}^{n(t)} \mathbf{1}_{\psi_{m_t}^k, \{\mathbf{p}_j\}}(\mathbf{o}_i, \mathbf{p}_j) \quad [8]$$

The Bayesian smoothed estimates are obtained analogously to Equation (5) with partitioned offences for $\widehat{\theta}_{m_t}$ 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 generate 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 τ 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

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

where $S(9, i)$ is the Stirling number of the second kind, giving the number of partitions of the offences $\{\mathbf{o}_1, \dots, \mathbf{o}_9\}$ 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:

$$\begin{aligned} \Pr(M_{m,t}|\mathbf{d}) &\propto \Pr(\mathbf{d}|M_{m,t}) \Pr(M_{m,t}) \\ &\approx \Pr(\mathbf{d}|\widehat{\theta}_{m,t}) \Pr(M_{m,t}) \\ &\propto \mathcal{L}(\widehat{\theta}_{m,t}) \cdot e^{-6\ell(m_t)} \\ &= e^{-\text{AIC}(M_{m,t})/2} \end{aligned} \quad [9]$$

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:

$$P_{m,t} = \Pr(M_{m,t}|\mathbf{d}) = \frac{e^{-\text{AIC}_{m,t}/2}}{\sum_m e^{-\text{AIC}_{m,t}/2}} \quad [10]$$

Figure 2 illustrates the transition diagram where the vertices represent a partitioned model, $M_{m,t}$, for time-block t and offence-partition ψ_m , which has edges to all possible models in time-block $t+1$. Thus, we allow a transition from $M_{m'}$ 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 τ . 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 *Python* code to perform the modelling is complex. In order to verify that the code performs as desired, code validation is employed with simulated data of known parameters, so that by running the model selection process the original parameters are recovered.

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.

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 the Old Bailey's jurisdiction changed from just London and Middlesex, to 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 see 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.

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, p. 68-70]. 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 council 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.

Longitudinal model selection. From the empirical frequencies, two classes of non-parametric models are generated where punishments are assumed to be either (1) dependent or (2) independent of offence for a range of time-block lengths, Δt . The AIC scores for a representative set of models with different values of Δt across the dependent and independent classes of models are ranked in Table 1.

The model ranking is invariant under the AICc criterion (results not shown). The best model by the AIC (or AICc) criterion is the dependent model with $\Delta t = 3$ (Dep 3yr), having significant support with a difference in AIC score to the next best model of 310. The dependence of punishments on offences is expected for data on a court of law, and $\Delta t = 3$ establishes that fundamental processes driving events at the Old Bailey change significantly every three years. Only a representative subset of all the Δt values considered in our model selection procedure are reported in Table 1.

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 Break-

ingPeace) 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 λ into a parametric penalty function λK to find an optimal $\lambda \in (1, 15)$ that maximizes the likelihood of the held-out validation data, for instance. The range for λ is chosen to conservatively include the AIC, AICc and the BIC penalty regimes for the data set. 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.

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 with meaning intrinsic to the data. However, models explored in this manner performed worse than all models which evenly partitioned the time period, further indicating that the three year period is fundamental to events at the Old Bailey. The best mechanism for dynamically detecting changes in underlying processes is *change point detection* [35], a method outside the scope of this study.

Exploratory data analysis. The best model found by AIC is thoroughly investigated graphically. Many different perspectives on $\hat{\theta}_{i,j,t}$ 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 $\delta(\alpha_i, \alpha'_i, t)$, showing the absolute difference between estimated probabilities of punishments for pairs of offence slices. Fig-

ure 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 $\delta(\cdot, \cdot, t)$ 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 $\delta(\alpha_7, \alpha_8, t)$ 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.

Partitioned offence model. Exploratory data analysis, specifically the $\delta(\alpha_i, \alpha'_i, t)$ 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 S1. 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 $\lceil 240/3 \rceil = 80$ 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 offence partitions from the perspectives of being punished by Death, Imprisonment or Transport in Tables S1 to S3, respectively. The second and third best models from the MCMC sampler are depicted in Tables S4 and S5 to show the agreement between of 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 S1 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 S4 and S5 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 dramatic 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 dig-

itized 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 straight-forward 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].

We hope that this study, along with its methods through publicly available code for acquiring, extracting, transforming and analyzing the OBO data, will contribute to the foundations being established for future collaboration among 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 Mohand Delampady, Elena Moltchonova and Amandine Veber for their valuable comments on nonparametric models and statistical methodology, and Tim Hitchcock, Bob Schumacher and Simon DeDeo for their valuable historical comments that greatly improved this manuscript. We are grateful to the team at Old Bailey Online for maintaining and providing the dataset.

- Jockers ML (2013) *Macroanalysis: Digital Methods and Literary History*. (University of Illinois Press), 1st edition edition.
- Ramsay S (2011) *Reading Machines: Toward an Algorithmic Criticism*. (University of Illinois Press), 1st edition edition.
- Moretti F (2005) *Graphs, Maps, Tree: Abstract Models for a Literary History*. (Verso, London ; New York).
- Mohr JW, Wagner-Pacifi R, Breiger RL, Bogdanov P (2013) Graphing the grammar of motives in National Security Strategies: Cultural interpretation, automated text analysis and the drama of global politics. *Poetics* 41(6):670–700.
- Underwood T, Black ML, Auvil L, Capitanu B (2013) Mapping Mutable Genres in Structurally Complex Volumes, arXiv e-print 1309.3323.
- Mimno D (2012) Computational historiography: Data mining in a century of classics journals. *J. Comput. Cult. Herit.* 5(1):3:1–3:19.
- Heuser R, Le-Khac L (2012) A Quantitative Literary History of 2,958 Nineteenth-century British Novels: The Semantic Cohort Method, (Stanford University, Stanford Literary Lab), Technical report.
- Sainte-Marie MB, Meunier JG, Payette N, Chartier JF (2011) The concept of evolution in the Origin of Species: a computer-assisted analysis. *Literary and Linguistic Computing* 26(3):329–334.
- Hitchcock T et al. (1674–1913) *The Old Bailey Proceedings Online* (www.oldbaileyonline.org). Accessed: 24 March 2012.
- Thornbury W (1878) *Old and New London*. (Petter & Galpin) Vol. 2, pp. 461–477.
- Klingenstein S, Hitchcock T, DeDeo S (2014) The civilizing process in London’s Old Bailey. *Proceedings of the National Academy of Sciences* 111(26):9419–9424.
- Cohen D et al. (2011) *Data Mining with Criminal Intent*, (University of Hertfordshire), Technical report.
- Archer D (2014) *Historical Pragmatics: Evidence From The Old Bailey*. *Transactions of the Philological Society* 112(2):259–277.
- Dedeo S, Hawkins R, Klingenstein S, Hitchcock T (2013) Bootstrap Methods for the Empirical Study of Decision-Making and Information Flows in Social Systems. *Entropy* 15(6):2246–2276.
- Elias N, Jephcott E (1982) *The civilizing process*. (Pantheon Books New York) Vol. 2.
- Stone L (1983) Interpersonal violence in English society 1300-1980. *Past and Present* pp. 22–33.
- Eisner M (2003) Long-term historical trends in violent crime. *Crime and Justice* pp. 83–142.
- Spierenburg PC (2001) Violence and the civilizing process: does it work? *Crime, Histoire & Sociétés/Crime, History & Societies* pp. 87–105.
- Mennell S (2007) *The American civilizing process*. (Polity).
- May AN (2003) *The bar and the Old Bailey, 1750-1850*. (UNC Press Books).
- Akaike H (1973) *Information Theory and an Extension of the Maximum Likelihood Principle in Selected Papers of Hirotugu Akaike*. (Springer), pp. 199–213.
- Python Software Foundation (2015) *Python* (<https://python.org>).
- McKinney W (2010) *Data structures for statistical computing in Python*. pp. 51–56.
- McKinney W (2012) *Python for data analysis: Data wrangling with Pandas, NumPy, and IPython*. (O’Reilly Media, Inc.).
- Wes McKinney & PyData Development Team (2015) *pandas* (<http://pandas.pydata.org/>).

26. O'Hagan A (1994) *Kendall's Advanced Theory of Statistics, Vol. 2b: Bayesian Statistics*. (John Wiley & Sons).
27. Burnham KP, Anderson DR (2002) *Model selection and multimodel inference: a practical information-theoretic approach*. (Springer Science & Business Media).
28. The National Archives (2015) *Trials in the Old Bailey and the Central Criminal Court* (<http://www.nationalarchives.gov.uk/records/research-guides/old-bailey.htm>).
29. Demographia (2015) *Greater London, Inner London & Outer London Population & Density History* (<http://www.demographia.com/dm-lon31.htm>).
30. Briggs J, Harrison C, McInnes A, Vincent D (2005) *Crime and punishment in England: an introductory history*. (Routledge).
31. Kercher B (2003) Perish or prosper: The law and convict transportation in the British Empire, 1700–1850. *Law and History Review* 21(03):527–584.
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:695.
33. Lemmings D (2005) Criminal trial procedure in eighteenth-century England: The impact of lawyers. *The Journal of Legal History* 26(1):73–82.
34. Emsley C (2014) *The English police: A political and social history*. (Routledge).
35. Killick R, Eckley I (2014) changepoint: An R package for changepoint analysis. *Journal of Statistical Software* 58(3):1–19.
36. Klingenstein S, Hitchcock T, DeDeo S (2014) The civilizing process in London's Old Bailey. *Proceedings of the National Academy of Sciences* 111(26):9419–9424.
37. Shoemaker RB (2008) The Old Bailey proceedings and the representation of crime and criminal justice in eighteenth-century London. *The Journal of British Studies* 47(03):559–580.
38. Devereaux S (1996) The city and the sessions paper: "Public justice" in London, 1770–1800. *The Journal of British Studies* 35(04):466–503.
39. Huber M (2007) The Old Bailey Proceedings, 1674–1834. Evaluating and annotating a corpus of 18th-and 19th-century spoken English. *Annotating Variation and Change* 1.
40. (2015) Metropolitan Police Act (<http://www.legislation.gov.uk/ukpga/Geo4/10/44>).
41. (1861) Bill to consolidate and amend Statute Law of England and Ireland relating to offences against the person. 19th Century House of Commons Sessional Papers III:627.

Tables and Figures

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

No.	Model	K	$\log(\mathcal{L})$	AIC
4	Dep 3yr	4960	-480030.60	969981.219
6	Depr 4yr	3720	-481425.54	970291.088
2	Dep 2yr	7440	-477773.84	970427.684
8	Dep 5yr	2976	-483284.22	972520.448
10	Dep 10yr	1488	-487093.85	977163.715
0	Dep 1yr	14880	-474199.77	978159.557
12	Dep 15yr	992	-490532.03	983048.072
14	Dep 20yr	744	-491854.43	985196.869
1	Indep 1yr	3360	-491684.85	990089.712
3	Indep 2yr	1680	-493694.84	990749.699
5	Indep 3yr	1120	-495180.65	992601.306
7	Indep 4yr	840	-496195.71	994071.436
16	Dep 25yr	620	-496804.44	994848.883
9	Indep 5yr	672	-497728.45	996800.919
18	Dep 30yr	496	-498682.93	998357.866
11	Indep 10yr	336	-500741.72	1002155.45
13	Indep 15yr	224	-503764.69	1007977.38
20	Dep 40yr	372	-504682.61	1010109.23
15	Indep 20yr	168	-504935.42	1010206.84
17	Indep 25yr	140	-509243.00	1018766.01
22	Dep 50yr	310	-509979.89	1020579.78
19	Indep 30yr	112	-510927.84	1022079.69
21	Indep 40yr	84	-517644.45	1035456.91
23	Indep 50yr	70	-522660.80	1045461.60
24	Dep 100yr	186	-528494.51	1057361.03
25	Indep 100yr	42	-540626.11	1081336.23
26	Dep 240yr	62	-573432.44	1146988.89
27	Indep 240yr	14	-587599.35	1175226.71

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.

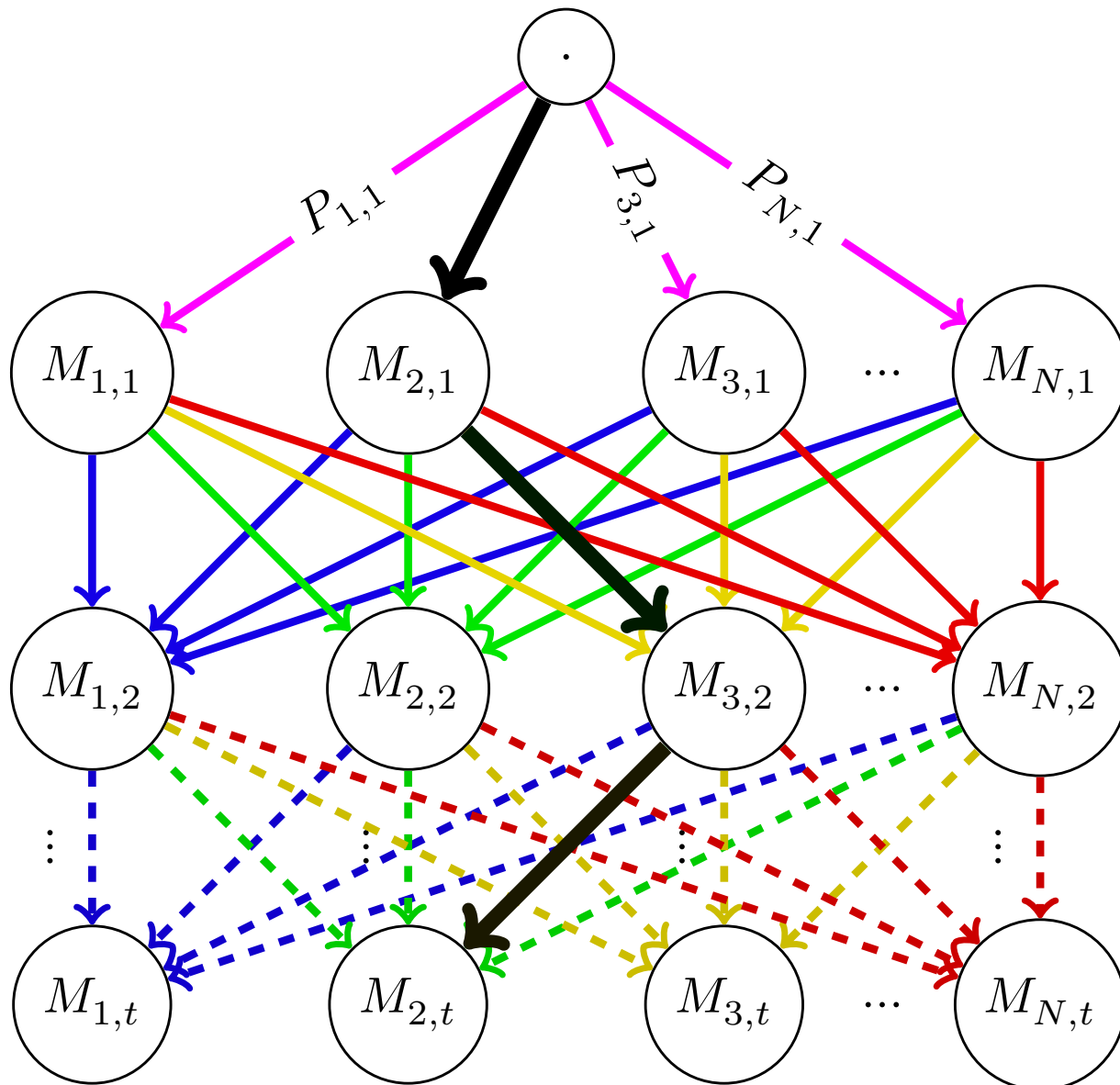


Fig. 2: The transition diagram of an MCMC sampler, with edge weights given by Equation (9) for models $M_{m,t}$ where t is the time-block, m ranges from 1 to 21,147 (N). Each vertex in time-block t has an edge to every vertex in $t + 1$ with the edge weight $\Pr_{t+1,m}$ from the connecting model in 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.

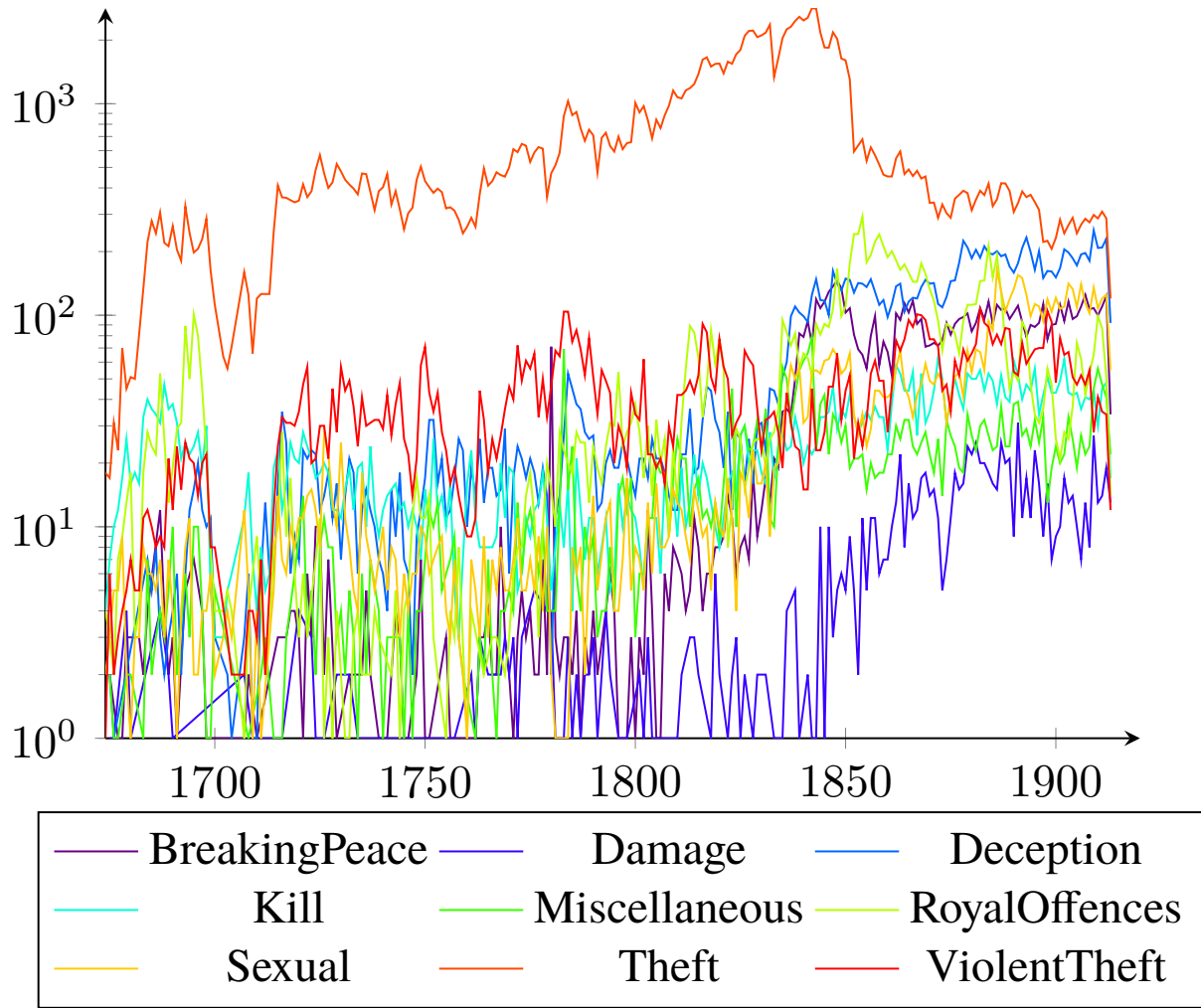


Fig. 3: Empirical frequencies per offence for the entire period of the Old Bailey on a log scale with frequency on the y-axis and date on the x-axis. Note the absence of data in 1701, the dramatic changes in Theft occurrences from 1840 to 1855 and increase in all other categories from 1840.

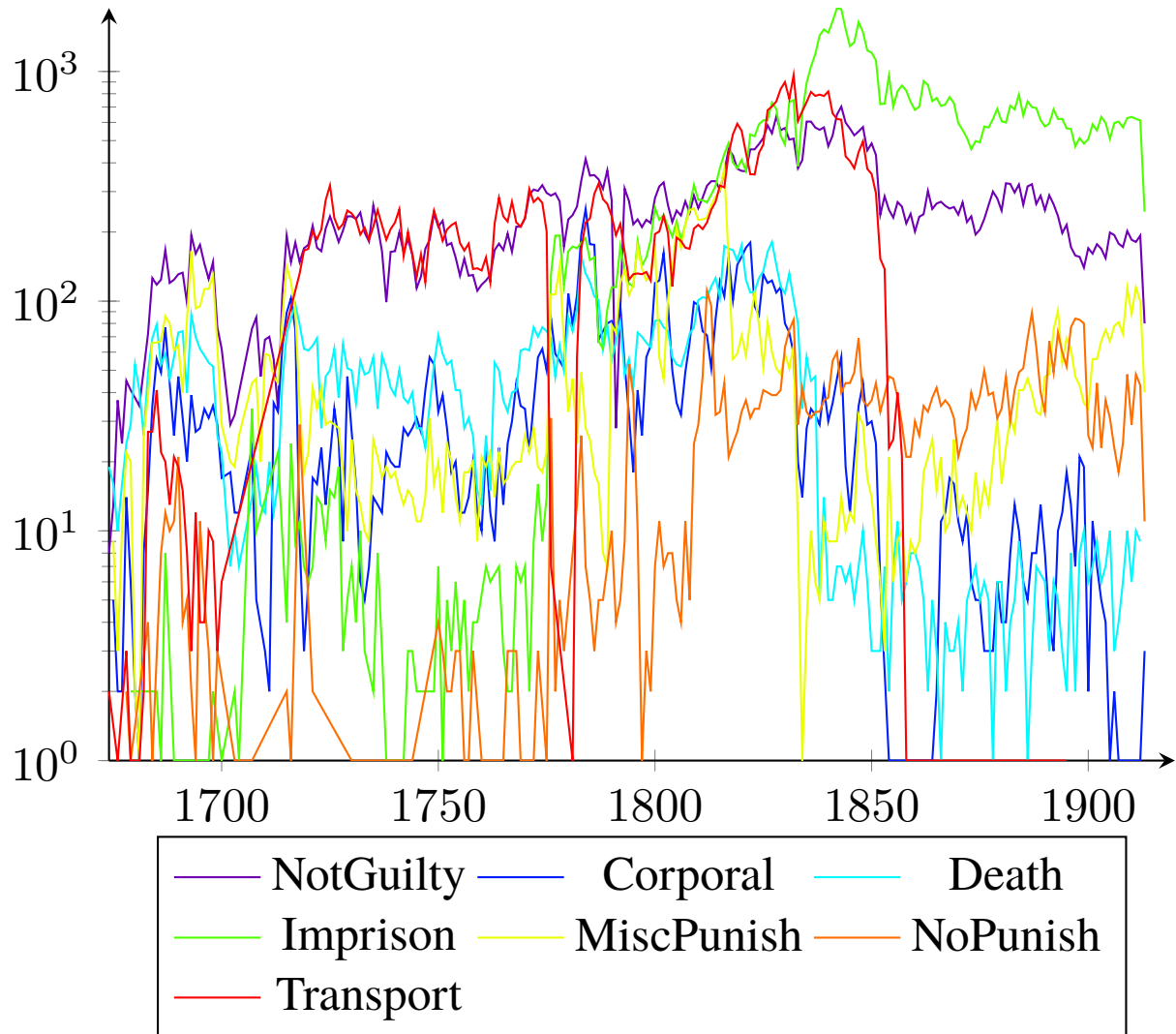


Fig. 4: Empirical frequencies per offence for the entire period of the Old Bailey on a log scale with frequency on the y-axis and date on the x-axis. Note the absence of data in 1701, the beginning of Transport in 1718 with a sudden break in 1780, ceasing in 1857, and the dramatic increase in Imprison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

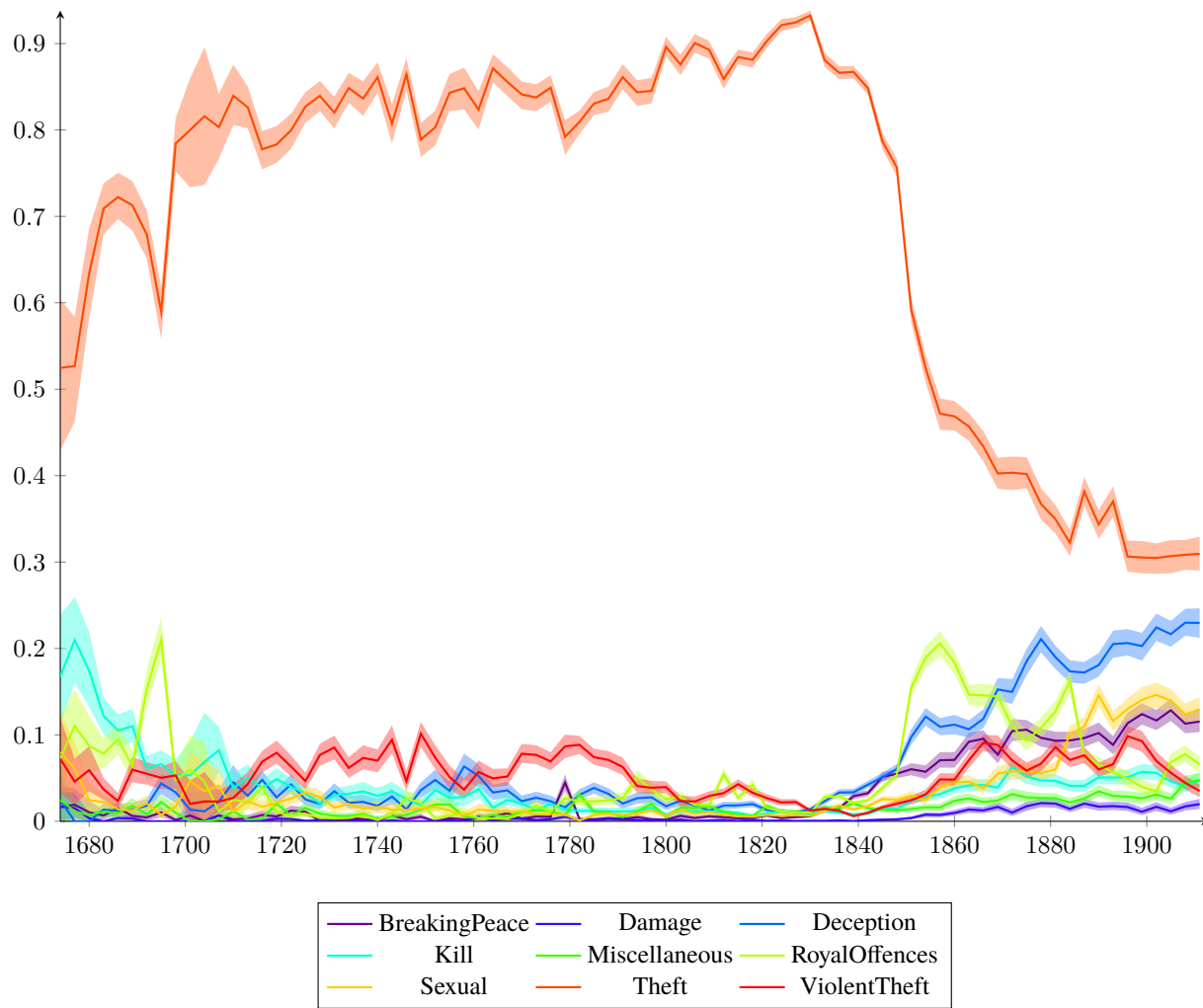


Fig. 5: Plot of offence probability estimates per offence for the best model (Dep 3yr)

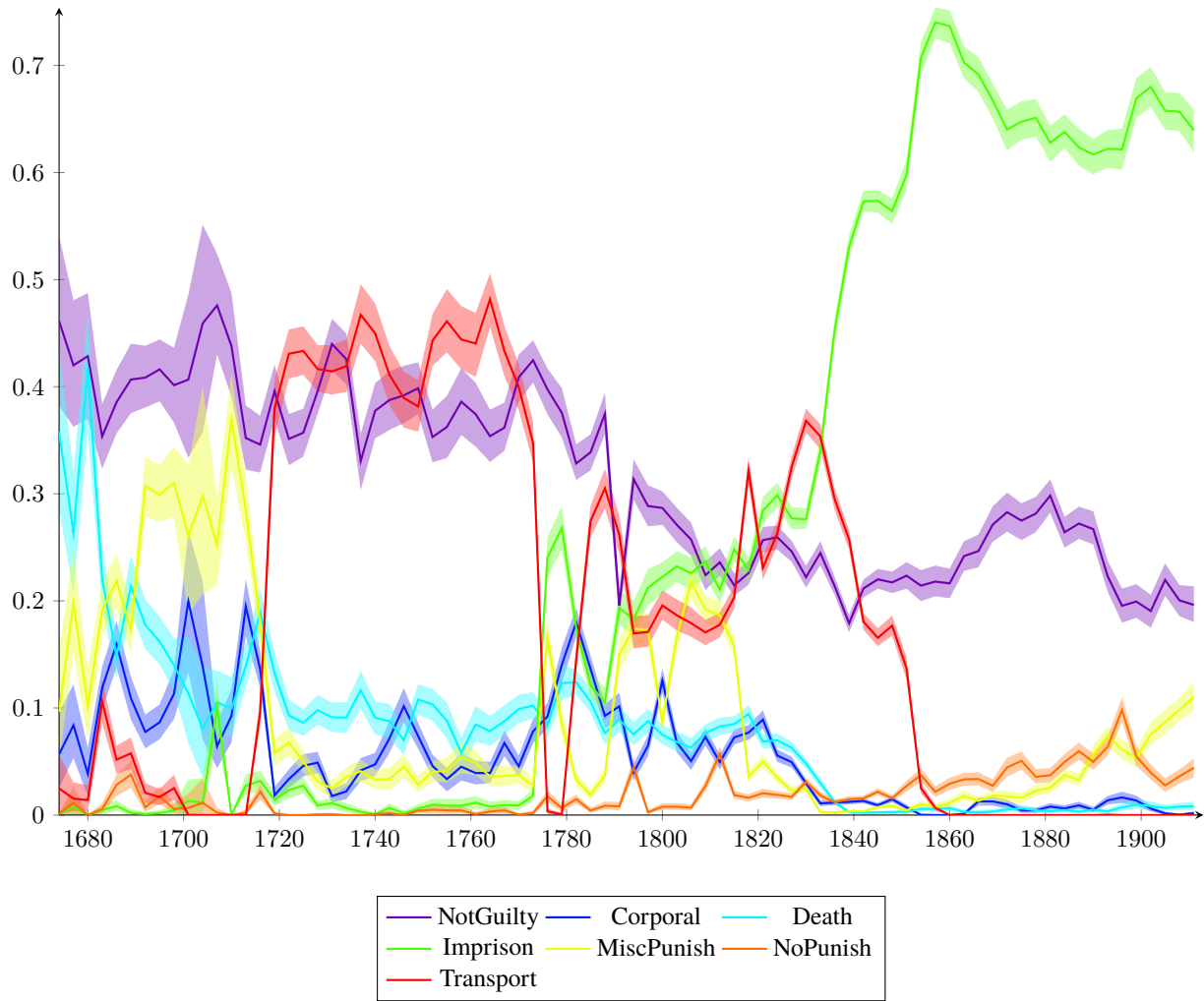


Fig. 6: Plot of punishment probability estimates per offence for the best model (Dep 3yr)

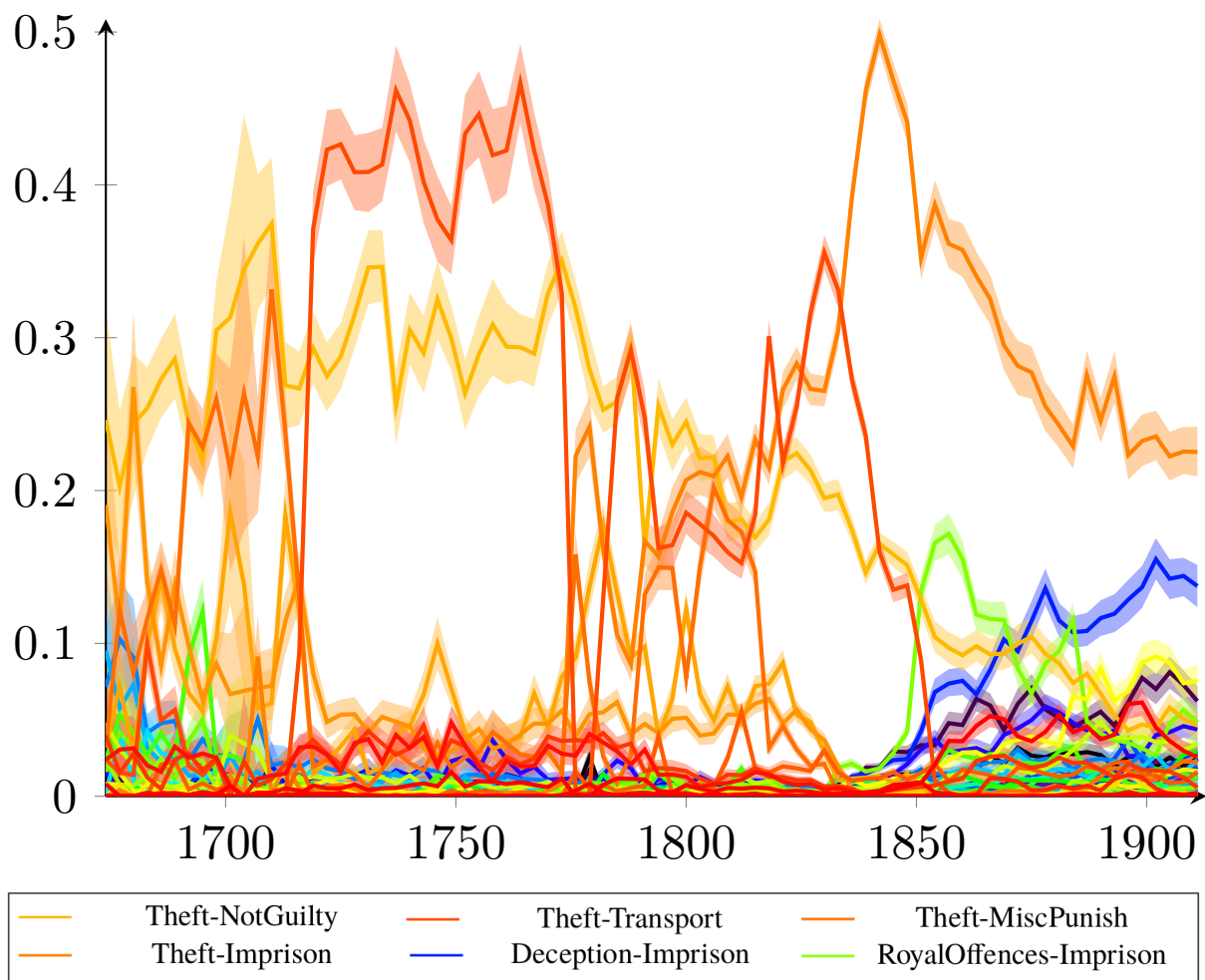


Fig. 7: Plot of offence-punishment pair probability estimates for the best model (Dep 3yr), with the dominant lines in the legend.

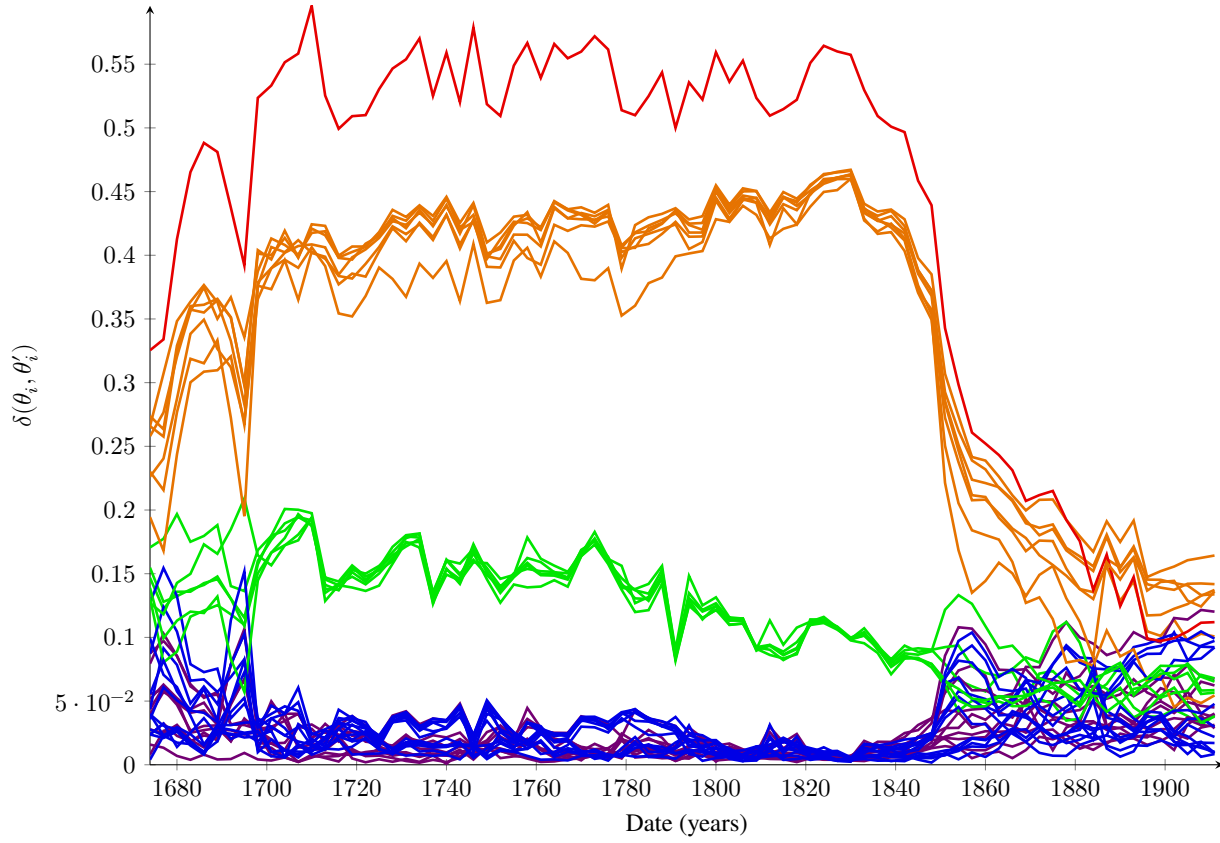


Fig. 8: Plot of $\delta(\cdot, \cdot, t)$ between punishment slices for all possible pairs of offences from dependent estimates with $\Delta t = 3$. The orange lines are Theft compared to all other offences, the green lines are Sexual offences compared with all other offences, the red line is Theft compared with Sexual, the violet lines are for the offences of BreakingPeace and Deception with all remaining offences, and the blue lines contain the δ values for the remainder of the offence pairs.

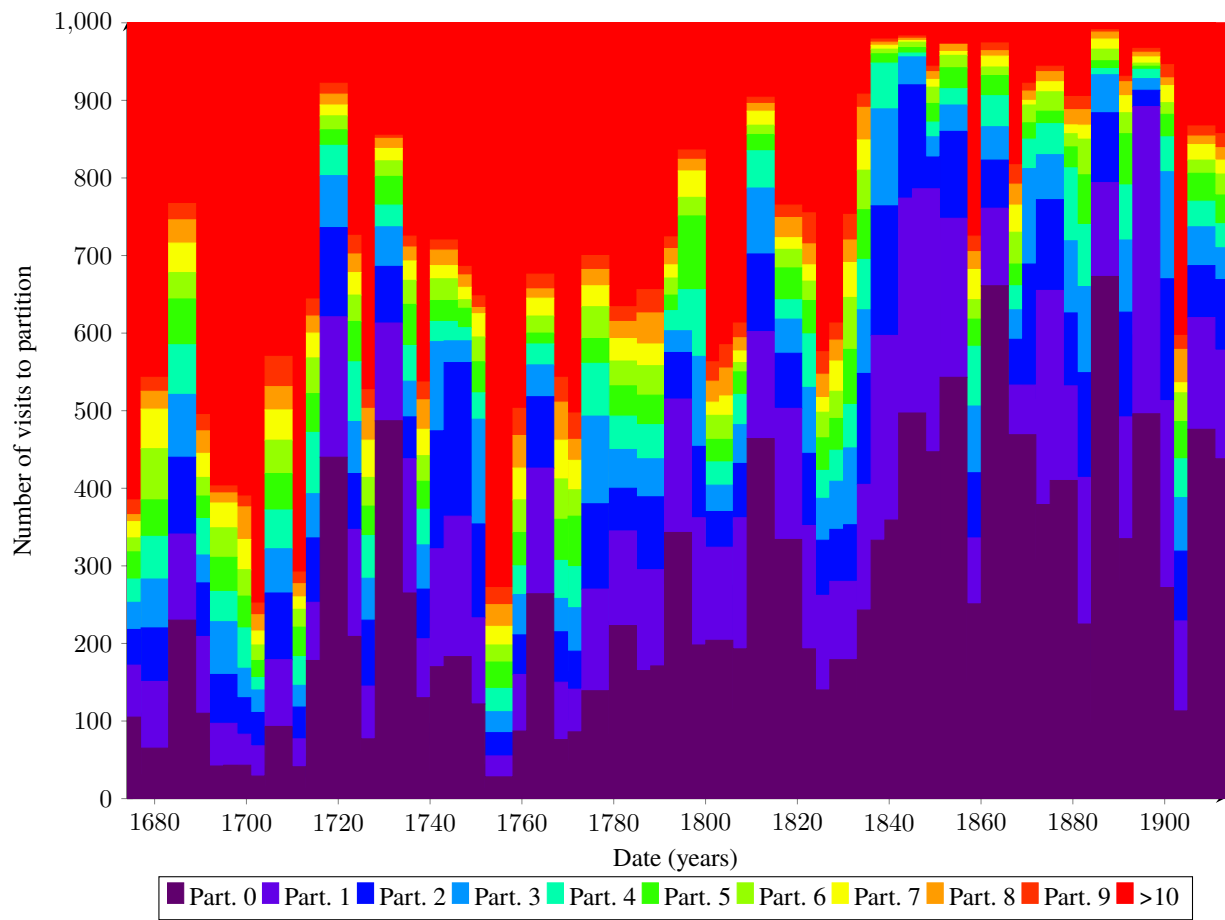


Fig. 10: Stacked bar plot of the number of times a partition is visited during 1,000 samples from the MCMC algorithm for the best ten partitions and the remaining.

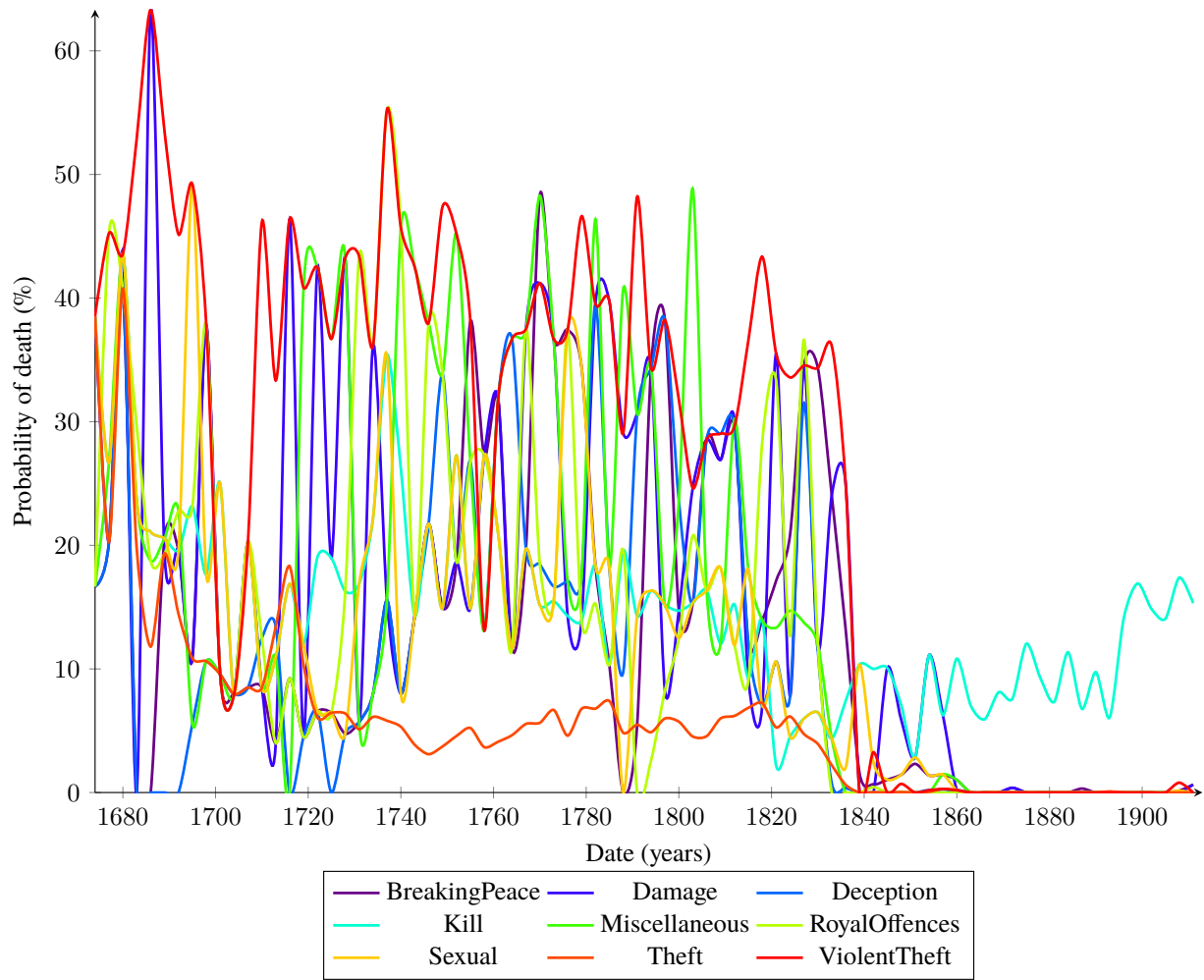


Fig. 11: The probability of receiving Death as a punishment for offences in the best partitioning by AIC.

Table S1: The partitions for the 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.

Date	Partition with blocks sorted by probability of death
1674	[[38.53, Damage, Kill, Sexual, Theft, ViolentTheft], [16.67, BreakingPeace, Deception, Miscellaneous, RoyalOffences]]
1677	[[45.24, RoyalOffences, ViolentTheft], [26.67, Kill, Miscellaneous, Sexual], [20.27, BreakingPeace, Damage, Deception, Theft]]
1680	[[43.59, Damage, Kill, Miscellaneous, Sexual, ViolentTheft], [40.76, BreakingPeace, Deception, RoyalOffences, Theft]]
1683	[[52.94, ViolentTheft], [29.58, RoyalOffences], [23.57, Kill, Miscellaneous, Sexual], [19.69, Theft], [0.00, BreakingPeace, Damage, Deception]]
1686	[[63.33, Damage, ViolentTheft], [21.09, Kill, Sexual], [18.75, Miscellaneous, RoyalOffences], [11.80, Theft], [0.00, BreakingPeace, Deception]]
1689	[[53.45, ViolentTheft], [20.57, BreakingPeace, Kill, Miscellaneous, Sexual], [19.33, Damage, RoyalOffences, Theft], [0.00, Deception]]
1692	[[45.16, ViolentTheft], [22.78, Miscellaneous, RoyalOffences], [19.59, BreakingPeace, Damage, Kill, Sexual], [14.46, Theft], [0.00, Deception]]
1695	[[49.21, Sexual, ViolentTheft], [23.15, Kill, RoyalOffences], [10.85, BreakingPeace, Damage, Theft], [5.56, Deception, Miscellaneous]]
1698	[[37.84, BreakingPeace, Damage, RoyalOffences, ViolentTheft], [17.65, Kill, Sexual], [10.62, Deception, Miscellaneous, Theft]]
1701	[[25.00, Kill, Sexual], [9.56, BreakingPeace, Damage, Deception, Miscellaneous, RoyalOffences, Theft, ViolentTheft]]
1704	[[7.95, BreakingPeace, Damage, Deception, Kill, Miscellaneous, RoyalOffences, Sexual, Theft, ViolentTheft]]
1707	[[20.27, Damage, Kill, Miscellaneous, RoyalOffences, Sexual, ViolentTheft], [8.49, BreakingPeace, Deception, Theft]]
1710	[[46.15, ViolentTheft], [12.50, Deception, RoyalOffences], [8.48, BreakingPeace, Damage, Kill, Miscellaneous, Sexual, Theft]]
1713	[[33.33, ViolentTheft], [13.41, Deception, Theft], [10.96, Kill, Miscellaneous, Sexual], [4.00, BreakingPeace, Damage, RoyalOffences]]
1716	[[46.39, Damage, ViolentTheft], [18.30, Theft], [16.88, Kill, Sexual], [9.23, BreakingPeace, RoyalOffences], [0.00, Deception, Miscellaneous]]
1719	[[40.85, Miscellaneous, ViolentTheft], [12.00, Kill, Sexual], [10.30, Theft], [4.48, BreakingPeace, Damage, Deception, RoyalOffences]]
1722	[[42.48, Damage, Miscellaneous, ViolentTheft], [19.05, Kill], [6.57, BreakingPeace, Deception, RoyalOffences, Sexual], [6.06, Theft]]
1725	[[36.70, Miscellaneous, ViolentTheft], [18.97, Damage, Kill], [6.45, BreakingPeace, RoyalOffences, Sexual], [6.45, Theft], [0.00, Deception]]
1728	[[43.24, Damage, Miscellaneous, ViolentTheft], [16.28, Kill, RoyalOffences], [6.39, Theft], [4.76, BreakingPeace, Deception, Sexual]]
1731	[[43.38, RoyalOffences, ViolentTheft], [16.90, Kill, Sexual], [5.71, BreakingPeace, Damage, Deception, Miscellaneous], [5.12, Theft]]
1734	[[36.19, Damage, RoyalOffences, ViolentTheft], [22.22, Kill, Sexual], [8.00, BreakingPeace, Deception, Miscellaneous], [6.12, Theft]]
1737	[[55.24, RoyalOffences, ViolentTheft], [35.38, Kill, Sexual], [15.56, BreakingPeace, Damage, Deception, Miscellaneous], [5.81, Theft]]
1740	[[45.71, Miscellaneous, RoyalOffences, ViolentTheft], [26.19, Kill], [8.00, BreakingPeace, Damage, Deception, Sexual], [5.32, Theft]]
1743	[[42.50, Miscellaneous, ViolentTheft], [14.29, BreakingPeace, Damage, Deception, RoyalOffences, Sexual], [14.29, Kill], [3.86, Theft]]
1746	[[38.00, Miscellaneous, RoyalOffences, ViolentTheft], [21.74, BreakingPeace, Damage, Deception, Kill, Sexual], [3.10, Theft]]
1749	[[47.40, ViolentTheft], [33.88, Deception, Miscellaneous, RoyalOffences], [14.93, BreakingPeace, Damage, Kill, Sexual], [3.73, Theft]]
1752	[[45.19, Miscellaneous, ViolentTheft], [27.27, Kill, Sexual], [18.67, BreakingPeace, Damage, Deception, RoyalOffences], [4.57, Theft]]
1755	[[38.10, BreakingPeace, ViolentTheft], [26.76, Deception, Miscellaneous, RoyalOffences], [14.89, Damage, Kill, Sexual], [5.22, Theft]]
1758	[[27.27, BreakingPeace, Damage, Kill, RoyalOffences, Sexual], [13.13, Deception, Miscellaneous, ViolentTheft], [3.67, Theft]]
1761	[[31.88, BreakingPeace, Damage, Deception, Miscellaneous, ViolentTheft], [20.97, Kill, RoyalOffences, Sexual], [4.11, Theft]]
1764	[[36.69, Deception, Miscellaneous, ViolentTheft], [11.67, BreakingPeace, Damage, Kill, RoyalOffences, Sexual], [4.64, Theft]]
1767	[[37.50, Damage, Miscellaneous, RoyalOffences, ViolentTheft], [19.70, BreakingPeace, Deception, Kill, Sexual], [5.55, Theft]]
1770	[[48.28, BreakingPeace, Miscellaneous], [41.21, Damage, ViolentTheft], [18.46, Deception, RoyalOffences], [15.15, Kill, Sexual], [5.64, Theft]]
1773	[[36.67, BreakingPeace, Damage, Miscellaneous, ViolentTheft], [16.67, Deception], [15.48, Kill, RoyalOffences, Sexual], [6.68, Theft]]
1776	[[37.44, BreakingPeace, RoyalOffences, Sexual, ViolentTheft], [17.11, Deception, Miscellaneous], [14.29, Damage, Kill], [4.60, Theft]]
1779	[[46.62, ViolentTheft], [34.44, BreakingPeace, Sexual], [17.50, Deception, Miscellaneous], [13.92, Damage, Kill, RoyalOffences], [6.77, Theft]]
1782	[[46.39, Miscellaneous], [39.52, Damage, Deception, ViolentTheft], [18.37, BreakingPeace, Kill, Sexual], [15.28, RoyalOffences], [6.79, Theft]]
1785	[[39.74, Damage, ViolentTheft], [18.18, Deception, Miscellaneous, Sexual], [10.34, BreakingPeace, Kill, RoyalOffences], [7.41, Theft]]
1788	[[40.74, Miscellaneous], [29.10, Damage, ViolentTheft], [19.57, Kill, RoyalOffences], [9.76, Deception], [4.83, Theft], [0.00, BreakingPeace, Sexual]]
1791	[[48.20, ViolentTheft], [30.65, Damage, Deception, Miscellaneous], [14.29, Kill, Sexual], [5.48, BreakingPeace, Theft], [0.00, RoyalOffences]]
1794	[[34.36, BreakingPeace, Damage, Deception, Miscellaneous, ViolentTheft], [16.33, Kill, Sexual], [4.88, Theft], [2.59, RoyalOffences]]
1797	[[38.22, BreakingPeace, Deception, ViolentTheft], [15.18, Kill, Miscellaneous, Sexual], [8.24, Damage, RoyalOffences], [5.99, Theft]]
1800	[[31.78, ViolentTheft], [23.61, Deception, Miscellaneous], [14.71, BreakingPeace, Damage, Kill], [12.50, RoyalOffences, Sexual], [5.73, Theft]]
1803	[[48.84, Miscellaneous], [24.64, Damage, ViolentTheft], [20.78, RoyalOffences], [15.33, BreakingPeace, Deception, Kill, Sexual], [4.53, Theft]]
1806	[[28.57, BreakingPeace, Damage, Deception, ViolentTheft], [16.35, Kill, Miscellaneous, RoyalOffences, Sexual], [4.56, Theft]]
1809	[[29.03, Deception, ViolentTheft], [26.92, BreakingPeace, Damage], [18.18, RoyalOffences, Sexual], [12.04, Kill, Miscellaneous], [5.99, Theft]]
1812	[[30.19, Damage, Miscellaneous], [29.61, BreakingPeace, Deception, ViolentTheft], [15.22, Kill], [11.92, RoyalOffences, Sexual], [6.20, Theft]]
1815	[[36.56, ViolentTheft], [18.06, Miscellaneous, Sexual], [12.07, BreakingPeace, Deception], [9.26, Damage, Kill], [9.03, RoyalOffences], [6.85, Theft]]
1818	[[43.35, ViolentTheft], [28.38, RoyalOffences], [13.95, BreakingPeace, Kill, Miscellaneous], [7.23, Theft], [7.04, Damage, Deception, Sexual]]
1821	[[35.51, Damage, ViolentTheft], [33.33, RoyalOffences], [17.14, BreakingPeace], [13.33, Miscellaneous], [10.58, Deception, Sexual], [5.27, Theft], [2.17, Kill]]
1824	[[33.58, ViolentTheft], [20.83, BreakingPeace], [14.71, Miscellaneous], [12.68, RoyalOffences], [7.79, Damage, Deception], [6.15, Theft], [4.50, Kill, Sexual]]
1827	[[36.62, RoyalOffences], [34.54, BreakingPeace, Damage, ViolentTheft], [31.51, Deception], [13.75, Miscellaneous], [5.98, Kill, Sexual], [4.73, Theft]]
1830	[[34.35, BreakingPeace, ViolentTheft], [11.97, Miscellaneous, RoyalOffences], [11.54, Damage, Deception], [6.48, Kill, Sexual], [3.93, Theft]]
1833	[[36.14, ViolentTheft], [24.05, BreakingPeace, Damage], [4.37, Kill, Miscellaneous, Sexual], [2.21, Theft], [0.75, Deception], [0.00, RoyalOffences]]
1836	[[25.00, Damage, ViolentTheft], [13.22, BreakingPeace], [7.41, Kill], [2.00, Sexual], [0.63, Theft], [0.37, Deception], [0.00, Miscellaneous], [0.00, RoyalOffences]]
1839	[[10.36, Kill, Sexual], [1.56, BreakingPeace], [0.00, Damage, Miscellaneous], [0.00, Deception, ViolentTheft], [0.00, RoyalOffences], [0.00, Theft]]
1842	[[10.00, Kill], [3.26, ViolentTheft], [2.45, Sexual], [0.66, BreakingPeace], [0.44, RoyalOffences], [0.03, Theft], [0.00, Damage, Miscellaneous], [0.00, Deception]]
1845	[[10.08, Damage, Kill], [1.04, BreakingPeace, Sexual], [0.00, Deception], [0.00, Miscellaneous], [0.00, RoyalOffences], [0.00, Theft], [0.00, ViolentTheft]]
1848	[[7.07, Kill], [5.88, Damage], [1.44, BreakingPeace, Sexual], [0.70, ViolentTheft], [0.04, Theft], [0.00, Deception, Miscellaneous], [0.00, RoyalOffences]]
1851	[[2.83, Damage, Kill, Sexual], [2.32, BreakingPeace], [0.00, Deception], [0.00, Miscellaneous], [0.00, RoyalOffences], [0.00, Theft], [0.00, ViolentTheft]]
1854	[[11.11, Damage, Kill], [1.38, BreakingPeace, Sexual], [0.19, Deception, ViolentTheft], [0.05, Theft], [0.00, Miscellaneous], [0.00, RoyalOffences]]
1857	[[25.5, Damage, Kill], [1.45, BreakingPeace, Miscellaneous, Sexual], [0.29, Theft, ViolentTheft], [0.00, Deception], [0.00, RoyalOffences]]
1860	[[10.83, Kill], [1.01, BreakingPeace, Miscellaneous], [0.19, Damage, Deception, Sexual], [0.19, Theft, ViolentTheft], [0.00, RoyalOffences]]
1863	[[7.04, Kill], [0.00, BreakingPeace], [0.00, Damage, Sexual, ViolentTheft], [0.00, Deception], [0.00, Miscellaneous], [0.00, RoyalOffences, Theft]]
1866	[[5.93, Kill], [0.00, BreakingPeace, Damage], [0.00, Deception, Sexual], [0.00, Miscellaneous], [0.00, RoyalOffences, Theft], [0.00, ViolentTheft]]
1869	[[8.11, Kill], [0.00, BreakingPeace], [0.00, Damage, Sexual], [0.00, Deception, Theft], [0.00, Miscellaneous], [0.00, RoyalOffences], [0.00, ViolentTheft]]
1872	[[7.59, Kill], [0.38, BreakingPeace, Damage], [0.00, Deception, Sexual, Theft], [0.00, Miscellaneous], [0.00, RoyalOffences], [0.00, ViolentTheft]]
1875	[[12.03, Kill], [0.00, BreakingPeace, Deception, Sexual], [0.00, Damage, ViolentTheft], [0.00, Miscellaneous], [0.00, RoyalOffences, Theft]]
1878	[[9.29, Kill], [0.00, BreakingPeace, Damage], [0.00, Deception, Sexual, Theft], [0.00, Miscellaneous], [0.00, RoyalOffences], [0.00, ViolentTheft]]
1881	[[7.33, Kill], [0.00, BreakingPeace, Sexual], [0.00, Damage, Miscellaneous], [0.00, Deception], [0.00, RoyalOffences, Theft], [0.00, ViolentTheft]]
1884	[[11.35, Kill], [0.00, BreakingPeace], [0.00, Damage, Miscellaneous], [0.00, Deception, RoyalOffences], [0.00, Sexual, Theft], [0.00, ViolentTheft]]
1887	[[6.77, Kill], [0.32, BreakingPeace], [0.00, Damage, Sexual], [0.00, Deception, RoyalOffences], [0.00, Miscellaneous], [0.00, Theft], [0.00, ViolentTheft]]
1890	[[9.74, Kill], [0.00, BreakingPeace], [0.00, Damage, Sexual], [0.00, Deception, RoyalOffences, Theft], [0.00, Miscellaneous], [0.00, ViolentTheft]]
1893	[[6.08, Kill], [0.09, Theft], [0.00, BreakingPeace], [0.00, Damage, Sexual], [0.00, Deception], [0.00, Miscellaneous], [0.00, RoyalOffences], [0.00, ViolentTheft]]
1896	[[14.06, Kill], [0.00, BreakingPeace], [0.00, Damage, Miscellaneous, Sexual], [0.00, Deception], [0.00, RoyalOffences, Theft], [0.00, ViolentTheft]]
1899	[[16.91, Kill], [0.00, BreakingPeace, Miscellaneous], [0.00, Damage, Sexual], [0.00, Deception], [0.00, RoyalOffences, Theft], [0.00, ViolentTheft]]
1902	[[14.86, Kill], [0.00, BreakingPeace], [0.00, Damage, RoyalOffences, Theft], [0.00, Deception], [0.00, Miscellaneous, Sexual], [0.00, ViolentTheft]]
1905	[[14.05, Kill], [0.00, BreakingPeace, Damage, Miscellaneous, Sexual], [0.00, Deception], [0.00, RoyalOffences, Theft], [0.00, ViolentTheft]]
1908	[[17.39, Kill], [0.80, ViolentTheft], [0.15, BreakingPeace, Sexual], [0.00, Damage, Deception, Miscellaneous], [0.00, RoyalOffences, Theft]]
1911	[[15.32, Kill], [0.64, BreakingPeace, Damage], [0.00, Deception], [0.00, Miscellaneous, Sexual], [0.00, RoyalOffences, Theft, ViolentTheft]]

Table S2: The partitions for the best offence-partitioned model that can vary every three years, with each block preceded by the probability of imprisonment within that block as a percentage.

Date	Partition with blocks sorted by probability of imprisonment
1674	[0.79, BreakingPeace, Deception, Miscellaneous, RoyalOffences], [0.13, Damage, Kill, Sexual, Theft, ViolentTheft]]
1677	[1.52, Kill, Miscellaneous, Sexual], [0.77, BreakingPeace, Damage, Deception, Theft], [0.34, RoyalOffences, ViolentTheft]]
1680	[0.18, Damage, Kill, Miscellaneous, Sexual, ViolentTheft], [0.07, BreakingPeace, Deception, RoyalOffences, Theft]]
1683	[6.02, BreakingPeace, Damage, Deception], [3.02, RoyalOffences], [0.82, Kill, Miscellaneous, Sexual], [0.42, ViolentTheft], [0.02, Theft]]
1686	[10.48, BreakingPeace, Deception], [2.81, Miscellaneous, RoyalOffences], [0.48, Damage, ViolentTheft], [0.41, Theft], [0.11, Kill, Sexual]]
1689	[8.16, Deception], [0.81, BreakingPeace, Kill, Miscellaneous, Sexual], [0.25, ViolentTheft], [0.02, Damage, RoyalOffences, Theft]]
1692	[1.18, BreakingPeace, Damage, Kill, Sexual], [0.68, Deception], [0.23, ViolentTheft], [0.08, Miscellaneous, RoyalOffences], [0.02, Theft]]
1695	[0.38, Kill, RoyalOffences], [0.23, Sexual, ViolentTheft], [0.20, Deception, Miscellaneous], [0.18, BreakingPeace, Damage, Theft]]
1698	[2.90, BreakingPeace, Damage, RoyalOffences, ViolentTheft], [0.28, Kill, Sexual], [0.19, Deception, Miscellaneous, Theft]]
1701	[1.58, BreakingPeace, Damage, Deception, Miscellaneous, RoyalOffences, Theft, ViolentTheft], [0.89, Kill, Sexual]]
1704	[1.30, BreakingPeace, Damage, Deception, Kill, Miscellaneous, RoyalOffences, Sexual, Theft, ViolentTheft]]
1707	[11.00, BreakingPeace, Deception, Theft], [5.60, Damage, Kill, Miscellaneous, RoyalOffences, Sexual, ViolentTheft]]
1710	[1.10, ViolentTheft], [0.45, Deception, RoyalOffences], [0.04, BreakingPeace, Damage, Kill, Miscellaneous, Sexual, Theft]]
1713	[12.57, BreakingPeace, Damage, RoyalOffences], [2.72, ViolentTheft], [2.60, Deception, Theft], [1.57, Kill, Miscellaneous, Sexual]]
1716	[27.91, BreakingPeace, RoyalOffences], [24.14, Deception, Miscellaneous], [2.21, Damage, ViolentTheft], [0.67, Theft], [0.19, Kill, Sexual]]
1719	[16.63, BreakingPeace, Damage, Deception, RoyalOffences], [3.62, Miscellaneous, ViolentTheft], [0.37, Theft], [0.14, Kill, Sexual]]
1722	[15.43, BreakingPeace, Deception, RoyalOffences, Sexual], [4.55, Damage, Miscellaneous, ViolentTheft], [0.82, Theft], [0.23, Kill]]
1725	[21.66, BreakingPeace, RoyalOffences, Sexual], [18.10, Deception], [6.55, Miscellaneous, ViolentTheft], [0.91, Theft], [0.25, Damage, Kill]]
1728	[6.12, BreakingPeace, Deception, Sexual], [2.66, Kill, RoyalOffences], [2.77, Damage, Miscellaneous, ViolentTheft], [0.57, Theft]]
1731	[18.78, BreakingPeace, Damage, Deception, Miscellaneous], [0.84, RoyalOffences, ViolentTheft], [0.25, Theft], [0.20, Kill, Sexual]]
1734	[18.29, BreakingPeace, Deception, Miscellaneous], [0.18, Kill, Sexual], [0.14, Damage, RoyalOffences, ViolentTheft], [0.09, Theft]]
1737	[6.98, BreakingPeace, Damage, Deception, Miscellaneous], [1.09, RoyalOffences, ViolentTheft], [0.22, Kill, Sexual], [0.01, Theft]]
1740	[2.29, BreakingPeace, Damage, Deception, Sexual], [0.34, Kill], [0.14, Miscellaneous, RoyalOffences, ViolentTheft], [0.01, Theft]]
1743	[11.63, BreakingPeace, Damage, Deception, RoyalOffences, Sexual], [0.34, Kill], [0.12, Miscellaneous, ViolentTheft], [0.01, Theft]]
1746	[1.11, BreakingPeace, Damage, Deception, Kill, Sexual], [0.14, Miscellaneous, RoyalOffences, ViolentTheft], [0.01, Theft]]
1749	[4.69, BreakingPeace, Damage, Kill, Sexual], [0.94, Deception, Miscellaneous, RoyalOffences], [0.46, Theft], [0.08, ViolentTheft]]
1752	[8.19, BreakingPeace, Damage, Deception, RoyalOffences], [6.68, Kill, Sexual], [1.59, Miscellaneous, ViolentTheft], [0.10, Theft]]
1755	[12.88, Deception, Miscellaneous, RoyalOffences], [2.43, Damage, Kill, Sexual], [0.23, BreakingPeace, ViolentTheft], [0.01, Theft]]
1758	[4.87, BreakingPeace, Damage, Kill, RoyalOffences, Sexual], [4.18, Deception, Miscellaneous, ViolentTheft], [0.27, Theft]]
1761	[13.13, Kill, RoyalOffences, Sexual], [1.55, BreakingPeace, Damage, Deception, Miscellaneous, ViolentTheft], [0.34, Theft]]
1764	[11.90, BreakingPeace, Damage, Kill, RoyalOffences, Sexual], [1.54, Deception, Miscellaneous, ViolentTheft], [0.24, Theft]]
1767	[10.71, BreakingPeace, Deception, Kill, Sexual], [0.14, Damage, Miscellaneous, RoyalOffences, ViolentTheft], [0.08, Theft]]
1770	[7.91, Deception, RoyalOffences], [7.39, BreakingPeace, Miscellaneous], [6.28, Kill, Sexual], [0.42, Theft], [0.09, Damage, ViolentTheft]]
1773	[14.46, Kill, RoyalOffences, Sexual], [13.57, Deception], [1.02, BreakingPeace, Damage, Miscellaneous, ViolentTheft], [0.95, Theft]]
1776	[26.50, Deception, Miscellaneous], [26.23, Theft], [14.58, Damage, Kill], [5.49, BreakingPeace, RoyalOffences, Sexual, ViolentTheft]]
1779	[30.52, Theft], [27.86, Deception, Miscellaneous], [21.70, Damage, Kill, RoyalOffences], [9.56, ViolentTheft], [3.49, BreakingPeace, Sexual]]
1782	[34.92, RoyalOffences], [19.89, Theft], [6.41, BreakingPeace, Kill, Sexual], [4.27, Miscellaneous], [3.00, Damage, Deception, ViolentTheft]]
1785	[22.54, BreakingPeace, Kill, RoyalOffences], [12.73, Theft], [10.88, Deception, Miscellaneous, Sexual], [0.92, Damage, ViolentTheft]]
1788	[27.33, Kill, RoyalOffences], [18.61, BreakingPeace, Sexual], [13.59, Deception], [10.33, Theft], [7.94, Miscellaneous], [2.72, Damage, ViolentTheft]]
1791	[51.90, RoyalOffences], [24.15, Kill, Sexual], [19.39, BreakingPeace, Theft], [17.97, Damage, Deception, Miscellaneous], [3.70, ViolentTheft]]
1794	[33.74, RoyalOffences], [20.70, Kill, Sexual], [18.60, Theft], [5.20, BreakingPeace, Damage, Deception, Miscellaneous, ViolentTheft]]
1797	[37.82, Damage, RoyalOffences], [21.99, Theft], [18.88, Kill, Miscellaneous, Sexual], [4.55, BreakingPeace, Deception, ViolentTheft]]
1800	[28.98, RoyalOffences, Sexual], [23.08, Theft], [15.48, Deception, Miscellaneous], [6.30, BreakingPeace, Damage, Kill], [5.54, ViolentTheft]]
1803	[30.06, RoyalOffences], [24.22, Theft], [16.76, BreakingPeace, Deception, Kill, Sexual], [11.96, Miscellaneous], [1.66, Damage, ViolentTheft]]
1806	[23.99, Kill, Miscellaneous, RoyalOffences, Sexual], [23.24, Theft], [7.63, BreakingPeace, Damage, Deception, ViolentTheft]]
1809	[26.98, Kill, Miscellaneous], [24.94, Theft], [15.58, RoyalOffences, Sexual], [4.40, BreakingPeace, Damage], [3.32, Deception, ViolentTheft]]
1812	[22.75, Theft], [17.25, Damage, Miscellaneous], [15.53, Kill], [15.05, RoyalOffences, Sexual], [3.49, BreakingPeace, Deception, ViolentTheft]]
1815	[28.04, Damage, Kill], [26.40, Theft], [22.32, RoyalOffences], [19.64, Miscellaneous, Sexual], [6.16, BreakingPeace, Deception], [4.03, ViolentTheft]]
1818	[24.38, Theft], [21.10, BreakingPeace, Kill, Miscellaneous], [15.83, RoyalOffences], [12.78, Damage, Deception, Sexual], [1.24, ViolentTheft]]
1821	[35.09, Kill], [32.38, Miscellaneous], [29.43, Theft], [23.26, RoyalOffences], [22.25, Deception, Sexual], [8.98, BreakingPeace], [3.00, Damage, ViolentTheft]]
1824	[41.05, RoyalOffences], [30.73, Theft], [28.76, Damage, Deception], [25.21, Miscellaneous], [21.75, Kill, Sexual], [1.60, ViolentTheft], [0.60, BreakingPeace]]
1827	[28.82, Theft], [25.76, Kill, Sexual], [23.93, Miscellaneous], [18.51, RoyalOffences], [13.89, Deception], [1.10, BreakingPeace, Damage, ViolentTheft]]
1830	[28.44, Theft], [25.45, Miscellaneous, RoyalOffences], [23.28, Kill, Sexual], [14.56, Damage, Deception], [2.40, BreakingPeace, ViolentTheft]]
1833	[66.96, RoyalOffences], [34.54, Theft], [29.76, Kill, Miscellaneous, Sexual], [16.52, Deception], [6.51, BreakingPeace, Damage], [2.58, ViolentTheft]]
1836	[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]]
1839	[76.42, RoyalOffences], [63.73, BreakingPeace], [53.20, Theft], [44.63, Kill, Sexual], [42.45, Deception, ViolentTheft], [36.10, Damage, Miscellaneous]]
1842	[75.17, RoyalOffences], [58.72, Theft], [57.52, BreakingPeace], [48.55, Sexual], [44.68, Deception], [35.18, Kill], [29.51, Damage, Miscellaneous], [25.16, ViolentTheft]]
1845	[73.92, RoyalOffences], [59.45, Theft], [53.07, BreakingPeace, Sexual], [46.35, Deception], [32.47, Miscellaneous], [28.22, ViolentTheft], [21.04, Damage, Kill]]
1848	[77.96, RoyalOffences], [58.29, Theft], [51.91, BreakingPeace, Sexual], [39.20, Deception, Miscellaneous], [32.97, ViolentTheft], [28.43, Kill], [0.84, Damage]]
1851	[87.01, RoyalOffences], [59.43, Theft], [55.27, BreakingPeace], [43.57, Deception], [42.16, Damage, Kill, Sexual], [35.37, ViolentTheft], [21.20, Miscellaneous]]
1854	[87.39, RoyalOffences], [73.80, Theft], [61.64, BreakingPeace, Sexual], [58.16, Deception, ViolentTheft], [29.56, Miscellaneous], [27.89, Damage, Kill]]
1857	[83.33, RoyalOffences], [76.43, Theft, ViolentTheft], [67.26, Deception], [63.65, BreakingPeace, Miscellaneous, Sexual], [43.08, Damage, Kill]]
1860	[84.47, RoyalOffences], [76.08, Theft, ViolentTheft], [66.12, Damage, Deception, Sexual], [63.22, BreakingPeace, Miscellaneous], [45.95, Kill]]
1863	[76.07, RoyalOffences, Theft], [69.06, BreakingPeace], [64.96, Damage, Sexual, ViolentTheft], [62.47, Deception], [43.76, Kill], [38.62, Miscellaneous]]
1866	[76.11, RoyalOffences, Theft], [68.03, Deception, Sexual], [61.78, BreakingPeace, Damage], [57.64, ViolentTheft], [38.62, Kill], [33.53, Miscellaneous]]
1869	[79.69, RoyalOffences], [71.57, Deception, Theft], [57.21, Damage, Sexual], [56.86, ViolentTheft], [56.03, BreakingPeace], [34.36, Kill], [24.84, Miscellaneous]]
1872	[75.65, RoyalOffences], [67.63, Deception, Sexual, Theft], [58.01, BreakingPeace, Damage], [55.30, ViolentTheft], [42.86, Kill], [33.07, Miscellaneous]]
1875	[69.52, RoyalOffences, Theft], [64.34, BreakingPeace, Deception, Sexual], [61.21, Damage, ViolentTheft], [40.71, Kill], [36.82, Miscellaneous]]
1878	[79.98, RoyalOffences], [67.39, Deception, Sexual, Theft], [63.02, ViolentTheft], [57.97, BreakingPeace, Damage], [37.68, Miscellaneous], [35.10, Kill]]
1881	[70.77, RoyalOffences, Theft], [61.96, ViolentTheft], [60.45, Deception], [55.15, BreakingPeace, Sexual], [40.49, Damage, Miscellaneous], [38.10, Kill]]
1884	[69.59, Sexual, Theft], [66.19, Deception, RoyalOffences], [58.91, ViolentTheft], [48.48, BreakingPeace], [47.70, Damage, Miscellaneous], [40.53, Kill]]
1887	[72.18, Theft], [63.35, Deception, RoyalOffences], [59.57, ViolentTheft], [55.99, Damage, Sexual], [51.81, BreakingPeace], [33.89, Miscellaneous], [30.93, Kill]]
1890	[68.35, Deception, RoyalOffences, Theft], [61.96, ViolentTheft], [56.76, Damage, Sexual], [53.94, BreakingPeace], [34.51, Kill], [31.56, Miscellaneous]]
1893	[74.11, Theft], [64.54, RoyalOffences], [60.70, ViolentTheft], [58.15, Deception], [53.76, Damage, Sexual], [51.99, BreakingPeace], [44.69, Kill], [35.05, Miscellaneous]]
1896	[72.48, RoyalOffences, Theft], [62.28, Deception], [61.03, ViolentTheft], [54.11, BreakingPeace], [52.79, Damage, Miscellaneous, Sexual], [39.96, Kill]]
1899	[76.37, RoyalOffences, Theft], [67.32, Deception], [65.97, ViolentTheft], [62.71, Damage, Sexual], [61.66, BreakingPeace, Miscellaneous], [33.19, Kill]]
1902	[76.83, Damage, RoyalOffences, Theft], [69.01, Deception], [64.21, ViolentTheft], [61.62, Miscellaneous, Sexual], [59.98, BreakingPeace], [47.39, Kill]]
1905	[71.87, RoyalOffences, Theft], [65.63, Deception], [64.04, ViolentTheft], [62.83, BreakingPeace, Damage, Miscellaneous, Sexual], [36.48, Kill]]
1908	[72.93, RoyalOffences, Theft], [66.51, ViolentTheft], [62.72, Damage, Deception, Miscellaneous], [61.90, BreakingPeace, Sexual], [37.52, Kill]]
1911	[72.86, RoyalOffences, Theft, ViolentTheft], [59.80, Deception], [59.16, Miscellaneous, Sexual], [54.50, BreakingPeace, Damage], [49.68, Kill]]

Table S3: The partitions for the best offence-partitioned model that can vary every three years, with each block preceded by the probability of transportation within that block as a percentage.

Date	Partition with blocks sorted by probability of transportation
1674	[2.88, Damage, Kill, Sexual, Theft, ViolentTheft], [0.79, BreakingPeace, Deception, Miscellaneous, RoyalOffences]]
1677	[2.80, BreakingPeace, Damage, Deception, Theft], [0.34, RoyalOffences, ViolentTheft], [0.19, Kill, Miscellaneous, Sexual]]
1680	[1.96, BreakingPeace, Deception, RoyalOffences, Theft], [0.18, Damage, Kill, Miscellaneous, Sexual, ViolentTheft]]
1683	[14.04, Theft], [6.30, ViolentTheft], [2.96, Kill, Miscellaneous, Sexual], [0.75, BreakingPeace, Damage, Deception], [0.20, RoyalOffences]]
1686	[6.89, Theft], [3.81, Damage, ViolentTheft], [0.89, Kill, Sexual], [0.48, BreakingPeace, Deception], [0.13, Miscellaneous, RoyalOffences]]
1689	[7.40, Damage, RoyalOffences, Theft], [1.02, Deception], [0.25, ViolentTheft], [0.10, BreakingPeace, Kill, Miscellaneous, Sexual]]
1692	[2.94, Theft], [1.84, ViolentTheft], [0.68, Deception], [0.15, BreakingPeace, Damage, Kill, Sexual], [0.08, Miscellaneous, RoyalOffences]]
1695	[4.99, Sexual, ViolentTheft], [2.35, BreakingPeace, Damage, Theft], [0.20, Deception, Miscellaneous], [0.05, Kill, RoyalOffences]]
1698	[2.89, Deception, Miscellaneous, Theft], [1.54, BreakingPeace, Damage, RoyalOffences, ViolentTheft], [0.28, Kill, Sexual]]
1701	[0.89, Kill, Sexual], [0.11, BreakingPeace, Damage, Deception, Miscellaneous, RoyalOffences, Theft, ViolentTheft]]
1704	[0.16, BreakingPeace, Damage, Deception, Kill, Miscellaneous, RoyalOffences, Sexual, Theft, ViolentTheft]]
1707	[0.19, Damage, Kill, Miscellaneous, RoyalOffences, Sexual, ViolentTheft], [0.04, BreakingPeace, Deception, Theft]]
1710	[1.10, ViolentTheft], [0.45, Deception, RoyalOffences], [0.04, BreakingPeace, Damage, Kill, Miscellaneous, Sexual, Theft]]
1713	[0.57, BreakingPeace, Damage, RoyalOffences], [0.34, ViolentTheft], [0.20, Kill, Miscellaneous, Sexual], [0.02, Deception, Theft]]
1716	[12.06, Theft], [2.21, Damage, ViolentTheft], [1.76, BreakingPeace, RoyalOffences], [0.20, Deception, Miscellaneous], [0.19, Kill, Sexual]]
1719	[47.35, Theft], [7.14, Miscellaneous, ViolentTheft], [3.20, BreakingPeace, Damage, Deception, RoyalOffences], [0.14, Kill, Sexual]]
1722	[52.88, Theft], [8.98, Damage, Miscellaneous, ViolentTheft], [1.56, BreakingPeace, Deception, RoyalOffences, Sexual], [0.23, Kill]]
1725	[51.54, Theft], [9.31, Miscellaneous, ViolentTheft], [2.54, Deception], [1.23, BreakingPeace, RoyalOffences, Sexual], [0.25, Damage, Kill]]
1728	[48.61, Theft], [8.88, Damage, Miscellaneous, ViolentTheft], [2.66, Kill, RoyalOffences], [0.17, BreakingPeace, Deception, Sexual]]
1731	[49.77, Theft], [4.52, RoyalOffences, ViolentTheft], [3.06, BreakingPeace, Damage, Deception, Miscellaneous], [1.61, Kill, Sexual]]
1734	[48.67, Theft], [7.76, Damage, RoyalOffences, ViolentTheft], [2.29, BreakingPeace, Deception, Miscellaneous], [0.18, Kill, Sexual]]
1737	[55.18, Theft], [6.80, RoyalOffences, ViolentTheft], [0.32, BreakingPeace, Damage, Deception, Miscellaneous], [0.22, Kill, Sexual]]
1740	[51.30, Theft], [10.61, Miscellaneous, RoyalOffences, ViolentTheft], [0.34, Kill], [0.29, BreakingPeace, Damage, Deception, Sexual]]
1743	[49.75, Theft], [8.45, Miscellaneous, ViolentTheft], [2.72, Kill], [0.20, BreakingPeace, Damage, Deception, RoyalOffences, Sexual]]
1746	[43.62, Theft], [14.14, Miscellaneous, RoyalOffences, ViolentTheft], [1.66, BreakingPeace, Damage, Deception, Kill, Sexual]]
1749	[46.09, Theft], [15.69, ViolentTheft], [2.60, Deception, Miscellaneous, RoyalOffences], [0.21, BreakingPeace, Damage, Kill, Sexual]]
1752	[53.94, Theft], [6.86, BreakingPeace, Damage, Deception, RoyalOffences], [6.77, Miscellaneous, ViolentTheft], [0.19, Kill, Sexual]]
1755	[52.89, Theft], [17.69, BreakingPeace, ViolentTheft], [6.69, Damage, Kill, Sexual], [4.43, Deception, Miscellaneous, RoyalOffences]]
1758	[49.39, Theft], [23.38, Deception, Miscellaneous, ViolentTheft], [0.32, BreakingPeace, Damage, Kill, RoyalOffences, Sexual]]
1761	[51.26, Theft], [14.60, BreakingPeace, Damage, Deception, Miscellaneous, ViolentTheft], [0.23, Kill, RoyalOffences, Sexual]]
1764	[53.53, Theft], [16.65, Deception, Miscellaneous, ViolentTheft], [0.24, BreakingPeace, Damage, Kill, RoyalOffences, Sexual]]
1767	[49.33, Theft], [12.64, Damage, Miscellaneous, RoyalOffences, ViolentTheft], [3.90, BreakingPeace, Deception, Kill, Sexual]]
1770	[45.98, Theft], [18.68, Deception, RoyalOffences], [7.97, Damage, ViolentTheft], [3.94, BreakingPeace, Miscellaneous], [0.22, Kill, Sexual]]
1773	[39.27, Theft], [20.24, Deception], [12.45, BreakingPeace, Damage, Miscellaneous, ViolentTheft], [2.55, Kill, RoyalOffences, Sexual]]
1776	[1.50, Deception, Miscellaneous], [0.34, Theft], [0.29, Damage, Kill], [0.07, BreakingPeace, RoyalOffences, Sexual, ViolentTheft]]
1779	[0.36, Deception, Miscellaneous], [0.18, Damage, Kill, RoyalOffences], [0.16, BreakingPeace, Sexual], [0.10, ViolentTheft], [0.09, Theft]]
1782	[19.73, Miscellaneous], [16.04, Theft], [7.57, Damage, Deception, ViolentTheft], [0.29, BreakingPeace, Kill, Sexual], [0.20, RoyalOffences]]
1785	[31.34, Theft], [15.87, Damage, ViolentTheft], [3.49, Deception, Miscellaneous, Sexual], [0.99, BreakingPeace, Kill, RoyalOffences]]
1788	[34.90, Theft], [15.42, Damage, ViolentTheft], [6.27, Deception], [0.53, Miscellaneous], [0.43, BreakingPeace, Sexual], [0.16, Kill, RoyalOffences]]
1791	[28.71, BreakingPeace, Theft], [17.37, ViolentTheft], [9.91, Damage, Deception, Miscellaneous], [0.34, Kill, Sexual], [0.24, RoyalOffences]]
1794	[19.22, Theft], [8.28, BreakingPeace, Damage, Deception, Miscellaneous, ViolentTheft], [2.33, Kill, Sexual], [0.12, RoyalOffences]]
1797	[19.43, Theft], [6.46, BreakingPeace, Deception, ViolentTheft], [4.59, Kill, Miscellaneous, Sexual], [0.17, Damage, RoyalOffences]]
1800	[20.70, Theft], [17.94, ViolentTheft], [8.53, Deception, Miscellaneous], [3.98, RoyalOffences, Sexual], [0.42, BreakingPeace, Damage, Kill]]
1803	[20.33, Theft], [19.05, Damage, ViolentTheft], [4.76, BreakingPeace, Deception, Kill, Sexual], [2.78, RoyalOffences], [2.66, Miscellaneous]]
1806	[18.90, Theft], [11.39, BreakingPeace, Damage, Deception, ViolentTheft], [7.01, Kill, Miscellaneous, RoyalOffences, Sexual]]
1809	[17.88, Theft], [17.51, Deception, ViolentTheft], [10.13, RoyalOffences, Sexual], [1.98, Kill, Miscellaneous], [0.55, BreakingPeace, Damage]]
1812	[25.05, RoyalOffences, Sexual], [17.75, Theft], [16.80, BreakingPeace, Deception, ViolentTheft], [2.16, Damage, Miscellaneous], [0.31, Kill]]
1815	[31.35, RoyalOffences], [20.79, Theft], [14.78, BreakingPeace, Deception], [14.09, Miscellaneous, Sexual], [13.28, ViolentTheft], [0.26, Damage, Kill]]
1818	[34.14, Theft], [23.49, RoyalOffences], [19.82, Damage, Deception, Sexual], [15.69, ViolentTheft], [3.65, BreakingPeace, Kill, Miscellaneous]]
1821	[30.91, Deception, Sexual], [24.12, Theft], [13.00, RoyalOffences], [11.70, Damage, ViolentTheft], [6.83, Miscellaneous], [3.27, BreakingPeace], [0.31, Kill]]
1824	[27.68, Theft], [16.52, ViolentTheft], [15.77, Damage, Deception], [8.24, Kill, Sexual], [7.56, Miscellaneous], [0.60, BreakingPeace], [0.20, RoyalOffences]]
1827	[34.18, Theft], [13.93, Miscellaneous], [12.52, Deception], [12.09, Kill, Sexual], [11.93, BreakingPeace, Damage, ViolentTheft], [8.65, RoyalOffences]]
1830	[38.13, Theft], [36.68, Damage, Deception], [21.93, Miscellaneous, RoyalOffences], [9.39, Kill, Sexual], [8.51, BreakingPeace, ViolentTheft]]
1833	[48.61, Deception], [37.61, Theft], [22.59, RoyalOffences], [13.43, ViolentTheft], [5.74, Kill, Miscellaneous, Sexual], [2.71, BreakingPeace, Damage]]
1836	[31.46, Theft], [31.16, Deception], [19.58, Damage, ViolentTheft], [18.97, RoyalOffences], [11.29, Kill], [7.37, Miscellaneous], [5.98, BreakingPeace], [4.76, Sexual]]
1839	[33.57, Deception, ViolentTheft], [27.22, Theft], [16.21, RoyalOffences], [7.87, BreakingPeace], [7.07, Damage, Miscellaneous], [3.70, Kill, Sexual]]
1842	[37.11, ViolentTheft], [20.44, Deception], [18.81, Theft], [14.29, RoyalOffences], [11.13, Sexual], [6.36, BreakingPeace], [5.98, Damage, Miscellaneous], [2.68, Kill]]
1845	[45.57, ViolentTheft], [21.35, Deception], [17.16, Theft], [10.50, RoyalOffences], [7.86, Damage, Kill], [7.68, BreakingPeace, Sexual], [6.20, Miscellaneous]]
1848	[65.55, Damage], [39.96, ViolentTheft], [18.93, Deception, Miscellaneous], [18.23, Theft], [13.28, Kill], [10.98, BreakingPeace, Sexual], [8.81, RoyalOffences]]
1851	[37.28, ViolentTheft], [17.98, Deception], [15.07, Theft], [14.75, Miscellaneous], [11.64, BreakingPeace], [10.18, Damage, Kill, Sexual], [3.96, RoyalOffences]]
1854	[6.76, Deception, ViolentTheft], [3.86, BreakingPeace, Sexual], [3.29, Damage, Kill], [1.97, Miscellaneous], [1.91, Theft], [0.17, RoyalOffences]]
1857	[1.24, BreakingPeace, Miscellaneous, Sexual], [0.77, Theft, ViolentTheft], [0.60, Deception], [0.32, RoyalOffences], [0.11, Damage, Kill]]
1860	[0.12, Kill], [0.05, BreakingPeace, Miscellaneous], [0.03, Damage, Deception, Sexual], [0.02, RoyalOffences], [0.01, Theft, ViolentTheft]]
1863	[0.16, Miscellaneous], [0.10, Kill], [0.05, BreakingPeace], [0.04, Deception], [0.03, Damage, Sexual, ViolentTheft], [0.01, RoyalOffences, Theft]]
1866	[0.20, Miscellaneous], [0.11, Kill], [0.05, ViolentTheft], [0.04, BreakingPeace, Damage], [0.03, Deception, Sexual], [0.01, RoyalOffences, Theft]]
1869	[0.21, Miscellaneous], [0.13, Kill], [0.07, BreakingPeace], [0.07, Damage, Sexual], [0.06, ViolentTheft], [0.04, RoyalOffences], [0.01, Deception, Theft]]
1872	[0.20, Miscellaneous], [0.10, Kill], [0.09, ViolentTheft], [0.06, RoyalOffences], [0.05, BreakingPeace, Damage], [0.01, Deception, Sexual, Theft]]
1875	[0.20, Miscellaneous], [0.11, Kill], [0.07, Damage, ViolentTheft], [0.02, BreakingPeace, Deception, Sexual], [0.01, RoyalOffences, Theft]]
1878	[0.18, Miscellaneous], [0.10, Kill], [0.07, ViolentTheft], [0.04, BreakingPeace, Damage], [0.04, RoyalOffences], [0.01, Deception, Sexual, Theft]]
1881	[0.10, Kill], [0.09, Damage, Miscellaneous], [0.05, ViolentTheft], [0.03, BreakingPeace, Sexual], [0.02, Deception], [0.01, RoyalOffences, Theft]]
1884	[0.12, Damage, Miscellaneous], [0.10, Kill], [0.06, ViolentTheft], [0.04, BreakingPeace], [0.01, Deception, RoyalOffences], [0.01, Sexual, Theft]]
1887	[0.17, Miscellaneous], [0.11, Kill], [0.06, ViolentTheft], [0.05, BreakingPeace], [0.03, Damage, Sexual], [0.02, Deception, RoyalOffences], [0.01, Theft]]
1890	[0.14, Miscellaneous], [0.09, Kill], [0.08, ViolentTheft], [0.05, BreakingPeace], [0.03, Damage, Sexual], [0.01, Deception, RoyalOffences, Theft]]
1893	[0.17, Miscellaneous], [0.11, Theft], [0.10, Kill], [0.09, RoyalOffences], [0.07, ViolentTheft], [0.06, BreakingPeace], [0.04, Damage, Sexual], [0.02, Deception]]
1896	[0.11, Kill], [0.06, ViolentTheft], [0.05, BreakingPeace], [0.03, Damage, Miscellaneous, Sexual], [0.03, Deception], [0.02, RoyalOffences, Theft]]
1899	[0.11, Kill], [0.06, ViolentTheft], [0.04, BreakingPeace, Miscellaneous], [0.04, Damage, Sexual], [0.03, Deception], [0.02, RoyalOffences, Theft]]
1902	[0.10, Kill], [0.08, ViolentTheft], [0.05, BreakingPeace], [0.03, Miscellaneous, Sexual], [0.02, Damage, RoyalOffences, Theft], [0.02, Deception]]
1905	[0.12, Kill], [0.10, ViolentTheft], [0.03, Deception], [0.02, BreakingPeace, Damage, Miscellaneous, Sexual], [0.01, RoyalOffences, Theft]]
1908	[0.12, Kill], [0.11, ViolentTheft], [0.02, BreakingPeace, Sexual], [0.02, Damage, Deception, Miscellaneous], [0.01, RoyalOffences, Theft]]
1911	[0.13, Kill], [0.05, BreakingPeace, Damage], [0.03, Deception], [0.03, Miscellaneous, Sexual], [0.02, RoyalOffences, Theft, ViolentTheft]]

Table S4: 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.

Date	Partition with blocks sorted by probability of death
1674	[38.52, Damage, Kill, Miscellaneous, Sexual, Theft, ViolentTheft], [14.29, BreakingPeace, Deception, RoyalOffences]]
1677	[45.58, RoyalOffences, ViolentTheft], [26.86, Deception, Kill, Miscellaneous, Sexual], [20.37, BreakingPeace, Damage, Theft]]
1680	[43.77, Damage, Deception, Kill, Miscellaneous, Sexual, ViolentTheft], [40.83, BreakingPeace, RoyalOffences, Theft]]
1683	[53.36, ViolentTheft], [29.78, RoyalOffences], [23.67, Damage, Kill, Miscellaneous, Sexual], [19.71, Theft], [0.75, BreakingPeace, Deception]]
1686	[65.93, ViolentTheft], [22.08, Damage, Kill, Sexual], [18.88, Miscellaneous, RoyalOffences], [11.82, Theft], [0.48, BreakingPeace, Deception]]
1689	[53.69, ViolentTheft], [21.59, Kill, Miscellaneous, Sexual], [19.47, Damage, Theft], [16.39, BreakingPeace, RoyalOffences], [1.02, Deception]]
1692	[45.39, ViolentTheft], [22.86, Miscellaneous, RoyalOffences], [19.73, BreakingPeace, Kill, Sexual], [14.48, Theft], [0.68, Damage, Deception]]
1695	[49.43, Sexual, ViolentTheft], [23.20, Kill, RoyalOffences], [10.87, BreakingPeace, Theft], [5.75, Damage, Deception, Miscellaneous]]
1698	[38.03, BreakingPeace, RoyalOffences, ViolentTheft], [17.93, Kill, Sexual], [10.65, Damage, Deception, Miscellaneous, Theft]]
1701	[25.89, Kill, Miscellaneous, Sexual], [9.66, BreakingPeace, Damage, Deception, RoyalOffences, Theft, ViolentTheft]]
1704	[8.53, Theft], [6.72, BreakingPeace, Damage, Deception, Kill, Miscellaneous, RoyalOffences, Sexual, ViolentTheft]]
1707	[19.67, BreakingPeace, Damage, Kill, Miscellaneous, RoyalOffences, Sexual, ViolentTheft], [8.60, Deception, Theft]]
1710	[47.25, ViolentTheft], [12.95, Damage, Deception, RoyalOffences], [8.51, BreakingPeace, Kill, Miscellaneous, Sexual, Theft]]
1713	[33.67, ViolentTheft], [13.42, Damage, Deception, Theft], [11.15, Kill, Miscellaneous, Sexual], [4.57, BreakingPeace, RoyalOffences]]
1716	[45.98, ViolentTheft], [18.39, Damage, Theft], [17.07, Kill, Sexual], [9.45, BreakingPeace, RoyalOffences], [0.20, Deception, Miscellaneous]]
1719	[40.83, ViolentTheft], [40.48, Miscellaneous], [12.14, Kill, Sexual], [10.31, Theft], [4.69, BreakingPeace, Damage, Deception, RoyalOffences]]
1722	[42.60, Damage, Miscellaneous, ViolentTheft], [19.27, Kill], [9.28, Deception, RoyalOffences], [6.07, Theft], [3.51, BreakingPeace, Sexual]]
1725	[36.83, Miscellaneous, ViolentTheft], [19.55, Kill], [6.61, BreakingPeace, RoyalOffences, Sexual], [6.46, Theft], [0.31, Damage, Deception]]
1728	[43.34, Damage, Miscellaneous, ViolentTheft], [14.29, Kill], [6.75, BreakingPeace, Deception, RoyalOffences, Sexual], [6.40, Theft]]
1731	[43.62, ViolentTheft], [18.61, Kill, RoyalOffences, Sexual], [5.92, BreakingPeace, Damage, Deception, Miscellaneous], [5.13, Theft]]
1734	[36.41, RoyalOffences, ViolentTheft], [22.79, Damage, Kill, Sexual], [8.29, BreakingPeace, Deception, Miscellaneous], [6.13, Theft]]
1737	[55.37, RoyalOffences, ViolentTheft], [35.06, Damage, Kill, Sexual], [16.23, BreakingPeace, Deception, Miscellaneous], [5.82, Theft]]
1740	[45.85, Miscellaneous, RoyalOffences, ViolentTheft], [26.53, Damage, Kill], [8.29, BreakingPeace, Deception, Sexual], [5.33, Theft]]
1743	[42.62, Miscellaneous, ViolentTheft], [14.70, BreakingPeace, Deception, RoyalOffences, Sexual], [14.29, Damage, Kill], [3.87, Theft]]
1746	[38.14, Damage, Miscellaneous, RoyalOffences, ViolentTheft], [21.95, BreakingPeace, Deception, Kill, Sexual], [3.11, Theft]]
1749	[47.48, ViolentTheft], [34.00, Damage, Deception, Miscellaneous, RoyalOffences], [15.14, BreakingPeace, Kill, Sexual], [3.74, Theft]]
1752	[45.29, Miscellaneous, ViolentTheft], [27.46, BreakingPeace, Damage, Kill, Sexual], [18.86, Deception, RoyalOffences], [4.59, Theft]]
1755	[38.32, BreakingPeace, ViolentTheft], [22.34, Damage, Deception, Kill, Miscellaneous, RoyalOffences, Sexual], [5.24, Theft]]
1758	[27.60, BreakingPeace, Kill, RoyalOffences, Sexual], [13.28, Damage, Deception, Miscellaneous, ViolentTheft], [3.69, Theft]]
1761	[33.44, BreakingPeace, Deception, Miscellaneous, ViolentTheft], [21.20, Kill, RoyalOffences, Sexual], [4.10, Damage, Theft]]
1764	[36.27, Damage, Deception, Miscellaneous, ViolentTheft], [12.32, BreakingPeace, Kill, RoyalOffences, Sexual], [4.65, Theft]]
1767	[37.64, Damage, Miscellaneous, RoyalOffences, ViolentTheft], [21.81, BreakingPeace, Kill, Sexual], [17.19, Deception], [5.56, Theft]]
1770	[48.77, BreakingPeace, Miscellaneous], [42.23, ViolentTheft], [18.51, Damage, Deception, RoyalOffences], [15.37, Kill, Sexual], [5.65, Theft]]
1773	[37.27, BreakingPeace, Miscellaneous, ViolentTheft], [16.90, Deception], [15.11, Damage, Kill, RoyalOffences, Sexual], [6.69, Theft]]
1776	[37.51, BreakingPeace, RoyalOffences, Sexual, ViolentTheft], [17.29, Deception, Miscellaneous], [16.23, Kill], [4.59, Damage, Theft]]
1779	[46.72, ViolentTheft], [37.85, BreakingPeace], [17.86, Deception, Miscellaneous], [14.29, Damage, Kill, RoyalOffences, Sexual], [6.78, Theft]]
1782	[46.54, Miscellaneous], [39.24, Damage, Deception, Sexual, ViolentTheft], [18.51, BreakingPeace, Kill], [15.48, RoyalOffences], [6.80, Theft]]
1785	[39.80, Damage, ViolentTheft], [20.87, Deception, Miscellaneous], [10.47, BreakingPeace, Kill, RoyalOffences], [7.42, Theft], [0.62, Sexual]]
1788	[29.18, Damage, ViolentTheft], [24.70, Kill, Miscellaneous, RoyalOffences], [9.93, Deception], [4.84, Theft], [0.43, BreakingPeace, Sexual]]
1791	[48.30, ViolentTheft], [27.73, Damage, Deception], [21.02, Kill, Miscellaneous, Sexual], [5.48, BreakingPeace, Theft], [0.24, RoyalOffences]]
1794	[34.45, BreakingPeace, Deception, Miscellaneous, ViolentTheft], [17.58, Damage, Kill, Sexual], [4.89, Theft], [2.71, RoyalOffences]]
1797	[38.31, BreakingPeace, Deception, ViolentTheft], [15.98, Miscellaneous, Sexual], [10.04, Damage, Kill, RoyalOffences], [6.00, Theft]]
1800	[29.07, Deception, Miscellaneous, ViolentTheft], [15.13, BreakingPeace, Damage, Kill], [12.64, RoyalOffences, Sexual], [5.74, Theft]]
1803	[49.17, Miscellaneous], [25.22, ViolentTheft], [20.96, RoyalOffences], [15.58, BreakingPeace, Damage, Deception, Kill, Sexual], [4.54, Theft]]
1806	[28.68, BreakingPeace, Damage, Deception, ViolentTheft], [16.64, Kill, Miscellaneous], [16.23, RoyalOffences, Sexual], [4.57, Theft]]
1809	[29.49, ViolentTheft], [27.47, BreakingPeace, Damage], [21.25, Deception, RoyalOffences, Sexual], [12.17, Kill, Miscellaneous], [5.99, Theft]]
1812	[30.58, Deception, ViolentTheft], [30.46, Damage, Miscellaneous], [16.88, BreakingPeace, Kill], [11.98, RoyalOffences, Sexual], [6.20, Theft]]
1815	[36.63, ViolentTheft], [16.54, BreakingPeace], [14.29, Deception, Miscellaneous, Sexual], [9.52, Damage, Kill], [9.13, RoyalOffences], [6.85, Theft]]
1818	[43.44, ViolentTheft], [28.44, RoyalOffences], [15.08, BreakingPeace, Kill], [8.13, Damage, Deception, Miscellaneous, Sexual], [7.23, Theft]]
1821	[35.66, ViolentTheft], [33.52, RoyalOffences], [18.80, BreakingPeace, Damage], [13.49, Miscellaneous], [10.71, Deception, Sexual], [5.27, Theft], [2.48, Kill]]
1824	[33.69, ViolentTheft], [21.43, BreakingPeace], [14.92, Miscellaneous], [12.88, RoyalOffences], [7.24, Deception], [6.15, Theft], [5.25, Damage, Kill, Sexual]]
1827	[34.61, BreakingPeace, Damage, ViolentTheft], [31.70, Deception], [24.76, Miscellaneous, RoyalOffences], [6.11, Kill, Sexual], [4.73, Theft]]
1830	[34.46, BreakingPeace, ViolentTheft], [12.14, Deception], [12.07, Miscellaneous, RoyalOffences], [6.61, Kill, Sexual], [3.93, Damage, Theft]]
1833	[36.32, ViolentTheft], [25.19, BreakingPeace], [4.37, Damage, Kill, Miscellaneous, Sexual], [2.21, Theft], [0.85, Deception], [0.09, RoyalOffences]]
1836	[27.42, ViolentTheft], [13.34, BreakingPeace], [7.58, Kill], [2.10, Sexual], [0.63, Theft], [0.42, Deception], [0.11, Damage, Miscellaneous], [0.06, RoyalOffences]]
1839	[10.44, Kill, Sexual], [1.62, BreakingPeace], [0.08, Miscellaneous], [0.08, RoyalOffences], [0.04, Damage, Deception, ViolentTheft], [0.00, Theft]]
1842	[10.18, Kill], [3.02, Damage, ViolentTheft], [2.54, Sexual], [0.71, BreakingPeace], [0.51, RoyalOffences], [0.10, Miscellaneous], [0.04, Deception], [0.03, Theft]]
1845	[10.19, Damage, Kill], [1.08, Miscellaneous, Sexual], [0.82, BreakingPeace], [0.12, ViolentTheft], [0.05, RoyalOffences], [0.04, Deception], [0.00, Theft]]
1848	[7.22, Kill], [1.46, BreakingPeace, Sexual], [1.35, Damage, ViolentTheft], [0.04, RoyalOffences], [0.04, Theft], [0.03, Deception, Miscellaneous]]
1851	[2.75, Damage, Miscellaneous], [2.37, BreakingPeace], [2.23, Kill, Sexual], [0.14, ViolentTheft], [0.03, Deception], [0.02, RoyalOffences], [0.01, Theft]]
1854	[11.22, Damage, Kill], [1.43, BreakingPeace, Sexual], [1.04, ViolentTheft], [0.25, Miscellaneous], [0.06, Theft], [0.03, Deception], [0.02, RoyalOffences]]
1857	[6.36, Damage, Kill], [1.48, BreakingPeace, Miscellaneous, Sexual], [1.34, ViolentTheft], [0.20, Theft], [0.04, Deception], [0.02, RoyalOffences]]
1860	[10.95, Kill], [1.06, BreakingPeace, Miscellaneous], [0.23, Deception, Sexual], [0.19, Damage, Theft, ViolentTheft], [0.02, RoyalOffences]]
1863	[7.14, Kill], [0.16, Miscellaneous], [0.05, BreakingPeace], [0.04, Deception], [0.03, Damage, Sexual, ViolentTheft], [0.03, RoyalOffences], [0.01, Theft]]
1866	[6.03, Kill], [0.20, Miscellaneous], [0.05, ViolentTheft], [0.03, BreakingPeace, Damage, Sexual], [0.03, RoyalOffences], [0.01, Deception, Theft]]
1869	[8.24, Kill], [0.21, Miscellaneous], [0.06, ViolentTheft], [0.05, BreakingPeace, Damage], [0.04, RoyalOffences], [0.01, Deception, Sexual, Theft]]
1872	[7.68, Kill], [0.47, BreakingPeace], [0.20, Miscellaneous], [0.09, ViolentTheft], [0.06, RoyalOffences], [0.01, Damage, Deception, Sexual, Theft]]
1875	[12.14, Kill], [0.20, Miscellaneous], [0.10, ViolentTheft], [0.02, BreakingPeace, Deception, Sexual], [0.01, Damage, RoyalOffences, Theft]]
1878	[9.39, Kill], [0.18, Miscellaneous], [0.07, ViolentTheft], [0.05, BreakingPeace], [0.04, RoyalOffences], [0.01, Damage, Deception, Sexual, Theft]]
1881	[7.43, Kill], [0.05, BreakingPeace], [0.05, ViolentTheft], [0.04, Damage, Miscellaneous, Sexual], [0.02, Deception], [0.01, RoyalOffences, Theft]]
1884	[11.45, Kill], [0.06, ViolentTheft], [0.04, BreakingPeace], [0.02, Damage, Deception, Miscellaneous], [0.01, RoyalOffences, Sexual, Theft]]
1887	[6.87, Kill], [0.29, BreakingPeace, Miscellaneous], [0.06, ViolentTheft], [0.03, Damage, Sexual], [0.02, Deception, RoyalOffences], [0.01, Theft]]
1890	[9.83, Kill], [0.14, Miscellaneous], [0.08, ViolentTheft], [0.05, BreakingPeace], [0.03, Damage, Sexual], [0.02, Deception, RoyalOffences], [0.01, Theft]]
1893	[6.18, Kill], [0.17, Miscellaneous], [0.11, Theft], [0.07, Damage, RoyalOffences], [0.07, ViolentTheft], [0.06, BreakingPeace], [0.04, Sexual], [0.02, Deception]]
1896	[14.17, Kill], [0.06, ViolentTheft], [0.04, BreakingPeace, Miscellaneous], [0.04, Damage, Sexual], [0.03, Deception], [0.02, RoyalOffences, Theft]]
1899	[17.02, Kill], [0.06, ViolentTheft], [0.03, Deception], [0.02, BreakingPeace, Miscellaneous, Sexual], [0.02, Damage, RoyalOffences, Theft]]
1902	[14.96, Kill], [0.08, ViolentTheft], [0.05, BreakingPeace], [0.03, Miscellaneous, Sexual], [0.01, Damage, Deception, RoyalOffences, Theft]]
1905	[14.17, Kill], [0.10, ViolentTheft], [0.03, Deception], [0.02, BreakingPeace, Miscellaneous, Sexual], [0.01, Damage, RoyalOffences, Theft]]
1908	[17.52, Kill], [0.91, ViolentTheft], [0.17, BreakingPeace, Sexual], [0.02, Deception, Miscellaneous], [0.01, Damage, RoyalOffences, Theft]]
1911	[15.44, Kill], [0.68, BreakingPeace, Damage], [0.03, Deception], [0.03, Miscellaneous, RoyalOffences, Sexual], [0.02, Theft, ViolentTheft]]

Table S5: The partitions for the third 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.

Date	Partition with blocks sorted by probability of death
1674	[[39.39, Kill, Sexual, Theft, ViolentTheft], [15.71, BreakingPeace, Damage, Deception, Miscellaneous, RoyalOffences]]
1677	[[45.58, Deception, RoyalOffences, ViolentTheft], [26.86, Kill, Miscellaneous, Sexual], [20.37, BreakingPeace, Damage, Theft]]
1680	[[43.04, Kill, Miscellaneous, Sexual, ViolentTheft], [41.11, BreakingPeace, Damage, Deception, RoyalOffences, Theft]]
1683	[[53.36, ViolentTheft], [29.78, Damage, RoyalOffences], [23.67, Kill, Miscellaneous, Sexual], [19.71, Theft], [0.75, BreakingPeace, Deception]]
1686	[[63.81, Damage, ViolentTheft], [21.37, Kill], [19.01, Miscellaneous, RoyalOffences, Sexual], [11.82, Theft], [0.48, BreakingPeace, Deception]]
1689	[[53.69, ViolentTheft], [20.67, BreakingPeace, Kill, Miscellaneous, Sexual], [19.47, Damage, Theft], [17.97, RoyalOffences], [1.02, Deception]]
1692	[[45.39, ViolentTheft], [22.86, Damage, Miscellaneous, RoyalOffences], [19.73, BreakingPeace, Kill, Sexual], [14.48, Theft], [0.68, Deception]]
1695	[[49.43, Sexual, ViolentTheft], [23.20, Damage, Kill, RoyalOffences], [10.87, BreakingPeace, Theft], [5.75, Deception, Miscellaneous]]
1698	[[38.03, BreakingPeace, RoyalOffences, ViolentTheft], [17.93, Damage, Kill, Sexual], [10.65, Deception, Miscellaneous, Theft]]
1701	[[25.89, Damage, Kill, Sexual], [9.66, BreakingPeace, Deception, Miscellaneous, RoyalOffences, Theft, ViolentTheft]]
1704	[[8.53, Miscellaneous, Theft], [6.72, BreakingPeace, Damage, Deception, Kill, RoyalOffences, Sexual, ViolentTheft]]
1707	[[20.74, Damage, Kill, RoyalOffences, Sexual, ViolentTheft], [8.51, BreakingPeace, Deception, Miscellaneous, Theft]]
1710	[[47.25, Damage, ViolentTheft], [12.95, Deception, RoyalOffences], [8.51, BreakingPeace, Kill, Miscellaneous, Sexual, Theft]]
1713	[[33.67, ViolentTheft], [13.42, Deception, Theft], [11.15, Damage, Kill, Miscellaneous, Sexual], [4.57, BreakingPeace, RoyalOffences]]
1716	[[45.98, ViolentTheft], [18.31, Theft], [18.13, Damage, Kill, Sexual], [9.45, BreakingPeace, RoyalOffences], [0.20, Deception, Miscellaneous]]
1719	[[40.83, ViolentTheft], [29.21, Miscellaneous, RoyalOffences], [12.14, Kill, Sexual], [10.31, Theft], [4.12, BreakingPeace, Damage, Deception]]
1722	[[43.25, Miscellaneous, ViolentTheft], [19.27, Kill], [7.19, BreakingPeace, Damage, Deception, RoyalOffences, Sexual], [6.07, Theft]]
1725	[[36.83, Miscellaneous, ViolentTheft], [19.21, Damage, Kill], [6.46, Theft], [4.48, BreakingPeace, Deception, RoyalOffences, Sexual]]
1728	[[43.34, Damage, Miscellaneous, ViolentTheft], [11.15, Kill, RoyalOffences, Sexual], [6.40, Theft], [3.17, BreakingPeace, Deception]]
1731	[[43.49, RoyalOffences, ViolentTheft], [23.71, Kill], [5.47, BreakingPeace, Damage, Deception, Miscellaneous, Sexual], [5.13, Theft]]
1734	[[35.50, Damage, ViolentTheft], [24.30, Kill, RoyalOffences, Sexual], [8.29, BreakingPeace, Deception, Miscellaneous], [6.13, Theft]]
1737	[[54.85, Damage, RoyalOffences, ViolentTheft], [35.60, Kill, Sexual], [16.23, BreakingPeace, Deception, Miscellaneous], [5.82, Theft]]
1740	[[45.85, Damage, Miscellaneous, RoyalOffences, ViolentTheft], [26.53, Kill], [8.29, BreakingPeace, Deception, Sexual], [5.33, Theft]]
1743	[[42.27, Damage, Miscellaneous, ViolentTheft], [14.70, BreakingPeace, Deception, RoyalOffences, Sexual], [14.63, Kill], [3.87, Theft]]
1746	[[38.14, Miscellaneous, RoyalOffences, ViolentTheft], [21.95, BreakingPeace, Deception, Kill, Sexual], [3.11, Damage, Theft]]
1749	[[47.48, ViolentTheft], [34.00, Deception, Miscellaneous, RoyalOffences], [15.14, BreakingPeace, Kill, Sexual], [3.74, Damage, Theft]]
1752	[[45.29, BreakingPeace, Damage, Miscellaneous, ViolentTheft], [27.46, Kill, Sexual], [18.86, Deception, RoyalOffences], [4.59, Theft]]
1755	[[32.23, BreakingPeace, Damage, Sexual, ViolentTheft], [24.65, Deception, Kill, Miscellaneous, RoyalOffences], [5.24, Theft]]
1758	[[27.60, BreakingPeace, Kill, RoyalOffences, Sexual], [13.28, Deception, Miscellaneous, ViolentTheft], [3.69, Damage, Theft]]
1761	[[35.62, BreakingPeace, Damage, Miscellaneous, ViolentTheft], [26.90, Deception], [21.20, Kill, RoyalOffences, Sexual], [4.12, Theft]]
1764	[[36.79, Deception, Miscellaneous, ViolentTheft], [12.32, BreakingPeace, Kill, RoyalOffences, Sexual], [4.64, Damage, Theft]]
1767	[[36.99, Damage, RoyalOffences, ViolentTheft], [21.81, BreakingPeace, Kill, Sexual], [20.80, Deception, Miscellaneous], [5.56, Theft]]
1770	[[52.80, Miscellaneous], [41.30, Damage, ViolentTheft], [18.68, Deception, RoyalOffences], [16.87, BreakingPeace, Kill, Sexual], [5.65, Theft]]
1773	[[37.27, BreakingPeace, Miscellaneous, ViolentTheft], [16.10, Damage, Deception], [15.65, Kill, RoyalOffences, Sexual], [6.69, Theft]]
1776	[[37.51, BreakingPeace, RoyalOffences, Sexual, ViolentTheft], [16.93, Deception, Kill, Miscellaneous], [4.59, Damage, Theft]]
1779	[[43.81, BreakingPeace, ViolentTheft], [17.86, Deception, Miscellaneous], [14.29, Damage, Kill, RoyalOffences, Sexual], [6.78, Theft]]
1782	[[46.54, Miscellaneous], [39.55, Damage, Deception, ViolentTheft], [19.88, BreakingPeace, Kill], [14.86, RoyalOffences, Sexual], [6.80, Theft]]
1785	[[39.80, Damage, ViolentTheft], [31.84, Miscellaneous], [14.89, Deception, Sexual], [10.47, BreakingPeace, Kill, RoyalOffences], [7.42, Theft]]
1788	[[38.66, Damage, Miscellaneous], [29.20, ViolentTheft], [19.72, Kill, RoyalOffences], [9.93, Deception], [4.84, Theft], [0.43, BreakingPeace, Sexual]]
1791	[[48.30, ViolentTheft], [28.15, BreakingPeace, Damage, Deception, Miscellaneous], [14.63, Kill, Sexual], [5.50, Theft], [0.24, RoyalOffences]]
1794	[[34.43, BreakingPeace, Damage, Deception, Miscellaneous, ViolentTheft], [6.79, Kill, RoyalOffences, Sexual], [4.89, Theft]]
1797	[[38.31, BreakingPeace, Deception, ViolentTheft], [15.17, Damage, Kill, Miscellaneous, Sexual], [8.50, RoyalOffences], [6.00, Theft]]
1800	[[31.17, Damage, ViolentTheft], [23.81, Deception, Miscellaneous], [16.59, BreakingPeace, Kill], [12.64, RoyalOffences, Sexual], [5.74, Theft]]
1803	[[46.13, Damage, Miscellaneous], [25.22, ViolentTheft], [20.96, RoyalOffences], [15.43, BreakingPeace, Deception, Kill, Sexual], [4.54, Theft]]
1806	[[28.90, BreakingPeace, Deception, ViolentTheft], [16.34, Damage, Kill, Miscellaneous, RoyalOffences, Sexual], [4.57, Theft]]
1809	[[29.49, ViolentTheft], [27.47, BreakingPeace, Damage], [20.85, Deception, Sexual], [15.90, Kill, Miscellaneous, RoyalOffences], [5.99, Theft]]
1812	[[30.58, Deception, ViolentTheft], [30.46, Damage, Miscellaneous], [19.73, BreakingPeace], [15.53, Kill], [11.98, RoyalOffences, Sexual], [6.20, Theft]]
1815	[[35.16, BreakingPeace, ViolentTheft], [14.29, Deception, Miscellaneous, Sexual], [9.52, Damage, Kill], [9.13, RoyalOffences], [6.85, Theft]]
1818	[[43.44, ViolentTheft], [28.44, RoyalOffences], [18.01, BreakingPeace], [12.72, Kill, Miscellaneous], [7.23, Theft], [7.14, Damage, Deception, Sexual]]
1821	[[33.52, RoyalOffences], [32.06, BreakingPeace, Damage, ViolentTheft], [13.49, Miscellaneous], [10.71, Deception, Sexual], [5.27, Theft], [2.48, Kill]]
1824	[[33.69, ViolentTheft], [21.43, BreakingPeace], [15.06, Damage, Miscellaneous], [12.88, RoyalOffences], [7.24, Deception], [6.15, Theft], [4.63, Kill, Sexual]]
1827	[[36.82, RoyalOffences], [34.63, BreakingPeace, ViolentTheft], [31.77, Damage, Deception], [13.93, Miscellaneous], [6.11, Kill, Sexual], [4.73, Theft]]
1830	[[37.54, BreakingPeace], [32.74, ViolentTheft], [12.07, Miscellaneous, RoyalOffences], [11.68, Damage, Deception], [6.61, Kill, Sexual], [3.93, Theft]]
1833	[[36.32, ViolentTheft], [25.19, BreakingPeace], [5.18, Kill, Sexual], [3.31, Damage, Miscellaneous], [2.21, Theft], [0.85, Deception], [0.09, RoyalOffences]]
1836	[[27.42, ViolentTheft], [13.34, BreakingPeace], [6.83, Damage, Kill], [2.10, Sexual], [0.63, Theft], [0.42, Deception], [0.12, Miscellaneous], [0.06, RoyalOffences]]
1839	[[10.44, Kill, Sexual], [1.62, BreakingPeace], [0.08, Miscellaneous], [0.08, RoyalOffences], [0.04, Deception, ViolentTheft], [0.00, Damage, Theft]]
1842	[[10.18, Kill], [3.42, ViolentTheft], [2.37, Damage, Sexual], [0.71, BreakingPeace], [0.51, RoyalOffences], [0.10, Miscellaneous], [0.04, Deception], [0.03, Theft]]
1845	[[10.19, Damage, Kill], [1.63, Sexual], [0.82, BreakingPeace], [0.14, Miscellaneous], [0.12, ViolentTheft], [0.05, RoyalOffences], [0.04, Deception], [0.00, Theft]]
1848	[[6.72, Damage], [3.86, Kill, Miscellaneous], [1.46, BreakingPeace, Sexual], [0.80, ViolentTheft], [0.04, RoyalOffences], [0.04, Theft], [0.03, Deception]]
1851	[[2.96, BreakingPeace, Damage], [2.23, Kill, Sexual], [0.23, Miscellaneous], [0.14, ViolentTheft], [0.03, Deception], [0.02, RoyalOffences], [0.01, Theft]]
1854	[[11.22, Damage, Kill], [1.29, BreakingPeace, Sexual, ViolentTheft], [0.25, Miscellaneous], [0.06, Theft], [0.03, Deception], [0.02, RoyalOffences]]
1857	[[6.36, Damage, Kill], [1.45, BreakingPeace, Miscellaneous], [0.44, Deception, Sexual], [0.30, Theft, ViolentTheft], [0.02, RoyalOffences]]
1860	[[10.95, Kill], [1.06, BreakingPeace, Miscellaneous], [0.22, Theft], [0.17, Damage, Deception, Sexual, ViolentTheft], [0.02, RoyalOffences]]
1863	[[7.14, Kill], [0.16, Miscellaneous], [0.05, BreakingPeace], [0.04, Damage, Deception], [0.04, Sexual, ViolentTheft], [0.01, RoyalOffences, Theft]]
1866	[[6.03, Kill], [0.20, Miscellaneous], [0.05, ViolentTheft], [0.04, BreakingPeace, Damage], [0.03, RoyalOffences], [0.01, Deception, Sexual, Theft]]
1869	[[8.24, Kill], [0.21, Miscellaneous], [0.09, Sexual], [0.06, ViolentTheft], [0.05, BreakingPeace, Damage], [0.04, RoyalOffences], [0.01, Deception, Theft]]
1872	[[7.68, Kill], [0.43, BreakingPeace, Damage], [0.20, Miscellaneous], [0.09, ViolentTheft], [0.01, Deception, RoyalOffences, Sexual, Theft]]
1875	[[12.14, Kill], [0.20, Miscellaneous], [0.10, ViolentTheft], [0.02, BreakingPeace, Damage, Deception, Sexual], [0.01, RoyalOffences, Theft]]
1878	[[9.39, Kill], [0.18, Miscellaneous], [0.07, ViolentTheft], [0.06, Damage, Sexual], [0.05, BreakingPeace], [0.04, RoyalOffences], [0.01, Deception, Theft]]
1881	[[7.43, Kill], [0.09, Damage, Miscellaneous], [0.05, BreakingPeace], [0.05, ViolentTheft], [0.02, Deception, Sexual], [0.01, RoyalOffences, Theft]]
1884	[[11.45, Kill], [0.06, ViolentTheft], [0.04, BreakingPeace], [0.03, RoyalOffences], [0.02, Damage, Deception, Miscellaneous], [0.01, Sexual, Theft]]
1887	[[6.87, Kill], [0.30, BreakingPeace, Damage], [0.17, Miscellaneous], [0.06, ViolentTheft], [0.04, Sexual], [0.02, Deception, RoyalOffences], [0.01, Theft]]
1890	[[9.83, Kill], [0.14, Miscellaneous], [0.08, ViolentTheft], [0.05, BreakingPeace], [0.02, Damage, RoyalOffences, Sexual], [0.01, Deception, Theft]]
1893	[[6.18, Kill], [0.17, Miscellaneous], [0.11, Theft], [0.07, ViolentTheft], [0.06, BreakingPeace], [0.03, Damage, RoyalOffences, Sexual], [0.02, Deception]]
1896	[[14.17, Kill], [0.06, ViolentTheft], [0.05, BreakingPeace], [0.04, Damage, Sexual], [0.02, Deception, Miscellaneous], [0.02, RoyalOffences, Theft]]
1899	[[17.02, Kill], [0.06, ViolentTheft], [0.03, Deception], [0.02, BreakingPeace, Damage, Miscellaneous, Sexual], [0.02, RoyalOffences, Theft]]
1902	[[14.96, Kill], [0.08, ViolentTheft], [0.05, BreakingPeace], [0.03, Miscellaneous, Sexual], [0.02, Damage, Deception], [0.02, RoyalOffences, Theft]]
1905	[[14.17, Kill], [0.10, ViolentTheft], [0.03, Deception], [0.02, Theft], [0.01, BreakingPeace, Damage, Miscellaneous, RoyalOffences, Sexual]]
1908	[[17.52, Kill], [0.91, ViolentTheft], [0.10, BreakingPeace, Damage, Deception, Miscellaneous], [0.04, Sexual], [0.01, RoyalOffences, Theft]]
1911	[[15.44, Kill], [0.43, BreakingPeace], [0.20, Damage, Deception], [0.03, Miscellaneous, Sexual], [0.02, RoyalOffences, Theft, ViolentTheft]]

Appendix

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 at various times), with renewal of the license not guaranteed. 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 evidence upon the trial of every prisoner, whether he or she shall be convicted or acquitted.” [38]. Despite these criticisms, historians agree that the Proceedings presents an accurate, if often incomplete, account of courtroom proceedings, especially from 1778 onward [37].

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 being 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 became 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, the Metropolitan Police Act of 1829 [40] saw a total of 3,000 police officers 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[28][20, 147]. 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 with their sentences 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 transport 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.
- 1827-1828 Acts changing offence categorisation.
- 1832 Removal of the death penalty for most forms of theft.
- 1834 The Old Bailey given jurisdiction over all of England.
- 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 previously, 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 1674-1834 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 1674 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 symbol 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 where the query is generated as the submitted URL. The web content returned is of a known format, to allow the information to be reliably parsed. Throughout this section Listing S1 will be used to illustrate a typical OBO web-API query.

Listing S1: OBO web-API example query.

```
http://www.oldbaileyonline.org/obapi/ob? term0=fromdate_1774
&term1=todate_1824 &term2=offcat_theft&term3=
defgen_female &breakdown=punecat&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_n*, where *n* is incremented with every term. *&* separates query 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 &term1=todate_1824*), breaking down the results by punishment category (*breakdown=punecat*). The

count=1 term returns only one trial id, otherwise all matching trial id's are returned.

Listing S2: OBO web-API example of returned data.

```
{ "total" : 10234,
  "breakdown" :
  [
  { "term" : "imprison", "total" : 3479 },
  { "term" : "miscPunish", "total" : 2239 },
  { "term" : "transport", "total" : 2038 },
  { "term" : "corporal", "total" : 867 },
  { "term" : "death", "total" : 533 },
  { "term" : "noPunish", "total" : 128 }
  ]
, "hits" :
  [
  { "t17740112-2"
  }
  ]
}
```

Referring to Listing S2, 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 given 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 S1, 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* 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 S3 shows the code for generating URL's.

Listing S3: URL generating code

```
def generateURLCategoryRange(StartDate, EndDate,
                             OffenceCategory, Gender='None'):
    """Generate an OBOApi url for insertion to an http 'GET'
    """
    if Gender is 'None':
        return 'http://www.oldbaileyonline.org/obapi/ob? term=
        fromdate_' + StartDate+'&term1=todate_' + EndDate
        + '&term2=offcat_' + OffenceCategory + '&breakdown=
        puncat&count=1'
    else:
        return 'http://www.oldbaileyonline.org/obapi/ob? term=
        fromdate_' + StartDate+'&term1=todate_' + EndDate
        + '&term2=offcat_' + OffenceCategory + '&term3=
        defgen_' + Gender + '&breakdown=punecat&count=1'

def generateURLSubCategoryNotGuiltyRange(StartDate, EndDate,
                                          Gender='None'):
    """Generate an OBOApi url for insertion to an http 'GET'
    """
    if Gender is 'None':
        return 'http://www.oldbaileyonline.org/obapi/ob? term=
        fromdate_' + StartDate+'&term1=todate_' + EndDate +
        '&term2=vercat_notGuilty &breakdown= offsubcat &
        count=1'
    else:
        return 'http://www.oldbaileyonline.org/obapi/ob? term=
        fromdate_' + StartDate+'&term1=todate_' + EndDate
        + '&term2=vercat_notGuilty&term3=defgen_' + Gender
        + '&breakdown=offsubcat&count=1'
```

Listing S4 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 S4: Socket opening and URL getting code

```
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=retry, timeout=timeout,
                          maxsize=5)
...
def URLtoJSON(URL, httpPool):
    import json
    return json.loads( httpPool.request('GET', URL).data.
                        decode('utf8' ) )
```

Listing S5 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.punecat` is an array containing the punishment categories (sim. `c2n.offcat`)

Listing S5: JSON processing

```
import CategoryToNumberAssignment as c2n
_TempCategories = [0]*(1+len(c2n.punecat))

for Totals in _Json['breakdown']:
    _TempCategories[c2n.punecat.index(Totals['term']) + 1]
    = Totals['total']
# Append _Temps to the Rows
Row = Row + _TempCategories

"""Get not guiltyies by category and subcategory for the
period"""
#Get not guiltyies:
_JsonNotGuiltyCat = URLtoJSON(
    generateURLCategoryNotGuiltyRange( sStartDate, sEndDate
    ,G ender), http)
_JsonNotGuiltySubCat = URLtoJSON(
    generateURLSubCategoryNotGuiltyRange( sStartDate,
    sEndDate, Gender), http)

# Place values associated with locations in Row:
for Totals in _JsonNotGuiltyCat['breakdown']:
    Row[ _Columns.index( upcase_first_letter( Totals['term']
    ) + 'NotGuilty' ) ] = Totals['total']

for Totals in _JsonNotGuiltySubCat['breakdown']:
    Row[ _Columns.index( upcase_first_letter( Totals['term']
    ) + 'NotGuilty' ) ] = Totals['total']
```

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 S1 where the i, j th element is the frequency or probability estimate of p_j given o_i . With $\Delta t = 1$ 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 S2 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, o_i , is found as a slice of the columns from $o_i p_1$ to $o_i p_n$, and the offences associated with a punishment, p_j , is the selection of columns $\{o_i p_j, i \in 1 \dots 9\}$. 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.

smoothed probability estimates on the entire offence-punishment pairs data with the following code:

Listing S8: Probability estimate conditional on offence ($\theta(p_j | o_i)$) code snippet from the function: `conditionPunishmentEstimatesOnOffence`.

```
p = 7
o = 9
#For all offences 'normalise' each offence block.
for c in range(0,o):
    CatModCond.iloc[:,c*p:(c+1)*p] = CatMod.iloc[:,c*p:(c+1)*p]
        .div( CatMod.iloc[:,c*p:(c+1)*p].sum(axis=1), axis
            =0)
```

Listing S9: Offence estimate conditional on punishment ($p_j | o_i$) code snippet from the function: `conditionOffenceEstimatesOnPunishment`.

```
ModLen = CatMod.shape[1]
p = 7
o = 9
#For all punishment 'normalise' each punishment in its
offence block, taking slice as a list of rows.
for c in range(0,p):
    CatModCond.iloc[:,list(range(c,ModLen,p))] = CatMod.iloc
       [:, list(range(c,ModLen,p))] .div( CatMod.iloc[:,
            list(range(c,ModLen,p))] .sum(axis=1), axis=0)
```

Combine offence-punishment pairs on offence or punishment. From the offence-punishment pairs estimates or raw data, subsidiary datasets on offences or punishments only can be achieved by slicing and summing the original data as in Listing S10.

Listing S10: Combining offence-punishment pairs on offence or punishment.

```
def combineOnOffence(DataFrame):
    offenceGroup = [ offence for offence in c2n.offcatUp for
        punishment in c2n.puncatFullUp]
    OffenceFrame = DataFrame.groupby(offenceGroup, axis=1,
        sort=False).sum()
    return OffenceFrame

def combineOnPunishment(DataFrame):
    punishmentGroup = [ punishment for offence in c2n.
        offcatUp for punishment in c2n.puncatFullUp]
    PunishmentFrame = DataFrame.groupby(punishmentGroup,
        axis=1, sort=False).sum()
    return PunishmentFrame
```

Log likelihood. Referring to the \ln term of Equation (6), the log-likelihood between the the empirical frequencies and the the probability estimate table are found table wide assuming that both tables are of the same size as in Listing S11.

Listing S11: The log-likelihood code snippet from the function: `loglikilyhood`.

```
( EmpiricalFrame*np.log(ModelFrame) ).sum( axis=1 ).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 S12.

Listing S12: The AIC code snippet from the function: `generateAICstandardModels`.

```
if 'Dependent' in filename:
    k = (9-1)*(7-1)*CatEmp.shape[0] #number of rows
elif 'Independent' in filename:
    k = (9+7-2)*CatEmp.shape[0]
ll = loglikilyhood(CatEmp, CatMod)
AIC = 2*k - 2*ll
```

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 dimensional 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 S13.

Listing S13: Bootstrap for confidence interval code: `generateBootStrapEstimateConfidenceIntervals`.

```
#Initialise empty Panel to hold Bootstrapped DataFrames
BootstrapPanel = pd.Panel(items=list(range(N)), major_axis=
    EmpFrame.index, minor_axis=EmpFrame.columns)
#Loop through every row in the DataFrame to generate dummy
empirical data
for row in range(DFlen):
    #Represent the events from the row as a list, with each
    index occuring the number of times as the event,
    such that our list is as long as the number of
    events. We need this to pull samples from.
    EmpCountList = [ idx for idx,y in enumerate(EmpFrame.
        iloc[row]) for X in range(int(y)) ]
    EmpCountLen = len(EmpCountList)
    #Sample from EmpCountList N times
    #Initilise Empty DataFrame to hold Bootstrap data for
    this row
    RandomSampleFrame = pd.DataFrame( np.random.randint(0,
        EmpCountLen, (EmpCountLen,N) ), columns = list(
        range(N)), index=list(range(EmpCountLen)) )
    #In one giant step turn this into a frame of event
    counts
    RandomSampleEvents = RandomSampleFrame.applymap(lambda x
        : EmpCountList[x] ).apply(lambda x: x.value_counts(
        bins=[-2,-1] + list(range(DFlen+1)), sort=False).
        shift(1)[2:] )
    #We must ensure that the index of the RandomSampleEvents
    frame is the same as BootstrapPanel munot,axis
    RandomSampleEvents.index = EmpFrame.columns
    BootstrapPanel.iloc[:,row,:] = RandomSampleEvents
#Apply period resampling to the Panel
BootstrapPanel = BootstrapPanel.apply(lambda x:
    generateDependentModelLaplace(x,Delta=Delta), axis
    =[1,2])
ConfLowFrame = BootstrapPanel.apply(lambda x: x.quantile(
    ConfLow), axis=0)
ConfHighFrame = BootstrapPanel.apply(lambda x: x.quantile(
    ConfHigh), axis=0)

return ConfLowFrame, ConfHighFrame
```

Total variation distance. Three different TVD's are used: Consecutive TVD between each adjacent t , TVD between a fixed t and all $t \in \{1, 2, \dots, [240/\Delta t]\}$, and the TVD between punishment distributions of all pairs of offences in the same time-block. 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. The code to perform these operations is in Listing S14 and Listing S15.

Listing S14: The TVD code snippet from the function: `generateTVD`.

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

Listing S15: The pairwise TVD code snippet from the function: `generatePairwiseTV`.

```
PairwiseIndex=0
for i in range(0,OffenceArrayLen):
    for j in range(1,OffenceArrayLen-i):
        #This captures all n choose 2 possibilities
        PairWiseTV.iloc[:,PairwiseIndex] = 0.5*(
            ModelFrame.iloc[:, i:PunLen:(i+1)*PunLen+1] -
            ModelFrame.shift(-(j*PunLen), axis=1).iloc[:, i
                *PunLen:(i+1)*PunLen+1] ).abs().sum( axis=1)
        PairwiseIndex += 1
```

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 remedy this the required 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.pyx`. 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 Δt . The punishment distribution is not dependent on offence. The dummy date is produced with the function `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 Δt .

This process 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 <https://github.com/spmp/OBO-APIextraction-and-Analysis>.

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/Models` 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 S16 running *iPython* from a terminal with the command:

```
ipython3 -i ipython3-startup.py
```

Listing S16: *iPython* startup script

```
import argparse, pathlib, re, glob
import pandas as pd
import numpy as np
from partitionsets import partition
get_ipython().magic('load_ext autoreload')
get_ipython().magic('autoreload 2')
import CategoryToNumberAssignment as c2n
import OBApiExtraction as oboA
import OBOModelling as oboM
import OBOPartitioning as oboP
import OBOValidation as oboV
```

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

```
python3 OBApiExtraction.py -g
```

Data can be selectively downloaded from within the *iPython* shell using the `generateDataFramesRangeSaving` function as:

```
oboA.generateDataFramesRangeSaving(16740000,
    19130000, 10000, 'None', SaveEvery=1,
    File='OBOextractNoGender1yr', Overwrite=0)
```

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 `OBOextractNoGender1yr.csv`

To load the empirical data on categories use:

```
CatEmp =
```

```
oboM.loadCatEmp('Data/Raw/OBOextractNoGender1yr.csv')
```

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:

```
CatDepMod = oboM.generateDependentModelLaplace(
    oboM.changePeriod(CatEmp, Delta=2))
```

For model selection, as many models as reasonable need to be generated. This is achieved *en masse* by the `generateModelFiles` as:

```
oboM.generateModelFiles(CatEmp, Type='All')
```

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:

```
AICtable = oboM.generateAICstandardModels(
    ReturnFrame=True)
```

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:

```

    ConfLowFrame, ConfHighFrame =
oboM.generateBootStrapEstimateConfidenceIntervals(
    CatEmp, Delta=2)
These confidence frames can be easily saved as:
    ConfLowFrame.to_csv('Data/Bootstrap/
OB00ffencePunishmnetConfidenceLow2yr.csv')
and
    ConfHighFrame.to_csv('Data/Bootstrap/
OB00ffencePunishmnetConfidenceHigh2yr.csv')
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:
    CatDepModOffence = oboM.combineOnOffence(CatDepMod)
and
    CatDepModPunishment =
    oboM.combineOnPunishment(CatDepMod)
The partitioning is best performed using the Cython version of
OB0Partitioning.py, OB0partitioning.pyx which is loaded
and run as:

```

```

import pyximport
pyximport.install(reload_support=True)
import OB0partitioning as oboPc
AICFrame, DeathProbFrame =
    oboPc.generateAICtable(CatEmp)
Once the partitioning frames have been generated they need
to be sorted before carrying out the MCMC run:
    SortedIndex, SortedValues, DeathFrameSorted =
    sortAICtable(AICFrame, DeathEstimateFrame =
    DeathProbFrame, SortDeathEstimate=True)
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:
    oboPc.mcmcGenerateAll(SortedValues, SortedIndex,
    DeathFrameSorted)
These steps will allow the complete reproduction of this study.

```