Extralegal Determinants of Criminal Sentencing: The Case of British Columbia, 1864-1913[*]

Kris Inwood

University of Guelph

kinwood@uoguelph.ca

Ian Keay

Queen's University

keay@queensu.ca

Blair Long

University of Wisconsin-Milwaukee

longbp@uwm.edu

Job Market Paper

Latest Version

October 25, 2022

Abstract

What factors determine criminal sentences? While legal factors such as crime and criminal history should affect punishment, judges may also incorporate extralegal factors when handing down sentences. In this paper, we study the role that extralegal determinants played in the sentencing of criminals in British Columbia (BC) between 1864 and 1913. Using prison admissions data, we document the sentencing behaviour of judges and find more leniency towards women, Indigenous, and Chinese individuals. We find harsher sentences for the lowest and highest social classes. Over time, we find that these biases shifted, concurrent with significant historical events. A sentiment analysis of historical BC newspapers shows that public sentiment mirrored this pattern. To distinguish between taste-based and statistical bias we first estimate prisoners' predicted future recidivism and incorporate this into our main specification. We find that statistical bias is present but small in comparison to the extralegal bias we initially observe. We next augment Becker's (1968) model of punishment to incorporate both

1

channels of bias. We test implications of the model in a triple difference specification using the building of the Canadian Pacific Railway as an exogenous decrease in the incentives for crime. We detect substantial taste-based-bias against Chinese prisoners.

1 Introduction

A prominent area of study in the economics of crime has focused on the determinants of judicial sentencing decisions (Levitt and Miles, 2007). In particular, researchers have explored the extent to which *legal* and *extralegal* factors are incorporated into sentencing decisions across time and institutional contexts (Shermer and Johnson, 2010; Radzinowicz and Hood; 1978; Bushway and Piehl, 2001). In this paper, we use data from historical provincial prison admission ledgers to document the criminal sentencing behaviour of judges in the Canadian province of British Columbia (BC) during the late nineteenth and early twentieth centuries. We combine novel data, methods from machine learning, and economic theory to evaluate the role of extralegal factors in sentencing and the extent to which these extralegal factors reflect taste-based versus statistical bias.

Since Gary Becker's (1957) ground-breaking work on discrimination in labour markets, economists have worked to broaden our understanding of bias (Phelps, 1972; Arrow, 1973; Bertrand et al., 2005; Bohren et al., 2022), and the analysis of discrimination has been extended into a wide range of social, economic, and political settings (Lang and Kahn-Lang Spitzer, 2020). One area of particularly intensive study has been bias in the criminal justice system. Much of this work focuses on sentencing biases related to race and ethnicity (Mustard, 2001; Blair et al., 2004; Spohn et al., 1981; Burch, 2015), gender (Godfrey et al., 2005; Starr, 2015; Sorensen et al., 2014), social or economic class (Vickers, 2016), and other observable physical characteristics (Leventhal and Krate, 1977; Johnson and King, 2017). The presence and impact of bias within the criminal justice system has been studied in the context of jury composition (Anwar and Fang, 2006; Anwar et al., 2012, 2019), judicial traits (Cohen and Yang, 2019; Philippe, 2017; Schanzenbach, 2005), bail decisions (Arnold et al., 2018), and prosecutorial behaviour (Rehavi and Starr, 2014).

We focus on two tasks. First, we document patterns in average sentences between 1864 and 1913, identifying extralegal biases in sentence severity while controlling for a broad set of legal and extralegal factors. Our empirical specifications are based on sentencing equations that are conceptually consistent with both rational choice (Becker, 1968) and behavioural (Gelbach, 2021; Teichman and Zamir, 2014) judicial decision-making models. We employ LASSO (Tibshirani, 1996) model selection criteria to determine the maximally predictive determinants of sentences. In doing so, we take a 'data-driven' approach to characterizing sentencing behaviour, rather than choosing explanatory variables based on (potentially biased) *a priori* modelling assumptions. Our sentencing equations reveal systemically shorter sentences for prisoners of Indigenous-origin and women over the full 1864-1913 period, even after controlling for admission year, crime class, institution, and other observable character-

1

istics. We also find that the Chinese received shorter sentences, unless they were admitted during the 1876-1886 railway building boom in BC. In contrast, individuals who reported working in the lowest *and* highest skilled occupations prior to their incarceration received relatively long sentences on average.

We explore how these biases in sentencing have changed over time by estimating our empirical specification as a rolling regression through nine-year estimation windows. The time varying estimates allow us to document how judges' marginal sentencing decisions changed over our period of study. When we do this several interesting time series patterns emerge. The tendency for judges to impose relatively short sentences on Indigenous prisoners is present over the entire period of study, but perhaps surprisingly, it becomes stronger after the passage of the *Indian Act* by the Canadian federal government in 1876. With respect to female prisoners, in a modern context it is generally accepted that women receive shorter sentences than men for any given crime (Godfrey et al.; 2005; Ahola et al., 2009; Starr, 2015; Sorensen et al., 2014).[1] In our historical setting, we confirm that on average women received significantly shorter sentences than men, but the strength of this bias rose and fell through the late nineteenth and early twentieth centuries, becoming statistically indistinguishable from zero by the turn of the twentieth century. We also find that the leniency shown to the Chinese can only be identified during the years after 1885. Prior to this – most notably during the construction of the Canadian Pacific Railway (CPR) – the Chinese prisoners in BC's jails faced sentences at least as harsh as the average prisoner. For inmates who reported holding either relatively low skilled *or* relatively high skilled occupations, we find that the marginal sentence effects fell during the colonial period (1864-1871), before rising sharply as modern criminal justice institutions emerged in BC, eventually settling near zero by the end of our period in 1913. We conduct a complimentary sentiment analysis of historical newspapers and conclude that these sentencing biases closely tracked changes in public sentiment over our sample period.

Next, we conduct tests to distinguish between taste-based and statistical bias. Our empirical setting allows us to address well-known flaws of outcome tests of bias (Becker, 1957). First, *omitted variable bias* occurs when we do not have data on relevant legal determinants of sentences, and thus may detect extralegal biases that arise from underlying correlations between legal and extralegal factors (Ayres and Waldfogel, 1993). Our dataset is rich in possible regressors, including education and occupation information, criminal history, and even physical characteristics. We employ LASSO model selection to identify relevant

---

[1]We note that this gender-based leniency is not necessarily universal – several studies have found that differences in sentencing outcomes for men and women can be context-dependent (Kruttschnitt and Savolainen, 2009; Crew, 1991)

sentencing determinants and this, we argue, should allay concerns about the effect of omitted variables. Further, the panel dimension of our data allows use to estimate predicted future recidivism at the individual level. We can incorporate this in our main specification — analogous to the risk scores that are being employed in modern courts (Sidu, 2015; Albright, 2019) — to control for, what we argue, is the most important unobservable legal factor. Another well-known issue with outcome tests is *inframarginality bias*, which occurs when we only observe average outcomes, but know that a decision maker is operating on the margin (Mechoulan and Sahuguet, 2015). In our setting, this concern is reduced, because we are directly estimating marginal sentencing effects on a continuous sentencing variable.

In our first test, we exploit the panel nature of our data to predict future recidivism for each prisoner based on their observable characteristics. We then include the predicted probability of future recidivism as an additional explanatory variable in our sentencing equations. This approach allows us to detect evidence of statistical bias, but on average its impact is small. The illustrative example of sailors demonstrates the potential for statistical bias to have a substantive effect on sentencing – their predicted future recidivism is low because of the transient nature of their residency in BC, and only after controlling for this source of statistical bias, do we find that sailors' sentences were indistinguishable from the average prisoner. For Chinese prisoners, sentencing leniency remains at conventional levels of statistical significance after controlling for predicted future recidivism, but the size of the bias is reduced by approximately 15 percent (2.3 log-points). For Indigenous prisoners the result is similar – including the probability of reoffending reduces their differential sentences by about 12 percent (3.5 log-points). For women and prisoners at both the bottom and top of the skill distribution in BC, we observe small *increases* in bias once we control for the fact that female prisoners were slightly more likely to reappear in the prison ledgers after their first offence, and those with relatively low and high skilled occupations were slightly less likely to reappear.

We next augment a Beckerian model of optimal punishment to include channels of taste-based and statistical bias in criminal sentencing which differ across groups. The model admits an equilibrium relationship between sentencing and recidivism, and we demonstrate that it is sensitive to changes in both the judicial taste-based bias and criminal incentives facing groups. We use the model's predictions to discipline an empirical analysis where we use the building of the CPR as an exogenous shock to criminal incentives. The model predicts that when incentives for crime are lower, we should observe a corresponding decline in sentences. Results from our triple difference (TD) design show the opposite — that during the building of the CPR, sentences rose substantially for Chinese as compared to

North American-born workers. This finding indicates the presence of taste-based bias that negatively affected Chinese workers.

There are some important limitations of note in our analysis. First, we conduct an outcome test of bias in a setting with a continuous decision variable. In fact, this decision variable is the direct outcome we study. In this regard, we cannot recast this setup as a Roy model (Canay et al., 2020). The implications of the work of Canay et al. for our setting are thus unclear, and we leave them as a suggestion for future work. Similarly, another alternative approach to test for taste-based preferences is the rank-order test developed by Anwar and Fang (2006) and employed by Park (2017). Unfortunately, because we cannot identify the race of judges in BC, and because there was almost certainly no racial variation available for identification, we cannot adopt this approach. This limitation also rules out applying the tests described by Knowles et al. (2001) and Antonovics and Knight (2009).

We view our contributions in this paper as threefold. First, we provide new evidence on the determinants of criminal sentencing, highlighting the role of extralegal factors. In doing so, we adopt a novel approach, relying on LASSO to let the data speak for themselves. We also contribute to furthering our understanding of discrimination in criminal justice in a unique historical setting. These applications are important because historical settings often differ from contemporary ones along lines of interest. For example, in our context, we are able to study bias in criminal sentencing in an era when judicial discretion was higher than it is today (Inwood et al., 2022). Similarly, the building of the CPR in BC during our sample period allows us to use exogenous variation in incentives for crime to identify statistical bias – a natural experiment typically unavailable in modern settings. Finally, we formalize the analysis of extralegal bias in criminal sentencing using Becker's canonical framework, setting the stage for future, perhaps structural, analyses.

The rest of the paper proceeds as follows: section 2 describes our data and empirical specification; section 3 documents baseline sentencing patterns, pooling data over our sample period; section 4 documents how these patterns have changed over time alongside public sentiment; section 5 tests for the presence of taste-based versus statistical discrimination; and section 6 concludes.

2   Data and Empirical Specification

In this section we describe the evidence that can be drawn from BC's provincial prison admission ledgers, sample selection issues are discussed, our main variables of interest are defined, summary statistics are provided, and we include a description of our sentencing

equations and estimation strategies.

2.1 British Columbia's Provincial Prison Admission Ledgers

Admission records have survived and been manually digitized from the original hand written, bound ledgers for the provincial prison in Victoria – the colonial and provincial capital – from its opening in 1864 until Canada entered World War 1 (WW1) in 1914.[2] For the provincial prison in New Westminster – now a suburb of Vancouver, on BC's 'lower mainland' – records are also available from its opening in 1875 to 1914. Other ledgers have been digitized from 1911 to 1919 for smaller prisons in Nanaimo, Saanich, and Quesnel, although the evidence presented in this paper only uses entries from Victoria and New Westminster up to 1910, with a small number of entries from Nanaimo added for some specifications from 1911-1915. In total 33,296 entries have been digitized, describing 28,200 prisoners. To ensure sample consistency over time, our baseline specifications use only those admission records that survive the imposition of three sample restrictions: (i) we drop all records for individuals who were not found guilty of a criminal offence (approximately 12 percent of entries); (ii) we drop admission years after 1913 (4.5 percent of entries); and (iii) we drop entries that include sentences that were less than seven days or more than two years (approximately 30.5 percent of entries).[3]

The first sample restriction removes records for individuals who found themselves in prison awaiting trial or bail, those being kept for safe-keeping, witnesses, and 'lunatics'. Over our period of study these types of entries become increasingly rare and none of these individuals received sentences for criminal offences. The second restriction is imposed because the number of prisoners admitted into BC's provincial prisons after the start of WW1 falls sharply, and the characteristics of those admitted after the initial wave of WW1 enlistment (and later forced conscription) change abruptly.[4] The third sample restriction is imposed in an effort to maintain a common domain for the distribution of sentences over our period of study. Kolmogorov-Smirnov tests for common sentence distributions in consecutive years between 1864-1913 confirm that there were significant changes in the upper tail (sentences greater than two years) during the late 1870s, and the lower tail (sentences less than seven days) during the mid-1880s. These shifts in the sentencing distributions coincide with the opening of a federal penitentiary in New Westminster in 1878, and the reallocation of all

---

[2]An extended discussion of the data included in the ledgers can be found in Inwood and Keay (2021), and Inwood et al. (2022).

[3]When we explore the heterogeneity in our estimated marginal sentencing effects, we reintroduce the records for prisoners with sentences in the bottom and top tails of the sentencing distribution – see Table 2.

[4]There is one exception to the pre-WW1 sample restriction – the last estimation windows in our rolling regressions add entries from Naniamo for 1914 and 1915.

prisoners who were sentenced by local Police or Provincial Courts from the provincial prisons to local police lock-ups during the 1880s (Inwood et al., 2022). The imposition of our three sample restrictions leaves us with 17,586 admission records for our baseline specifications. These records comprise the universe of prisoners in British Columbia who were serving sentences longer than seven days, but less than two years between 1864 and 1913. Summary statistics for sentences, recidivism, and observable prisoner characteristics are reported in Appendix Table A1 for the full and restricted samples.

We recognize the value of combining prison admission records with police arrest records, court records, or prison log books, but in our historical context these other data sources are simply unavailable. We do not know anything about individuals who interacted with the police in BC during the late nineteenth and early twentieth centuries if they did not end up in prison. We do not have access to any crime or trial-specific details, including information about the presiding judge, prosecutor, or jury composition. Except in unusual circumstances, we do not even know about prisoner's behaviour in prison, or their release dates. The absence of these other sources of information limit our ability to fully characterize systemic bias as described by Bohren et al. (2022) throughout the criminal justice system in BC during our period of study. However, the granular detail in the admission ledgers does allow us to explore aspects of criminal sentencing, including taste-based and statistical discrimination in sentencing, that cannot be identified in even the richest modern evidentiary settings.

2.2 Hard Labour-Equivalent Sentences

86 percent of all admission records for BC's provincial prisons from 1864-1913 record a 'sentence'. Virtually all entries that do not include a sentence are for individuals who have not been found guilty of a crime, and they have been dropped from our restricted sample. For the prisoners with a sentence recorded, there is a remarkable range of highly detailed information provided, including the value of costs, damages and fines imposed, length of imprisonment, hard labour requirements, life sentences, death by execution, and various other forms of punishment, such as flogging, bread-and-water, solitary confinement, and half rations.

To combine the ledgers' sentencing details into a single, consistently defined measure of sentence severity that can be compared across prisoners, crimes, and time, we convert all sentences into *hard labour-equivalent days* of imprisonment. To derive hard labour-equivalents we first group sentences into categories that include: costs, damages or a fine; imprisonment; hard labour; some form of punishment; and life in prison or death sentences. Within these categories, we then distinguish sentences that include more than one of these penalties (AND

sentences), and sentences that include options (OR sentences).

The OR sentences gave prisoners a choice among penalties, and they reveal the relative values, in terms of equivalent severity, that judges in BC placed on the specific penalties and sentence components. Judges in BC, for example, consistently viewed one day of hard labour as approximately equivalent to two days of regular imprisonment; each stroke from the lash or whip was equivalent to about 4.5 days of hard labour; and each day of solitary confinement, bread-and-water, or half rations was worth about two days of hard labour.[5] We consider the 41 life and death sentences recorded in the ledgers to have been equivalent to 35 years of hard labour, based on the average age at admission for those receiving life and death sentences – approximately 30 years – and an assumed average life expectancy of 65 years.

Unfortunately, using OR sentences to value costs, damages, and fines in equivalent days in prison is not possible because these options become extremely rare during the second half of our period, and even during the 1860s and 1870s the fine-per-day equivalents were not consistent across crimes or admission years. We adopt an approach to assigning hard labour-equivalents to dollar values that is similar to that used to value a statistical life (Ajzenstadt, 2002; Goldin and Lewis, 1975). Specifically, we use low-skilled, semi-skilled, and high-skilled daily wages from Vancouver and Victoria to value costs, damages, and fines in terms of forgone days of labour.[6]

The sentences depicted in Figure 1 represent five-year moving averages of unconditional annual hard labour-equivalent days of incarceration for each admission year from 1864 to 1913 for all prisoners included in our restricted sample, female prisoners, Indigenous prisoners, and those who held low skilled occupations prior to their incarceration. Between 1864 and 1885, the average prisoner in BC's provincial prisons was sentenced to the equivalent of 70.8 days of hard labour. This rose to 114.7 days of hard labour between 1886-1913. When valued in terms of nominal skill-specific wages, over the whole period the average sentence is equivalent to $236.62 in forgone income, or slightly more than double Canadian GDP per capita in 1885. We also note that the overall average sentences before 1885 are significantly lower than the average sentences after 1885; there is no significant linear time trend in sentences before,

---

[5]Women, Indigenous prisoners, and Chinese prisoners rarely received sentencing options, in part because their sentences rarely included hard labour or punishment. To confirm that our qualitative conclusions are not dependent on our use of penalty equivalents derived from non-Indigenous male prisoners' sentencing options, we also use total days of incarceration (without converting to hard labour-equivalents), and regular days of incarceration (dropping all sentences that include hard labour or punishment). Full results using these alternate sentence measures are available upon request from the authors.

[6]Skill and city specific wages have been calculated from Emery et al. (2007); Inwood and Irwin (2002)); Urquhart and Buckley (1965); Belshaw (2009); and Harris et al. (2015).

or after 1885; and the annual coefficients of variation across prisoner-specific sentences are significantly higher during the years before 1885 (Inwood et al., 2022). When averaged over all crimes, prisoner characteristics, and admission years, female prisoners' sentences were nearly 33 percent lower than the average prisoner, and Indigenous prisoners' sentences were 32 percent lower, while the relatively low skilled prisoners' sentences were 7 percent higher. In what follows we estimate these sentencing differentials conditional on admission year, crime class, gaol, and observable prisoner characteristics.

2.3 Classifying Criminal Offences

In addition to the remarkable level of granular detail in the ledger entries describing prisoners' sentences, just over 96 percent of the entries also provide a reason for admission. To organize the information in the ledgers that describes an 'offence' or reason for incarceration, we first categorize crimes that fall within one of 12 broad classes that roughly match the offence categories specified in the 1892 federal *Criminal Code* (Inwood and Keay, 2021; Inwood and Roberts, 2020). We also include an 'All Other' broad crime class to capture all offences specified in the ledgers that do not obviously fall within any of the *Code's* categories, and we include an additional group for all offences that in some way specify Indigeneity (mostly related to violations of the *Indian Act*). The most common broad crime classes specified in the ledgers include offences related to alcohol and drugs, which account for over 10 percent of the prisoners serving between seven days and two years; property crimes, which account for about one quarter of all entries; and crimes against public order, which make up 22.5 percent of our restricted sample of ledger entries. We note that even when we include sentences in the top tail of the sentencing distribution (those longer than 2 years), crimes involving violence are quite rare, accounting for just slightly more than 2,000 prison admissions.

Of course, 14 broad categories do not come close to exploiting all the information included in the ledgers relating to prisoners' crimes. In our sentencing equations we further disaggregate the broad crime classes into 76 much more narrowly defined classes. To provide an illustrative example consider the broad category of 'property' crimes, which we divide into six narrower classes that include: arson; attempted theft; damage to property; possession of stolen property; theft; and property crime with violence. While it is possible that significant variation in the severity of prisoner-specific crimes remains even within our narrow classes, without police or trial records we cannot generate more disaggregated measures of crime severity.[7] We do note, however, that the level of crime class detail available in the BC

---

[7]We note that it does not appear that women or Indigenous prisoners necessarily committed less serious

provincial prison admission ledgers far exceeds what is typically documented in studies of historical judicial sentencing (Bodenhorn, 2009; Bindler and Hjalmarsson, 2019; Vickers, 2016).

2.4 Prisoners' Observable Characteristics

In addition to the finely detailed, prisoner-specific information about crimes and sentences, BC's historical prison admission ledgers also provide us with a wealth of information about prisoners' observable characteristics. Without photographs or other means of biometric identification, detailed written descriptions of prisoners' physical appearance, temperament, and occupation were the only source of identification available to authorities during our period of study (Antonie et al., 2021).

Birthplace, nationality, or ethnicity is recorded for over 90 percent of all prisoners in the admission ledgers. Indigenous prisoners, for example, make up nearly 40 percent of all admissions before 1885, but that share drops to just seven percent after 1885. The average new inmate in BC's provincial prisons was about 32 years of age, and well over 90 percent of the prisoners were men. There was a small decline in the proportion of female prisoners after the early 1880s that is coincident with the decline in Indigenous prisoners – 56 percent of all female prisoners were Indigenous.

Height in feet and inches is recorded for nearly 90 percent of the prisoners in our sample, and we find that on average the BC prisoners were approximately the same height as their criminal peers in Ireland, England, and Pennsylvania (Bodenhorn et al., 2012; Steckel, 2008). Perhaps a bit surprisingly, literacy was quite high among the prisoners in BC during the late nineteenth and early twentieth centuries, averaging slightly over 60 percent. About half of the prisoners were identified as Protestant, 25 percent Catholic, and nearly 20 percent claimed no religious affiliation.[8]

The ledger entries often include some indicator of the prisoner's health when admitted. The average prisoner was typically described as being healthy, however, over time the entries suggest some worsening in health status – 'grey', 'sallow', or 'haggard', become much more common descriptors near the end of our period. The prisoners' skin tone, hair, and eyes are also described, and we find that the share of prisoners with a light complexion, blond hair,

---

crimes either across, or within crime classes. Consider the example of property crimes – the costs and damages assessed for female and Indigenous prisoners within this crime class exceed those assessed for the average prisoner by $48 and $17, respectively.

[8]'No religion' is associated with an explicit entry in the religion field in the ledgers – typically 'atheist', 'agnostic', or 'none'. Blank ledger entries for all descriptive variables are coded as 'NA'.

and blue eyes was rising, while the decline in Indigenous prisoners coincided with a drop in prisoners described as having dark skin, black eyes, and black hair.[9]

More general descriptions of the prisoners' bodies are included in listings describing their 'proportion' or 'build'. The prisoners' conduct is consistently recorded as either 'temperate' or, for those prisoners who were known to have problems with alcohol or drugs, 'intemperate'.[10] The ledges also list prisoners' distinguishing marks – a category dominated by descriptions of scars and tattoos.[11]

A 'trade' is listed for about 80 percent of the admission entries. We group these trades into occupations based on the categories used in the 1901 Canadian census (Inwood and Keay, 2021), and we aggregate the occupations into skill groups that roughly correspond to social classes – high skilled, semi-skilled, and low-skilled (Bodenhorn, 2009; Reiss, 1965). The prisoners are split fairly evenly between low-skilled and semi-skilled occupations, although the lower skilled appear more often early in the period. Among the specific occupations recorded in the ledgers, the most common include low skilled labourers, domestics, and those with 'no occupation'. We do, however, see an increase in the share of semi-skilled fishermen, loggers, miners, and manufacturing workers after 1885, and among the high skilled prisoners recorded in the ledgers, merchants are the most common. Occupations are rarely recorded for the female prisoners, and the Indigenous and Chinese prisoners are overwhelmingly low skilled.

A final characteristic that can be identified from the ledgers is recidivism. On occasion repeat offenders are explicitly noted in the ledgers, but we rely on prisoner identification numbers, names, and time invariant prisoner characteristics – including birth year, gender, and ethnicity – to identify individuals who appear in the ledgers more than once. In total, nearly one-quarter of all prisoners are recorded multiple times. On average recidivists appear in the ledgers 2.6 times, and one remarkably recalcitrant prisoner – Lawrence M., an Irish-American who made a career from theft and selling alcohol to Indigenous individuals, thereby violating the *Indian Act* – appears a 44 times. We record how many times each prisoner appears in the ledgers in total, how many times each new entrant has already appeared (past recidivism), and how many times each new entrant will appear at a later date (future recidivism).

---

[9]We use Block (2018) to group these 'colonial complexion' descriptors.

[10]On occasion the descriptor 'intemperate' was also used for prisoners who were disruptive, but this was often recorded explicitly in the ledgers.

[11]We group distinguishing marks by both specific body parts, and by 'sympathetic marks', such as disabilities or injuries, and 'unsympathetic marks', such as pock marks, scars, or bad teeth.

2.5 Empirical Specification and Estimation

To assess conditional sentencing patterns across crime classes, prisoner characteristics, and time, we specify sentencing equations similar to those used in similar historical contexts by Howard Bodenhorn (2009), and Anna Bindler and Randi Hjalmarsson (2020). Our equations allow us to identify the impact of three sets of sentencing determinants on judicial decision-making. The first set of determinants include formal, codified legal factors, such as crime class, past recidivism, and sentence structure – fines, damages, punishment, or hard labour, for example. The second set of covariates includes the extralegal determinants that judges might have used as proxies or signals for future recidivism risk, danger to the public, or the denunciation and deterrence effect of their sentences. These extralegal determinants include prisoner characteristics, such as ethnicity, gender, height, occupation or social class, and physical or distinguishing features. With this set of determinants we also include gaol fixed effects to control for prison-specific factors, such as capacity constraints or prison conditions, that may affect sentencing across all crimes and admission years. The third set of determinants are captured by admission year fixed effects, and they include all time varying, system-wide factors, such as fiscal constraints or economic conditions, that affect all criminals in a given year. In some specifications we interact crime class and admission year fixed effects to control for time varying shifts in average sentences across crime classes. Our sentencing equations take the form:

$$log(Sentence_{it}) = \psi_t + \beta_t \mathbf{Y_t} + \delta_t \mathbf{LD_{it}} + \gamma_t \mathbf{XLD_{it}} + \epsilon_{it} \qquad (1)$$

Where: i indicates an admission record; t indicates an admission year; $\mathbf{Y}$ are admission year fixed effects capturing system-wide determinants that vary over time, but (proportionately) affect all sentencing decisions in a given year equally; $\mathbf{LD}$ are legal determinants that include fixed effects for 76 narrow crime classes, sentence structure, and past recidivism; $\mathbf{XLD}$ are extralegal determinants, including fixed effects for each prison, and up to 80 observable prisoner characteristics; and ϵ represents a residual that captures unobservable idiosyncratic sentence determinants that are uncorrelated with Y, LD, or XLD. The parameters to be estimated in our sentencing equations $(\psi, \beta, \delta, \gamma)$ have time subscripts because some specifications use rolling regressions with nine-year estimation windows to identify time-varying marginal sentencing effects.

All sentences are measured in hard labour-equivalent days in prison. We use the natural logarithm of sentence length for the dependent variable because even with our sample restrictions, sentences are skewed towards longer terms. All legal and extralegal sentence

determinants are categorical, taking the value 1 for prisoners with the given characteristic, crime class, gaol, or sentence structure, 0 otherwise (Inwood et al., 2022). All sentencing equations are estimated by OLS, with standard errors clustered at the intersection of crime class and admission year – 337 clusters when estimated over all admission years, with 31 admission records in the median cluster.[12]

The parameter estimates from our sentencing equations allow us to identify the marginal sentencing impact of each legal, extralegal, and system-wide determinant. Omitted prisoner characteristics are grouped into the constant ψ_t, such the estimated marginal impact of any included factor is measured relative to the 'average prisoner', as characterized by the omitted observable characteristics. This structure means that the marginal sentencing effect for any included determinant may depend on the inclusion of up to 80 extralegal covariates that can be extracted from the admission ledgers. Of course, many of the prisoners' observable characteristics are strongly collinear – Indigeneity, dark skin, black hair, and black eyes, for example, are all strongly correlated across prisoners. Other characteristics are uncorrelated with sentencing – the marginal effect of literacy on prisoners' sentences, for example, is both small and statistically indistinguishable from zero. As a result of these considerations, our preferred specification uses a parsimonious version of our sentencing equation that includes only the observable prisoner characteristics that are both relatively uncorrelated with the other right-hand-side variables, and strongly predictive of prisoner-specific sentencing outcomes.

To select covariates for inclusion in our baseline, parsimonious sentencing equation, we rely on a data-driven approach that uses a least absolute shrinkage and selection operator (LASSO) to pick out prisoner characteristics with relatively high predictive power and low covariance with the other sentencing determinants (Derenoncourt, 2019; Tibshirani, 1996). After partialling out crime class, admission year, and gaol fixed effects, our LASSO estimates guide us in the selection of five legal determinants and 18 extralegal determinants for inclusion in our parsimonious specification. Again, we emphasize that for our baseline estimates many prisoner characteristics are grouped into the omitted category captured by the constant term in the equation. For our purposes, aggregating characteristics in this way expands the characterization of the 'average prisoner', but it does not affect any of the qualitative conclusions related to the marginal sentencing effects that are of interest to us – specifically, sentencing differentials related to gender, ethnicity, and social class.

---

[12]Following MacKinnon et al. (2022) standard errors have been clustered along multiple dimensions and levels of aggregation. Our qualitative conclusions regarding the prisoner characteristics of interest are not dependent on clustering strategy. Standard errors using alternate clustering strategies are available upon request from authors.

3   Documenting Extralegal Bias in Criminal Sentencing

We turn now to the results from our estimation of variations of the specification described by Equation (1). The parameter estimates reported in Table 1 represent fixed marginal sentencing effects, estimated over all 1864-1913 admission years. Of particular interest are the marginal sentencing effects associated with prisoners who were Indigenous, Chinese – differentially identifying the impact of Chinese during the 1876-1886 railway building boom in BC – female, sailors, and those who reported being employed in either a low skilled or high skilled occupation prior to incarceration.[13] In the first column we include no additional conditioning information. In column (2) we control for differences in average sentences across crimes, time, and prison, by including narrow crime class, admission year, and gaol fixed effects. Column (3) adds crime class × admission year interaction terms, which control for changes in crime-specific average sentences over our period of study. The last column in Table 1 reports the results from our preferred, parsimonious specification, which includes crime class, admission year, and gaol fixed effects, and controls for the prisoners' observable characteristics – the legal and extralegal sentence determinants identified by our LASSO model selection criteria. The consistency in the signs and magnitude of the marginal sentencing effects across columns (2)-(4) give us confidence in our interpretation of the results. The parameter estimates reported in column (4) illustrate several noteworthy instances of bias in criminal sentencing, even after controlling for differences average sentences across narrowly defined crime classes, admission years, gaol, legal determinants, and prisoners' other observable characteristics.

Similar to much of the modern evidence on discriminatory criminal sentencing, we find that even during the late nineteenth and early twentieth centuries, women in BC's provincial prisons received sentences that were approximately 15 percent $(= e^{-0.158}-1)$ shorter than the average prisoner with similar characteristics, for a given crime. This gender-based leniency is typically interpreted as evidence of paternalism in sentencing, and it is consistent with the notion that judges in BC viewed women as somehow 'more delicate' than men, and as such, the welfare 'cost' of each day in prison would have been considered relatively high for women (Bindler and Hjalmarsson, 2020; Daly, 1989; Moulds, 1978; Gruhl et al., 1984; Chesney-Lind, 1977).[14] A related idea found in the literature on modern sentencing is that single men receive harsh sentences because they have no wives or families to support, so the welfare cost of their sentences will be relatively low. We also observe this form of reverse-paternalism

---

[13]The 'sailors' also include a small number of military occupations. We note that 'no occupation' is coded as low skilled, and it is associated with a specific entry in the ledgers, such as 'vagrant' or 'unemployed'.

[14]Consistent with our finding for all Indigenous prisoners, when we separately identify Indigenous women, we find that they received significantly lower sentences than non-Indigenous women.

in the parameters reported in Table 1 – unmarried prisoners in BC's jails received sentences that were nearly 4 percent longer than the average prisoner.

Given the casual racism expressed in many of the ledger entries, we surprisingly find that, like the female prisoners, Indigenous prisoners received much shorter sentences relative to the average prisoner with similar characteristics, for a given crime. The marginal sentencing coefficient associated with being Indigenous implies a reduction in hard labour-equivalent days of incarceration of approximately 3.0 log-points, which is equivalent to about 26 percent of the average sentence. Similarly, the marginal sentencing effect associated with being Chinese reveals another interesting pattern. While these individuals also received shorter sentences on average, (about 15 percent fewer days), during the building the Canadian Pacific Railway when temporary migration of Chinese railway workers into BC was very high, Chinese prisoners received sentences that were more than 28 percent *longer* than the average prisoner.

We suggest that these unexpected marginal sentencing effects associated with Indigenous and Chinese prisoners reflect a distinct form of historical paternalism.[15] Similar to the narrative typically associated with the sentencing of women, our results are consistent with the notion that between 1864 and 1913, judges in BC felt that the net social value of imposing harsh sentences on Indigenous and Chinese prisoners was relatively low. This may have been because judges considered the hardship of prison time to be particularly high for these individuals, or perhaps the denunciation and deterrence effect of their sentences may have been perceived to be particularly low.

Other possible instances of paternalism in the sentencing patterns revealed in the last column in Table 1 include significantly shorter sentences for prisoners described in the ledgers as short, small or thin, bald, and intemperate – a descriptor typically associated with drug or alcohol addiction. These marginal sentencing effects are again consistent with leniency directed towards those who were, or at least appeared to be, older, less healthy, or socially disadvantaged – in other words, those for whom the social value of a harsh sentence could have been perceived to be particularly low. Along yet another dimension, reflecting a typical life-cycle profile of criminality (Bindler and Hjalmarsson, 2017), we also find that younger prisoners received sentences that were significantly longer than the average prisoner with similar characteristics.

Turning to measures of occupational skill, which we interpret as a proxy for social class,

---

[15]The paternalism shown towards Indigenous and Chinese prisoners in BC echoes the marginal sentencing effect for black prisoners in Pennsylvania that Bodenhorn (2009: 294) found during the late nineteenth century.

we find that relative to the omitted semi-skilled category, prisoners who reported holding either a low skilled or high skilled occupation prior to incarceration, received sentences that were approximately seven percent longer than the average prisoner.[16] It seems that paternalism in sentencing did not extend into the labour market or social hierarchy at the end of the nineteenth and beginning of the twentieth centuries in British Columbia – clerks, merchants and white collar workers who had 'fallen from grace', and those who worked in jobs at the very bottom of the skill distribution faced similar, relatively harsh, sentencing biases.[17]

A final result we wish to highlight from Table 1 foreshadows an interesting illustrative example we will be describing in more detail below – after controlling for admission year, crime class, gaol, legal determinants, and other observable characteristics, sailors received sentences that were over seven percent shorter than the average prisoner. Given their reputation for disruptive behaviour and their low position in the social hierarchy, this result seems surprising. However, as we shall see, after we account for sailors' likelihood of reoffending, our results align more closely with our expectations.

The sentencing patterns that can be identified from the estimates reported in Table 1 do not necessarily imply that the BC criminal justice system was not discriminatory during our period, but rather that discrimination took complex and subtle forms, often reflecting a deep sense of ethnic superiority on behalf of BC's judges and the system as a whole. There were clear limits on the paternalistic leniency shown towards Indigenous and Chinese prisoners – this leniency was context-dependent and applied only to certain types of crimes, observable characteristics, and sentence structures. We also emphasize that although we cannot observe it in the prison admission records, reverse-paternalism may have been at work in other ways and at other stages in the administration of criminal justice in BC. Biases in sentencing may have simply been a reaction to other forms of discrimination.[18]

In Table 2 we explore the boundaries of sentencing bias and paternalism in our historical setting by reporting on heterogeneity in the marginal sentencing effects for sailors, Indigenous, Chinese, female, and high and low skilled prisoners. We consider heterogeneity

---

[16]The results related to prisoners' occupations should be interpreted with some caution because, relative to the other characteristics of interest, there are fewer entries recording an occupation. Our qualitative conclusions are robust to the exclusion of entries with no occupation recorded.

[17]When we interact low skilled occupations with Indigeneity, we again find that paternalism in sentencing did not extend into this class of prisoner – low skilled Indigenous prisoners' sentences were significantly higher than the average sentences for all Indigenous prisoners.

[18]Access to police or court records would allow us to identify biases that are unobservable in the admission ledgers. Because we only observe prisoners, we cannot assess ethnicity-based bias in policing, the decision to prosecute, or findings of guilt or innocence, for example.

along multiple dimensions, including in columns (1)-(3), variation in marginal sentencing across the more common crime classes – property crimes, crimes involving drugs or alcohol, and the most severe crimes that involved harm or violence.[19] In columns (4)-(7) we identify differences in marginal sentencing for prisoners across the distribution of sentences, looking specifically at those above and below the median sentence in our restricted sample, and the prisoners in the upper (sentences greater than two years) and lower (sentences less than or equal to seven days) tails who are dropped from the restricted sample that we use in our baseline specifications.[20] In the final column in Table 2 we report marginal sentencing effects for all prisoners (including those with sentences less than seven days or more than two years) for the admission years during which the Victoria Gaol was the only prison in BC. This unrestricted sample used in column (8) includes the entire universe of prison admissions in BC from 1864 until 1875, and as such, it is free of possible sample selection issues that might arise from our reliance on a restricted sample of admission records in our baseline estimates.

In general, the marginal sentencing patterns hold within the largest crime classes. Among property crimes, Indigenous and Chinese prisoners' sentences are again relatively short, and high skilled prisoners' sentences are relatively long, but the marginal sentencing effect of the other characteristics – women, low skilled, sailors – are less precisely estimated. For the crimes involving drugs or alcohol, Indigenous prisoners and women have shorter sentences than the average prisoner, and the low skilled have longer sentences, but in this case the effect of Chinese, high skill, and being a sailor is harder to identify statistically. An interesting contrast arises when we consider the marginal sentencing effects for violent or sexual crimes (column 3), and for sentences longer than two years (column 7). Among the most serious crimes, only female prisoners' still receive significantly shorter sentences than the average prisoner, while among the very longest sentences recorded in the admission ledgers, none of the prisoners' observable characteristics significantly affect sentencing. Judges' insensitivity to ethnicity and social class (and virtually all other extralegal determinants) reflects a lack of judicial discretion in sentencing for the most serious crimes. To put it bluntly, in BC during the late nineteenth and early twentieth centuries, there was no room for paternalism when sentencing violent male offenders. Interestingly, paternalism directed towards Indigenous or Chinese prisoners also did not extend to petty crimes – in column (4) we see that for those receiving sentences in the bottom tail of the sentencing distribution, Indigenous and Chinese prisoners spent significantly *more* time in prison, while

---

[19]The most severe crime classes include crimes causing harm (stabbing, shooting, and assault, for example), homicide, sexual crimes, and property crimes involving violence (theft with a weapon, or arson, for example).

[20]Recall that by the mid-1880s, administrative changes within the BC prison system had removed virtually all prisoners in the upper and lower tails of the sentence distribution from the Victoria and New Westminster provincial prisons.

the high skilled prisoners enjoyed relatively short sentences. The results reported in Columns (5) and (6) confirm that among the sentences covered by our restricted sample (greater than seven days, but less than two years), the marginal sentencing effects are consistent above and below the median, with the possible exception of Chinese prisoners, for whom there is no identifiable sentencing bias among the sentences below the 50th percentile.

Concerns that our sample restrictions, in conjunction with the reallocation of prisoners out of BC's provincial prisons and into local police lock-ups after 1885, and the opening of the federal penitentiary in New Westminster after 1878, may introduce sample selection bias into our baseline estimates, are somewhat allayed by our results using all ledger entries for the earliest pre-1875 admission years. This set of admission records covers the period before any other prison was in operation in BC, and it includes *every* prison admission in the province between 1864 and 1875. The marginal sentencing effects reported in column (8) are not as precisely estimated, but the signs and magnitudes are similar to our baseline specification.

4 Changes in Criminal Sentencing and Public Sentiment

Our period of study covers admission years from 1864 to 1913. During this era, British Columbia experienced institutional and cultural changes that transformed the legal and social environment in the province (Inwood et al., 2022; Valverde, 2008). It is reasonable to expect that the sentencing patterns we observe by pooling our data over the half century following the opening of the Victoria Gaol in 1864, may not have been constant over the entire period. Certain sub-periods characterized by exceptional episodes of change may drive the fixed marginal sentencing effects reported in Table 1. Our sentencing equations include admission year fixed effects, but they can not account for the possibility that judges' marginal sentencing behaviour may have shifted in response to the changes that were occurring in BC's institutional and cultural environment. To assess how judicial decision-making may have changed over the 1864-1913 period, we again estimate Equation (1), but rather than pooling all admission years, we estimate a series of rolling regressions through nine-year estimation windows. To further characterize cultural change and public sentiment during this period, we rely on analysis of historical newspapers from British Columbia and conduct sentiment analysis of their text to identify positive and negative sentiment associations with our regression results. Specifically, we compare the trends in regression coefficients from our rolling regressions to those of our sentiment polarity indices. In doing so, we are able to comment on the extent to which the extralegal sentencing determinants we observe in Section 3 are consistent with taste-based bias. In what follows, we first describe the newspapers we

use and our processing of these into data. We then describe the sentiment analysis we conduct and its parameters. Finally, we present and discuss together the results of our rolling regression analysis and our analysis of public newspapers.

4.1 Historical Newspaper Data and Sentiment Analysis

Analysis of historical newspapers has become a growing means of analysis in economic history. Improvements in archival access and the digitization of these documents, combined with increased availability of Optical Character Recognition (OCR) software allow researchers to use historical documents for a wide variety of empirical applications. Beach and Hanlon (2022) provide a recent survey of the use of historical newspapers in economics and outline unique challenges of these data and best practices for researchers.

We rely on two data sources which include a range of historical papers. Both sources are based at the University of Victoria libraries. Primarily, we use *The British Colonist* (later called *The Daily Colonist* (University of Victoria Libraries). For this daily paper, the universe of articles between 1858 and 1980 are available as images. Figure 2 depicts the first page of *The Daily Colonist* from September 10th, 1907.

We perform OCR on these papers, selecting only papers published on the 20th day of each month, an ostensibly random sample.[21] We use *Layout Parser*, a Python library developed by Shen et al. (2021) to OCR and parse the article text in these papers, leaving us with a dataset of 675,971 unique text phrases from the paper over the course of our sample period. We perform compound-aware spelling correction of each phrase using the SymSpell algorithm developed by Garbe (2012) to improve the accuracy of the OCR results.

Beach and Hanlon (2022) stress the importance of accounting for the selection of historical newspapers into databases. In order to appear in our data, newspapers need to i) survive as archival materials, ii) be chosen for digitization, iii) be digitized at a requisite quality, iv) be selected for inclusion in a database. To allay these concerns as best as is possible, we also rely on *Historical Victoria Newspapers: 1858 - 1836* (Taylor and Mindenhall, 2007) so that we incorporate additional perspectives.[22] A selection of 16,662 articles from these papers were manually digitized on a wide range of topics from Victoria's social history. We parse these articles into 39,086 unique text phrases. Our full sample then includes 715,037 individual text phrases. Table 5 demonstrates that relative to *The Daily Colonist*, these papers

---

[21]We select this subsample in light of both software and hardware limitations.

[22]These papers include *Daily News*, *Evening Express*, *The (Daily) Press*, *Vancouver Daily Evening Post*, *Victoria Daily Chronicle*, *Victoria Daily Standard*, *Victoria Gazette*, *Victoria Times*, *Victoria Times & Evening Express*.

contain more negative sentiment towards Chinese, Indigenous, and female individuals.[23]

We wish to classify sentences according to their sentiment. To do so, we rely on *TextBlob* (Loria, 2018), a natural language processing library in Python. We employ a Naive Bayes Classifier (Rish et al., 2001; Leung, 2007; Murphy et al., 2006) to classify individual text phrases as being positive or negative.[24] We report the sentiment polarity index derived from these classifications. We construct this index for the subsets of phrases that contain references to keywords referring to Chinese, Indigenous or female individuals.[25]

In our analysis, we present the share of mentions per year for each group for which the sentiment polarity index is negative. We use this measure of sentiment to account for annual and within-paper variation in the frequency with which keywords are mentioned. We view this approach as analogous to the methods of Ferrara et al. (2022) and Albright et al. (2021), who normalize mentions of keywords relative to ubiquitous words. Our approach is a natural extension of this for sentiment analysis.

4.2 Trends in Criminal Sentencing and Public Sentiment

The time-varying marginal sentencing effects for Indigenous, Chinese, and female prisoners are presented in Figures 3, 4, 5, along with our measures of public sentiment toward these groups. These figures also include a measure of public sentiment directed toward individuals with each of these characteristics – the five- year moving average of their share of negative mentions in BC's newspapers. From these figures we can immediately see that, to varying degrees, the marginal sentencing effects were not constant over our period of study and changes in these effects track public sentiment very closely. More specifically, as the share of negative mentions in BC's newspapers rises (falls) for Indigenous individuals, those of Chinese, and women, sentence severity relative to prisoners with otherwise similar characteristics also rises (falls) for Indigenous, Chinese, and female prisoners.

From Figure 3 we can see that Indigenous prisoners' were significantly shorter than the average prisoners with similar observable characteristics over the entire 1864-1913 period. However, we wish to draw attention to the decline in sentence severity for these prisoners immediately after the passage of the *Indian Act* in 1876 and its amendments in the early

---

[23]Robustness exercises that exclude the Victoria newspapers preserve our results in virtually every regard and are available upon request.

[24]TextBlob contains a pre-trained Naive Bayes Classifier for sentiment analysis. Although this classifier is trained on a set of film reviews, we verify it does accurately classify the sentiment of phrases in our historical newspapers as being positive or negative.

[25]Following Lutz et al. (2013), keywords for Chinese include: *chinaman, chinamen, celestial, chinese, orientals*; for Indigenous: *indian, native*; and for female: *women, girl, woman, lady, ladies, daughter, her, she, mrs, wife, miss, sister, abbess*

1880s.[26] This tendency towards leniency may reflect a judicial response to the provisions in the act that limited contact between Indigenous and non-Indigenous communities in BC, thereby reducing the likelihood of recidivism among Indigenous prisoners – a connection we explore in greater detail below. Only during the decade just prior to WW1 did Indigenous prisoners' sentences rise back towards the average. The coefficients from our sentencing equations track public sentiment very closely until the early twentieth century, when they diverge sharply. From Figure 3 we can also see that public sentiment toward Indigenous individuals, as expressed in BC's historical newspapers, follows a remarkably similar time series pattern. Negative mentions of our Indigenous keywords as a share of all mentions, rises through the 1860s and 1870s, before falling steadily over the next 20 years, only to again climb quite sharply around the turn of the twentieth century.[27]

Female prisoners also have a very distinct pattern in sentencing over our period of study. While the marginal effect of being female, estimated from the pooled admission years, indicates that women received much shorter sentences than men for a given crime, our rolling marginal effects demonstrate that this leniency is driven by much lower, and sharply falling sentences over the first 15 years and last 10 years of our period. Through the 1890s and early 1900s, female prisoners' marginal sentences rose steeply, such that for a few years around 1900 they were *higher* than sentences served by the average prisoner with similar characteristics. This pattern is consistent with shifting social norms and attitudes towards 'gendered crimes' that have been documented in BC (Dunae 2009, 2008). Further, these coefficients closely mirror the evolution of public sentiment. When we compare the effect of being female across crime classes, we find that the average impact of gender on sentencing moved with the share of women incarcerated for drug and alcohol offences – we can see this in the crime-specific marginal sentencing effects for women reported in column (3) in Table 2.

The sentencing trends for prisoners of Chinese also exhibit large fluctuations over our period study. The fixed marginal sentencing effect reported in Table 1 indicates that prisoners

---

[26]We note that although this post-1880 decline in average sentence severity for Indigenous prisoners is visually apparent from the figure, it is not statistically distinguishable from zero when estimated over all admission records.

[27]When we use a five-year moving average, a cubic time trend, or the non-stationary series components (derived from a Hodrick-Prescott filter) to smooth public sentiment and the time-varying marginal sentencing effects depicted in Figure 3, we can statistically confirm what our visual inspection of the Figure suggests – the non-linear time trends are closely and positively correlated over our period of study. This chronological correlation between trends in sentencing and sentiment is consistent with a narrative that includes some taste-based bias in judicial decision making. We emphasize that our common trend tests are descriptive – public sentiment toward ethnicity and gender tracked sentence severity in BC over our period of study, but this co-movement cannot be taken as evidence of a causal connection.

of Chinese received comparatively short sentences relative to the average prisoner. The rolling marginal effects depicted in Figure 5 illustrate that this is driven mainly by increasing leniency through the 1890s and first years of the twentieth century. Prior to the 1890, there is no discernible sentencing effect (or possibly a small positive effect) associated with being Chinese in BC's provincial prison ledgers. Recalling the harsh sentences imposed on those of Chinese during the railway building boom, as reported in Table 1, it is noteworthy that once the CPR was completed in 1886, and temporary migration from China slowed to a trickle, sentences for Chinese individuals became significantly lower than the average prisoner. The railway boom also coincided with a wave of anti-Asian sentiment in BC and on the west coast of the United States (Lee, 1889). During the decade before WW1, sentences for Chinese prisoners began to increase again, coincident with an era of rising anti-Asian sentiment in British Columbia. Similar to the case of Indigenous peoples and women, public sentiment towards the Chinese closely tracks our estimated marginal sentencing effects over the sample period.

The evidence we have provided describes chronological correlations linking shifts in public sentiment to changes in sentencing biases associated with ethnicity and gender. These correlations have no causal interpretation, but they are strongly consistent with the presence of taste-based biases in judicial decision making. Even after controlling for all other observable characteristics, as public sentiment turned against women and those of Chinese during the 1880s, or Indigenous individuals during the 1860s and early 1870s, judges imposed harsher sentences on prisoners with these same characteristics. We now turn to an exploration of another potential source of bias in criminal sentencing in BC during the late nineteenth and early twentieth centuries. In what follows, we present a stylized model that allows us to understand the interaction of taste-based and statistical bias in criminal sentencing. Following that, we use results from this model to test for the presence of statistical discrimination.

5 Distinguishing Taste-Based and Statistical Bias

Our sentencing equations reveal evidence of both positive and negative biases in criminal sentencing in BC during the late nineteenth and early twentieth centuries. When paired with sentiment analysis of historical BC newspapers, we find that public sentiment and these sentencing biases moved in tandem over our sample period, suggestive that judicial tastes played a role in determining sentence severity. We turn now to the task of developing a deeper understanding of these biases through an exploration of taste-based (Becker, 1957) and statistical discrimination in our historical setting (Arrow, 1973; Phelps, 1972). Origi-

nally conceived as an explanation for unequal labour market outcomes between groups, we map these alternate explanations for discrimination into the criminal justice space in an effort to document the presence of, and relative importance of each factor in determining the sentencing patterns we have documented in Table 1.

Statistical bias in the context of criminal sentencing would arise from greater proclivity for criminal activity for one group versus another, which we stress is not attributable to innate characteristics of different groups, but merely differences in their economic and social circumstances. For example, it could be that the opportunity cost of crime for some groups is lower because they earn relatively low wages in legal activities, and as a result, the probability that they reoffend might be unusually high. It could also be that a particular individual is unlikely to reappear in the criminal justice system of that particular jurisdiction.

Taste-based bias, on the other hand, would arise purely from the tastes of a biased judiciary. For simplicity, we focus not on the total supply of offences, as Becker did, but on the individual decision to commit one particular crime. Our results easily generalize to the case of many different possible criminal offences.

In what follows we conduct two separate tests that aim to distinguish the taste-based and statistical channels of bias. Each test requires us to construct an estimate of future recidivism risk at the individual level, and we first describe the method in which we do this. We then conduct our first test, which follows logically from our baseline specification. We include predicted future recidivism as an explanatory variable in the LASSO-selected specification of Equation (1). Implicit in this is a test for the presence of statistical discrimination. If, by including predicted future recidivism, we observe reductions in the biases we see in Table 1, then this suggests a portion of this bias is attributable to the statistical bias channel. Further, the statistical significance of predicted future recidivism would indicate that judges do take their own estimation of recidivism risk into account when determining sentences.[28]

The second test we conduct approaches the distinction between channels of bias from a theoretical perspective. We first formalize how judicial tastes and incentives for crime affect criminal sentencing in an augmented Becker (1968) model of criminal punishment. In the model, groups differ in their incentives to commit crime as well as the taste-based bias they face from a representative judge. They endogenously choose to commit criminal offences and

---

[28]Our approach to controlling for statistical discrimination assumes judicial foresight – the ability to predict *future* recidivism. If instead we assume that judges only use information about recidivism in the *past*, we find very similar patterns in statistical discrimination, but the impact on the marginal sentencing effects for women, and the relatively low and high skilled prisoners, becomes both smaller and less precisely estimated. Full results from this alternate approach are available upon request from the authors.

the judge consequently chooses a criminal sentence that minimizes social loss. The model admits comparative statics for the optimal criminal sentence with respect to judicial bias and criminal incentives. We test these predictions in a TD framework using an exogenous shock to incentives for crime rooted in a major event in BC's history, the building of the CPR.

5.1 Predicted Future Recidivism

To quantify prisoner-specific reoffending risk, we exploit a useful feature of our data – the repeat admissions of a sizeable portion of the prisoners listed in the ledgers. We observe approximately 25 percent of the prisoners more than once, and 2.4 percent of the prisoners appear more than five times. This means that we can observe both past and future recidivism for each prisoner. Using an indicator of future recidivism for each admission entry, we can estimate the marginal probability of reoffending based on individuals' observable characteristics.[29] We can then use the predicted likelihood of future recidivism as an additional sentencing determinant in Equation (1). The size and statistical strength of predicted future recidivism in our sentencing equations provides us with an estimate of the impact of statistical discrimination (at least with respect to the likelihood of reoffending), while any significant sentencing bias that remains leaves open the possibility that taste-based discrimination also played a role.

We execute our empirical strategy by first predicting the probability of future recidivism with a probit regression that includes an indicator for future recidivism as the dependent variable, and prisoners' observable characteristics (as selected by our LASSO criteria for inclusion in Equation (1)) as the explanatory variables.[30] In doing so, we draw a comparison to recent work by Mullainathan and Obermeyer (2022), who use machine learning tools to first predict health outcome risk scores, and include these in a regression of doctors' testing decisions. We stress the analogy between their findings of doctor under (over)-testing and judicial under (over)-sentencing. Further, we are implicitly incorporating a risk score into a judge's sentencing decision, something that has become a reality in contemporary courtrooms (Sidu, 2015; Albright, 2019).

The impact of the prisoners' gender, ethnicity, and social class, on predicted future recidivism, including indicators for past recidivism and natural resource employment, are shown in Table 3. Here, we find intuitively satisfying results, and importantly, parameter estimates that are conducive to our empirical strategy. First, we note that past recidivism

---

[29]We note that our admission year fixed effects control for changes in average recidivism rates over time.

[30]Results when we use a linear probability model rather than a probit are very similar to those reported.

(meaning those who were already repeat offenders) is a strong predictor of future recidivism. This result is analogous to results about path dependence from Bindler and Hjalmarsson (2019), who study jury decisions from London's *Old Bailey* during the nineteenth century. We also find that sailors were much less likely to reoffend than the average prisoner, as were Indigenous prisoners, prisoners of Chinese, and high and low skilled prisoners. Interestingly, gender has only a small and statistically insignificant effect on the likelihood of reoffending. Our identification of statistical discrimination relies on the strength of the impact of natural resource employment on the probability of future recidivism, which we find to be statistically distinguishable from 0 and, as expected, large and negative.

Of course, if we only include our sentencing equation determinants in the first stage probit, the impact of predicted future recidivism cannot be separately identified when included as an additional regressor in Equation (1) (Heckman, 1976, 1979). At least one determinant of future recidivism that is not directly correlated with sentence severity is needed to identify the impact of the likelihood of reoffending on sentencing. LASSO does not pick natural resource employment (farming, fishing, logging, and mining) for inclusion in our sentencing equations because (conditional on other observable characteristics) these occupations were not significantly related to sentence severity – an indicator for natural resource employment has a statistically insignificant estimated coefficient of -0.030 in Equation (1), with a standard error of 0.030 (see Table 4). However, because the semi-skilled occupations in the rapidly growing natural resource sector in BC in the years around 1900 were relatively well-paid, and located in geographically remote regions, we can plausibly expect that the probability of reoffending for farmers, fishers, loggers, and miners would be relatively low. As a result of these considerations, we can include an indicator for natural resource employment as an additional explanatory variable to predict future recidivism in our first stage probit.[31]

5.2 Incorporating Predicted Future Recidivism as a Regressor

Turning to the impact of predicted recidivism on sentencing, the first column in Table 4 includes the baseline marginal sentencing effects for the key observable characteristics (as reported in Table 1) – Indigenous, Chinese, female, and high and low skilled. We also include the estimated effect for sailors, and we report the marginal sentencing effect for natural resource occupations. In column (2) we report the estimated coefficients for the same key prisoner characteristics from a sentencing equation that includes predicted future

---

[31] Although natural resource occupations were disproportionately held by non-Indigenous men in BC during our period of study, some Indigenous, Chinese, and female prisoners are recorded in the ledgers as farmers, fishers, loggers, or miners prior to their incarceration, and even for these prisoners, natural resource occupations are negatively correlated with the likelihood of reoffending.

recidivism as an additional regressor.[32] The final column reports the results from a χ^2 test of the null hypothesis that the coefficients are the same in both specifications.

First, we note that the results confirm that after controlling for other observable characteristics, natural resource employment does not significantly affect sentence length. Predicted recidivism, on the other hand, has a very large and strongly statistically significant effect on sentencing – if a judge was certain that a particular prisoner would reoffend in the future, their sentence would be 74 log-points, or approximately 70 percent longer than the average prisoner with similar characteristics, for a given crime.

The results for sailors illustrate the impact that controlling for predicted future recidivism can have on marginal sentencing. From Table 3 we see that sailors were significantly less likely to reoffend, and from Table 4 we see that when we control for this fact, the negative sentencing differential enjoyed by sailors in our baseline specification drops by more than half and becomes statistically indistinguishable from zero. To explain this, we note that Victoria and New Westminster (Vancouver) were important transport hubs around the turn of the twentieth century, and Victoria was home to a large and active naval base. While this helps to explain why so many sailors appear in the admission ledgers, it also means that most of these sailors would not have been permanent BC residents – they were transient workers in Victoria and New Westminster. As such, they would have left BC upon their release from prison, and their probability of reoffending (in BC at least) would be low. Thus, optimal sentences for sailors, structured in part to reduce the probability of future recidivism, should be lower, holding all else constant. In short, the favourable sentencing bias BC judges showed towards sailors during the late nineteenth and early twentieth centuries appears to have been due to statistical discrimination – sailors were much less likely to reoffend in BC, so it was rational to sentence them to shorter terms in prison. After controlling for sailors' low predicted future recidivism, there is no evidence of additional taste-based bias in their favour – sailors' sentences become indistinguishable from the average prisoner.

Although less extreme, we find similar effects for Indigenous and Chinese prisoners, both of whom were significantly less likely to reoffend than the average prisoner. In the case of Indigenous prisoners, controlling for their predicted future recidivism reduces the positive bias in their sentencing by 12 percent (3.6 log-points) – the χ^2 statistic reported in the last column in Table 4 confirms the statistical significance of this change. To rationalize this, consider the impact of the *Indian Act* – and in particular the amendments adopted during

---

[32]The standard errors reported in the second column in Table 4 are clustered by crime class $\times$ admission year. If we use a 2SLS approach to adjust the errors to account for the inclusion of an estimated right-hand-side variable (predicted recidivism), statistical significance as reported in the table is unaffected.

the early 1880s – on the geographic mobility of Indigenous communities in BC.[33] By limiting the mobility of those of Indigenous-origin, the *Act* reduced the opportunity for interaction between those deemed to be 'Indian' by the Canadian government and the colonial criminal justice system. In this case, a judge might reasonably expect the likelihood of reoffending for Indigenous prisoners to have been significantly reduced after these amendments came into effect. In both of these examples, a rational judge would optimally introduce bias in their sentencing that has nothing to do with their own preferences.

For the Chinese, we estimate a (statistically significant) reduction in bias of nearly 14 percent (2.3 log-points). In contrast, for women, and high and low skilled workers, controlling for predicted recidivism actually raises (in absolute value) the estimated sentencing bias. Women were slightly more likely to reoffend than the average prisoner, while those in the tails of the skill distribution were significantly less likely to appear in the admission ledgers in the future – after controlling for these tendencies the residual sentencing biases grow. Again, these changes are statistically distinguishable from zero.

We interpret our findings as evidence that the bias we observe in criminal sentencing in BC between 1864 and 1913 originated in part from statistical sources. However, we also emphasize that although the fixed marginal sentencing effects associated with gender, ethnicity, and social class that are reported in Table 4 are significantly affected by the inclusion of an additional control for predicted future recidivism, the residual sentencing biases remain large. This suggests that the combined effects of public sentiment, paternalism, and taste-based discrimination also had an important role to play in the differential sentencing we observe. The point we wish to make here is simply that the marginal sentencing effects associated with extra-legal sentence determinants are significantly altered, but not fully accounted for, when we control for predicted future recidivism. This result leads us to conclude that both statistical and taste-based factors were influencing judicial decision making in BC during the late nineteenth and early twentieth centuries.

We have shown that we can partly explain away some of the bias we observe in the baseline results of Table 1. We turn now to another exercise, one that formalizes the connection of future recidivism to criminal sentencing using an established theoretical framework. This allows to study analytically and empirically how an exogenous shock to the incentives

---

[33]The *Indian Act* introduced the 'reserve system', after which most Indigenous people resided on what were typically geographically remote reserves. An 1881 amendment to the *Act* introduced Indian agents – state officials who served as liaisons between federal and provincial governments and the reserve population, while also acting as magistrates for a variety of legal disputes and petty crimes. Following other amendments in 1884 and 1885, Indigenous individuals' freedom to leave their home reserve without a 'pass' – written permission of their local Indian agent – was further curtailed.

for criminal offences can identify the taste and statistical channels of extralegal bias.

5.3 A Model of Extralegal Bias in Criminal Sentencing

In this section, we present a stylized model of criminal sentencing inspired by Becker (1968). We use the model to study the comparative statics of criminal sentencing with respect to judicial tastes and incentives for crime. Our analysis admits a simple test to separate taste-based and statistical channels of discrimination, provided we observe an exogenous shock to the incentives for crime.

We consider the decision of a representative member of a group indexed by $j \in \mathcal{J}$ with $|\mathcal{J}| = J$. Group membership is generalized so that Chinese people, taller than average people, or taller than average Chinese people, for example, would all comprise distinct groups.

5.3.1 Criminal Offences and the Crime Rate

We model the decision to commit a particular crime as a binary decision, $o \in \{0, 1\}$. This crime yields a private benefit to the criminal of v, which is unaffected by being caught, arrested, convicted and sentenced. We model enforcement as being distinct from the courts or a judge's decision and is represented by μ. μ equates to the probability of being caught and convicted of the crime. In Becker's original work, this factor would be endogenous alongside punishment. However, in our setting, we treat it as exogenous from the perspective of a judge, being determined by other state apparatuses such as the police force.

The penalty, or sentence, corresponding to the conviction of an individual from group j is f_j and does constitute a decision variable for a judge. We distinguish between groups here only in their unobservable *incentive for crime*: an idiosyncratic shock, η_j, that determines the net benefits of committing the offence. η_j reflects unobservable factors in a group's economic or social circumstances that determine its sensitivity to the net benefit of committing the crime. Assume that $\eta_j \sim \mathcal{N}(0, \sigma_{\eta_j})$. We treat this shock as having mean zero to reflect no systematic differences between groups. The variance of this shock, however, we treat as reflecting incentives for crime across groups. Denoting $\mathbb{I}(\cdot)$ as the indicator function the decision of a member of group j is:

$$o = \mathbb{I}\{v - \mu f_j + \eta_j > 0\}$$

Writing the cdf of η_j as $1 - p_j(\eta_j)$, the crime rate of group j is:

$$p_j(\mu f_j - v)$$

and the total supply of crime from group j is:

$$Q_j(f_j) = p_j(\mu f_j - v)J \tag{2}$$

5.3.2   Social Harm, Enforcement Costs, and Private Costs

Following Becker and others, we construct a standard *social loss function* associated with crime from three components. First, there is the social damage of the crime itself, the function $\delta(\cdot)$, which is increasing in the total amount of crime (falling in the magnitude of punishment). That is, $\frac{\partial \delta}{\partial Q_j(\cdot)} > 0$ and $\frac{\partial \delta}{\partial f_j} < 0$. The total damage done to society by crimes from group j is $\delta\Big(Q_j(f_j)\Big)$.

Second, there is the cost of implementing the punishment f_j, $c(\cdot)$. This cost function is convex in the standard way and reflects the resources employed in imprisoning an individual or the bureaucratic costs of levying fines, etc. The total cost of punishing a group j is then $\mu c\Big(Q_j(f_j)f_j\Big)$.

Finally, there is the social cost of punishing the individual — the punishment itself — which is translated into social cost by the parameter b_j so that the social cost (from the perspective of a judge) of punishing the individual is $\mu b_j f_j Q_j(f_j)$. We might wonder how statistical discrimination may interact with the taste-based bias of a judge. The parameters $\{b_j\}_{j\in\mathcal{J}}$ reflect these tastes. A value of $b_j = 1$ indicates no bias - the social loss to group j of punishment is exactly its own private loss; a value of $b_j < 1$ indicates a negative bias - the judge discounts the loss to society of punishing group j; and a value of $b_j > 1$ indicates positive bias - the judge overstates the social loss of punishing group j. Combining each of these terms and aggregating across groups, the social loss function associated with all crime is:

$$\Lambda(\mathbf{f}) = \sum_{j\in\mathcal{J}} \left(\delta\Big(Q_j(f_j)\Big) + \mu c\Big(Q_j(f_j)f_j\Big) + \mu b_j f_j Q_j(f_j) \right) \tag{3}$$

5.3.3   Optimal Sentencing

A judge chooses an optimal, group-specific punishment to minimize the social loss of crime. As Equation (3) illustrates, they do this while incorporating the endogenous response of criminal activity to punishment, f_j. A judge's optimization problem is:

$$\min_{\{f_j\}_{j\in\mathcal{J}}} \Lambda(\mathbf{f}) = \sum_{j\in\mathcal{J}} \left(\delta\Big(Q_j(f_j)\Big) + \mu c\Big(Q_j(f_j)f_j\Big) + \mu b_j f_j Q_j(f_j) \right) \tag{4}$$

28

The solution to (4) - the vector $\mathbf{f}^*$ - is analogous to solving an individual sentencing problem corresponding to each group, and it implies a minimum value $\Lambda(\mathbf{f}^*)$. Note that by segmenting the market for crime, a judge can achieve a superior optimum in terms of the social loss function. We stress the analogy of this result to profit maximization by price discrimination (Stigler, 1966; Varian, 1989). This implies that when there are unobservable factors that influence criminal activity that vary across groups, optimal sentencing should reflect this. This reflects rational statistical discrimination. Further, the solution to (4) admits an equilibrium schedule, $Q_j^*(f_j^*)$ which we stress is analogous to a supply curve. This endogenous relationship forms the basis of the test we conduct to distinguish the channels of bias. Figure 6 illustrates this relationship by the judicial taste and incentive parameters of a group. Each parameter has implications for the nature of this relationship. First, a higher (lower) value of b_j, or positive (negative) taste-based bias corresponds to a lower (higher) optimal sentence for any $p(f_j)$. Second, an increase (decrease) in the criminal incentive of a group leads to a lower (higher) optimal sentence. Each of these parameters affects the equilibrium relationship between sentencing and the supply of crime as a supply shifter.

Combining each of these implications yields a prediction that is testable in our data. If we observe a shock that reduces the unobservable incentives for crime (an increase in σ_{η_j}), we should expect the statistical discrimination channel to reduce the optimal sentence. If we do not observe this decrease, and instead see no change or an increase in sentence severity, then we must conclude that there is a corresponding increase in taste-based bias. We state this formally below by means of Lemma 1 and Proposition 1.

Lemma 1. *A increase in the criminal incentive parameter of group j, σ_{η_j}, will decrease the optimal sentence of that group, f_j^*.*

Proposition 1. *If we observe an increase in the criminal incentive parameter of group j, σ_{η_j}, and an increase in the optimal sentence of that group f_j^*, then this increase must come from the taste-based channel, b_j.*

In what follows, we test Proposition 1 in our data. To do so, we rely on an exogenous shock to incentives for crime stemming from the building of the CPR in British Columbia.

5.3.4 Building the Canadian Pacific Railway

A stipulation of BC's joining confederation in 1871 was that Eastern and Western Canada be connected by rail within 10 years. Geological surveys began immediately and though the project was delayed, by 1876 the planning for the Winnipeg-BC leg of the railway

had begun (Hayes, 2010). The re-election of John A. Macdonald's conservative party in 1878 brought a renewed focus on the railway and its pacific connection. By 1879, the first contracts were awarded to Andrew Onderdonk, an American contractor, for building the CPR in BC. For our purposes, we use the years 1878 through 1886 as the period during which the CPR was being planned, surveyed and constructed. It was incorporated in 1881 and completed in 1886.

Building the CPR brought a significant economic expansion to BC, as well as an enormous influx of labour to the province. In particular, Chinese immigrants were brought to BC to work on the railway, as they could be paid less and treated worse than domestic or European workers (Munro, 1971). Onderdonk insisted that due to labour shortages, the project could not proceed without importing Chinese Labour (Li, 1995). Figure 7 illustrates that beginning in the 1881, there is a surge in the prison population of both individuals born in China and Europe, as well as those born in the Americas. While influxes of young, single, low-skilled males during economic booms are often thought increase crime rates, improvements in local economic conditions have been shown to reduce criminality and recidivism (Bartik et al., 2019; Yang, 2017). We rely on the economic expansion and greater labour market opportunities associated with the railway boom as an exogenous shock that decreased the incentive for crime.

Coincident with the re-election of Macdonald's conservatives began a wave of anti-Chinese sentiment in which "...relations between the mainstream population of British Columbia and the Chinese became increasingly hostile" (Warburton, 1999). In 1878, the Workingman's Protective Association formed, with its stated purpose being:

> "... the mutual protection of the working classes of British Columbia against the great influx of Chinese; to use all legitimate means for the suppression of their immigration..."

The Association successfully lobbied for a law banning the Chinese from working on publicly funded projects, but this lasted less than a year as it was met with strikes by Chinese working in other sectors (Ward, 2002).

5.3.5 Criminal Sentencing During the Canadian Pacific Railway Years

Proposition 1 implies that when the economic circumstances of a group improve for an exogenous reason, we should see a decrease in the average criminal sentence they receive via the statistical discrimination channel. In what follows, we use a TD specification that relies on the building of the CPR as a source of exogenous variation in incentives for crime. This

empirical specification is shown in Equation (5):

$$log(Sentence_{it}) = \alpha + \delta\mathbf{LD_{it}} + \gamma\mathbf{XLD_{it}} + \rho PFR_{it} + \eta D_i$$
$$+ \mu PFR_{it} \times D_i + \phi RR_t \times D_i + \lambda PFR_{it} \times RR_t \qquad (5)$$
$$+ \psi RR_t \times PFR_{it} \times D_i + \epsilon_{it}$$

where D_i is an indicator of group membership for {female, chinese, NA-born, foreign-born}, PFR_{it} is the predicted future recidivism of individual i in year t, and RR_t is an indicator of the years during the railway boom. The coefficient ψ is our coefficient of interest — the average treatment effect on the treated (Olden and Møen, 2022). $\hat{\psi} \neq 0$ would indicate that there is a differential impact of predicted future recidivism across groups. Given that their marginal sentencing effect is proportional to a prisoner's risk, observing differences across groups will imply that judicial tastes affect bias in sentencing. For this treatment effect to be interpreted as causal, the well-known parallel trends assumption must be met. Figure 8 illustrates this trend and gives us confidence in our estimates, as there is no significant difference from zero in this trend prior to 1878.

Table 6 depicts estimates of $\hat{\psi}$. Column (1) depicts these treatment effects for the Chinese, females, and other foreign-born prisoners. The estimated treatment effect for the Chinese is enormous — it suggests that during the railway boom, the marginal sentencing effect on predicted future recidivism for the Chinese was over one and a half times what it was for North-American born prisoners. Similarly, for the primarily European foreign-born workers, we observe no difference compared to domestic prisoners. Importantly, for females, we also observe no effect. This supports our narrative that the CPR shock should affect the criminal incentives for males. Column (2) illustrates the treatment effect of our estimation of Equation (6) using the indicator of Asian ethnicity instead of Chinese. This category includes Japanese and Southeast Asian prisoners as well. Still, we observe a very large treatment effect, though it is somewhat lessened. This result illustrates that the bias we observe primarily affected Chinese prisoners. Column (3) presents the results if we use an alternative characterization of the railway boom — the years 1881-1886. As noted above, these were the years in which most of the actual construction was done. Interestingly, we find no significant difference in the impact of predicted recidivism on sentencing here. Further, Figure 8 illustrates that the treatment effect we detect in Column (1) of Table 6 is driven by these pre-construction years. Interpreting these results in a historical light, it appears as though the wave of anti-Chinese sentiment during the late 1870s did influence criminal sentencing. By the early 1880s, however, labour shortages and pressure to complete the CPR may have reduced this bias. The results in Table 6 suggest taste-based discrimination with

respect to the Chinese during the early phase of building the CPR.

6 Concluding Remarks

In this paper, we present evidence that extralegal factors influenced criminal sentencing in British Columbia between 1864 and 1913. The fact that judges incorporate these factors can be seen as a source of bias in criminal sentencing. We first document these biases across the dimensions of gender, ethnicity, and social class. We find that women, Indigenous prisoners, and prisoners of Chinese all received sentences that were shorter than the average prisoner for a given crime, while those who reported working in the highest and lowest skilled occupations prior to incarceration received relatively long sentences.

Over time, we find that Indigenous prisoners received shorter sentences over our entire period of study, but this bias deepened after the passage of the *Indian Act* in 1876. For women, the bias we observe when pooling our data is driven mainly by the years at the very start and very end of our period. Chinese prisoners also received shorter sentences on average, but this is driven by later years – specifically, after the building of the CPR. During BC's railway building boom we observe longer sentences for Chinese prisoners. We find that public sentiment, as measured in historical newspapers from our sample period, closely tracks these changes in criminal sentencing.

We then employ two empirical tests to identify the presence of rational statistical biases in judicial decision-making in BC. Both tests rely on a prisoner's probability of re-offending in the future, which we predict using the panel dimension of our data.

Our first test relies on the idea that predicted future recidivism should be a determinant of sentence length, and may be correlated with the prisoner's observable characteristics. Without conditioning on this factor, we may mistakenly attribute observed aggregate sentencing bias entirely to taste-based discrimination. When we include predicted future recidivism as a regressor in our main empirical specification, we detect evidence of statistical discrimination, but in most cases, much of the initial bias we observe remains. This exercise attenuates the bias for Indigenous persons by over 3.5 percentage points and for Chinese individuals origin by over 2.5 percentage points. Including predicted future recidivism in our preferred specification exacerbates the bias we observe for both women and prisoners reporting high and low skilled occupations, suggesting that taste-based and statistical discrimination worked in opposite directions.

We next characterize extralegal bias in a Beckerian model of optimal criminal sentencing. The model admits an equilibrium relationship between sentence length and the

recidivism rate. Simulation results illustrate that when a group's incentives for criminal behaviour decrease, so should their sentence length. Similarly, the larger the bias a judge has against a group, the longer their sentence will be.

We test these implications using a triple difference specification where building the CPR represents an exogenous decline in criminal incentives. We identify longer sentences for Chinese prisoners that are disproportionate to their predicted future recidivism. These results demonstrate that during the early years of building the CPR, the criminal justice system — like the broader population of BC — exhibited taste-based bias against the Chinese.

References

Ahola, Angela S., Sven Å Christianson, and Åke Hellström (2009), "Justice needs a blindfold: Effects of gender and attractiveness on prison sentences and attributions of personal characteristics in a judicial process." *Psychiatry, Psychology and Law*, 16, S90–S100.

Ajzenstadt, Mimi (2002), "Racializing prohibitions: Alcohol laws and racial/ethnic minorities in British Columbia, 1871–1927." *Regulating Lives: Historical Essays on the State, Society, the Individual, and the Law*, 97–119.

Albright, Alex (2019), "If you give a judge a risk score: Evidence from Kentucky bail decisions." Harvard University: John M. Olin Fellow's Discussion Paper.

Albright, Alex, Jeremy A Cook, James J Feigenbaum, Laura Kincaide, Jason Long, and Nathan Nunn (2021), "After the burning: The economic effects of the 1921 tulsa race massacre." Technical report, National Bureau of Economic Research.

Antonie, L., K. Inwood, and H. Maxwell-Stewert (2021), *Prison Bodies: The Height and Weight of Men in Canadian Prisons, 1874–1935*. Unpublished Manuscript.

Antonovics, Kate and Brian G. Knight (2009), "A new look at racial profiling: Evidence from the Boston Police Department." *The Review of Economics and Statistics*, 91, 163–77.

Anwar, Shamena, Patrick Bayer, and Randi Hjalmarsson (2012), "The impact of jury race in criminal trials." *The Quarterly Journal of Economics*, 127, 1017–1055.

Anwar, Shamena, Patrick Bayer, and Randi Hjalmarsson (2019), "A jury of her peers: The impact of the first female jurors on criminal convictions." *The Economic Journal*, 129, 603–50.

Anwar, Shamena and Hanming Fang (2006), "An alternative test of racial prejudice in motor vehicle searches: Theory and evidence." *American Economic Review*, 96, 127–51.

Arnold, David, Will Dobbie, and Crystal S. Yang (2018), "Racial bias in bail decisions." *The Quarterly Journal of Economics*, 133, 1885–932.

Arrow, Kenneth J. (1973), *The Theory of Discrimination, Discrimination in Labor Markets*. Princeton University Press.

Ayres, Ian and Joel Waldfogel (1993), "A market test for race discrimination in bail setting." *Stan. L. Rev.*, 46, 987.

Bartik, Alexander W, Janet Currie, Michael Greenstone, and Christopher R Knittel (2019),

"The local economic and welfare consequences of hydraulic fracturing." *American Economic Journal: Applied Economics*, 11, 105–55.

Beach, Brian and W Walker Hanlon (2022), "Historical newspaper data: A researcher's guide and toolkit."

Becker, Gary S. (1957), *The Economics of Discrimination.* University of Chicago Press.

Becker, Gary S. (1968), "Crime and punishment: An economic approach." In *The Economic Dimensions of Crime*, 13–68, Springer.

Belshaw, John (2009), *Becoming British Columbia: A Population History.* UBC Press.

Bertrand, Marianne, Dolly Chugh, and Sendhil Mullainathan (2005), "Implicit discrimination." *American Economic Review*, 95, 94–8.

Bindler, Anna and Randi Hjalmarsson (2017), "Prisons, recidivism and the age–crime profile." *Economics Letters*, 152, 46–9.

Bindler, Anna and Randi Hjalmarsson (2019), "Path dependency in jury decision making." *Journal of the European Economic Association*, 17, 1971–2017.

Bindler, Anna and Randi Hjalmarsson (2020), "The persistence of the criminal justice gender gap: Evidence from 200 years of judicial decisions." *The Journal of Law and Economics*, 63, 297–339.

Blair, Irene V., Charles M. Judd, and Kristine M. Chapleau (2004), "The influence of Afrocentric facial features in criminal sentencing." *Psychological Science*, 15, 674–79.

Block, Sharon (2018), *Colonial Complexions: Race and Bodies in Eighteenth-Century America.* University of Pennsylvania Press.

Bodenhorn, H., C. Moehling, and G. Price (2012), "Short criminals: Stature and crime in early America." *Journal of Law and Economics*, 55, 393–419.

Bodenhorn, Howard (2009), "Criminal sentencing in nineteenth century Pennsylvania." *Explorations in Economic History*, 46, 287–98.

Bohren, J Aislinn, Peter Hull, and Alex Imas (2022), "Systemic discrimination: Theory and measurement." Technical report, National Bureau of Economic Research.

Burch, Traci (2015), "Skin color and the criminal justice system: Beyond black-white disparities in sentencing." *Journal of Empirical Legal Studies*, 12, 395–420.

Bushway, Shawn D. and Anne M. Piehl (2001), "Judging judicial discretion: Legal factors and racial discrimination in sentencing." *Law and Society Review*, 35, 733–764.

Canay, Ivan A, Magne Mogstad, and Jack Mountjoy (2020), "On the use of outcome tests for detecting bias in decision making." Technical report, National Bureau of Economic Research.

Chesney-Lind, Meda (1977), "Judicial paternalism and the female status offender: Training women to know their place." *Crime & Delinquency*, 23, 121–30.

Cohen, Alma and Crystal S. Yang (2019), "Judicial politics and sentencing decisions." *American Economic Journal: Economic Policy*, 11, 160–91.

Crew, B. Keith (1991), "Sex differences in criminal sentencing: Chivalry or patriarchy?" *Justice Quarterly*, 8, 59–83.

Daly, Kathleen (1989), "Rethinking judicial paternalism: Gender, work-family relations, and sentencing." *Gender & Society*, 3, 9–36.

Derenoncourt, Ellora (2019), *Can you move to opportunity? Evidence from the Great Migration.* Unpublished Manuscript.

Dunae, Patrick A. (2008), "Geographies of sexual commerce and the production of prostitutional space: Victoria, British Columbia, 1860–1914." *Journal of the Canadian Historical Association*, 19, 115–142.

Dunae, Patrick A. (2009), "Sex, charades, and census records: Locating female sex trade workers in a Victorian city." *Social History*, 42, 267–297.

Emery, J.C. Herbert, Kris Inwood, and Henry Thille (2007), "Hecksher-Ohlin in Canada: New estimates of regional wages and land prices." *Australian Economic History Review*, 47, 22–48.

Ferrara, Andreas, Joung Yeob Ha, and Randall Walsh (2022), "Using digitized newspapers to refine historical measures: The case of the boll weevil." Technical report, National Bureau of Economic Research.

Garbe, W (2012), "Symspell: Symmetric delete algorithm."

Gelbach, Jonah B. (2021), *Testing Economic Models of Discrimination in Criminal Justice.* Unpublished Manuscript.

Godfrey, Barry S., Stephen Farrall, and Susanne Karstedt (2005), "Explaining gendered sen-

tencing patterns for violent men and women in the late-Victorian and Edwardian period." *British Journal of Criminology*, 45, 696–720.

Goldin, Claudia D. and Frank D. Lewis (1975), "The economic cost of the American civil war: Estimates and implications." *The Journal of Economic History*, 35, 299–326.

Gruhl, John, Susan Welch, and Cassia Spohn (1984), "Women as criminal defendants: A test for paternalism." *Western Political Quarterly*, 37, 456–67.

Harris, Richard, Ian Keay, and Frank D. Lewis (2015), "Protecting infant industries: Canadian manufacturing and the national policy, 1870–1913." *Explorations in Economic History*, 56, 15–31.

Hayes, Derek (2010), *Historical atlas of the North American railroad.* University of California Press Berkeley, CA.

Heckman, James J (1976), "The common structure of statistical models of truncation, sample selection and limited dependent variables and a simple estimator for such models." In *Annals of economic and social measurement, volume 5, number 4*, 475–492, NBER.

Heckman, James J (1979), "Sample selection bias as a specification error." *Econometrica: Journal of the econometric society*, 153–161.

Inwood, Kris and Jim Irwin (2002), "Land, income and regional inequality: New estimates of provincial incomes and growth in Canada, 1871–1891." *Acadiensis*, 31, 157–84.

Inwood, Kris and Ian Keay (2021), *Indigenous Prisoners in British Columbia's Jails, 1864-1916: Evidence of Bioeconomic Well-Being and Sample Selection.* Unpublished Manuscript.

Inwood, Kris, Ian Keay, and Blair Long (2022), *Institutional Change and Criminal Sentencing: Evidence from British Columbia's Jails, 1864-1913.* Unpublished Manuscript.

Inwood, Kris and Evan Roberts (2020), "Indians are the majority of the prisoners? historical variations in incarceration rates for indigenous women and men in British Columbia." *The Howard Journal of Crime and Justice*, 59, 350–69.

Johnson, Brian D. and Ryan D. King (2017), "Facial profiling: Race, physical appearance, and punishment." *Criminology*, 55, 520–47.

Knowles, John, Nicola Persico, and Petra Todd (2001), "Racial bias in motor vehicle searches: Theory and evidence." *Journal of Political Economy*, 109, 203–29.

Kruttschnitt, Candace and Jukka Savolainen (2009), "Ages of chivalry, places of paternalism: Gender and criminal sentencing in Finland." *European Journal of Criminology*, 6, 225–47.

Lang, Kevin and Ariella Kahn-Lang Spitzer (2020), "Race discrimination: An economic perspective." *Journal of Economic Perspectives*, 34, 68–89.

Lee, Joseph (1889), "Anti-chinese legislation in british america." *The Quarterly Journal of Economics*, 3, 359–364.

Leung, K Ming (2007), "Naive bayesian classifier." *Polytechnic University Department of Computer Science/Finance and Risk Engineering*, 2007, 123–156.

Leventhal, Gloria and Ronald Krate (1977), "Physical attractiveness and severity of sentencing." *Psychological Reports*, 40, 315–18.

Levitt, Steven D. and Thomas J. Miles (2007), "Empirical study of criminal punishment." *Handbook of Law and Economics*, 1, 455–95.

Li, Chris (1995), "The chinese: a valuable asset to the canadian pacific railway or an" evil" to white labourers?" *British Columbia History*, 28, 9.

Loria, Steven (2018), "textblob documentation." *Release 0.15*, 2.

Lutz, John, Patrick Dunae, Jason Gilliland, Don Lafreniere, and Megan Harvey (2013), "Turning space inside out: Spatial history and race in victorian victoria." *Historical GIS Research in Canada, Jennifer Bonnell and Marcel Fortin, editors*, 1–26.

MacKinnon, James, Morten Nielsen, and Matthew Webb (2022), "Cluster-robust inference: A guide to empirical practice." *Journal of Econometrics*, XXX, 1–28.

Mechoulan, Stéphane and Nicolas Sahuguet (2015), "Assessing racial disparities in parole release." *The Journal of Legal Studies*, 44, 39–74.

Moulds, Elizabeth F. (1978), "Chivalry and paternalism: Disparities of treatment in the criminal justice system." *Western Political Quarterly*, 31, 416–30.

Mullainathan, Sendhil and Ziad Obermeyer (2022), "Diagnosing physician error: A machine learning approach to low-value health care." *The Quarterly Journal of Economics*, 137, 679–727.

Munro, John A (1971), "British columbia and the" chinese evil": Canada's first anti-asiatic immigration law." *Journal of Canadian Studies*, 6, 42–51.

Murphy, Kevin P et al. (2006), "Naive bayes classifiers." *University of British Columbia*, 18, 1–8.

Mustard, David B. (2001), "Racial, ethnic, and gender disparities in sentencing: Evidence from the US federal courts." *The Journal of Law and Economics*, 44, 285–314.

Olden, Andreas and Jarle Møen (2022), "The triple difference estimator." *The Econometrics Journal*, 25, 531–553.

Park, Kyung H. (2017), "Do judges have tastes for discrimination? evidence from criminal courts." *Review of Economics and Statistics*, 99, 810–23.

Phelps, Edmund S. (1972), "The statistical theory of racism and sexism." *American Economic Review*, 62, 659–61.

Philippe, Arnaud (2017), "Gender disparities in criminal justice." Technical report, Toulouse School of Economics.

Radzinowicz, Leon and Roger Hood (1978), "Judicial discretion and sentencing standards: Victorian attempts to solve a perennial problem." *University of Pennsylvania Law Review*, 127, 1288–349.

Rehavi, M. Marit and Sonja B. Starr (2014), "Racial disparity in federal criminal sentences." *Journal of Political Economy*, 122, 1320–54.

Reiss, Albert J. (1965), *Occupations and Social Status*. Free Press.

Rish, Irina et al. (2001), "An empirical study of the naive bayes classifier." In *IJCAI 2001 workshop on empirical methods in artificial intelligence*, volume 3, 41–46.

Schanzenbach, Max (2005), "Racial and sex disparities in prison sentences: The effect of district-level judicial demographics." *The Journal of Legal Studies*, 34, 57–92.

Shen, Zejiang, Ruochen Zhang, Melissa Dell, Benjamin Charles Germain Lee, Jacob Carlson, and Weining Li (2021), "Layoutparser: A unified toolkit for deep learning based document image analysis." *arXiv preprint arXiv:2103.15348*.

Shermer, Lauren O'Neill and Brian D. Johnson (2010), "Criminal prosecutions: Examining prosecutorial discretion and charge reductions in US federal district courts." *Justice Quarterly*, 27, 394–430.

Sidu, Dawinder S. (2015), "Moneyball sentencing." *Boston College Law Review*, 56, 671–731.

Sorensen, Todd A., Surpriya Sarnikar, and Ronald L. Oaxaca (2014), "Do you receive a

lighter prison sentence because you are a woman or a white? an economic analysis of the federal criminal sentencing guidelines." *The BE Journal of Economic Analysis & Policy*, 14, 1–54.

Spohn, Cassia, John Gruhl, and Susan Welch (1981), "The effect of race on sentencing: A re-examination of an unsettled question." *Law & Society Review*, 16, 71–88.

Starr, Sonja B. (2015), "Estimating gender disparities in federal criminal cases." *American Law and Economics Review*, 17, 127–59.

Steckel, Richard H. (2008), "Biological measures of the standard of living." *Journal of Economic Perspectives*, 22, 129–52.

Stigler, George Joseph (1966), *Theory of price*.

Taylor, Leona and Dorothy Mindenhall (2007), "Index of historical victoria newspapers." https://web.uvic.ca/vv/newspaper.

Teichman, Doron and Eyal Zamir (2014), "Judicial decision-making: A behavioral perspective." In *Oxford Handbook of Behaviorial Economics and the Law*, 664–702, Oxford University Press.

Tibshirani, Robert (1996), "Regression shrinkage and selection via the LASSO." *Journal of the Royal Statistical Society: Series B (Methodological)*, 58, 267–88.

University of Victoria Libraries (????), "Bristish colonist online: 1858 - 1980." https://britishcolonist.ca.

Urquhart, M.C. and K. Buckley (1965), *Historical Statistics of Canada*. MacMillan.

Valverde, Mariana (2008), *The age of light, soap, and water: Moral reform in English Canada, 1885–1925*. University of Toronto Press.

Varian, Hal R (1989), "Price discrimination." *Handbook of industrial organization*, 1, 597–654.

Vickers, Chris (2016), "Socioeconomic status and judicial disparities in England and Wales, 1870–1910." *Explorations in Economic History*, 61, 32–53.

Warburton, Rennie (1999), "The workingmen's protective association, victoria, bc, 1878: Racism, intersectionality and status politics." *Labour/Le Travailleur*, 43, 105–120.

Ward, W Peter (2002), *White Canada forever: Popular attitudes and public policy toward Orientals in British Columbia*. McGill-Queen's Press-MQUP.

Yang, Crystal S (2017), "Local labor markets and criminal recidivism." *Journal of Public Economics*, 147, 16–29.
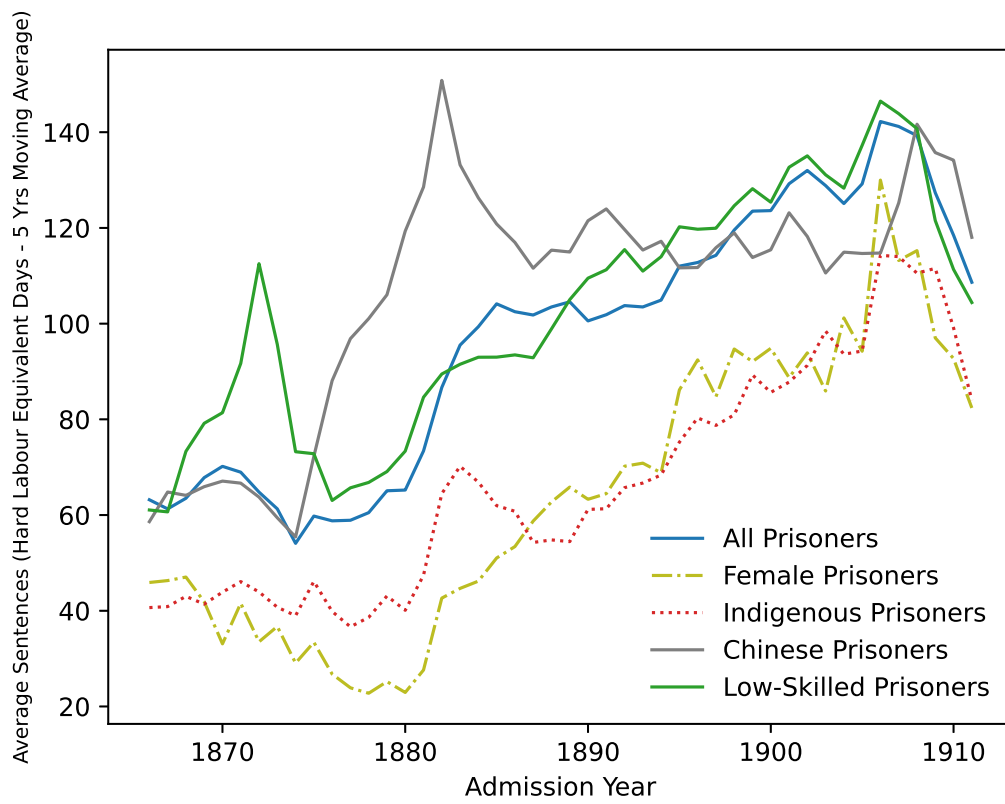
Figure 1: Average Hard Labour-Equivalent Sentences (5-Year Moving Average)

Figure 2: Front Page of *The Daily Colonist*, September 10th, 1907
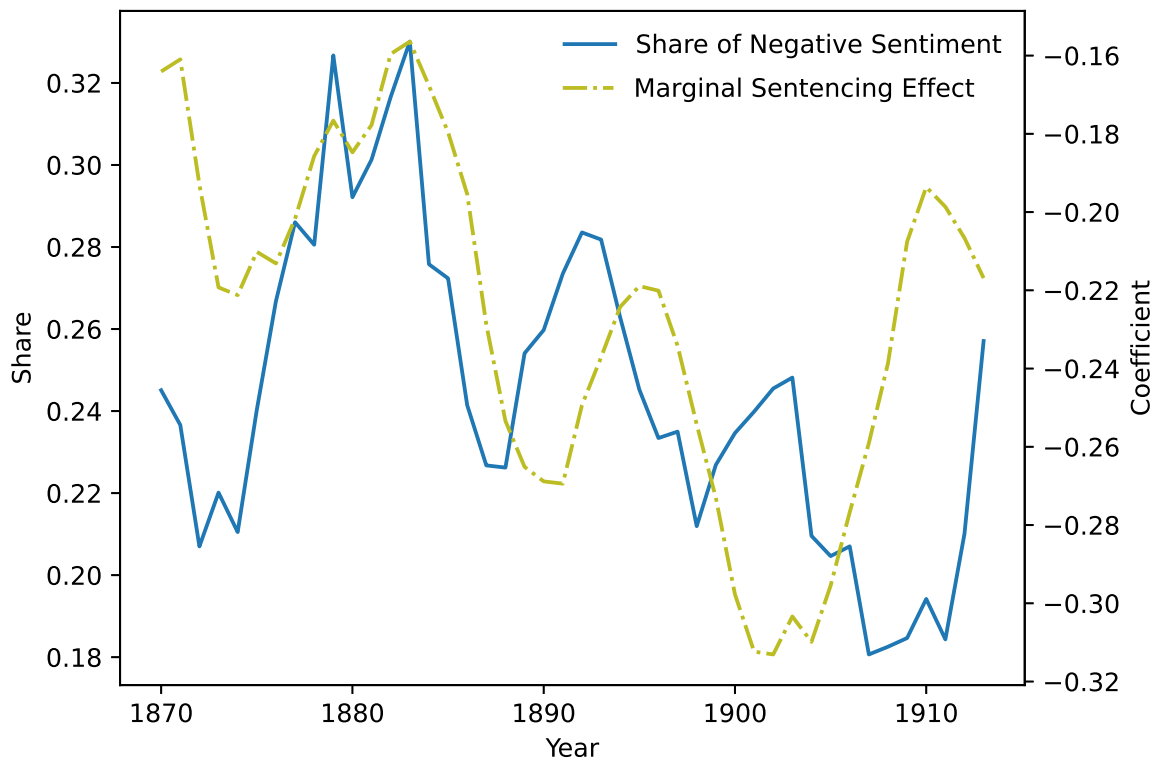
Figure 3: Trends of Share of Negative Sentiment Statements - Indigenous
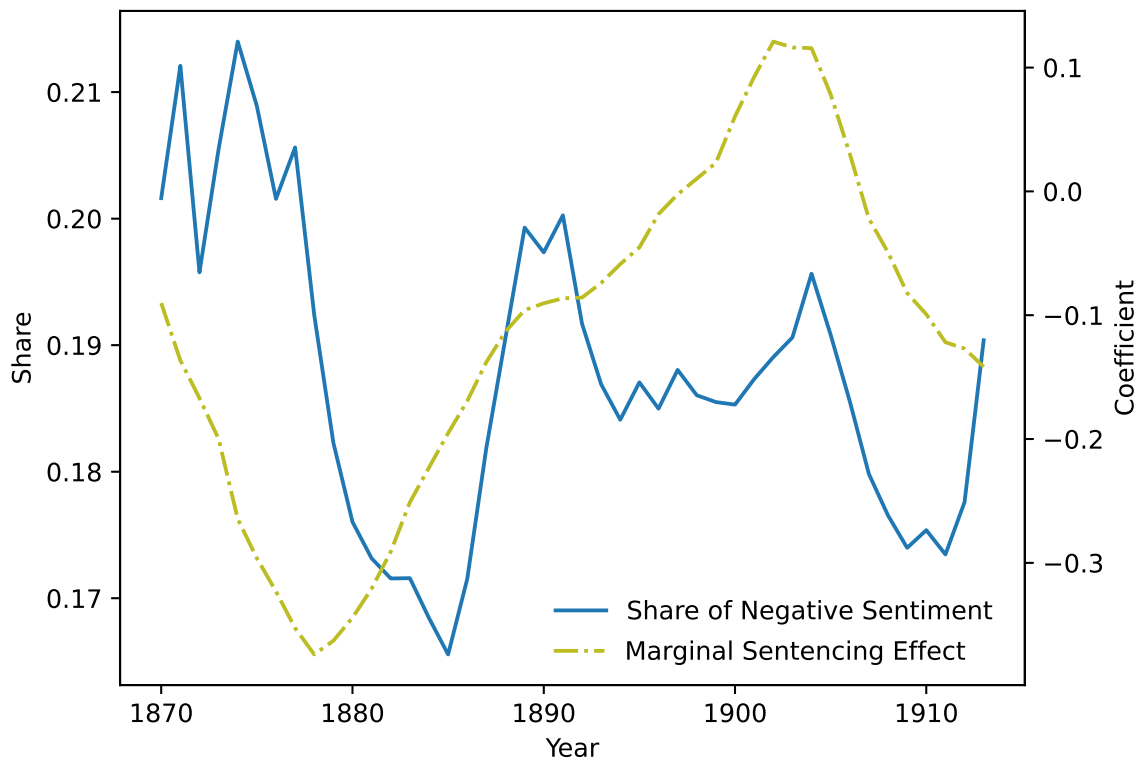
Figure 4: Trends of Sentencing Bias and Share of Negative Sentiment Statements - Female
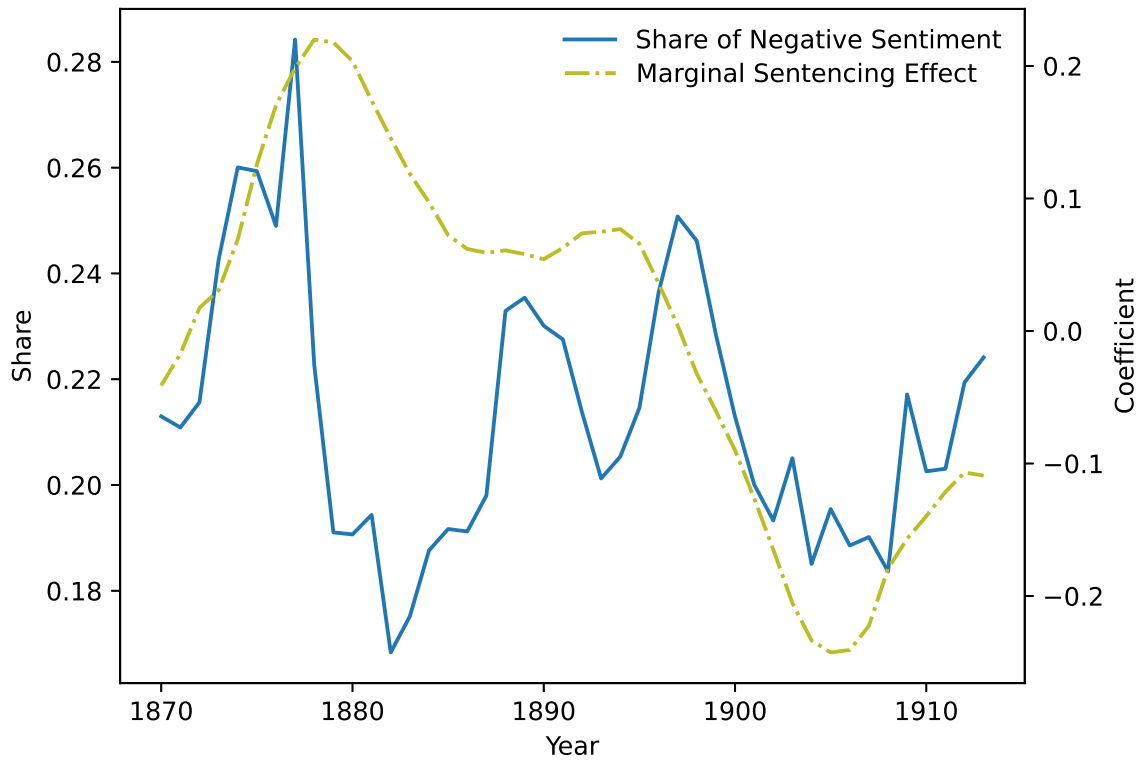
Figure 5: Trends of Sentencing Bias and Share of Negative Sentiment Statements - Chinese
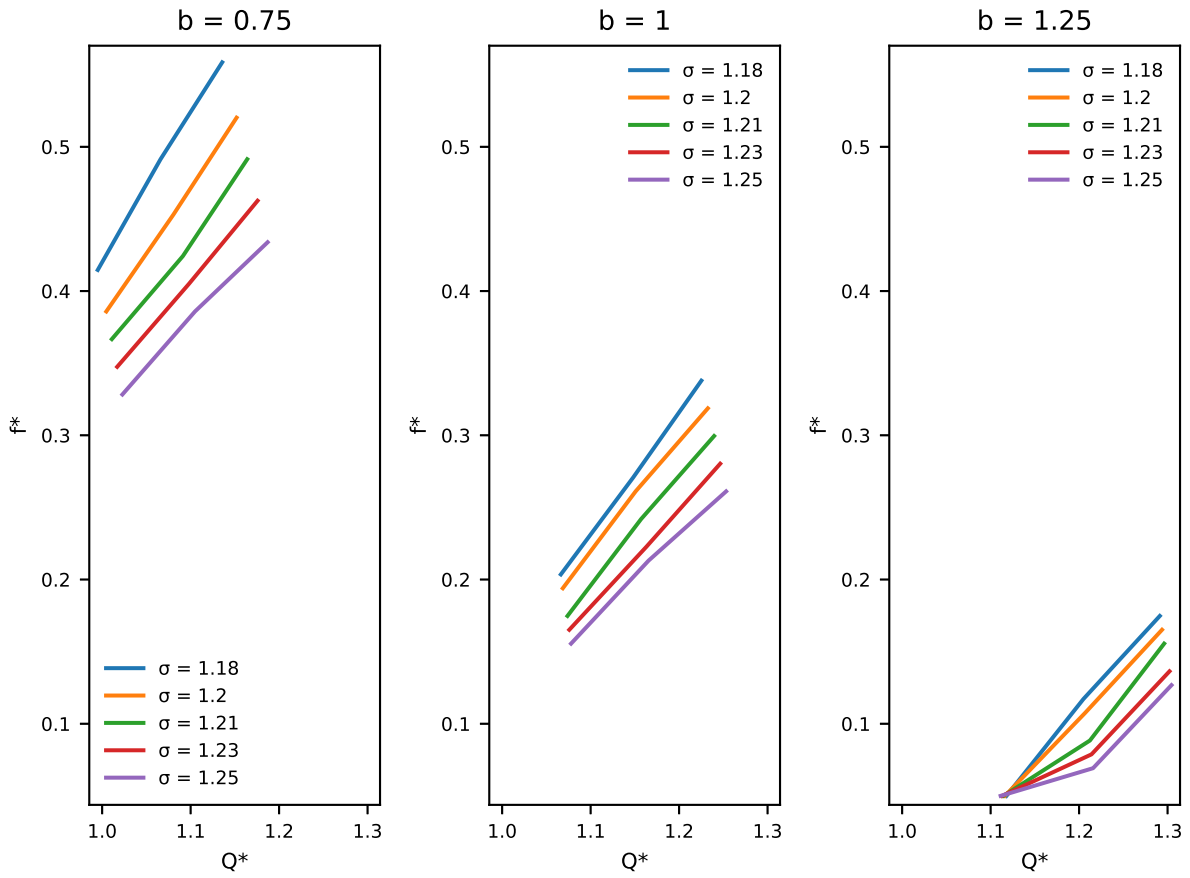
Figure 6: The Equilibrium Relationship Between Optimal Sentencing and the Supply of Crime
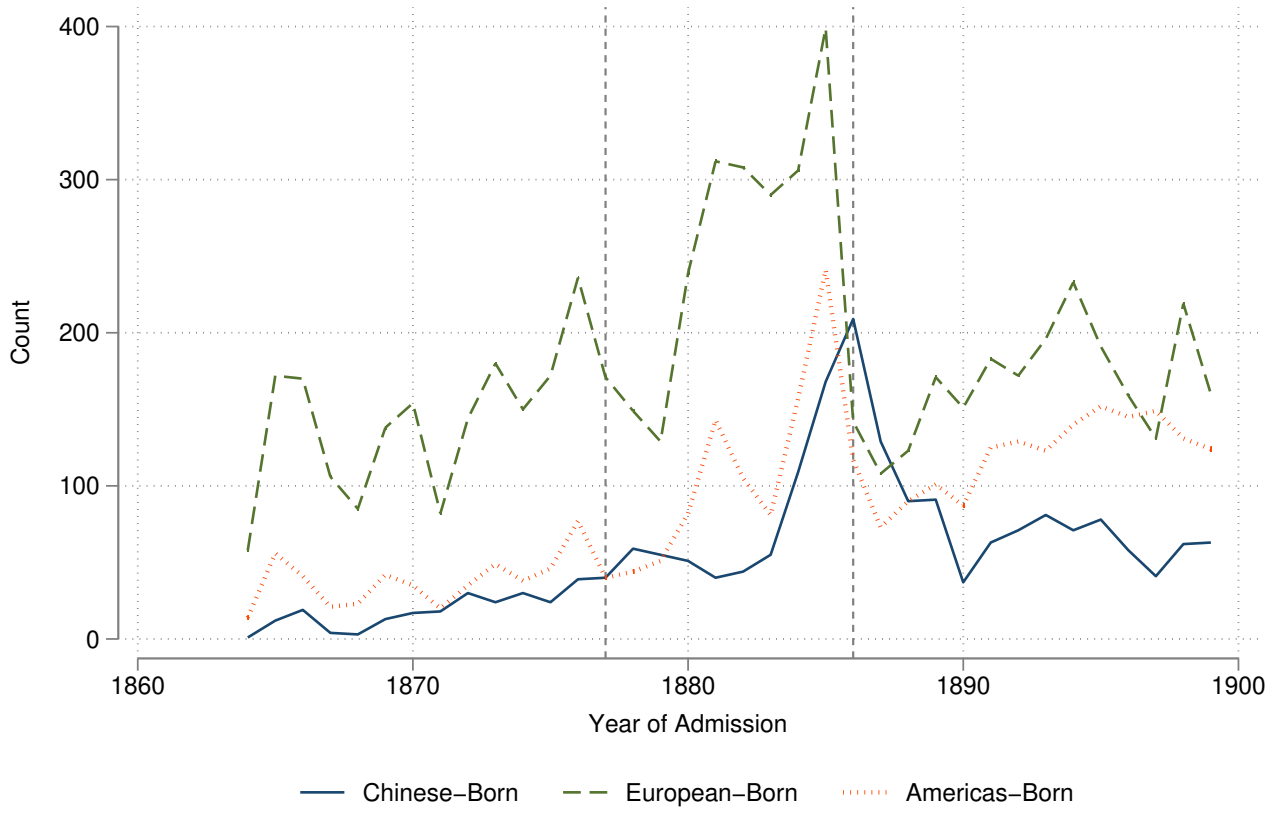
Figure 7: Number of Admissions by Year by Origin
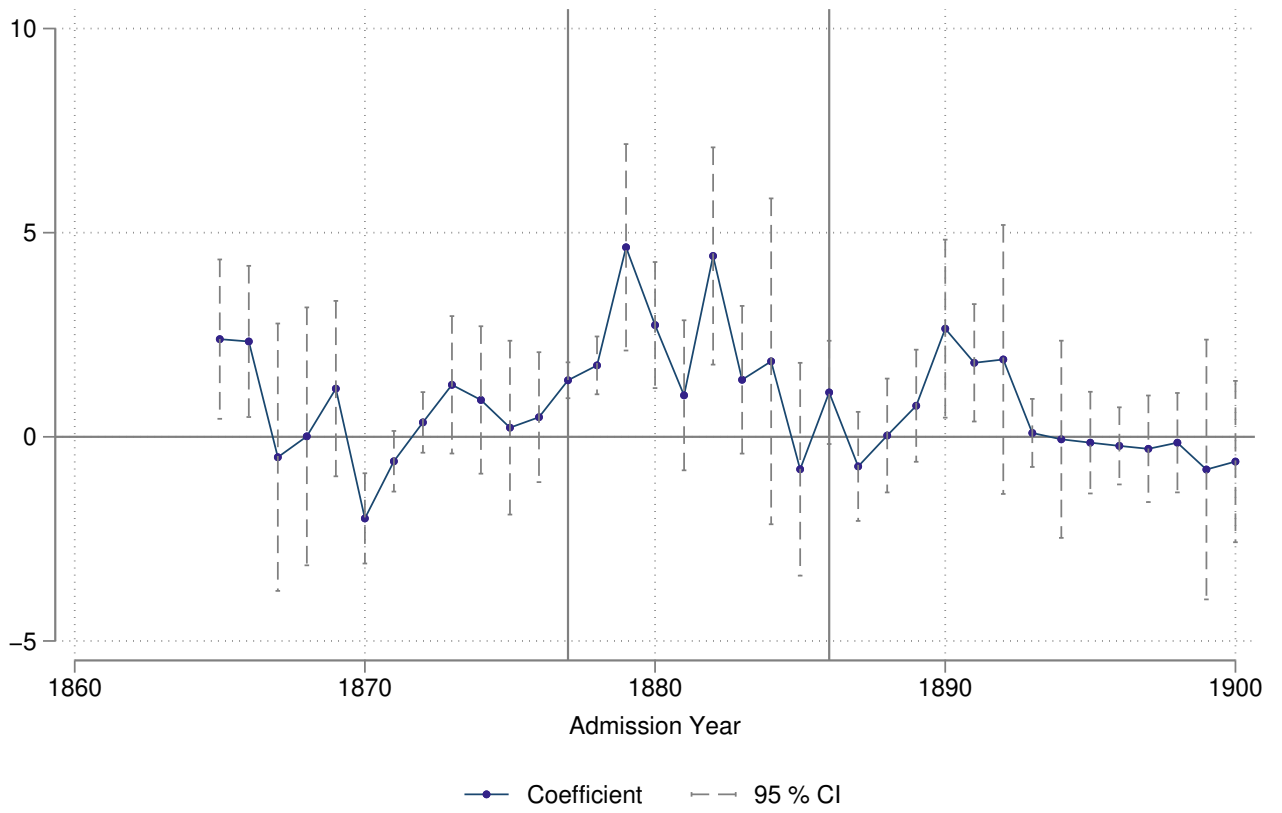
Figure 8: Pre-Trends in Equation 6

Table 1: Marginal Sentencing Effects, 1864-1913

Sentence Determinant	Dep. Variable = ln(Hard Labour-Equivalent Days)			
	(1)	(2)	(3)	(4)
Past Recidivism				0.1842***
				(0.019)
Incl. HL				0.5395***
				(0.031)
Incl. Cost or Fine				-0.8242***
				(0.179)
AND Sentence				0.2323***
				(0.052)
Fall Admission				0.0444**
				(0.021)
UK-Origin				-0.0348*
				(0.020)
Young				0.0416**
				(0.017)
Big and Tall				0.0451***
				(0.016)
Small and Short				-0.0505***
				(0.016)
Uses Alias				0.1073***
				(0.035)
Single				0.0372*
				(0.021)
Other Religion				0.1795***
				(0.044)
Unhealthy Complexion				0.0914***
				(0.030)
Bald				-0.1130**
				(0.054)
Intemperate				-0.1011***
				(0.019)
Unflattering Marks				0.0364**
				(0.017)
Indigenous	-0.7744***	-0.3564***	-0.2685***	-0.2998***
	(0.084)	(0.038)	(0.037)	(0.034)
Chinese	0.0484	-0.0447	-0.0279	-0.1584***
	(0.055)	(0.036)	(0.036)	(0.042)
Chinese×RR Building	0.0459	0.2115**	0.2766***	0.2931***
	(0.091)	(0.104)	(0.100)	(0.108)
Female	-0.2420***	-0.2011***	-0.1848***	-0.1320***
	(0.042)	(0.038)	(0.038)	(0.036)
High-Skilled	0.2049***	0.0781**	0.0681*	0.0791**
	(0.071)	(0.039)	(0.039)	(0.037)
Low-Skilled	0.1357***	0.0777***	0.0750***	0.0684***
	(0.045)	(0.025)	(0.025)	(0.023)
Sailor	-0.1775***	-0.0664**	-0.0617*	-0.0736**
	(0.055)	(0.033)	(0.033)	(0.032)
Narrow Crime Class FE		Yes	Yes	Yes
Admission Year FE		Yes	Yes	Yes
Gaol FE		Yes	Yes	Yes
Crime Class × Year FE			Yes	

Notes: $N = 17,586$ (singletons dropped). R^2 for Col. (1) = 0.091; Col. (2) = 0.324; Col. (3) = 0.377; Col. (4) = 0.375. Mean dep. var. = 4.147. Restricted sample drops sentences less than 7 days and more than 2 years. Standard errors (in parentheses) are clustered by crime class×admission year. *, **, *** indicates statistical significance with 90%, 95%, and 99% confidence, respectively.

Table 2: Heterogeneity in Marginal Sentencing Effects

	Dep. Variable = ln(Hard Labour-Equivalent Days)							
					Sentence Distribution			Victoria
Sentence Determinants	Only Property Crimes	Crimes Involving Drugs or Alcohol	Violent and Sexual Crimes	≤ 7 Days	Below Median	Above Median	> 2 Years	Gaol Only
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Indigenous	-0.2118***	-0.5155***	-0.0750	0.3360***	-0.0615**	-0.1630***	0.1327	-0.2526**
	(0.078)	(0.061)	(0.076)	(0.078)	(0.026)	(0.028)	(0.116)	(0.099)
Chinese	-0.2504***	0.1328	0.0507	1.0337***	0.0157	-0.1643***	0.0238	0.1036
	(0.067)	(0.092)	(0.161)	(0.258)	(0.037)	(0.033)	(0.147)	(0.251)
Female	-0.0672	-0.2998***	-0.1677*	-0.0822	-0.0795***	-0.0646**	-0.0655	-0.4190***
	(0.083)	(0.054)	(0.095)	(0.066)	(0.031)	(0.032)	(0.175)	(0.116)
High-Skilled	0.1485**	-0.0477	-0.0705	-0.5096***	0.0184	0.0691**	-0.0003	0.4490*
	(0.066)	(0.133)	(0.159)	(0.193)	(0.033)	(0.027)	(0.103)	(0.259)
Low-Skilled	-0.0008	0.1504***	0.0765	0.0209	0.0743***	0.0116	0.0496	0.0545
	(0.049)	(0.055)	(0.064)	(0.094)	(0.020)	(0.016)	(0.066)	(0.090)
Sailor	-0.0463	0.0909	-0.0594	-0.1430*	0.0062	-0.0623**	0.0468	-0.2530**
	(0.062)	(0.082)	(0.110)	(0.074)	(0.030)	(0.028)	(0.151)	(0.103)
Crime Class FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Admission Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Gaol FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Legal and Extra-legal Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
N	4,534	3,549	2,692	4,907	8,865	10,601	443	2,006

Notes: Col. (1)-(2), and (5)-(6) use restricted sample (sentences more than 7 days and less than or equal to 2 years); Col. (3) uses all sentences; Col. (4) uses only sentences less than or equal to 7 days; Col. (7) uses only sentences more than 2 years; and Col. (8) uses all sentences for admission years 1864-1875 (unrestricted universe of *all* prisoners incarcerated in BC). All columns use broad crime class FE. Standard errors (in parentheses) are clustered by narrow crime class×admission year. *, **, *** indicates statistical significance with 90%, 95%, and 99% confidence, respectively.

Table 3: Determinants of Future Recidivism

	Probit: Dep. Variable = $\mathbf{1}$\{Future Recidivist\}
Recidivism Determinant	(1)
Past Recidivism	0.8059***
	(0.042)
Indigenous	-0.1833***
	(0.046)
Chinese	-0.1606**
	(0.064)
Female	0.0674
	(0.067)
High-Skilled	-0.2093***
	(0.068)
Low-Skilled	-0.0686**
	(0.030)
Sailor	-0.2780***
	(0.050)
Natural Resource Occupations	-0.1145**
	(0.051)
Narrow Crime Class FE	Yes
Admission Year FE	Yes
Gaol FE	Yes
Legal and Extra-legal Controls	Yes

Notes: $N = 17,468$. $R^2 = 0.131$. Mean dep. var. $= 0.150$. Restricted sample drops sentences less than 7 days and more than 2 years. Standard errors (in parentheses) are clustered by crime class×admission year. *, **, *** indicates statistical significance with 90%, 95%, and 99% confidence, respectively.

Table 4: Impact of Statistical Bias – Controlling for Future Recidivism

| | Dep. Variable = ln(Hard Labour-Equivalent Days) | | |
Sentence Determinant	Baseline Sentencing Equation (1)	Incl. Predicted Recidivism (2)	H_0 : Equal Marginal Effects (χ^2) (3)
Indigenous	-0.2998***	-0.2641***	4.800**
	(0.034)	(0.038)	(0.029)
Chinese	-0.1584***	-0.1357***	3.690*
	(0.042)	(0.043)	(0.055)
Female	-0.1320***	-0.1428***	3.360*
	(0.036)	(0.037)	(0.067)
High-Skilled	0.0791**	0.0978**	4.150**
	(0.037)	(0.039)	(0.042)
Low-Skilled	0.0684***	0.0739***	5.590**
	(0.023)	(0.023)	(0.018)
Sailor	-0.0736**	-0.0332	5.060**
	(0.032)	(0.038)	(0.024)
Natural Resource Occupations	-0.0296		
	(0.030)		
Predicted Recidivism		0.7418**	
		(0.322)	
Narrow Crime Class FE	Yes	Yes	
Admission Year FE	Yes	Yes	
Gaol FE	Yes	Yes	
Legal and Extralegal Controls	Yes	Yes	

Notes: N for Col. (1) = 17,584; Col. (2) = 17,468. R^2 for Col. (1) = 0.375; Col. (2) = 0.374. Mean dep. var. = 4.147. Restricted sample drops sentences less than 7 days and more than 2 years. Col. (1) and (2) standard errors (in parentheses) are clustered by crime class×admission year. H_0 in Col. (3) = χ^2 test for equal marginal sentencing effects (p-values in parentheses). *, **, *** indicates statistical significance with 90%, 95%, and 99% confidence, respectively.

Table 5: Differences in Sentiment Between Newspapers

	Dep. Variable = All BC Newspapers					
	Indigenous		Chinese-Origin		Female	
	Lin. Probability Sentiment< 0 (1a)	Sentiment Polarity Index (1b)	Lin. Probability Sentiment< 0 (2a)	Sentiment Polarity Index (2b)	Lin. Probability Sentiment< 0 (3a)	Sentiment Polarity Index (3b)
Daily Colonist	-0.0605***	0.0389***	-0.0023	0.0298***	-0.0737***	0.0520***
	(0.013)	(0.006)	(0.026)	(0.010)	(0.007)	(0.003)
Month×Year FE	Yes	Yes	Yes	Yes	Yes	Yes
N	3,792	3,792	2,876	2,876	47,490	47,490

Notes: All BC newspapers include: *Victoria Daily Colonist, Daily News, Evening Express, The (Daily) Press, Vancouver Daily Evening Post, Victoria Daily Chronicle, Victoria Daily Standard, Victoria Gazette, Victoria Times, Victoria Times and Evening Express.* All columns estimated by OLS. Standard errors (in parentheses) are clustered by newspaper×year. All specifications include year and month FE. *, **, *** indicates statistical significance with 90%, 95%, and 99% confidence, respectively.

Table 6: Triple Difference Treatment Effects

	Dep. Variable = ln(Hard Labour-Equivalent Days)		
	(1)	(2)	(3)
$Chinese \times PFR \times CPR$	1.661***		
	(0.5175)		
$Asian \times PFR \times CPR$		1.426***	
		(0.5175)	
$Chinese \times PFR \times CPR_{short}$			0.609
			(0.6171)
$Female \times PFR \times CPR$	-0.211		
	(0.3787)		
$Foreign - Born \times PFR \times CPR$	0.036		
	(0.2414)		
Narrow Crime Class FE	Yes	Yes	Yes
Admission Year FE	Yes	Yes	Yes
Gaol FE	Yes	Yes	Yes
Legal and Extralegal Controls	Yes	Yes	Yes

Notes: Standard errors (in parentheses) are clustered by crime class×admission year. F-statistic reported for test of equal marginal sentencing effects. *, **, *** indicates statistical significance with 90%, 95%, and 99% confidence, respectively.

Appendix Table A1: Summary Statistics

Admissions, Sentences, Crimes, and Characteristics	All Records Mean (1864-1919)	Restricted Sample Mean (1864-1913)	Prisoner Characteristics (cont.)	All Records Mean (1864-1919)	Restricted Sample Mean (1864-1913)
Admissions			**Age Group Shares**		
Total	33,296	17,586	Under 20	0.071	0.069
Per Year	594.57	351.72	20s	0.415	0.427
			30s	0.283	0.274
Sentence Length			40s	0.147	0.142
			50s	0.071	0.072
Hard Labor: All Yrs	123.12	102.89	Over 60	0.013	0.016
Pre-1885	67.787	70.775	Average Age	31.926	31.981
Post-1885	156.78	114.73			
$ Amount	279.75	236.62	Age Recorded	0.905	0.941
Sentence Recorded	0.859	1	**Gender**		
Sentence Structure			Male	0.942	0.954
			Female	0.058	0.046
Costs /Damage	0.001	0			
Fine	0.004	0.003	Gender Recorded	0.986	0.989
Regular Term	0.151	0.102			
Hard Labour	0.412	0.645	**Height**		
AND Sentence	0.090	0.090			
OR Sentence	0.310	0.245	Average: All (in.)	66.478	66.588
Most Severe	0.028	0	Tallest Q1	69.429	69.503
			Shortest Q4	62.829	62.995
Broad Crime Class					
			Height Recorded	0.866	0.894
Alcohol and Drugs	0.202	0.100			
Military Discipline	0.064	0.082	**Literacy**		
Causing Harm	0.100	0.108			
Causing Death	0.018	0.007	Read Only	0.013	0.015
Specifies Indigeneity	0.096	0.152	Write Only	0.001	0.001
Fraud	0.036	0.034	Read and Write	0.594	0.630
Involving Animals	0.003	0.003	None	0.392	0.354
Sexual	0.025	0.018			
Property	0.226	0.262	Literacy Recorded	0.888	0.921
Against Public Order	0.165	0.225			
Safe-Keeping	0.053	0	**Marital Status**		
Other	0.012	0.008			
			Single	0.776	0.777
Crime Recorded	0.964	0.990	Married	0.208	0.204
			Widow	0.016	0.019
Alias Used	0.030	0.035			
			Marital Status Recorded	0.998	0.998
Recidivism					
			Religion		
Repeat Offenders	0.272	0.245			
Average # Convictions	1.424	1.498	Protestant	0.474	0.519
			Catholic	0.242	0.255
Birthplace / Ethnicity			Other	0.070	0.066
			None	0.214	0.161
Indigenous	0.222	0.172			
Canada (Non-Indigenous)	0.133	0.146	Religion Recorded	0.849	0.872
USA	0.110	0.119			
UK-Ireland	0.300	0.319	**Complexion-Health**		
Europe	0.092	0.097			
China	0.109	0.115	Healthy	0.705	0.666
Japan	0.012	0.007	Unhealthy	0.295	0.334
Other Asia	0.008	0.007			
Africa-Caribbean	0.004	0.004	**Complexion-Colour**		
Australia-NZ	0.005	0.006			
Latin America	0.006	0.007	Dark	0.436	0.399
			Light	0.450	0.490
Birthplace Recorded	0.905	0.932	Medium	0.098	0.089
			Other	0.015	0.022
			Complexion Recorded	0.829	0.839

Appendix Table A1: Summary Statistics (cont.)

Prisoner Characteristics (cont.)

	All Records Mean (1864-1919)	Restricted Sample Mean (1864-1913)
Body / Proportion		
Big	0.010	0.010
Stout	0.166	0.174
Medium	0.623	0.611
Small	0.078	0.074
Thin	0.064	0.076
Other	0.059	0.055
Proportion Recorded	0.748	0.756
Conduct		
Temperate	0.378	0.362
Moderate	0.031	0.015
Intemperate	0.589	0.620
Addict/Drug Use	0.002	0.002
Conduct Recorded	0.809	0.854
Distinct Marks		
Blindness	0.019	0.018
Poor Eye Sight	0.012	0.009
Disabled / Injured	0.051	0.050
Facial Mark	0.057	0.057
Fingers	0.034	0.039
Teeth	0.013	0.012
Scar	0.348	0.333
Tattoo	0.261	0.270
Other	0.100	0.059
Distinct Marks Recorded	0.221	0.256
Occupation/Skill/Social Class		
Low-Skilled / Lower Class	0.480	0.465
Semi-Skilled / Middle Class	0.473	0.490
High Skilled / Upper Class	0.047	0.044
Sailor	0.138	0.152
Natural Resources	0.120	0.104
Occupation Recorded	0.728	0.802

Notes: See text for detailed description of variables and sources. Restricted sample drops sentences less than or equal to 7 days and more than 2 years. Recorded entries do not include illegible or nonsensical entries. 'Sentence equivalents' value days in prison, fines, and costs using skill-specific wages/day. 'Most severe' includes life in prison, death, punishment, and hard labour sentences above the 95th percentile of all sentences. 'No religious affiliation' and 'no occupation' do not include blank entries. Skill/social class categories based on Duncan-index Reiss (1965).