

2022

The Right to Counsel: Criminal Prosecution in 19th Century London

Bryan McCannon

West Virginia University, bryan.mccannon@mail.wvu.edu

Zachary Porreca

West Virginia University

Follow this and additional works at: https://researchrepository.wvu.edu/faculty_publications



Part of the [Economic History Commons](#)

Digital Commons Citation

McCannon, Bryan and Porreca, Zachary, "The Right to Counsel: Criminal Prosecution in 19th Century London" (2022). *Faculty & Staff Scholarship*. 3109.

https://researchrepository.wvu.edu/faculty_publications/3109

This Article is brought to you for free and open access by The Research Repository @ WVU. It has been accepted for inclusion in Faculty & Staff Scholarship by an authorized administrator of The Research Repository @ WVU. For more information, please contact beau.smith@mail.wvu.edu.

The Right to Counsel: Criminal Prosecution in 19th Century London

Bryan C. McCannon*
West Virginia University

Zach Porreca†
West Virginia University

November 30, 2021

Abstract

Exploiting a novel data set of criminal trials in 19th century London, we evaluate the impact of an accused’s right to counsel on convictions. While lower-level crimes had an established history of professional representation prior to 1836, individuals accused of committing a felony did not, even though the prosecution was conducted by professional attorneys. The Prisoners’ Counsel Act of 1836 remedied this and first introduced the right to counsel in common law systems. Using a difference-in-difference estimation strategy we identify the causal effect of defense counsel. We find the surprising result that the professionalization of the courtroom led to an increase in the conviction rate, which we interpret as a consequence of jurors feeling that the trial became fairer. We go further and employ a topic modeling approach to the text of the transcripts to provide suggestive evidence on how the trials changed when defense counsel was fully introduced.

Keywords: defense counsel; Latent Dirichlet Allocation; Old Bailey; prosecution; right to counsel; text analysis

1 Introduction

The ability of the criminal justice system to provide effective outcomes is a central policy issue today. We want an institution that deters crime, is responsible with the public’s resources, recognizes the consequences to both the victim and the accused, mitigates errors, and is applied equally to all. Among the many institutional features of concern is the quality of the defendant’s legal representation. In the United States, the right to defense counsel is included explicitly in the Sixth Amendment to the U.S. Constitution, and the landmark Supreme Court case of *Gideon v Wainwright* (1963) extended this protection to the requirement that defense counsel be publicly provided to those lacking financial resources. The right to counsel was incorporated into the Napoleonic Code in 1808 (Zacharis, 2008). Civil law countries have included it since. The U.K. introduced the right first in 1836 and adopted a publicly-funded legal aid society in 1949.¹ A growing literature questions the quality of the publicly-provided defense in the U.S.² Since all have access

*Corresponding author: Department of Economics, John Chambers College of Business and Economics, 1601 University Avenue, Morgantown WV 26506, bryan.c.mccannon@gmail.com. We thank Adam Nowak for discussions on text analysis methods, along with Anna Bindler and Toke Aidt for conversations about the Old Bailey and 19th Century English institutions.

†Department of Economics, John Chambers College of Business and Economics, 1601 University Avenue, Morgantown WV 26506, zjp00003@mix.wvu.edu

¹It was established in the Legal Aid and Advice Act of 1949. It replaced the Poor Prisoners Defence Act of 1930, which was more limited in scope.

²We do not know of any empirical investigations into defense counsel in other countries or legal systems.

to defense counsel, these investigations can only evaluate how outcomes differ as relevant factors, such as compensation, vary. They are unable to identify the full consequence of having a right to defense counsel as they do not observe outcomes arising without this right.

To evaluate the impact of defensive legal representation, we leverage a unique data set only recently digitized and made public – the criminal trials at the Old Bailey in London. The Old Bailey served as the central courthouse for the prosecution of crimes in London during the 1700s and 1800s. This is an important period for the development of common law as institutional features that exist in common law countries today were invented there and then.

Prior to this period, the accused stood trial before a judge and jury without any legal counsel. The Prisoners’ Counsel Act, passed by Parliament in 1836, introduced the right to counsel for felony cases. Ironically (to the modern observer), misdemeanor cases had an established history of being allowed the use of defense counsel. Many misdemeanor crimes involved business interests with civil implications. Consequently, misdemeanor prosecution had evolved to allow business owners to hire legal counsel in these cases. Those accused of committing felony crimes did not have this right. This legal environment provides the opportunity to estimate a difference-in-difference specification to identify the causal effect that the introduction of a right to counsel had on criminal trials.

This is what we do. We document a counter-intuitive outcome. Felony crimes, relative to misdemeanors, experience an *escalation* in the conviction rate once the defendant has the full right to defensive counsel. While potentially surprising, we view our result as being in line with Bindler and Hjalmarsson (2018a) who also study convictions at the Old Bailey. They document an increase in the conviction rate when the severity of the sanction imposed is reduced. They interpret their result as evidence that jurors respond to the harshness of the punishment. When a conviction results in the death penalty, as was common in the 1700s, jurors are unwilling to convict. We extend this logic to the lop-sided trials prior to 1836. Prosecution could use legal professionals, but the defendant could not. Our results suggest that in an “unfair” institutional arrangement such as this, jurors are again hesitant to convict. Once the criminal trial became professionalized, English jurors felt free to convict, on the margin. We show that this effect arises regardless of the sanction’s harshness. Further, the effect is similar for both cases that utilize defense counsel and for those defendants who opt not to hire representation. Thus, the change represents a shift to the legal institution’s functioning.

We go further and evaluate the text of the trial transcripts. First, using a dictionary method approach, we identify which cases likely involved defense counsel and show that four crimes in particular experienced large increases after the Act’s implementation: fraud, pickpocketing, theft, and animal theft. Evaluating each subset of thousands of trial transcripts, we next employ a popular topic modeling method, Latent Dirichlet Allocation, to uncover hidden themes in the trials. We find that answers related to the exact timing of events exhibit a noticeable increase in prevalence after 1836. This suggests that professional representation improved the testimony’s exactness.

We are not the first to study trials at the Old Bailey. Bindler and Hjalmarrson (2017; 2018a; 2019; 2020; 2021) provide a series of analyses utilizing the data created by the Old Bailey project. Our main result uses the same data set as Bindler and Hjalmarrson (2018a). There, as stated previously, the authors examine the effect that punishment severity has on jury verdicts. They use a difference-in-difference method to link reductions in punishment severity with increases in conviction rates. The authors made use of the halt in penal transportation during the American Revolutionary War and the offense-specific abolition of capital punishment as natural experiments.

Bindler and Hjalmarrson (2017) examines trends in age amongst those convicted at the Old Bailey. Their findings are in line with the bulk of criminology literature in that the largest share of those convicted were in their 20s. Bindler and Hjalmarrson (2019) examine path dependency amongst jury verdicts. In these proceedings, juries would hear and decide on several cases in the same session. The authors examine whether the decision made by a jury on one case affected their decisions on subsequent cases heard in that session. The authors find a positive correlation between the previous verdict and that jury's subsequent decisions. Bindler and Hjalmarrson (2020) ask whether the gender disparity in sentencing observed in modern data (Starr, 2015) has been a consistent phenomenon. They show that it has. Hence, demographics of the accused matter.

Finally, in their most recent work, Bindler and Hjalmarrson (2021) make use of the Old Bailey proceedings to examine how the development of the London Metropolitan Police Force, the world's first professional police force, impacted crime rates. This was done by manually geocoding locations into treated and untreated areas from the transcripts and from daily police reports. Using this data, the 1829 Metropolitan Police Act serves as a natural experiment to ask whether the development of a professional police force deterred crime. They find significant reductions in violent crimes with a particularly large reduction in burglaries.³

Legal historians have also studied trials at the Old Bailey. Our historical background provided in the next section relies heavily on archival work by May (2003). Legal historians have labeled the Prisoners' Counsel Act of 1836 as "arguably the most significant development in criminal trial procedure during the nineteenth century" (Griffiths, 2014, p.28). The Act has been credited with taking an important step towards the creation of the adversarial court system and that the "cut-and-thrust nature of adversarial procedure in turn gave rise to many features of the ordinary trial that we now take for granted" (Gallanis, 2006, p.163).

As mentioned, work on the modern U.S. legal system has primarily explored the contrasting outcomes generated by different types of defense counsel. For example, Roach (2014) identifies gaps between outcomes achieved by salaried, full-time public defenders and court-appointed attorneys who receive a fee for each case handled. Abrams and Yoon (2007) look within a public defender office and show that attorney experience correlates strongly with case outcomes. Roach (2017) considers a case study of New York state increasing assigned counsel's hourly rate and shows that the effect interacts with the underlying poverty rate of the district they work in. Agan, Freedman, and Owens (2021) are able to compare case outcomes achieved by

³Similar to this, Voth (1998) uses references to employment as a measurement of time use by laborers.

attorneys working as court-appointed attorneys, for relatively low fee, to those achieved by the same attorneys when working privately for defendants. While a substantial amount of the discrepancy in outcomes is due to case selection, much is left to differences in un-measurable effort exertion. Shem-Tov (2021) exploits within case variation of co-defendants whose cases are handled by different attorneys. He shows that defendants represented by public defenders obtain more favorable outcomes than court-appointed attorneys. Differences between defense attorneys in experience and compensation are important, but they cannot measure the impact that the right to counsel has on outcomes.

Our work also connects to the Bayesian jury literature which explores the consequences of the jury collecting information, updating their beliefs, and making conviction decisions to maximize their objective functions. The theme in these papers is that the information created by the criminal justice institutions will affect their willingness to convict, which will cycle back and affect decisions made by other legal actors. For example, Friedman and Wickelgren (2006) evaluate how juror’s imperfect information creates a lower bound on the amount of deterrence that can be achieved even when the potential sanction is unbounded. Bjerck (2007) illustrates how juror updating of beliefs frustrates the ability of plea bargaining to fully screen the innocent from the guilty. Relatedly, theoretical models of jury voting highlight the impact of unanimous voting rules on the ability of the jury to reach the correct verdict (Feddersen and Pesendorfer, 1998; Duggan and Martinelli, 2001). Coupled with recent documentation that gender (Hoekstra and Street, 2021) and racial (Flanagan, 2018) composition of the jury matters, our result points to the importance of juror’s preferences and incentives in understanding legal outcomes.

We execute our analysis of the right to counsel in a series of steps. First, in Section 2 we provide a brief historical background of criminal prosecution in London and describe the specifics of the institutional change. In Section 3 we examine the trial transcripts for evidence of increases in defense counsel presence and the validity of our identification strategy. The data is described in Section 4. Section 5 presents our main result. In Section 6 we dig into the transcripts of the trial and apply topic modeling methods to provide suggestive evidence of how the professionalization of the trials affected the courtroom.

2 Historical Background

2.1 Criminal Prosecution in London

The criminal justice system of 18th century England was largely characterized by penalties much harsher than those doled out today. A complex and non-centralized system of statutes, acts, and decrees established sentencing guidelines for most offense categories. At the start of the 18th century, more than 200 crimes were punishable by execution (Bindler and Hjalmarsson, 2018a; Wade, 2009). The 18th century saw gradual reform in this “Bloody Code”, with less serious offenses (such as pickpocketing for instance) no longer treated as capital crimes. The era’s distinction between misdemeanor and felony was often blurred. The four Offense Against a Person Acts of 1828, 1837, 1861, and 1875 provide some degree of the formalization to the criminal code (England and Wales Law Commission, 2009).

Those accused of crimes in London would have their cases heard in either the Old Bailey or at a lower magisterial court. Felony charges, as well as more serious misdemeanors, were tried at the Old Bailey's higher courts. Each session at the Old Bailey would consist of a series of cases tried before the same judge and that session's juries.⁴ Every case would be heard in front of a jury particular to the crime's location, either the city of London or Middlesex County (Cockburn and Green, 1988). Contemporaneous writings estimate that the average trial "never ... exceeded eight and a half minutes" (Wontner, 1833).

The accused would often find him/herself jailed in Newgate Prison; a prison originally built in the 12th century out of a gate of the ancient Roman London Wall. The Old Bailey courthouse was built connected to this prison, allowing for the quick movement of prisoners between confinement, trial, and execution (Halliday, 2007). Conditions inside the prison are recorded as being particularly harsh. Those sentenced to death spent their final days chained and shackled underground in open sewer dungeons. The remaining prisoners lived under the control of gaolers: Private administrators hired by the sheriffs managed the prison. Poor hygiene and medical care allowed illness to run rampant through the prison. Dozens of inmates died annually due to the poor conditions of confinement (Halliday, 2007).

2.2 Proceedings

Since the 17th century the actions of notorious criminals were commonly published and enjoyed by the people of England. In January 1679 the Court of Alderman for London ordered that accounts of the proceedings at the Old Bailey can only be published with approval of the Lord Mayor. This acted to monopolize the market into an annual publication known as the *Old Bailey Proceedings*. In 1787 publication became subsidized and copies of the Proceedings were made for public officials. The use of it as an official record of criminal trials in the city of London and Middlesex County continued until the early 1900s.⁵ The digitization and posting of the Proceedings was made possible through recent funding from Arts and Humanities Research Council and the Economic and Social Research Council affiliated with the British government.

The Proceedings provide a unique opportunity to explore the testimony, questioning, and rulings at the Old Bailey. Entries include the identify of the judge and accused, provide a transcript of the testimony given at trial, and present the ruling and sentence (if applicable). We will use data taken from these transcripts in our analysis.

2.3 Defense Counsel Prior to 1836

Prior to the 18th century the administration of criminal justice in London can be categorized as "straining a system designed for rural communities rather than metropolitan conurbation" (May, 2003, p.14). Those

⁴Juries were made up solely of wealthy men, with the Juries Act of 1825 requiring that all jurors own property and be between the ages of 21 and 60. Magisterial courts handled the least-serious offenses. Each session of the court would begin with the assembling of two separate twelve-man juries; one composed of men from London and the other composed of men from greater Middlesex County. Positions as judges were primarily awarded to wealthy and educated men, either as acknowledgment of their legal expertise or as a reward for political favors (Cockburn and Green, 1988). See Hanly (2021) for a further discussion of jury selection procedures.

⁵See oldbaileyonline.org and Shoemaker (2008) for historical accounts.

harmful by the crime prosecuted their cases directly, and those accused of committing crimes were forced to defend themselves. As discussed by Bindler and Hjalmarsson (2018a), the most serious crimes came with the possibility of capital punishment. One could gain the “benefit of clergy” or a royal pardon to avoid this harsh punishment. May (2003) summarizes the administration of criminal justice in the eighteenth century as being “underpinned by the dual premises of terror and mercy” (p.13).

In the 18th century, attorneys, known as solicitors, were introduced to prosecute criminal cases. The professionalization of prosecution was driven initially by government agencies defending the state’s interest. For example, the Royal Mint, Treasury, and Bank of England were able to use professional solicitors to prosecute counterfeiting and coinage violations. Consequently, an industry of trained solicitors arose to prosecute crimes. In addition, the city of London created a City Solicitor position to represent the city in criminal prosecutions. See Koyama (2012) for a detailed discussion of private prosecution services during this period of time.

Those accused of committing a felony crime, though, were prohibited from having attorneys represent them at trial. It was a common viewpoint that the criminal trial’s objective was a search for the truth and that defense counsel would simply act to distort the truth. The dominant belief was that the judge offered the best protection of the prisoner’s rights during a trial (May, 2003). Defendants should be motivated to tell the truth and did not have need for counsel.

Ironically, while felony crimes prohibited professional representation, misdemeanor crimes had developed a tradition of having professional solicitors provide defense services. Often misdemeanor crimes were closely connected with civil harms, such as fraud or forgery. Consequently, businesses were often connected with the defense and business owners hired professional attorneys to represent their interests.⁶ May (2003) argues that the introduction of lawyers first arose to help businesses prepare their defense in these more minor cases.

Thus, leading into the 19th century, the felony courtroom in London was a rather lop-sided undertaking. An experienced judge and a jury of affluent property owners determined guilt and sentencing. The prosecution of the crime was conducted by a professional attorney, yet the defendant was alone to defend him/herself in the courtroom.⁷

2.4 Prisoners’ Counsel Act of 1836

Prior to 1836, the courthouses of England were not entirely absent defense counsel. At the discretion of judges, a solicitor could be used to perform a cross-examination of witnesses on behalf of the accused (Langbein, 1978). Similarly, defense counsel could address the court on issues of the law, again, if allowed by

⁶Treason cases also allowed for defense counsel, but no such case arises in the Proceedings during the time period we study.

⁷As an illustration, a popular publication at the time, intended to be a guidebook, encouraged prosecuting counsel that they “ought to confine himself to a simple detail of the facts he expects to prove, because the prisoner has no opportunity of laying his case before the jury by his counsel; and even the privilege of stating circumstances, however dryly, in such order and direction as may tend most directly to a particular conclusion is, of itself, no small advantage accorded to the prosecutor, and certainly should be exercised with great forbearance and caution ... In cases of misdemeanour, the prosecuting counsel is not thus restricted, because here the defendant is allowed to make a real defence by his counsel and, therefore, here the counsel for the prosecutor may not only state his facts, but reason upon them, and anticipate any line of defence which his opponent may probably adopt” (Dickenson and Talfourd, 1929).

the judge (Beattie, 1991). These two privileges do not seem to be frequently awarded or taken advantage of. Importantly, prior to 1836, counsel representing the accused was unable to address the jury, which included the inability to make opening or closing speeches.

While defense counsel's contribution to the trial was limited, case preparation was still possible prior to the trial (Langbein, 1999). However, defense counsel could not gain access to defendants detained at Newgate. In fact, neither the defendant nor a solicitor even had the right to a copy of either the formal charges or the depositions taken.⁸ When a connection between a defendant and a solicitor was made prior to the trial, the solicitor's actions tended to be restricted to helping the defendant author a written statement which s/he could then read aloud at court. In the rare circumstances where solicitors were employed, they were only done so in the moments immediately before the trial and were typically only then informed of the charges levied and the evidence present (May, 2003). The inability of a solicitor to defend the accused to the jury or to prepare a defense prior to the trial date, coupled with the not-insignificant fees charged, led to there being little incentive for defendants to hire professional help in felony cases.

Reformers worked to remedy this deficiency. Attempts to pass an act through Parliament granting a full right to counsel failed several times during the early 19th century (Bentley, 2003).⁹ Proponents of this sort of reform argued that it was not logical that a defendant be "allowed counsel to defend him for a twopenny trespass but denied the like privilege where his life was at stake" (Bentley, 2003). Legal periodical press at the time referred to the Act as promoting "fair-play speech" referring to the fact that prior to the 1836 only the prosecution could have a legal professional speak to the jury on the matter. The question at hand was whether professional advocates would promote the search for truth. Summarizing the opposition, Serjeant Spankie commented during the Parliamentary debates that

it appears to me ... the effect of allowing counsel in felonies would totally destroy the temper, moderation, and sobriety of the administration of criminal justice. The counsel for the prosecution would follow the bad example of the counsel for the prisoner. In a public contest the strict discharge of duty would yield to the fame of eloquence and the ardour of victory; the authority of the judge would be despised; the jury would be exposed to the corruptions of the worst arts of the forum (Grant, 1838).¹⁰

Thus, it was uncertain to policymakers at the time whether rectifying the lop-sided nature of felony trials by providing a right to counsel would promote the truth or aid the guilty's escape from punishment. In fact, some parliamentarians suggested eliminating the use of professional attorneys for the prosecution. Thomas Wontner, writing on developments in the courts and law in Great Britain in 1832, reports that an honorable member of parliament would be willing to join a bill that would "take away the privilege of counsel from

⁸May (2003) explains that professional etiquette prohibited a solicitor from entering the gaol. Thus, there would not be a chance to interview a client.

⁹This time period, more generally, was one of many reforms to political, as well as legal, reforms. See, as an example, Aidt and Franck's (2015) analysis of revolutionary riots and extensions of the franchise in the early 1830s.

¹⁰Stated more succinctly, James Harmer testified that "the generality of profession are of opinion that counsel ought not to be allowed to speak for a prisoner, and that the practice would be injurious to the accused" (quoted in Wood (2003)).

the prosecutor, so that both parties may be placed on equal footing” (Wontner, 1833, p.322). Another parliamentarian opposed efforts to provide a right to counsel arguing that it would “lead to a trial of skill and the prosecutor, being the richer party, would generally succeed” (p.323). Wontner himself argues against such logic stating that a “ridiculous fiction is resorted to of the judge being counsel for the prisoner” (p.324). Clearly, political leaders in Great Britain at the time recognized the unfair, lop-sided nature of felony criminal trials and debated the best institutional change.

Finally, in 1836 Parliament passed the Prisoners’ Counsel Act (hereafter Act). The Act changed the dynamic of the criminal trial in England by allowing all defendants accused of a felony the right *to make full Answer and Defence thereto, by Counsel learned in the Law* (Beattie, 1991). In practice, this act extended the right to a defense counsel from misdemeanor offenses to all accused. The Act solidified the defense counsel’s involvement in felony cases, taking the final step in allowing defense counsel to directly and fully address the jury (Langbein, 1983). Being able to address the jury meant that defense counsel was now able to directly question and debate the veracity of the evidence presented against their client (Langbein, 1999).

Still, despite the Act granting this right to all accused, many defendants still represented themselves at trial due to an inability to pay for private representation. Bentley (2003) points out that at the turn of the 20th century, more than sixty years after the Act’s implementation, defendants represented by counsel were still a minority. The end of the 19th century saw numerous attempts to address this issue. Eventually, in 1903 the Poor Prisoner’s Act was passed allowing prisoners to plead *in forma pauperis* and to be assigned counsel by the court (Bentley, 2003).

Our identification strategy relies on a control group of crime types that previously had access to defense counsel prior to the Act’s implementation. Legal historians assert that a comprehensive, well-articulated criminal code did not exist in the common law structure of 19th century London (Wade, 2009). “Misdemeanor” crimes were allowed to have defense counsel, but the delineation of which crimes this includes is not clear. The most minor offenses were handled by magisterial courts. Some crimes one would normally associate with misdemeanor offenses, such as libel, are prosecuted at the Old Bailey. For the purpose of this study, we define our control group of less-severe crimes using two lists. First, from the Offence Against Persons Act we use a list of crimes clearly articulated there as misdemeanors. Any crime type on this list that also arises at the Old Bailey is treated as a control crime. Second, as stated previously, over 200 crimes came with the possibility of a death sentence during the Bloody Code of the 18th century. Very few crimes escaped this severe sanction. Any crime that at no point in the 1700s or 1800s had the possibility of capital punishment is also included in our control group. The treated crimes, then, are those previously subject to the Bloody Code and were not those formally defined as misdemeanors.

3 Evidence of Defense Counsel at the Old Bailey

The Act, as stated, provided a greater role for defense attorneys in felony criminal cases in London. By being able to be involved in the case before trial and being able to directly question witnesses and address

the jury, defense counsel presumably provides more value to the accused. Recognizing many defendants will lack resources, hiring a defense attorney comes at a significant opportunity cost. Thus, not all defendants will pay for the service. At the margin, then, we would expect more to be hired. We look to the transcripts from the trials at the Old Bailey to verify that we indeed see an increase in the prevalence of defense counsel.

To do this, we first scrape all transcripts from criminal trials at the Old Bailey using the online portal, *The Proceedings of the Old Bailey*.¹¹ We focus on the 80-year period centered on 1836. The transcripts are formatted into a question-then-answer format. Each question and corresponding answer is separated into distinct paragraphs. The identity of the questioner is typically noted. For example, if a question comes from the judge, the question will start with “Judge Q: ...”. The attorney for the prosecution, the defense counsel (if present), the judge, and even members of the jury can ask questions.

When a question is asked by a defense attorney to a prosecutor’s witness, the question is recorded as “Cross-examine: ...”. We search the text of each case for the presence of the phrase “cross examine” (with punctuation and capitalization eliminated) and different tenses (i.e., “cross examined”). Further, we add searches for “solicitor”. An indicator variable is created if any of these terms exists in the text.

We take this indicator variable as our measurement of the presence of defense counsel. We emphasize, though, that this measurement is not perfectly precise. The court recorder may fail to identify a defense attorney when one is hired.¹² A question may not be assigned an author or a transcript can be incomplete.¹³ Thus, our measure should be thought of as a lower bound to the actual prevalence of a defense attorney.¹⁴ Further, this measurement does not account for the new, important activities allowed by the Act; namely the ability to directly address the jury and have access to charging information, witnesses, and the accused prior to trial. Nevertheless, it should measure fairly well changes in the prevalence of defense counsel at the Old Bailey. Figure 1 compares the proportion of cases that record the existence of defense counsel before and after the 1836 Act.

Prior to the Act’s implementation (the light gray bars), the less-serious crimes had a substantially higher rate of having defense counsel presence, as measured by the keyword search of the trial transcripts. Control crimes come with substantially less severe punishments, so the incentive to hire a costly attorney is lower, *ceteris paribus*. The fact that they have (measurable) representation at almost a 50% higher rate confirms that solicitor’s value was restricted in felony cases. The growth in the use of defense counsel is substantial for the treated crimes, while only a negligible change arises for the control crimes (dark gray bars).

Figure 1 shows that defense counsel grows in prevalence for felony crimes, overall, but not for the less-serious crimes. This is evidence in line with our argument that the less-serious crimes can be used as an

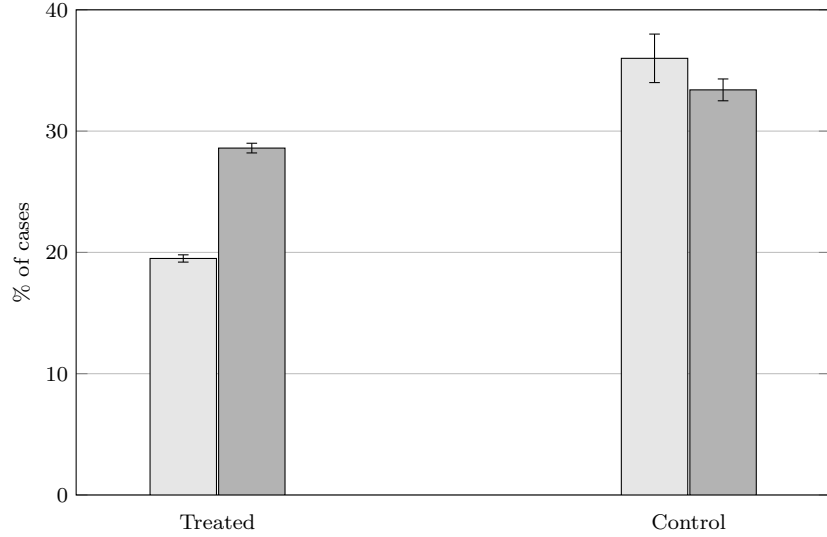
¹¹www.oldbaileyonline.org

¹²Recognizing the Proceedings’s initial target audience, legal details that may have been of less interest to the general population were often omitted (Gallanis, 2006). This could very well include recognizing the defense counsel.

¹³To this point, Langbein (1983) compares a judge’s personal notes to the official record of the *Old Bailey Session Papers* and found seven instances of the failure to note counsel’s presence. There exists a general consensus that the *Session Papers* under-report the frequency of counsel for both the defense and prosecution (Beattie, 1991).

¹⁴It is also conceivably possible that a defendant brings his/her own witnesses that are cross-examined by the prosecutor’s attorney. While typically the name of the attorney representing the prosecutor is used for each question, it could occasionally be the case that the cross-examination noted is done by a prosecutor.

Figure 1: Use of Defense Counsel Before and After the 1836 Act



Each column depicts the proportion of observations where a defense counsel is present in the transcript of the trial. The left two columns are for our set of treated crimes and the right two columns are for the control crimes. The light gray columns are the proportion in the years prior to the Act’s implementation, while the dark gray columns are for the years where the Act is in place. The 95% confidence intervals are depicted.

un-treated, control group. Thus, a difference-in-difference estimation strategy will be used to obtain a causal estimate of the impact of defense counsel on trial outcomes.

4 Data

Our primary data source is the quantification of the records of the Old Bailey. Specifically, we start with the same data set as used in Bindler and Hjalmarsson (2018a).¹⁵ Basic information is recorded for each criminal case between 1715 and 1900. The data set is at the offender level. That is, a defendant may have more than one charge levied against him/her in an observation in the data set. If two or more individuals are accused of committing a crime together, each defendant is a separate observation. There are 265,662 observations in the full data set.

For each, the crime the defendant is charged with, whether s/he is convicted, and (if convicted) the sentence imposed is recorded. Basic demographic information such as the defendant’s gender and age are provided. The Old Bailey handled criminal cases arising in the city of London as well as Middlesex County. This distinction is measured when available. Basic descriptive statistics are provided in Table 1.

We use indicator variables to record when a trial transcript does not provide information on age, gender, or jury identity. Gender has almost full reporting with just less than 80% of the defendants being male. Since Bindler and Hjalmarsson (2017) show the changing age profile at the Old Bailey and Bindler and Hjalmarsson (2020) highlight gender disparities, it will be important to include these controls in our estimations. In our

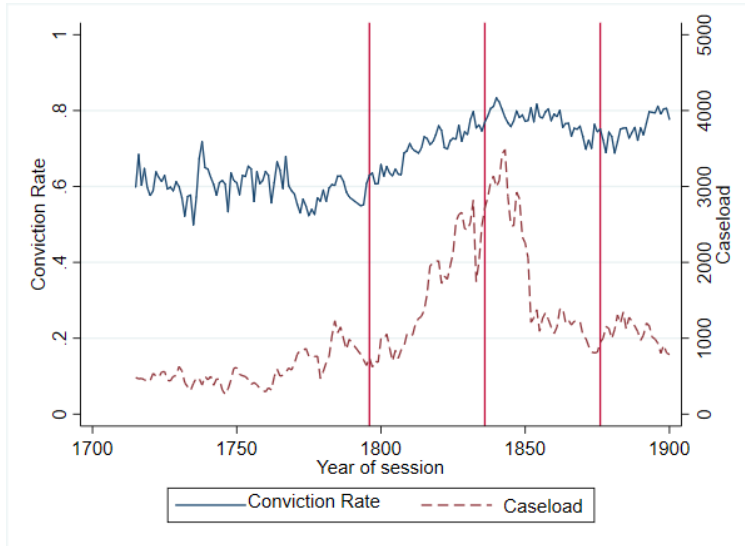
¹⁵This data is made available on the *AEJ*’s website (Bindler and Hjalmarsson, 2018b). We greatly appreciate these authors making the data available.

Table 1: Descriptive Statistics

Control	Mean
Age less than 18	0.1277
Age missing	0.2347
Male defendant	0.7953
Gender missing	0.0024
Number of defendants	1.463
London jury	0.0482
Middlesex jury	0.1890

The baseline time period (1796 to 1876) is considered here; $N = 136,144$. The omitted group for the jury composition is unknown compositions (i.e., missing). All variables are binary except the number of defendants. For it, the standard deviation is 1.366 with a median of 1 and maximum value of 35. Of the 333 observations with gender missing, 82 are also missing the age.

Figure 2: Conviction Rate over Time



time frame, 69.9% of cases have only one defendant and the bulk of the observations do not identify the jury’s identify ($\approx 75\%$). London went through numerous changes between 1715 and 1900. Institutions, economic well-being, and cultural identity adjusted. Wars were fought and the Industrial Revolution took hold. Therefore, we choose in our baseline specification to limit the time range to the forty years before and the forty years after the Act’s implementation. This mitigates some of the significant structural changes in the city. We will, of course, assess the sensitivity of our results to the cutoffs. Thus, in our baseline analysis we consider criminal cases at the Old Bailey between 1796 and 1876. There are 135,365 observations.

Regarding the outcome variable of interest, for the baseline period of 1796 to 1876 (± 40 years around the Act’s implementation), 75.2% of the observations had a conviction. The crime prosecuted falls within our treated crime types for 89.4% of the observations. Figure 2 provides the time series for the convictions. We collapse the data so that the figure reports the conviction rate for each year.

The three vertical lines define the time periods used in study. The baseline specifications will, as stated,

consider crimes between 1796 and 1876. As one can see, the conviction rate grows rapidly during our selected sample period until the early 1840s (solid line). It is then relatively flat after that. This figure does not account for the mix of crimes tried at the Old Bailey or the changing demographics of the pool of defendants. It also illustrates that we must carefully account for the time trend in prosecutions prior to the Act’s implementation.

In addition, Figure 2 illustrates the number of cases handled at the Old Bailey each year (dashed line). Similar to the conviction rate, it illustrates a steady increase into the 1840s, but the caseload experiences a substantial decline reaching its pre-boom level by the 1850s.

By omitting cases heard at the Old Bailey in the 1700s, periods of relative low activity and convictions are excluded from the analysis. The time period considered focuses on the period of heightened activity when pressure to dispose of cases (among other concerns) is highest.

5 Difference-in-Difference Results

The primary econometric model to estimate is

$$Conviction_{icy} = \beta_1 Post_{iy} \times Treated_{ic} + \tau_{iy} + \kappa_{ic} + X_{icy}\beta + \epsilon_{icy}. \quad (1)$$

The dependent variable, $Conviction_{icy}$, as previously described is an indicator variable equal to one if case i , of crime type c in year y , resulted in a conviction. $Post_{iy}$ is an indicator variable equal to one if $y > 1836$, the year of the Act’s implementation. The indicator $Treated_{ic}$ is equal to one if observation i is a treated crime type. The indicator variables τ_{iy} are a set of year fixed effects, and κ_{ic} are crime controls. X_{icy} is a set of control variables (those provided in Table 1). The coefficient β_1 is the difference-in-difference coefficient of interest. If $\beta_1 < 0$, then felonies treated with the Act experience a (relative) reduction in the conviction rate after the Act is passed, while if $\beta_1 > 0$, then the treated felonies experience an increase in the conviction rate. Alternatively, if $\beta_1 = 0$, then the creation of the right to counsel had no effect on trial outcomes.

There are thirty-two crime categories included in the crime fixed effects – nine of which make up the controls and twenty-three crimes which make up the treated group. Standard errors will be clustered at the crime by year level. Since sanctions are crime specific but can vary over time as legislation changes each crime’s punishment severity, we feel that this is the appropriate level for the experiment (Abadie *et al.*, 2017).

5.1 Parallel Trends

The estimation strategy, summarized in equation (1), can be interpreted as providing a causal identification of the Act’s impact on convictions at the Old Bailey so long as the treated and control crime types are following parallel trends prior to the Act’s implementation. If so, then divergence between the two types of cases after the Act is put into place capture the effect of it on trial outcomes. If the difference in the

Table 2: Test of (Linear) Pre-Act Parallel Trends

	<i>Baseline</i>	<i>Alternative Time Windows</i>		
coverage:	±40 years	±30 years	±50 years	BH window
years:	[1796-1876]	[1806-1866]	[1786-1886]	[1803-1871]
	[1]	[2]	[3]	[4]
Year ($\hat{\alpha}_1$)	0.0048 *** (0.0013)	0.0044 * (0.0024)	0.0040 (0.0009)	0.0044 ** (0.0019)
Year x Treated	-0.0005 (0.0012)	-0.00003 (0.00236)	0.0004 (0.0010)	0.0002 (0.0018)
Crime Fixed Effects?	Yes	Yes	Yes	Yes
R^2	0.055	0.053	0.057	0.054
AIC	149,514	126,621	172,145	136,502
N	135,696	126,621	172,145	125,635
# clusters	81	61	98	69

Dependent variable is equal to one if the case resulted in a conviction. Each specification differs in the time range considered. Controls include an indicator variable for being over the age of 18, age missing, male, gender missing, trial using a London jury, and the trial using a Middlesex jury. We include the number of defendants involved in the case. There are 32 crime indicator variables included. Only $\hat{\alpha}_1$ and $\hat{\alpha}_2$ are reported here, but the full model specified in (2) is estimated. Standard errors presented in parentheses are clustered at the crime by year level; *** 1%, ** 5%, * 10% level of significance.

conviction rates do not change, and the two are on parallel paths, then the Act can be argued to be ineffective at changing courtroom outcomes.

To provide a test for parallel trends, we consider a linear time trend in the pre-treatment period. Specifically, we estimate

$$\begin{aligned}
Conviction_{icy} = & \alpha_1 Year_{iy} + \alpha_2 Year_{iy} \times Treated_{ic} \\
& + \alpha_3 Post_{iy} \times Year_{iy} + \alpha_4 Post_{iy} \times Treated_{ic} \times Year_{it} \\
& + \alpha_5 Post_{iy} + \alpha_6 Post_{iy} \times Treated_{ic} + \kappa_{ic} + \epsilon_{icy}.
\end{aligned} \tag{2}$$

The coefficient α_1 captures the linear time trend in convictions for the non-treated crime types and α_2 measures whether the treated crimes follow a different time trend in the pre-Act period. Testing whether $\hat{\alpha}_2 = 0$ is our assessment of parallel trends. The coefficients α_3 and α_4 allow for the time trends to be distinct after the Act takes effect. The coefficients α_5 and α_6 allow for jumps at the Act's implementation. Table 2 provides the results.

The table assesses parallel trends over various time frames. For each, while there is a slight, gradual increase in the conviction rate prior to the Act's implementation, the treated crimes and the untreated crimes do not follow different time paths. For example, in our baseline specification, [1], the p -value on the interaction is 0.714. For the most part, the conviction rate for both felonies and the control crimes are flat over time.¹⁶

¹⁶While not presented in Table 2 (since the focus is on the pre-Act time period), the time trend diverges after the Act's implementation. A level shift up in the time trend occurs ($\hat{\alpha}_5 = 8.03$; $p < 0.01$) and the time trend for the treated crimes after the treatment does not differ from the non-treated crimes ($\hat{\alpha}_6$ has $p > 0.6$).

In addition, it is also worthwhile to test for nonlinear discrepancies between the treated and control as the assumption of linear time trends may be too restrictive. Section 5.3.3 evaluates the dynamic effects of our main results by breaking up the 40-year window prior to the Act’s implementation into eight pentads (i.e., five-year groups) and the eight pentads following it. These are each interacted with the treatment indicator and the set replaces $Year \times Treated$ and $Post \times Year \times Treated$ in equation (2).

The results are presented there and provide additional confirmation of common trends in the pre-Act periods. Thus, our identification strategy permits causal identification.

Another concern regarding the identification strategy is that the stable unit treatment value assumption (SUTVA) holds (Rubin, 1978). What is essential is that the treatment on the felony crimes do not spill over and affect the prosecution of the misdemeanor crimes. SUTVA would be violated if, for example, the jury’s willingness to convict in the misdemeanors change after defendants in trials for felony crimes receive the right to have full use of defense counsel. While unable to test directly, this seems unlikely. SUTVA would also be violated if the legal institution’s change affected either the charging decisions or the population’s willingness to commit crimes in the first place. If either occurred, then the “control” crimes after the Act’s implementation would no longer match the “control” crimes prior to the Act so that, as a consequence, the estimate would no longer serve as a reliable counterfactual and the treatment effect’s estimate would be biased. Thus, SUTVA would be violated if the mix of cases arising at the Old Bailey changes with the Act’s implementation. Our supplemental appendix explores the sensitivity of the results to changes in the case mix and shows that the main result is not driven by case mix.

5.2 Main Result

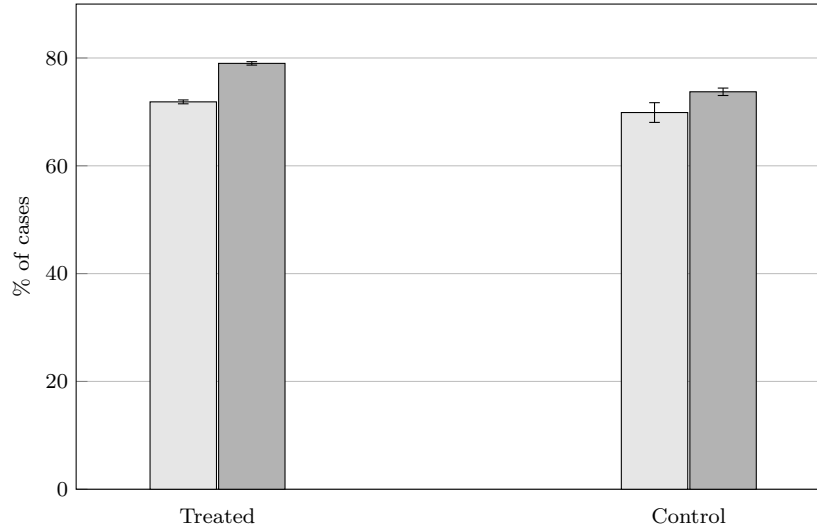
The primary question we ask here is how did the introduction to the right to counsel influence convictions of cases at trial. We first partition our data set into those treated crimes and the untreated ones and compare the mean conviction rate prior to the Act’s implementation to the mean after. Figure 3 depicts them.

Prior to the Act’s implementation, the conviction rate for the two types of crimes were strikingly similar. While the control crimes experienced a slight increase, the conviction rate jumps up substantially for the treated crimes. Here, the conviction rate increases by 9.9%. Rather, the net expansion in the conviction rate was an increase of 3.89 percentage points.

We now turn to the estimation of equation (1) to establish whether the observation in Figure 3 is robust. Table 3 presents the results.

The first column provides our main result. That treated felonies, relative to the control crimes, experience an increase in the conviction rate suggests that the right to counsel corresponds to an increase in convictions. Not only does the statistical significance remain when alternative time periods are considered, but the magnitude of the difference-in-difference coefficient is stable. It achieves its largest value for the time period used in Bindler and Hjalmarsson (2018a) of 1803-71; column [4]. Using our baseline specification, [1], at the mean the conviction rate increases for felonies by 3.0%, which is in line with the observation in Figure 3.

Figure 3: Convictions Before and After the 1836 Act



Each column depicts the proportion of observations with a guilty conviction. The left two columns are for our set of treated crimes and the right two columns are for the misdemeanor crimes (full definition). The light gray columns are the proportion in the years prior to the Act's implementation, while the dark gray columns are for the years where the Act is in place.

Table 3: Main Result

	<i>Baseline</i>	<i>Alternative Time Windows</i>		
coverage:	±40 years	±30 years	±50 years	BH window
years:	[1796-1876]	[1806-1866]	[1786-1886]	[1803-1871]
	[1]	[2]	[3]	[4]
Post x Treated	0.0250 *** (0.0064)	0.0239 *** (0.0063)	0.0190 ** (0.0080)	0.0263 *** (0.0063)
Crime Fixed Effects?	Yes	Yes	Yes	Yes
Year Fixed Effects?	Yes	Yes	Yes	Yes
Controls?	Yes	Yes	Yes	Yes
R^2	0.514	0.554	0.458	0.528
AIC	59,165	38,097	86,860	48,981
N	135,365	117,229	153,046	125,315
# clusters	81	61	98	69
DV μ	0.7523	0.7617	0.7440	0.7583

Dependent variable is equal to one if the case resulted in a conviction. Each specification differs in the time range considered. Controls include an indicator variable for being over the age of 18, age missing, male, gender missing, trial using a London jury, and the trial using a Middlesex jury. We include the number of defendants involved in the case. There are 32 crime indicator variables included. Standard errors presented in parentheses are clustered at the crime by year level; *** 1%, ** 5%, * 10% level of significance.

5.3 Additional Results

Previous scholarship on criminal trials at the Old Bailey has documented important disparities and institutional changes. We explore whether these differences influence our main result.

5.3.1 London Metropolitan Police

The London Metropolitan Police Force was one of the world’s first professional police forces. It was established initially by the Metropolitan Police Act of 1829, and its range would be expanded by the Metropolitan Police Act of 1839. Prior to the formation of this professional force, law enforcement was conducted by unpaid appointed constables or by the army itself. The force initially consisted of slightly under 1000 officers of varying rank, divided into divisions and tasked with patrolling the city’s seventeen newly defined police territories (Emsley, 1991).¹⁷ By many accounts, the police force was initially quite unpopular and constables were often subject to verbal and physical assault (Emsley, 1991). Nevertheless, this police force patrolled London, investigated crimes, apprehended known criminals, and brought them forth for criminal proceedings.

Bindler and Hjalmarsson (2021) evaluate the effect of the creation of the police force on crime. In our analysis, one can be concerned that the cases arising at the Old Bailey may be affected. More thorough investigations could potentially improve the evidence and lead to more convictions for example. The concern is that our estimation strategy is simply capturing changes in law enforcement practices rather than defense counsel.

Since our difference-in-difference estimation strategy includes year fixed effects, which will capture year by year differences in the caseload, the threat to our identification arises if the effect of the police force differs between the treated and control crimes.

To address this concern, we create an indicator variable, $Police_{iy}$, which is equal to one if $y > 1829$, interact it with the treatment indicator, and add it to equation (1). Table 4 presents the results.

Column [1] only includes the interaction between the introduction of the police force and the treated crimes. Notably, the coefficient is smaller and statistically insignificant. Column [2] provides better information. Once the introduction of the police force is separated from the introduction of the right to counsel, contrasting effects arise. Criminal convictions in the decade prior to the Act’s implementation see a decrease in conviction rates, relative to the untreated control crimes. Once this effect is removed, the magnitude of the right to counsel’s impact grows. Therefore, not only does the creation of a police force not explain away our result, but controlling for its introduction strengthens our finding.

5.3.2 Capital Punishment

As analyzed by Bindler and Hjalmarsson (2018a), this time period in British criminal law represents an important transition between a Draconian style criminal justice system, where most felony crimes came with the death penalty, to one where less-severe sanctions were used. They show that the jury’s convictions adjust

¹⁷See Davies (2002) for a further discussion of private policing in London.

Table 4: Metro Police

coverage:	±40 years	±40 years
years:	[1796-1876]	[1796-1876]
	[1]	[2]
Police x Treated	0.0131 (0.0098)	-0.0150 (0.0104)
Post x Treated		0.0317 *** (0.0060)
Year Fixed Effects	Yes	Yes
Crime Controls?	Yes	Yes
Controls?	Yes	Yes
R^2	0.529	0.529
AIC	54,754	54,744
clusters	81	81
N	135,363	135,363

Dependent variable is equal to one if the case resulted in a conviction. Controls include an indicator variable for being over the age of 18, age missing, male, gender missing, trial using a London jury, and the trial using a Middlesex jury. We include the number of defendants involved in the case. There are 32 crime indicator variables included. Standard errors presented in parentheses are clustered at the crime by year level; *** 1%, ** 5%, * 10% level of significance.

with the dropping of capital punishment. To that end, we estimate the following econometric model:

$$\begin{aligned}
Conviction_{icy} = & \gamma_1 Post_{iy} \times Treated_{icy} \times NotDeathEligible_{icy} + \gamma_2 Post_{iy} \times Treated_{ic} \\
& + \gamma_3 Post_{iy} \times NotDeathEligible_{icy} + \gamma_4 Treated_{ic} \times NotDeathEligible_{icy} \\
& + \gamma_5 NotDeathEligible_{icy} + \tau_{iy} + \kappa_{ic} + X_{icy}\gamma + \epsilon_{icy}.
\end{aligned} \tag{3}$$

This triple-difference specification identifies whether there is a systematic discrepancy between the effect of the expansion to a right to counsel for felony crimes currently eligible for capital punishment and those not. The indicator variable *Not Death Eligible* is equal to one if the crime came with the possibility of capital punishment earlier in British courts, but at the time of the observation has had that maximum possible sanction reduced. Since crimes that were never eligible for the death penalty are included in the control group, the triple difference interaction captures whether the increase in the conviction rate varies by the current severity of the crime's sanction. Table 5 presents the estimation results.

Consistently, the triple-difference coefficient is statistically insignificant. Thus, there does not seem to be a relationship between whether a felony crime is currently eligible for the death penalty and the impact of the right to counsel on the jury's willingness to convict.¹⁸

5.3.3 Is It the Solicitors?

The results up to this point can be thought of as an intent-to-treat analysis. That is, we evaluate how the addition of a right to counsel affects convictions. We have not considered whether defendants actually hire defense counsel after the Act's implementation and whether our effect is limited to only those cases where

¹⁸If, alternatively, the difference-in-difference term in equation (1) is disaggregated into eligible and not eligible crimes, both are positive, statistically significant at the 1% level, and relatively similar in size.

Table 5: Triple Difference

	<i>Baseline</i>	<i>Alternative Time Windows</i>		
coverage:	±40 years	±30 years	±50 years	BH window
years:	[1796-1876]	[1806-1866]	[1786-1886]	[1803-1871]
	[1]	[2]	[3]	[4]
Post x Treated x Not Death Eligible	-0.1150 (0.1328)	-0.1250 (0.1411)	-0.0409 (0.1522)	-0.1217 (0.1395)
Crime Fixed Effects?	Yes	Yes	Yes	Yes
Year Fixed Effects	Yes	Yes	Yes	Yes
Controls?	Yes	Yes	Yes	Yes
R^2	0.529	0.569	0.475	0.544
AIC	54,735	34,040	81,995	44,677
clusters	81	61	98	69
N	135,363	117,299	153,046	125,315

Dependent variable is equal to one if the case resulted in a conviction (either through a trial conviction or a guilty plea). Each specification differs in the time range considered. The full model presented in equation (3) is estimated, but only the triple difference coefficient is presented. Standard errors clustered by crime are presented in parentheses; *** 1%, ** 5%, * 10% level of significance.

one is present. To expand our analysis, we introduce our indicator variable *Counsel*, which equals one if the dictionary method discussed in Section 3 identifies the presence of a defense attorney, to the analysis. This is the same variable used to identify the change in the defense counsel’s presence, as depicted previously in Figure 1.

To explore the difference between cases where a defense attorney is present from cases where one cannot be detected from the text of the transcript, we estimate the following econometric model:

$$\begin{aligned}
Conviction_{icy} = & \delta_1 Post_{iy} \times Treated_{icy} \times Counsel_{icy} + \delta_2 Post_{iy} \times Treated_{ic} \\
& + \delta_3 Post_{iy} \times Counsel_{icy} + \delta_4 Treated_{ic} \times Counsel_{icy} \\
& + \delta_5 Counsel_{icy} + \tau_{iy} + \kappa_{ic} + X_{icy}\delta + \epsilon_{icy}.
\end{aligned} \tag{4}$$

The triple difference coefficient, δ_1 , identifies whether the treatment effect differs between cases where defense counsel is used and cases where they do not seem to be present. Table 6 presents the estimation results.

The first column simply adds the indicator variable for the presence of defense counsel to the main specification presented previously. The difference-in-difference coefficient is not affected by this variable’s inclusion. Further, a negative and statistically significant coefficient on this indicator exists. This suggests that either defense counsel is successful at avoiding convictions for their clients, or that there is selection in which cases result in the defendant willing and able to hire an attorney.

The triple difference estimation presented in [2] produces an estimated effect ($\widehat{\delta}_1$) that is not statistically different from zero. Thus, there is not a measurable difference between the heightened conviction rate produced after the Act’s implementation between those cases where a defense counselor is utilized and those cases where the defendant chooses not to hire an attorney. This result is important as it implies that it is not necessarily a signal being sent to the jury that a defendant is willing/unwilling to hire professional

Table 6: Use of Defense Counsel

	[1]	[2]
Post x Treated x Counsel		0.0249 (0.0199)
Post x Counsel		0.0659 *** (0.0213)
Treated x Counsel		-0.0251 (0.0179)
Post x Treated	0.0261 *** (0.0066)	0.0137 (0.0114)
Counsel	-0.0849 *** (0.0051)	-0.0354 * (0.0197)
Crime Fixed Effects?	Yes	Yes
Year Fixed Effects	Yes	Yes
Controls?	Yes	Yes
R^2	0.536	0.536
AIC	53,111	52,986

Dependent variable is equal to one if the case resulted in a conviction (either through a trial conviction or a guilty plea). Each specification differs in the time range considered. The full model presented in equation (3) is estimated, but only the triple difference coefficient is presented. Standard errors clustered by crime are presented in parentheses; *** 1%, ** 5%, * 10% level of significance.

representation. Both cases with defense counsel and those without experience a heightened probability of conviction. Thus, the introduction of the right to counsel represents an institutional change.

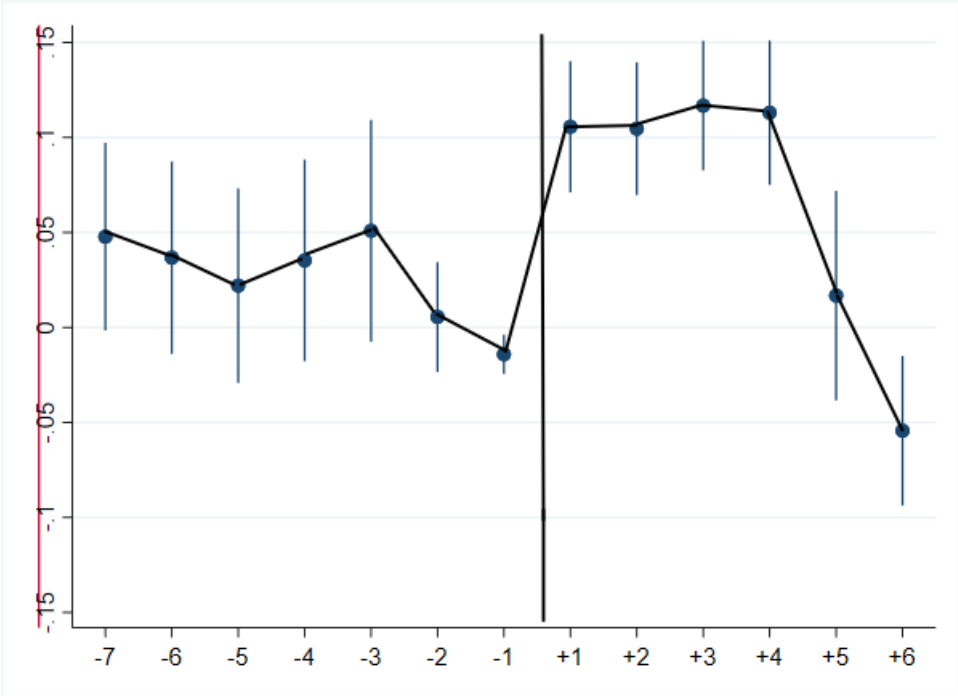
5.3.4 Dynamic Effects

Did the introduction to the full right to counsel have an immediate effect on the jury? The analysis up to this point presumes that the change in the conviction rate is abrupt and constant over time. Here, we consider the Act's dynamic effects.

To do so, we divide the time into pentads. The pentads leading up to the Act's implementation and the pentads following are interacted with the treatment indicator variable replacing $Post \times Treated$ in equation (1). Figure 4 presents the point estimates and the 95% confidence intervals of the six pentads following implementation (periods +1 through +6) and the six pentads prior to implementation (periods labeled -7 through -2).

As one can see, the effect occurs in the first pentad following the Act's implementation and has a relatively stable, positive effect for twenty years after the Act's implementation. By five pentads after its implementation, which corresponds to the 1856 to 1860 time period, the effect is essentially zero. That the effect is relatively immediate and stable for two decades points to English juries having a notable correction to their willingness to convict. While it is uncertain exactly why this effect eventually zeroes out, individuals on juries 30 to 40 years after the Act's implementation may simply not be aware of the unbalanced, lop-sided

Figure 4: Dynamic Effects



Each coefficient presented is a pentad interacted with the treatment indicator. The interactions are numbered with 0 representing the pentad immediately prior to the Act's implementation (1831-1835), interacted with the treatment indicator. All control variables specified in equation (1) are included but not depicted. An indicator variable for being after 1836 is used, rather than separate year fixed effects, and an indicator variable for being a treated crime, rather than separate crime controls, are included but not depicted. The results are estimated on the 1786 to 1886 time period, but extreme pentads are not reported as by +5 the effect is statistically zero. Hence, the six pentads following the implementation (+1 through +6) and the six pentads prior (-7 through -2) are shown. The 95% confidence intervals are shown for each.

Table 7: Heterogeneous Treatment Effects

	[1]	[2]
Post x Treated x Male	0.0300 *** (0.0064)	
Post x Treated x Female	0.0033 (0.0075)	
Post x Treated x Age less than 18		0.0378 *** (0.0071)
Post x Treated x Age greater than 18		0.0223 *** (0.0050)
Crime Fixed Effects	Yes	Yes
Year Fixed Effects	Yes	Yes
Controls	Yes	Yes
R^2	0.529	0.155
AIC	54,702	7147
N	135,363	103,593

Dependent variable is equal to one if the case resulted in a conviction. Controls include an indicator variable for being over the age of 18 (in [1] only), age missing (in [1] only), male (in [2] only), gender missing (in [2] only), trial using a London jury, and the trial using a Middlesex jury. We include the number of defendants involved in the case. There are 32 crime indicator variables included. Standard errors presented in parentheses are clustered at the crime by year level; *** 1%, ** 5%, * 10% level of significance.

nature of trials prior to the creation of the right to counsel and may no longer be correcting for the improved professionalization.

Also, as one can see, the pentads leading up to the Act’s implementation consistently include zero in the 95% confidence interval. This is further evidence supporting the finding of parallel trends and the causal interpretation of the result.

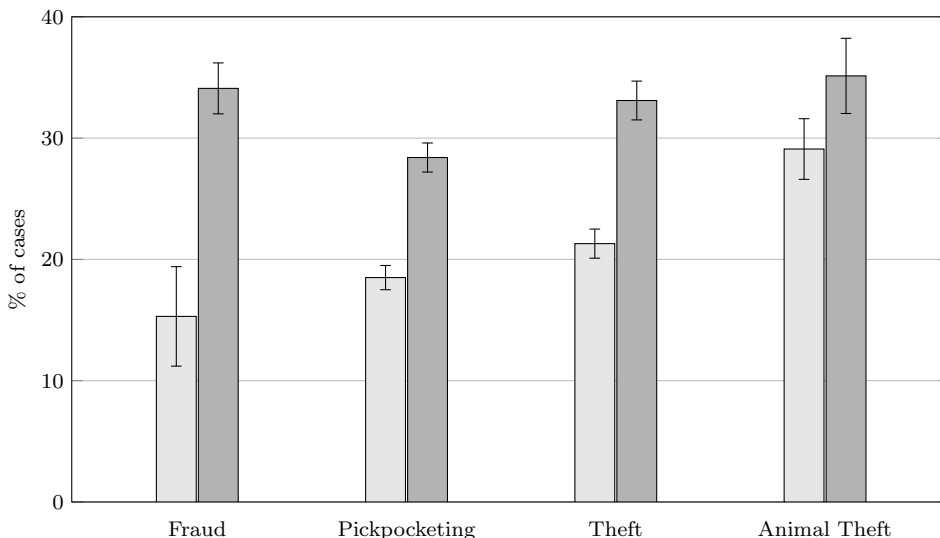
The possible exception is the pentad just prior to the Act’s implementation (*Pentad* -1). The slight reduction in the conviction rate for the (soon-to-be) treated crimes is modest. Further, the coefficient has the opposite sign as the estimated treatment effect. Thus, even if it was taken to suggest that the treated crimes were beginning to diverge from the controls, this effect would bias our results towards zero and we would be underestimating the right to counsel’s effect.

5.3.5 Heterogeneous Treatment Effects

Past research on trials at the Old Bailey, as mentioned, focused on gender and age of the defendant. Here, we explore whether the right to counsel had heterogeneous treatment effects based on the demographic factors. For each, we simply disaggregate the difference-in-difference coefficient. Table 7 provides the results.

The increased conviction rate is concentrated in male defendants. Similar-sized effects, though, are experienced by the young and old alike.

Figure 5: Before and After the 1836 Act



Each column depicts the proportion of observations where a defense counsel is present in the transcript of the trial. The light gray columns are the proportion in the years prior to the Act’s implementation, while the dark gray columns are for the years where the Act is in place. The 95% confidence intervals are depicted as well.

6 Text Analysis

We now turn to an analysis of the transcripts of the Old Bailey criminal trials. We use a popular topic modeling method in text analysis. The objective is to uncover the “themes”/“topics” covered within each trial. Our goal is to provide evidence of whether the topics covered during the courtroom trials adjusted after defense counsel was able to fully participate in criminal trials. This can shed light on the possible mechanism behind our result.

First, as previously described, we record the prevalence of defense counsel by identifying the existence of related keywords in the text depicted in Figure 1 previously. We did this for the full data set. Expanding this work, we partition the data set by the crime committed and identify which crimes exhibit the most dramatic increases in the use of defense counsel. Doing so, four crimes record clear, measurable increases: fraud, pickpocketing, theft¹⁹, and animal theft. Figure 5 illustrates.

The jumps are substantial for each ranging from a 21% increase (animal theft) to a 123% increase (fraud) in the measured rate of defense counsel presence. The appendix provides the time series for each of these four crimes as well. Each exhibits an upward trend after 1836 with a rather flat trend prior. Hence, we will focus on the four crimes that are shown in Figure 5 to exhibit large changes in the presence of defense counsel and investigate the trial transcript’s text for each subset. While topic modeling methods are subject to interpretation, we use it to provide depth to the causal analysis already conducted.

¹⁹Specifically, the crime “theft from place”, which is distinct from petty larceny, simple larceny, grand larceny, and “stealing from master”.

6.1 Latent Dirichlet Allocation

Latent Dirichlet Allocation (hereafter LDA) is a computational linguistic algorithm (Blei, Ng, and Jordan, 2003). It allows for the creation of communication measures based on topic models, which is a class of machine learning algorithms for natural language processing.

LDA allows for automatic clustering of any kind of textual documents into a user chosen number of themes, known as *topics*. It uses a probabilistic model of text data. The logic is that when authors write about a particular topic, they tend to use the same words. Hence, in texts about the same topic, similar words tend to co-occur. LDA describes each topic as a probability distribution over words, and each document as a probability distribution over topics.

Formally, each document d in a set of documents D is described as a probabilistic mixture of T topics. A document topic vector, θ_d , describes the document. Rather,

$$\theta = \begin{pmatrix} \theta_1 \\ \vdots \\ \theta_D \end{pmatrix} = \begin{pmatrix} P(t=1|d=1) & \cdots & P(t=T|d=1) \\ \vdots & \ddots & \vdots \\ P(t=1|d=D) & \cdots & P(t=T|d=D) \end{pmatrix}. \quad (5)$$

Here, $P(t|d)$ is the probability weight put on topic t within document d . Each topic t in the set of topics T is described by a probability distribution over the vocabulary of V words present in all documents. Rather,

$$\phi = (\phi_1 \quad \cdots \quad \phi_T) = \begin{pmatrix} P(w=1|t=1) & \cdots & P(w=1|t=T) \\ \vdots & \ddots & \vdots \\ P(w=V|t=1) & \cdots & P(w=V|t=T) \end{pmatrix}. \quad (6)$$

Here, $P(w|t)$ is the probability weight put on word w in topic t .

With this, the presumption of LDA is that a document is written according to the following process. First, each topic's distribution over words in the vocabulary is drawn according to the Dirichlet distribution function; $\phi \sim Dir(\beta)$. For each document in the corpus each's mixture of topics is drawn again according to the Dirichlet distribution; $\theta_d \sim Dir(\alpha)$. Thus, the researcher must only select the number of topics to organize the documents into, T , and the two hyperparameters, α and β . The likelihood of the corpus is

$$\prod_{d=1}^D P(\theta_d|\alpha) \left(\prod_{n=1}^{N_d} \sum_{z_{d,n}} P(z_{d,n}|\theta_d) P(w_{d,n}|z_{d,n}, \phi) \right). \quad (7)$$

Gibbs sampling is used to estimate the conditional probabilities that best explain the corpus of documents.²⁰

From this process, LDA calculates for each document the probability it falls within each topic. This probability distribution classifies the document as it measures the mixture of topics contained within.

LDA does not require the researcher's pre-knowledge or rely on his/her discretion when it comes to knowing which words to search for within the text. It provides a way of uncovering hidden themes in text without having to link themes to particular word lists prior to estimation. LDA is valuable when the researcher does not know a priori which words are the important ones to track. This allows one to avoid subjective judgments and to account for context.

²⁰The variable $z_{d,n}$ records the word's position in the document.

The use of LDA in economics is rather new. One notable exception is the work of Hansen, McMahon, and Prat (2018) who evaluates Federal Open Market Committee transcripts. They use the variation in topics discussed, assessed using LDA, as the dependent variable to appreciate how experience on the FOMC and transparency interact.²¹ Another application is Larsen and Thorsrud (2019) who use LDA to model topics covered in a Norwegian business newspaper and use the prevalence of the topics in a time series analysis of a measurement of asset prices. McCannon (2020) uses it to evaluate descriptions of wine adding the probability weights created by the LDA process as explanatory variables in a hedonic price regression.²²

Outside of economics, LDA is popular. It has been used effectively in related fields such as political science. Grimmer (2010) and Quinn *et al.* (2010) use LDA to analyze press releases from U.S. Senators. It has been used in marketing to evaluate online discussions of products (Tirunillai and Tellis, 2014), customer reviews in tourism management (Guo, Barnes, and Jia, 2017), and public discourse during the COVID-19 pandemic (Xue *et al.*, 2020). In finance it has also proven to be useful identifying trends in 10-K disclosures (Dyer, Lang, and Stice-Lawrence, 2017) and gauging investor complaints to the Consumer Financial Protection Bureau (Bastani, Namavari, and Shaffer, 2019).

As a baseline analysis, separately for each of the four crimes we organize the trial transcripts into ten topics ($T = 10$). As with the main result previously identified, we consider trials that occurred in the 80-year window centered on the Act’s implementation. Our supplemental appendix provides details on the data cleaning process. For example, common stopwords (such as “and” and “the”) are eliminated, as is punctuation, capitalization, numbers, and proper names. LDA is conducted for each subsample. As described, this process produces a probability distribution over the ten topics for each observation. In addition, the probability distribution over words in the vocabulary is estimated for each topic.

6.2 Analysis

To illustrate the output, consider one topic, which for convenience we label as *Topic 1*, created within the transcripts of trials involving those accused of fraud. Across the 2328 fraud trials between 1796 and 1876 the probability weight put on Topic 1 has a mean of 0.061 with a standard deviation of 0.166. Figure 6 represents the time series of Topic 1’s prevalence. In Figure 6, for illustration purposes, we collapse the data to an annual data set depicting each year’s mean and 95% confidence intervals.

As one can see, the prevalence of Topic 1 is both flat and essentially zero into the 1850s. Only after 1850 does the confidence interval begin to lie above zero. Exploring which words dominate this topic, Table 8 lists selected words that have a large probability of arising in a document that involves Topic 1.²³ In Table 8 we rank the words according to their probabilities derived from the LDA estimation.

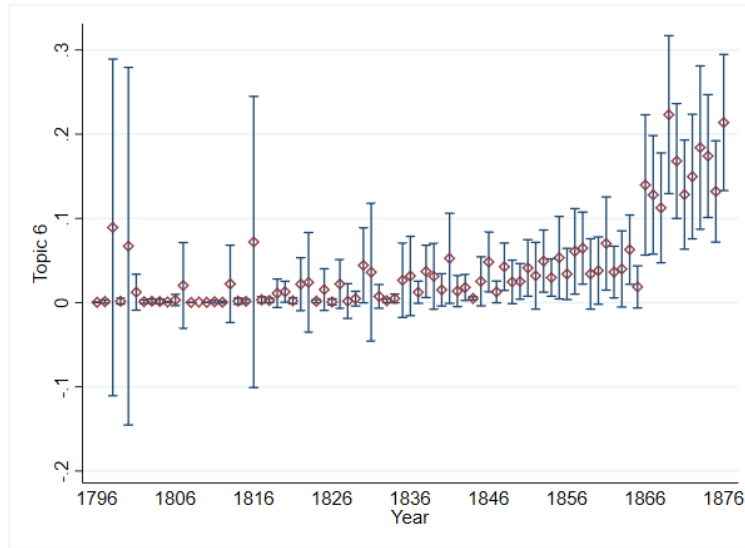
Clearly, words related to mail show up as being prominent in Topic 1. Looking at the other nine topics, these words are less likely to arise in trials not involving this topic. Topic 1 can be relabeled as *Mail Fraud*.

²¹Edison and Carcel (2021) also analyzes FOMC texts using LDA.

²²See Grajzl and Murrell (2021) for an interesting text analysis of British common law and Nowak and Smith (2017) for a text analysis of home listings. These papers, while related, use a different but similar text analysis method.

²³We do not report common words such as “said” and “man” which arise with a large probability in most topics.

Figure 6: *Topic 1 (Mail Fraud)* over Time



The probability weight put on this topic (from the LDA process) for each observation is collapsed to the annual level. The mean is depicted with the 95% confidence interval. Only those observations where fraud is the crime are considered for the 80-year window; 1796 to 1876.

Table 8: Differentiating Words in *Fraud: Mail*

	<i>Fraud: Mail</i>	<i>Other 9 Topics</i>	
word	word rank	mean	median
letter	2	175.4	110
street	3	212.4	102
name	4	29.3	17
received	5	22.1	17
letters	8	308.7	338
office	9	120.2	81
sent	13	38.7	35
signed	14	187.1	105
road	19	326.3	358
wrote	26	180.9	163
writing	27	101.6	97
address	37	334.7	501
place	38	98.1	76
paper	40	97.4	83
average rank	17.5	110.5	149.5

The first column provides the ranking of the word in the list of words associated with Topic 1. The second and third columns consider the word rankings in the other nine topics reporting the mean and median, respectively. The final row provides the average rank for the selected words in Topic 1 and the average of the averages and medians for the other nine topics. When a word does not arise in the top 500 words for a particular topic, then it receives a ranking of 501.

Mail-related fraud is a potential source for fraud. Investigating British history, the first postage stamp was introduced in Great Britain in 1840. Previously, mail was paid for by the recipient and the fee was a function of the distance traveled, along with its weight. Not too infrequently, then, the recipient rejected the delivery. To simplify the mail service and expand its use in the population, the pre-paid stamp was introduced. It had a flat fee (regardless of distance traveled) below previous postal rates. This presumably reduced rejection rates allowing for economies of scale to bring down the mail’s average cost. As a consequence, mailing became common and we would expect that mail-related fraud became more prevalent. LDA is able to detect the rise of mail fraud in the 1850s, 60s, and 70s.²⁴ While unrelated to the introduction of the right to counsel, it does demonstrate validity to the analysis method.

Setting aside mail fraud, it turns out that one topic within each set of ten topics consistently experiences a noticeable increase in prevalence, and this change coincides approximately with the introduction to the right to counsel in 1836. We refer to these topics as *Timing*. Table 9 provides the selected words that dominate this topic for each of the four crimes.

For each crime words related to the timing of events are prominent. Thus, we call these topics *Timing*. For each crime, the LDA estimation identifies timing-related words as a distinct topic. Figure 7 depicts the time series of each.

For each of the four crime categories, the selected *Timing* topic visually exhibits a change in the time trend. That is, the prevalence of that particular topic seems rather flat but then begins a gradual escalation. This escalation begins near the 1836 date – the date of the Act’s implementation. There are not any noticeable changes in the prevalence of any other topic near 1836. The appendix provides their time series. Once again, this suggests that the increased prevalence of defense attorneys coincides with testimony that highlights the exact timing of events which we interpret as an effect of the professionalization of the courtroom.

On the other hand, if timing-related words become more frequent in the treated crimes after the Act is implemented, then it should also be the case that in the control crimes a topic with timing-related words should *not* change with the Act’s implementation. To study this, we conduct the LDA estimation on the subset of trials that fall within the control category. Doing this, we once again find one topic that is related to the timing of events.²⁵ Figure 8 presents the time series of this topic in the trial transcripts of the control crimes.

Clearly, the prevalence is quite flat. A difference-in-means *t*-test comparing those control crimes prior to 1836 to those after generates a statistically insignificant difference. Therefore, while timing-related testimony becomes more prevalent in the treated, felony cases it shows no change in the control crimes.

Rather than rely on visual evaluation, we ask whether we can find statistically significant evidence that

²⁴In fact, the word “post” is ranked 74 and “post-office” is ranked 185 in *Mail Fraud*. The latter is not even in the top 500 for any of the other nine topics.

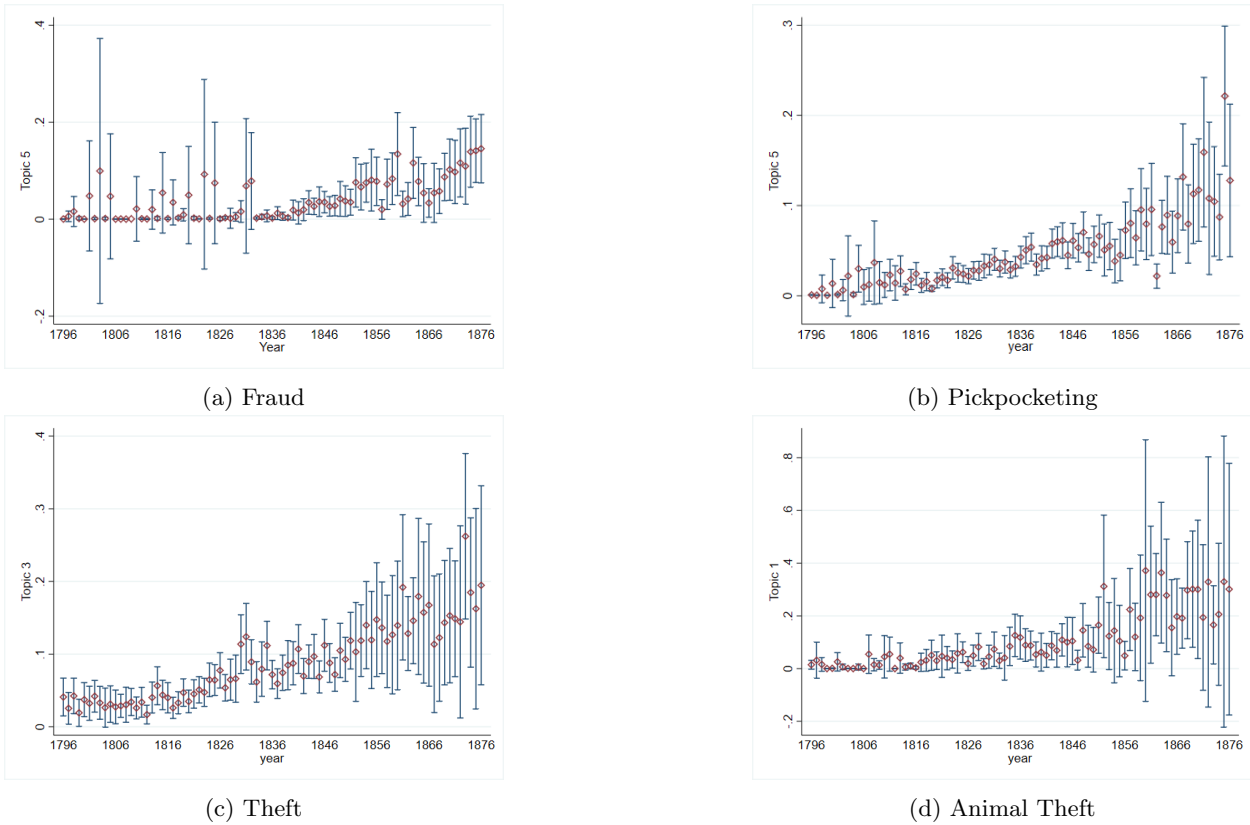
²⁵We provide the selected list of differentiating words and their prevalence in the other nine topics in the supplemental appendix.

Table 9: *Timing* Across Crimes

<i>Fraud</i>			<i>Pickpocketing</i>		
word	rank	avg rank (other 9)	word	rank	avg rank (other 9)
time	2	11.9	o'clock	5	17.2
day	17	23.4	night	15	102.0
never	18	15.3	time	16	81.2
last	41	39.78	morning	17	133.3
afterwards	50	44.6	never	23	69.4
date	51	342.8	day	30	115.0
months	52	56.1	minutes	31	61.2
since	53	101.9	months	33	121.4
now	54	98.3	afterwards	34	117.3
year	56	244.7	years	41	63.1
present	58	139.4	evening	47	75.7
remember	75	192.1	half-past	48	146.7
days	104	100.9	since	57	315.0
years	118	90.9	hour	87	166.1
times	133	127.4			
always	134	173.3			
avg rank	63.5	112.7	avg rank	34.6	113.2
<i>Theft</i>			<i>Animal Theft</i>		
word	rank	avg rank (other 9)	word	rank	avg rank (other 9)
o'clock	6	85.3	o'clock	8	9.8
time	8	57.6	day	11	30.7
day	14	72.8	morning	13	11.9
never	15	60.7	time	22	41.7
morning	16	29.2	afterwards	30	47.2
minutes	20	169.1	minutes	31	143.8
evening	31	82.4	night	40	27.0
years	36	47.3	years	41	50.7
afterwards	37	91.0	evening	45	58.2
night	39	78.4	months	49	103.0
hour	51	156.0	half-past	51	160.4
half-past	58	225.3	hour	64	145.4
months	60	86.6			
clock	70	432.8			
afternoon	88	216.0			
avg rank	36.6	126.0	avg rank	33.8	69.1

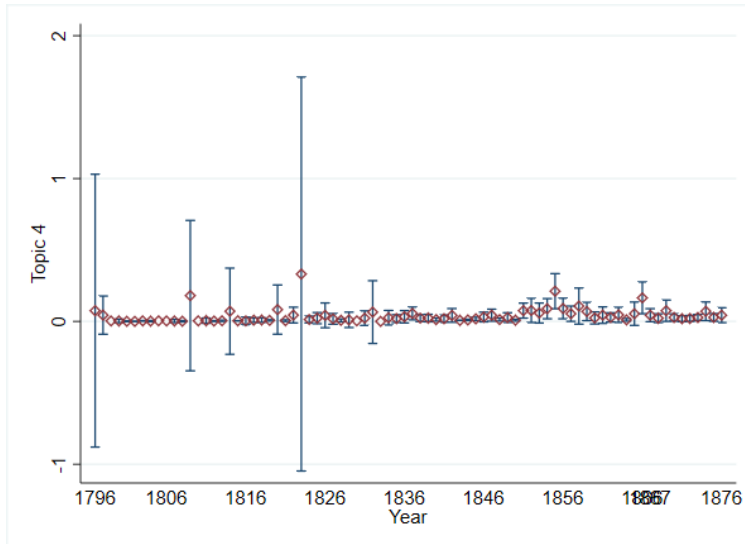
In each panel, the first column provides the ranking of the word in the list of words associated with *Timing*. The second column provides the average of the word ranking in the other nine topics. The final row provides the average rank for the selected words in both *Timing* and the average of the averages for the other nine topics. When a word is not in a topic's top 500 words, it is assigned a ranking of 501.

Figure 7: *Timing over Time*



The probability weights put on these topics (from the LDA process) for each observation are collapsed to the annual level. The means are depicted with the 95% confidence intervals. Only those observations where fraud is the crime are considered for the 80-year window; 1796 to 1876.

Figure 8: *Timing in Misdemeanor Cases*



The probability weights put on this topic (from the LDA process) for each observation are collapsed to the annual level. The mean is depicted with the 95% confidence interval. Only those observations for the 80-year window (1796 to 1876) is considered.

Table 10: Break in Time Trends for *Timing*

crime:	Fraud [1]	Pickpocketing [2]	Theft [3]	Animal Theft [4]	Control Crimes [5]
constant ($\hat{\zeta}_0$)	0.00777 (1.30434)	-1.67518 *** (0.31958)	-3.09613 *** (0.39133)	-3.37149 *** (0.84928)	-0.04080 (1.55630)
Post ($\hat{\zeta}_1$)	-5.28525 *** (1.41390)	-1.58490 *** (0.43390)	-2.26303 *** (0.67407)	-7.47237 *** (1.40535)	-1.25927 (1.65438)
Year ($\hat{\zeta}_2$)	0.00001 (0.00072)	0.00093 *** (0.00081)	0.00173 *** (0.67407)	0.00188 *** (1.40535)	0.00004 (0.00085)
Post x Year ($\hat{\zeta}_3$)	0.00287 *** (0.00077)	0.01107 *** (0.00024)	0.00086 *** (0.00037)	0.00406 *** (0.00076)	0.00068 (0.00091)
R^2	0.0477	0.0383	0.0405	0.0968	0.0035
F	38.8 ***	144.1 ***	118.1 ***	76.4 ***	2.1 *
N	2328	10,864	8396	2141	1855

Dependent variable is the probability weight put on that particular topic. Standard errors presented in parentheses; *** 1%, ** 5%, * 10% level of significance.

the time trend changes course after 1836. To that end, for each topic we estimate the following model:

$$Topic_{iy} = \zeta_0 + \zeta_1 Post_{iy} + \zeta_2 Year_{iy} + \zeta_3 Post_{iy} \times Year_{iy} + \epsilon_{iy}. \quad (8)$$

The dependent variable is the probability weight put on the topic for observation i , which arose in year y . The constant ζ_0 is the intercept, while ζ_2 captures the time trend for the topic's prevalence prior to the Act's implementation. The variable $Year_{iy}$ is a linear time trend where every observation in year y takes the value equal to that year. The indicator variable $Post_{iy}$ is equal to one if $y > 1836$. Hence, ζ_1 allows for a jump at the implementation of the right to counsel and ζ_3 measures how (and whether) the time trend adjusts when defense counsel is fully introduced into criminal trials at the Old Bailey.

If the Act causes a break at 1836 shifting up the presence of the topic (as a level-effect), then $\hat{\zeta}_1$ will be positive and large but $\hat{\zeta}_3$ will be expected to be equal to zero. If, on the other hand, the effect of the Act grows gradually over time then $\hat{\zeta}_3$ will be positive. Table 10 presents the results for each of the four crime category subsamples.

Each topic exhibits a slow increase in its prevalence prior to the Act's implementation, followed by a sharp upward rotation in the slope. The results indicate that there was a gradual effect of the Act on the criminal trials at the Old Bailey, as measured by the text of the trial transcripts.²⁶ The final column considers the control crimes and illustrates no change in the time trend around the time of the Act's implementation.

While not presented here, we replicate our text analysis using twenty topics ($T = 20$). Once again, one topic selects timing-related words in each crime subsample, and again experiences an increase in prevalence

²⁶The large, negative coefficients on *Post* are not a concern as they are an artifact of the units of measurement for the time trend. The value of *Year* comes from the interval [1796, 1876]. Thus, as an example, consider Animal Theft, which has a large and negative coefficient on *Post*. Moving from the year 1836 to 1839 brings the fitted value back up to its previous level.

after 1836. Avoiding redundancy, we do not present the extra results here. In addition, we replicate the structural break estimation for the other nine topics within each of the four crime types. Few other topics even have a positive rotation in the time period after 1836.

7 Conclusion

We provide the first causal identification of the impact of the introduction to the right to counsel on criminal trials. This is done by taking advantage of a unique situation in 1800s British courts. Recognizing that minor offenses had a history of allowing defense counsel but felonies did not, we employ a difference-in-difference estimation strategy to identify the impact of the introduction of the right to counsel for felony crimes. We show that conviction rates increased. We interpret this result as suggesting that when jurors feel that the criminal proceedings are unfairly stacked against the defendant, they are unwilling to convict. By professionalizing the courtroom, jurors no longer have to correct for the imbalance and convictions increase.²⁷

We employ a popular topic modeling approach to text analysis to provide some suggestive information on how the trial's change. Investigating four crimes that are recognized to have substantial increases in the prevalence of defense counsel after the Act's implementation, we show that words related to the timing of events becomes a more important part of the trial's testimony.

One potential implication to modern criminal justice institutions could be the issue of the quality of publicly-provided defense counsel. While indigent defense is commonly provided to those who lack financial resources, strong evidence points to worse outcomes arising by those who cannot pay for their own attorney. Further, full-time, salaried public defenders outperform attorneys who are appointed at a case-by-case basis. Our results may suggest that improving the quality of the defense counsel may lead to increased convictions. The normative implication of this is unclear as one cannot predict whether these increased convictions came with heightened or mitigated rates of type I errors and, relatedly, whether it enhances the deterrent effect.

We use one popular topic modeling approach known as Latent Dirichlet Allocation to analyze the text. It is by no means the only computational linguistic algorithm that can be used. Future researchers may be able to learn more about the impact of the introduction to the right to counsel by applying other methods. Similarly, LDA at most provides suggestive evidence as the meaning of the topics are open to interpretation. Nevertheless, we feel that applying these techniques complements the more-formal econometric analysis by providing information on the underlying mechanism.

Our investigation is best thought of as an intent-to-treat analysis as we explore how the introduction to the right to counsel affects the conviction rate, regardless of whether the defendant hired an attorney or not. Along with only having a noisy measurement of the use of a defense counsel, there is no reason to believe that the introduction of the right to counsel only affects those cases where one is used.

A challenge to all empirical investigations is external validity. We do not know, for example, if modern

²⁷This result was anticipated by the historian David Bentley. In his discussion of the Prisoners' Counsel Act Bentley (2003) predicted that "[r]ather than leading to wrongful acquittals, the Bill would be likely to result in more convictions, for juries would no longer acquit out of sympathy excited by the prisoner's inability to defend himself" (Bentley 2003; p. 108).

jurors respond similarly to those serving in the 18th and 19th century in London. Also, the investigation explores procedures in a common law country. It is unclear how the right to counsel affects outcomes in civil law systems. Our work, though, makes a unique contribution to the small, growing literature investigating how the incentives of legal actors affect criminal justice outcomes.

8 References

Abadie, Alberto, Athey, Susan, Imbens, Guido W., and Woolridge, Jeffrey (2017), When Should You Adjust Standard Errors for Clustering?, NBER Working Paper No. 24003.

Abrams, David and Yoon, Albert (2007), The Luck of the Draw: Using Random Case Assignment to Investigate Attorney Ability, *University of Chicago Law Review* 74: 1145-1177.

Aidt, Toke S. and Franck, Raphael (2015), Democratization Under the Threat of Revolution: Evidence from the Great Reform Act of 1832, *Econometrica* 83(2): 505-547.

Agan, Amanda, Matthew Freedman, and Emily Owens (2021), Is Your Lawyer a Lemon? Incentives and Selection in the Public Provision of Criminal Defense, *Review of Economics and Statistics* 103(2): 294-309.

Bastani, Kaveh, Namavari, Hamed, and Shaffer, Jeffrey (2019), Latent Dirichlet Allocation (LDA) for Topic Modeling of the CFPB Consumer Complaints, *Expert Systems with Applications* 127: 256-271.

Beattie, John M. (1991), Scales of Justice: Defense Counsel and the English Criminal Trial in the Eighteenth and Nineteenth Centuries, *Law and History Review* 9(2): 221-267.

Bentley, David (2003), *English Criminal Justice in the 19th Century*, London: Hambledon Press.

Bindler, Anna and Hjalmarsson, Randi (2017), Prisons, Recidivism, and the Age-Crime Profile, *Economics Letters* 152: 46-49.

Bindler, Anna and Hjalmarsson, Randi (2018a), How Punishment Severity Affects Jury Verdicts: Evidence from Two Natural Experiments, *American Economic Journal: Economic Policy* 10(4): 36-78.

Bindler, Anna and Hjalmarsson, Randi (2018b), Replication Data For: How Punishment Severity Affects Jury Verdicts: Evidence from Two Natural Experiments, Nashville, TN: American Economic Association [publisher]. Ann Arbor, MI: Inter-university Consortium for Political and Social Research.

Bindler, Anna and Hjalmarsson, Randi (2019), Path Dependency in Jury Decision Making, *Journal of the European Economic Association* 17(6): 1971-2017.

Bindler, Anna and Hjalmarsson, Randi (2020), The Persistence of Criminal Justice Gender Gap: Evidence from 200 Years of Judicial Decisions, *Journal of Law and Economics* 63(2): 297-339.

- Bindler, Anna and Hjalmarsson, Randi (2021), The Impact of the First Professional Police Forces on Crime, *Journal of the European Economic Association*, forthcoming.
- Bjerk, David (2007), Guilt Shall Not Escape Nor Innocent Suffer? The Limits of Plea Bargaining when Defendant Guilt is Uncertain, *American Law and Economics Review* 9(2): 305-329.
- Blei, David, Ng, Andrew Y., and Jordan, Michael I. (2003), Latent Dirichlet Allocation, *Journal of Machine Learning Research* 3(1): 993-1022.
- Cockburn, J. S. and Green, Thomas A. (1988), *Twelve Good Men and True: The Criminal Trial Jury in England, 1200-1800*, Princeton: Princeton University Press.
- Davies, Stephen (2002), The Private Provision of Police During the Eighteenth and Nineteenth Centuries, in David T. Beito, Peter Gordon, and Alexander Tabarrok (eds.), *The Voluntary City: Choice, Community, and Civil Society*, Ann Arbor: University of Michigan Press, 191-224.
- Dickenson, William and Talfourd, Thomas Noon (1829), *Practical Guide to the Quarter Sessions and Other Sessions of the Peace*, London: S. Sweet.
- Duggan, John and Martinelli, Cesar (2001), A Bayesian Model of Voting in Juries, *Games and Economic Behavior* 37(2): 259-294.
- Dyer, Travis, Lang, Mark, and Stice-Lawrence, Lorein (2017), The Evolution of 10-K Textual Disclosures: Evidence from Latent Dirichlet Allocation, *Journal of Accounting and Economics* 64(2-3): 2209-22245.
- Edison, Hali and Carcel, Hector (2021), Text Data Analysis Using Latent Dirichlet Allocation: Application to FOMC Transcripts, *Applied Economics Letters* 28(1): 38-42.
- Emsley, Clive (1991), *The English Police: A Political and Social History*, London: Harvester Wheatsheaf.
- England and Wales Law Commission. (2009, April). *Newsletter Spring 2009*. https://webarchive.nationalarchives.gov.uk/20090416025930/http%3A//www.lawcom.gov.uk/docs/newsletter_spring_2009.pdf.
- Feddersen, Timothy and Pesendorfer, Wolfgang (1998), Convicting the Innocent: The Inferiority of Unanimous Jury Verdicts under Strategic Voting, *American Political Science Review* 92(1): 23-35.
- Flanagan, Francis X. (2018), Race, Gender, and Juries: Evidence from North Carolina, *Journal of Law and Economics* 61(2): 189-214.
- Friedman, Ezra and Wickelgren, Abraham L. (2006), Bayesian Juries and the Limits to Deterrence, *Journal of Law, Economics, and Organization* 22(1): 70-86.

- Gallanis, T.P. (2006), The Mystery of Old Bailey Counsel, *Cambridge Law Journal* 65(1): 159-173.
- Grajzl, Peter and Murrell, Peter (2021), A Machine-Learning History of English Caselaw and Legal Ideas Prior to the Industrial Revolution I: Generating and Interpreting the Estimates, *Journal of Institutional Economics* 17(1): 1-19.
- Griffiths, Cerian Charlotte (2014), The Prisoners' Counsel Act 1836: Doctrine, Advocacy and the Criminal Trial, *Law, Crime and History* 2: 28-47.
- Grimmer, Justin (2010), A Bayesian Hierarchical Topic Model for Political Texts: Measuring Expressed Agendas in Senate Press Releases, *Political Analysis* 18(1): 1-35.
- Guo, Yue, Barnes, Stuart J., and Jia, Qiong (2017), Mining Meaning from Online Rating and Reviews: Tourist Satisfaction Analysis Using Latent Dirichlet Allocation, *Tourism Management* 59: 467-483.
- Halliday, Stephen (2007), *Newgate: London's Prototype of Hell*, The History Press.
- Hanly, Conor (2021), Jury Selection in Victorian England, *American Journal of Legal History* 61(1): 36-55.
- Hansen, Stephen, McMahon, Michael, and Prat, Andrea (2019), Transparency and Deliberation with in the FOMC: A Computational Linguistics Approach, *Quarterly Journal of Economics* 133(2): 801-870.
- Hoekstra, Mark and Street, Brittany (2021), The Effect of Own-Gender Jurors on Conviction Rates, *Journal of Law and Economics*, forthcoming.
- Jones, Mark and Johnstone, Peter (2012), *History of Criminal Justice*, New York: Routledge.
- Koyama, Mark (2012), Prosecution Associations in Industrial Revolution England: Private Providers of Public Goods?, *Journal of Law and Economics* 41: 95-130.
- Langbein, John H. (1978), The Criminal Trial before the Lawyers, *University of Chicago Law Review* 45(2): 263-316.
- Langbein, John H. (1983), Shaping the Eighteenth-Century Criminal Trial: A View from the Ryder Sources, *University of Chicago Law Review* 50(1): 1-136.
- Langbein, John H. (1999), The Prosecutorial Origins of Defence Counsel in the Eighteenth Century: The Appearance of Solicitors, *Cambridge Law Journal* 58(2): 314-365.
- Larsen, Vegard H. and Thorsrud, Leif A. (2019), The Value of News for Economic Development, *Journal of Econometrics* 210(1): 203-218.

- McCannon, Bryan C. (2020), Wine Descriptions Provide Information, *Journal of Wine Economics* 15(1): 203-218.
- Nowak, Adam and Smith, Patrick (2017), Textual Analysis in Real Estate, *Journal of Applied Econometrics* 32(4): 896-918.
- Quinn, Kevin M., Monroe, Burt L., Colaresi, Michael, Crespin, Michael H., and Radev, Dragomir R. (2010), How to Analyze Political Attention with Minimal Assumptions and Costs, *American Journal of Political Science* 54(1): 209-228.
- Roach, Michael A. (2014), Indigent Defense Counsel, Attorney Quality, and Defense Outcomes, *American Law and Economics Review* 16(2): 577-619.
- Roach, Michael A. (2017), Does Raising Indigent Defender Pay Rates Improve Defendant Outcomes?, *Applied Economics Letters* 23(14): 1025-1030.
- Rubin, Donald B. (1978), Bayesian Inference for Causal Effects: The Role of Randomization, *Annals of Statistics* 6(1): 34-58.
- Shoemaker, Robert B. (2008), The Old Bailey Proceedings and the Representation of Crime and Criminal Justice in Eighteenth-Century London, *Journal of British Studies* 47(3): 559-580.
- Shem-Tov, Yotam (2021), Make-or-Buy? The Provision of Indigent Defense Services in the U.S., *Review of Economics and Statistics*, forthcoming.
- Smith, Oliver (2019), ‘Strike, Man, Strike!’ – On the Trail of London’s Most Notorious Public Execution Sites, *The Telegraph*, January 17, <https://www.telegraph.co.uk/travel/destinations/europe/united-kingdom/england/london/articles/Londons-most-notorious-execution-sites/>
- Starr, Sonja (2015), Estimating Gender Disparities in Federal Criminal Cases, *American Law and Economics Review* 17(1): 127-159.
- Tirunillai, Seshadri and Tellis, Gerard J. (2014), Mining Marketing Meaning from Online Chatter: Strategic Brand Analysis of Big Data Using Latent Dirichlet Allocation, *Journal of Marketing Research* 51(4): 463-479.
- Voth, Hans-Joachim (1997), Time and Work in Eighteenth-Century London, *Journal of Economic History* 58(1): 29-58.
- Wontner, Thomas (1833), *Old Bailey Experience*, reprint (2018), Forgotten Books.

Xue, Jia, Chen, Junxiang, Chen, Chen, Zheng, Chengda, Li, Sijia, and Zhu, Tingshao (2020), Public Discourse and Sentiment During the COVID-19 Pandemic: Using Latent Dirichlet Allocation for Topic Modeling on Twitter, *PLOS One* 15(9): e02394441.

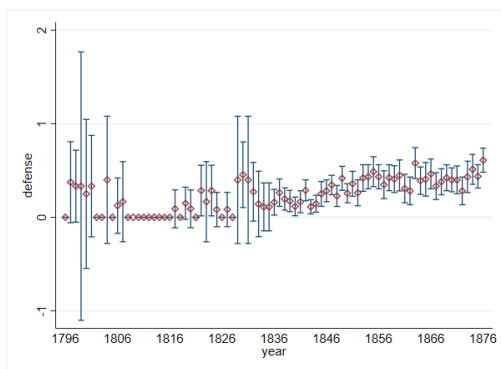
Wade, Stephen (2009), *Britain's Most Notorious Hangmen*, South Yorkshire: Wharncliffe.

Zacharis, Clemence (2008), The Code D'Instruction Criminelle 1808, *History of the Two Empires*, Foundational Napoleon 2021, Napoleonica-Research Foundation Digital Collection, napoleon.org/en/history-of-the-two-empires/articles/the-code-dinstruction-criminelle-1808.

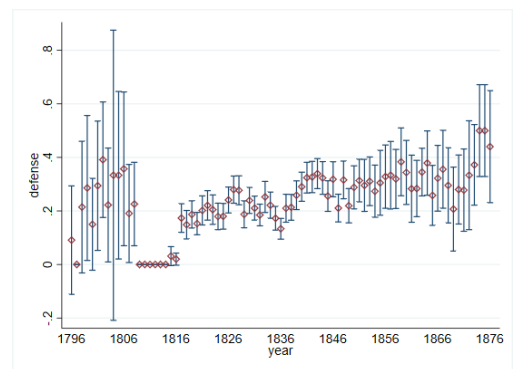
9 Appendix

The time series for the presence of defense counsel in the four selected crimes are presented in Figure 9. The time series for the other topics not presented in the main body, for each of the four crimes, are provided in Figures 10, 11, 12, 13.

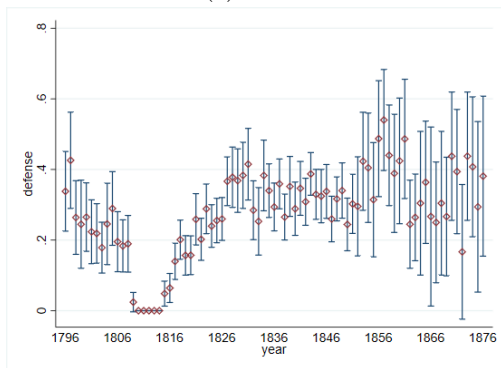
Figure 9: Defense Counsel Presence



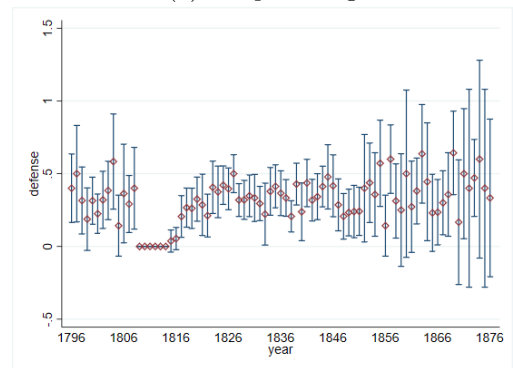
(a) Fraud



(b) Pickpocketing

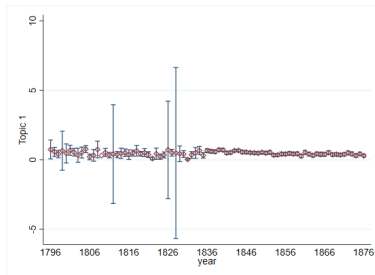


(c) Theft

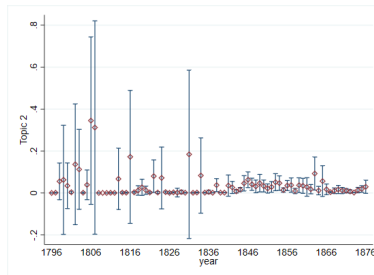


(d) Animal Theft

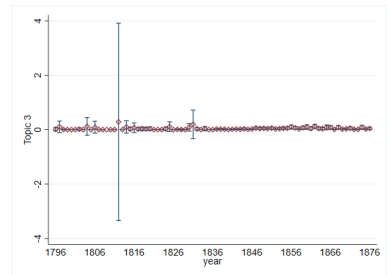
Figure 10: Fraud Topics Over Time



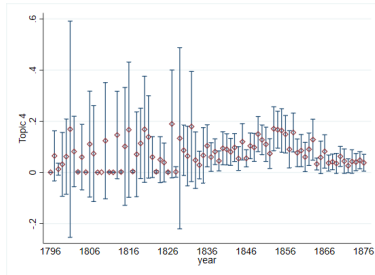
(a) Topic 1



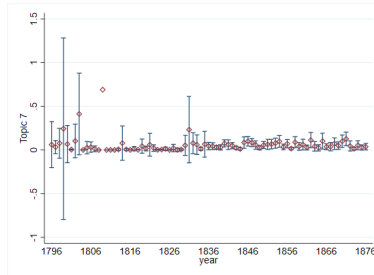
(b) Topic 2



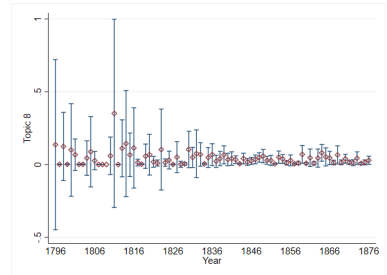
(c) Topic 3



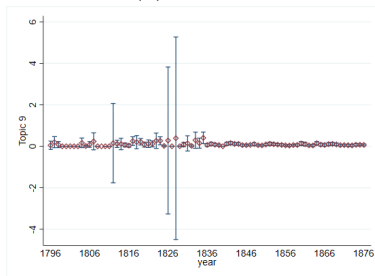
(d) Topic 4



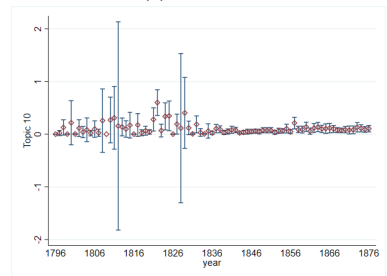
(e) Topic 7



(f) Topic 8



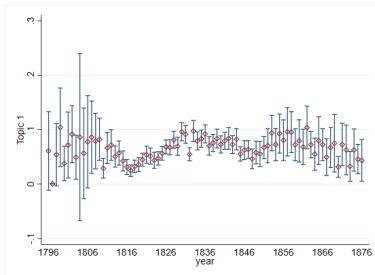
(g) Topic 9



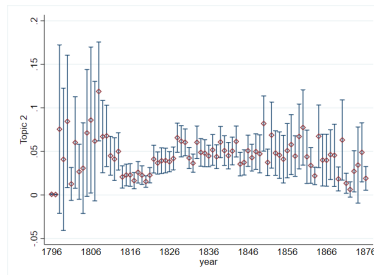
(h) Topic 10

The topic numbered “Topic 5” is *Timing* and “Topic 6” is *Mail Fraud*. They are presented in the text.

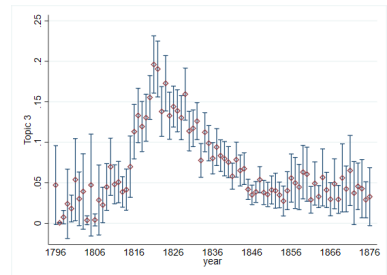
Figure 11: Pickpocketing Topics Over Time



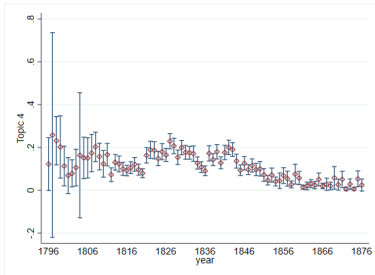
(a) Topic 1



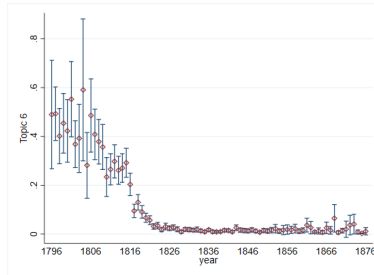
(b) Topic 2



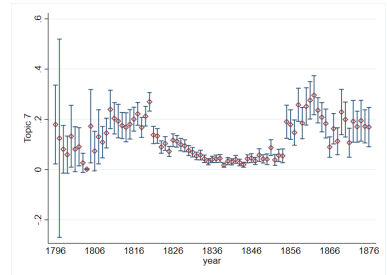
(c) Topic 3



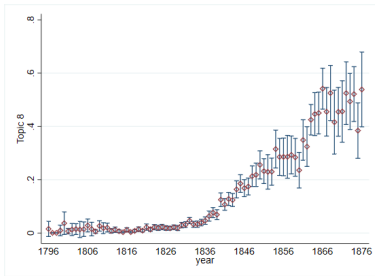
(d) Topic 4



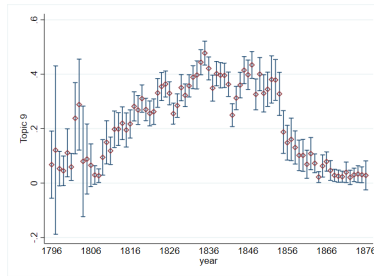
(e) Topic 6



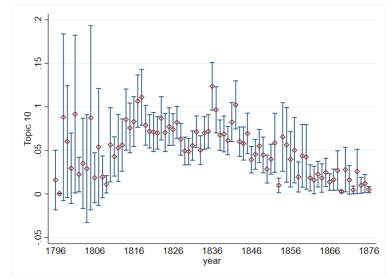
(f) Topic 7



(g) Topic 8



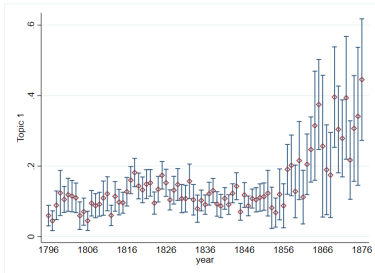
(h) Topic 9



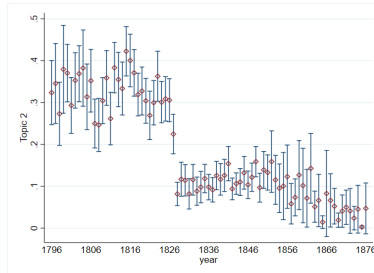
(i) Topic 10

The topic numbered “Topic 5”, *Timing*, is presented in the text.

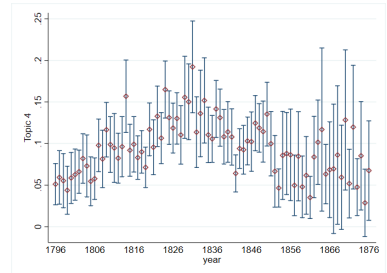
Figure 12: Theft Topics Over Time



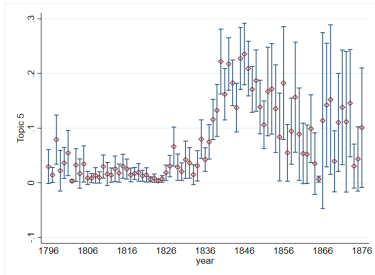
(a) Topic 1



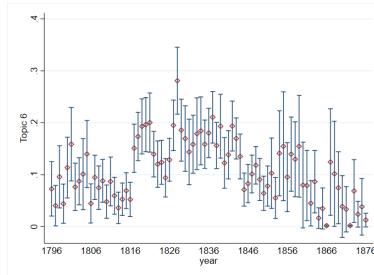
(b) Topic 2



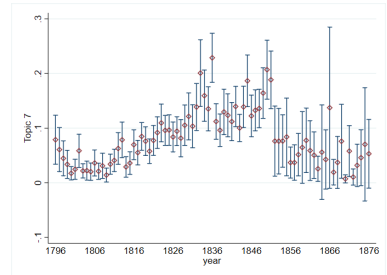
(c) Topic 4



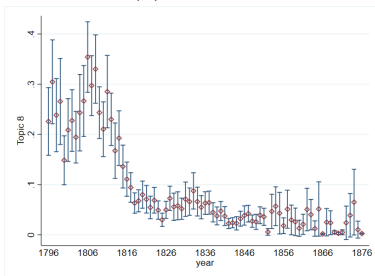
(d) Topic 5



(e) Topic 6



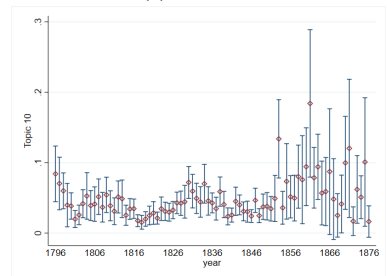
(f) Topic 7



(g) Topic 8



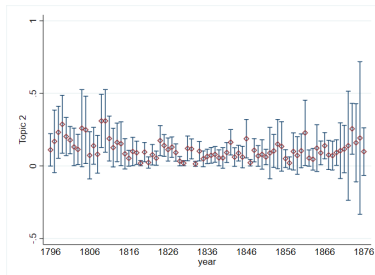
(h) Topic 9



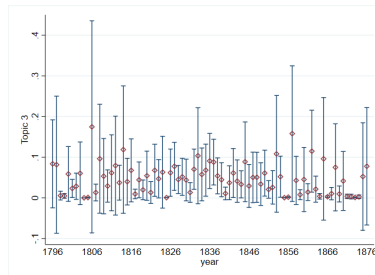
(i) Topic 10

The topic numbered “Topic 3”, *Timing*, is presented in the text.

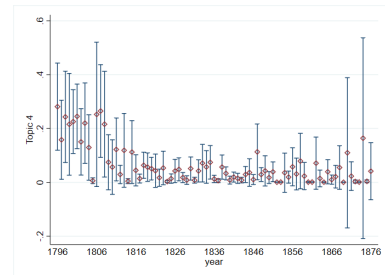
Figure 13: Animal Theft Topics Over Time



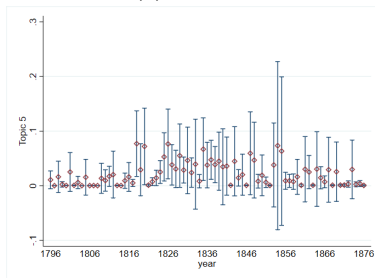
(a) Topic 2



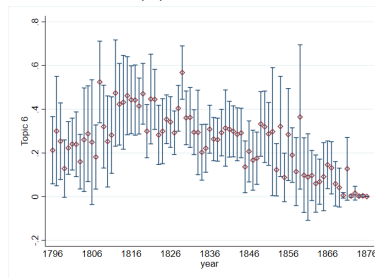
(b) Topic 3



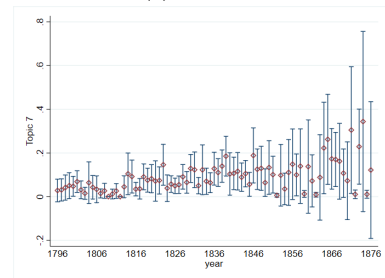
(c) Topic 4



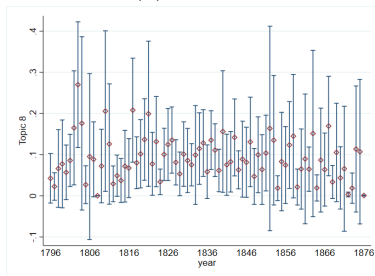
(d) Topic 5



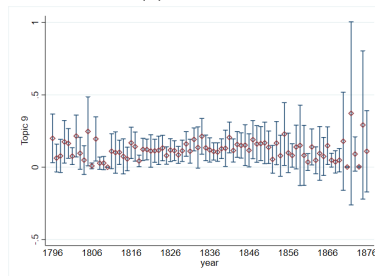
(e) Topic 6



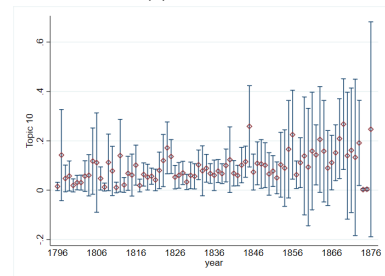
(f) Topic 7



(g) Topic 8



(h) Topic 9



(i) Topic 10

The topic numbered “Topic 1”, *Timing*, is presented in the text.