2023

Essays on the Economics of Law and Crime

Zachary J. Porreca
West Virginia University, zjp00003@mix.wvu.edu

West Virginia University
Morgantown, West Virginia, Spring 2023

Essays on the Economics of Law and Crime

Zachary Porreca

Dissertation submitted to the

John Chambers College of Business and Economics

at West Virginia University

in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

in

Economics

Bryan McCannon, Ph.D., Committee Chairperson

Brad Humphreys, Ph.D.

Kole Reddig, Ph.D.

Jakub Lonsky, Ph.D.

Department of Economics

Morgantown, West Virginia. February 2023

Keywords: Crime, Urban, Law

Abstract

Essays on the Economics of Law and Crime

Zachary Porreca

The first chapter examines the connection between gentrification and urban violence. I demonstrate a positive and plausibly causal relationship between urban redevelopment and gun violence in Philadelphia. As the underlying mechanism, I focus on gentrification's displacement effect on local drug markets. Treating the city as a spatial network of city blocks and using two-way fixed effects differences-in-differences estimators, I show the gentrification of one block increases violence across the surrounding neighborhood. I find that some 2,400 (8%) of Philadelphia's shootings between the years 2011 and 2020 can be attributed to spillover effects from the gentrification of drug blocks. This effect is nearly ten times stronger than that observed on blocks without high levels of drug crime. This study also contributes a new empirical measurement of gentrification drawn primarily from property sales, along with building, zoning, and alteration permit issuance and utilizes a novel nearest-neighbor network approach to identify spatial spillover effects.

The second chapter formalizes the synthetic difference-in-differences estimator for staggered treatment adoption settings, as briefly described in Arkhangelsky et al. (2021). To illustrate the importance of this estimator, I use replication data from Abrams (2012). I compare the estimators obtained using SynthDiD, TWFE, the group time average treatment effect estimator of Callaway and Sant'Anna (2021), and the partially pooled synthetic control method estimator of Ben-Michael et al. (2021) in a staggered treatment adoption setting. I find that in this staggered treatment setting, SynthDiD provides a numerically different estimate of the average treatment effect. Simulation results show that these differences may be attributable to the underlying data generating process more closely mirroring that of the latent factor model assumed for SynthDiD than that of additive fixed effects assumed under traditional difference-in-differences frameworks.

The third chapter is joint work with Dr. Bryan McCannon. In it, we exploit a novel data set of criminal trials in 19th century London to 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 At of 1836 remedied this imbalance 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 perceiving the trial as being 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.

Acknowledgements

First, I would like to extend my gratitude towards my dissertation chair, Dr. Bryan McCannon. For taking the chance and getting me into the economics department in the first place, bearing with my less than fully thought out ideas, and pointing me in the right directions in research and academia broadly I am forever grateful. I am extremely thankful for the guidance of my other committee members, Dr. Brad Humphreys, Dr. Kole Reddig, and Dr. Jakub Lonsky. I owe almost everything to my undergraduate dean, Dr. James Brown for his work pushing the Second Chance Pell Grant program at Bloomsburg University and by doing so (and with his efforts after) giving me the opportunity to succeed in obtaining an education. Thank you to the Center for Free Enterprise, Institute for Humane Studies, and Mercatus Center for the generous funding that afforded me the opportunities to push my research forward and for the wonderful discussions that continue to inspire me. Thank you to all of the graduate students and faculty (past and present) that I have interacted with at WVU and learned from in our conversations; in particular: Alex Cardazzi, Josh Martin, Alex Marsella and Dr. Adam Nowak. Thank you to Hillary Probst, for your unending support and love, for being understanding when I'm lost in the computer for days, for believing in me when I decided to leave construction behind, and for trying to make sense of my nonsense. From Linecom to the future, I couldn't have done it without you. A special thank you to my best friend, Cannon, for always sitting under my desk or nearby watching me work, licking my hand, and for taking me on long walks when it's time to slow down and think for a bit. I need to thank my friends who have always supported me, especially Abdourhamane Ali, Kevin Corcoran, Paco Ramos, Dylan Kinsey, Rob DiLuzio, Billy Traynor, and Frank Grazulis for pushing me to pursue this path; and to my friends that didn't make it to see me make something good happen. Tommy, Mikey, Josh, Spencer, Danny and Will, I'll miss you guys forever. We always were some scrappy damned bastards. Thank you to my aunts, uncles, cousins, and grandparents that have always been in my corner. Lastly, I need to extend a special thank you to my mom, dad, and brother, Ray, for your support, love, and belief in me even when it was not easy. Your love and support has meant the world to me and none of this would be possible without you.

Contents

List of Figures

List of Tables

1 Gentrification, Gun Violence, and Drug Markets

The published version of this chapter is available in the Journal of Economic Behavior and Organization, Volume 207, pages 235-256.

1.1 Introduction

Gentrification and rapid urban development change the landscape of a city, bringing new amenities and new residents into traditionally neglected neighborhoods. However, this development also has the potential to bring with it spatial spillovers and negative externalities, as crime may be pushed away from newly developed blocks outward into the surrounding neighborhood. In this paper, I seek to address this question empirically: does gentrification create spatial spillovers that cause increases in crime in the surrounding, yet to be redeveloped, neighborhood?

Across urban America, competition among criminal actors and organizations visibly manifests itself through incidents of gun violence. Despite the existence of incentives for criminals to coordinate with one another, organized outcomes often to fail to emerge. I argue that this failure to obtain cooperative outcomes is a product of the instability of officially unenforceable property rights in illegal criminal markets. Gentrification contributes to this instability, as the redevelopment of city blocks that once supported such illicit markets pushes criminals towards more intensive competition for the remaining blocks that are capable of sustaining their illegal activity.

In this study I demonstrate the linkage between gentrification and gun violence. Using city-block level data from Philadelphia, I specify a two-way fixed effects differences-in-differences estimator to empirically validate the relationship between the gentrification of one block and levels of gun violence across the surrounding neighborhood. I find a robust result showing that on average gentrification increases levels of gun violence in neighbor blocks. This effect is more pronounced when the gentrified block has a history of drug crime, with an average increase of nearly nine shootings in the surrounding neighborhood. Back of the envelope calculations suggest that roughly 5,800 shootings, or 21% of the city's shootings across the ten year study window can be attributed to spillover effects from gentrification, while 2,400 (or 8%) can be attributed to the gentrification of drug blocks. I find that gentrification has a statistically and economically significant effect on rates of gun violence in the neighborhood surrounding a newly gentrified block. I then document a potential mechanism for this effect, showing that following a block's gentrification drug crime tends to increase on surrounding blocks and decrease on the newly gentrified block. Intuitively, this indicates that drug crime that existed on a block pre-gentrification is pushed into the surrounding neighborhood by the new development. This in turn increases the intensity of competition in the surrounding neighborhood and leads to the increases in gun violence documented herein.

While I am not the first to explore the relationship between gentrification and crime, this paper is indeed the first to study the impact of gentrification on crime displacement. O'Sullivan (2005) provides a theoretical model positing a causal relationship between exogenously falling crime rates and an influx of higher income residents. The author went further, to assert that gentrification is self-reinforcing in that the influx of higher income residents will further decrease crime rates, which further incentivizes an even greater displacement of low income residents in favor of those with a higher income. Similarly, Ellen et al. (2019) finds that reductions in crime rates invite increases in rates of wealthy and educated residents moving into the central city. While these works examine how crime effects the residential decisions that lead to gentrification, they do no consider how gentrification can effect existing neighborhood dynamics. As such, I examine the effect that gentrification will have on crime rates on the city's "frontiers"; blocks that are newly beginning the process of gentrification. My approach is unique in its focus on gentrification's impact on high drug crime blocks. Gentrification causes increasingly tense competition in concentrated drug markets. This competition in turn leads to increasing levels of neighborhood violence.

Numerous authors have empirically explored the relationship between gentrification and crime. Papachristos et al. (2011) regresses homicide and robbery rates on a distinct proxy for gentrification: the number of coffee shops in a neighborhood. The authors find that their markers of gentrification are associated with declining homicide rates, but increasing robbery rates. Smith (2012) utilizes a similar approach, regressing gang related homicides on neighborhood demographic factors, coffee shops, and an indicator for the demolition of public housing. The author finds decreases in homicides as markers of gentrification grow more prevalent. More recently, Autor et al. (2017) finds that rapid gentrification precipitated by the end of rent control in Cambridge, Massachusetts led to significant decreases in the overall crime rate. My result instead utilizes a novel nearest-neighbor network approach to focus on the spatial spillovers of crime wrought by gentrification. While I similarly find that gentrification reduces drug crime on the immediate gentrified block, there is a spatial spillover which causes adjacent non-gentrified areas to experience spikes in both drug crime and shootings.

Neighborhood, census tract, or other artificially drawn borders and boundaries do not necessarily capture the reality of a city. A city is a collection of streets, the lattice-work of their intersections creating the individual blocks that serve as the base level points from which change across a city can be observed. Violence, gentrification, and crime begin as extremely localized phenomena. As such, I examine block level variation in relevant markers and leverage this as a unique setting for identification: taking block level gentrification as an exogenous treatment so as to allow plausibly causal inferences to be made regarding the effect that gentrification has on gun violence.

This study provides a counter-intuitive result that runs against the major current of gentrification and crime related literature. Gentrification imposes a trade-off between the positive externalities of increased economic activity and the negative externality of increased gun violence in the surrounding community. These spikes in new real estate interest push violence outward and into the surrounding neighborhood. Much of this increase in violence may be able to be explained as the result of competition in the illegal drug markets that exist within pockets of the urban environment. The peculiarities of this sort of competition have been largely unexamined in the literature. As such, I offer a new insight into the spatial dynamics of gun violence with implications for urban planning, development, and policing.

Skaperdas (2001), Varese (2010), and Skarbek (2016) have all posited similarly that organized criminal governance is more likely to arise in settings in which official forms of governance fail to provide for the interests of all sub-populations. However, the presence of such an official failing does not guarantee that the organized outcome of criminal governance will arise. Skaperdas (2001) provides a reason for the failure to cooperate: that criminal competition more closely resembles an “arms race", where market share can be taken through violent means that decrease overall market welfare. In this study, I provide empirical backing to Skaperdas’s proposition. I demonstrate that in illegal drug markets, lacking official governance, a destabilizing shock leads to a ramping up of this “arm’s race" which in turn causes eruptions in violence as the market finds a new cooperative equilibrium.

What is often viewed as an improvement to an area is able to cause increases in violence. I theorize that the failure to cooperate in this competitive environment is a product of the ability of a criminal organization to gain market share through aggressive and militaristic tactics. The likelihood that all firms across a market should elect not to engage in such strategies decreases as the physical extent of the market to be competed for decreases. Exogenous shocks such as gentrification are responsible for this market concentration and give rise to the increases in gun violence observed.

Several other recent studies have taken a similar approach in examining spatial dimensions of competition among criminal organizations [1]. DeAngelo (2012) builds a theoretical model of spatial competition, and with this model predicts that increases in law enforcement can deter lower productivity criminals’ entrance into the market. Sobrino (2019) examines Mexican drug cartel violence as a competition for control over lucrative territory. She finds empirically that violence coincided with the entrance of additional cartels into a municipality. Castillo et al. (2020) also examines Mexican cartel violence, modeling municipal level competition as a contest for revenue. They find that scarcity caused by cocaine seizures led to significant increases in levels of violence. Bruhn (2021) examines gang presence across the city of Chicago. He finds that as a gang enters new territory, levels of crime (including violent crime) in that area increase. While these studies focus on the entrance and exit of competitors and the impact of supply shocks on the illegal market, I document the impact that changes in the availability of officially unenforceable territorial claims has on the dynamics of competition.

This paper is the first to document gentrification’s displacement effect on urban crime. This result, while running contrary to related literature that posits a negative relationship between gentrification and crime, suggests that the extension of urban development to low income areas may create negative externalities as drug and violent crime is displaced into increasingly concentrated adjacent yet to be redeveloped neighborhoods. I further contribute a new easily replicable empirical measurement of gentrification, drawn primarily from property sales along with building, zoning, and alteration permit issuance. This new measurement is able to capture gentrification at its narrowest and most realistic resolution: the

[1] Skaperdas (1992) and Polo (2005) provide important theoretical contributions regarding violence as a competitive strategy among criminal organizations and in settings absent formal property rights

individual block level.

1.2 Hypothesis

The markets for illegal goods are notably absent formal regulation. There is no protection of property rights and no legal enforcement of contracts or territorial claims. This unregulated market allows criminal organizations to compete for profit both through the traditional vehicles of price and quantity as well as through competitive strategies that are unparalleled in traditional market dynamics. The forcible seizure of a rival's assets or territory is a very real possibility.

Particular areas of a city are more capable than others of sustaining a drug market. This may be due to physical characteristics, ease of customer access, or even simply a lack of formal oversight (Eck 1994). Because of the inherent limits to the potential area of a drug market, any shock that reduces its total area will be a disruptive and destabilizing force that increases the value of the city blocks that make up the remainder of the viable set for the market. An organization that loses its territory will likely attempt to secure new territory elsewhere in the remaining blocks, rather than exit the market entirely. This can lead to increases in violence through either the direct engagement of a rival organization in an attempt to seize territory through military effort or by the increase in tension that can arise from the closer spatial proximity of rival organizations.

Gentrification, which entails the arrival of new residents of higher socio-economic standing to a traditionally lower income neighborhood, is this sort of destabilizing force. Ellen, Horn & O'Regan (2012) found that gentrified neighborhoods grow wealthier, more educated, exhibit higher rates of home ownership, and experience significant racial demographic changes. This can be inferred to represent a replacement of many of the neighborhood's original residents; thus making it more difficult for a criminal organization to find the level of community support necessary to operate openly. Further, empirical evidence has demonstrated that gentrification leads to increased policing and the adoption of more punitive policing practices (Laniyonu, 2017). It is reasonable to believe that increases in policing would deter criminal activity, as famously predicted by Becker (1968). Regardless of the actual mechanism at play, it is reasonable to assume that gentrification will make a block less suitable for drug competition. As such it is expected that the gentrification of a block within a drug market will push criminal organizations to the surrounding neighborhood, with the potential to cause spikes in violence.

Due to the geographic limitations on the portions of a city in which a drug market is able to be sustained, it is expected that a displaced criminal organization will be forced to find new territory within the immediate surrounding blocks as this area has already proven itself as capable of sustaining drug activity. Local knowledge, some level of community support, and access to a proven clientele contribute to this locality constraint and increase the likelihood that the gentrification of former drug blocks will lead to violence in the surrounding neighborhood.

1.3 Data

1.3.1 Description of Data

The City of Philadelphia has made the majority of the data utilized in this study publicly available through their OpenDataPhilly portal. I downloaded Shapefiles representing the city's Census Block Groups, neighborhoods, and street corner intersections independently and mapped them to one another. I further utilized the Philadelphia Police Department's (PPD) crime incidents data set, the city's Department of Licensing and Inspections' (DLI) building and zoning permits data set, and the city's Office of Property Assessment's (OPA) property data set. Data from the years 2010-2020 are used. [2]

The OPA property assessment data includes geocoded locations of each property and the date of that property's last sale. Each property is mapped to the city block it is found on.

The DLI building and zoning permit data contains records of each permit issued within the city. The permits consist of structural alterations, demolitions, new construction, and change of zoning status permits. Each record has the date of the permit's issuance, it's permit type, and the geocoded location of the property that the permit pertains to. Each permit is mapped and assigned to its constituent city block.

The PPD crime incidents data contains incident dates, locations, and a brief descriptor of the type of incident. Here, only those incidents described as "aggravated assault-firearm", "homicide- criminal", and "narcotic/ drug law violations" were kept. Each crime is mapped to the city block that it occurred on. There are a total of 25113 individual city blocks examined over a period of ten years.

It is of note that each year roughly fifty homicides occur in the City of Philadelphia that are not firearm related. As such, as much as 3% of what are here identified as shootings may be mislabeled.[3] However, it is likely that more shootings than this go unreported each year. [4] The extent of this under-reporting is unclear, as shootings without wounded victims may not attract police attention when on average five people are shot in the city each day.

The map below shows the locations of Philadelphia's highest gun violence areas. These are those blocks experiencing a level of neighborhood shooting total greater than three standard deviations above the city-wide mean. The locations of these blocks are overlaid on a map of the city's officially designated neighborhoods, the boundaries of which are primarily the city's main surface streets.

[2] Only the years 2011-2020 were actually made use of in the study, so as to limit the study period to an even ten years. 2010 data was used solely for determining lags and deltas of variables

[3] https://data.philly.com/philly/crime/homicides/

[4] https://www.brookings.edu/blog/up-front/2016/04/27/gun-violence-in-major-u-s-cities-is-massively-underreported/

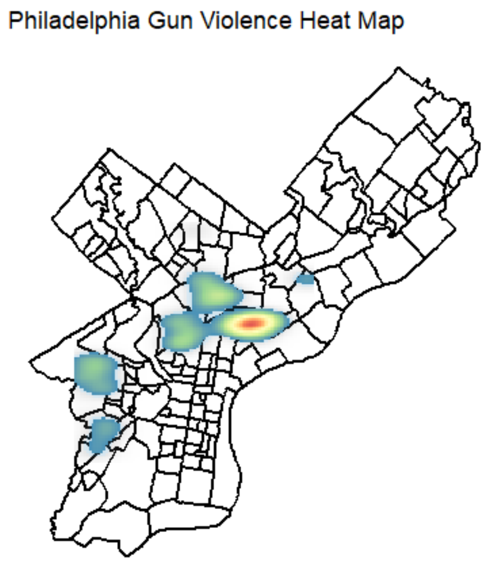


Figure 1.1: Heat map depicting the city's highest gun violence clusters

Summary statistics for these series are displayed in the table below.

Statistic	N	Mean	St. Dev.	Min	Max
Number of Shootings	251130	0.109	0.461	0	69
Number of Drug Crimes	251130	0.386	2.244	0	163
Property Sales	251130	0.174	0.653	0	32
Building, Zoning, and Alteration Permits	251130	1.967	5.269	0	461

Table 1.1: Summary statistics for block level variables.

Note that each observation in the table is for an individual block.

1.3.2 Dyads and the Directional Network of City Blocks

From this block level data, I build a directional spatial network in order to avoid having the same incidents counted multiple times. Each block is linked to its 24 closest neighbors. This is equivalent to a two block radius surrounding each block. The 24 block radius is utilized to present the best representation of a block's immediate surrounding neighborhood while managing the high computational cost of these directional networks. The figure below depicts the manner in which response blocks will be linked to a typical signal block.

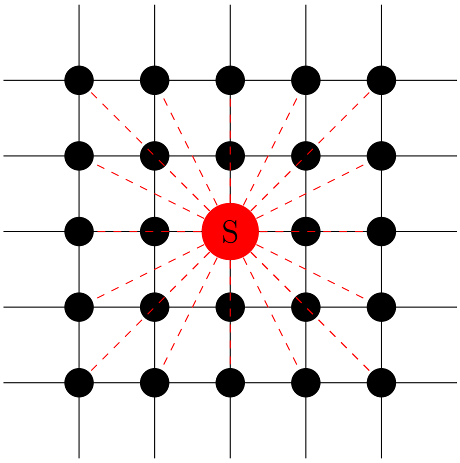


Figure 1.2: Diagram of spatial network set up

The above diagram illustrates the k-nearest neighbor spatial network (with k=24). Each node represents a city block. The central red node (labeled S) is the “signal" node, linked to each of the black nodes with a red dashed line representing the dyad connection. In this instance black nodes represent the response nodes of the dyads. Each dashed red line dyad connection represents a separate observation.

Block $a's$ link to block b is one observation, while block $b's$ link to block a is a separate observation. In total, this provides 602,712 unique linkages, herein referred to as “dyads". The first node will be referred to as the “signal block". The node it is linked to will be referred to as the “response block". Treatments will occur to the signal block, while outcomes will be measured at the response block. In practice, this means that a treatment to block a will have its effect measured at some block b in its immediate radius. The figure below depicts the bi-directional nature of these dyad linkages.

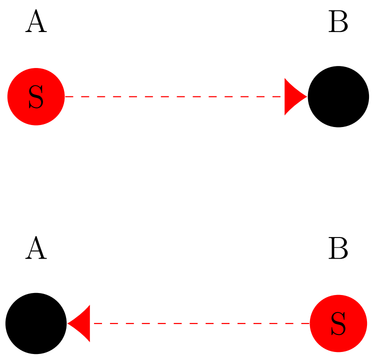


Figure 1.3: Diagram illustrating bi-directional nature of dyads

This diagram illustrates the bi-directional nature of dyads in the spatial network. Block a links to block b in one dyad with a as the “signal node" (the block that is capable of receiving treatment) and b as the “response node" (the block at which treatment effect is measured). In a separate dyad, block a is connected to block b in the opposite relationship, with b as the “signal node" and a as the “response node". Each node pairing provides two separate dyad observations.

In total, I index the data in dyad by year panels. Each of the 25,113 city blocks is linked to its 24 closest neighbors and charted over the course of ten years, for a total of 6,027,120 observations. As such, the primary specifications of the models specified will have over 600,000 dyad specific fixed effects. At the fine level of resolution at which data is made use of here, reliable covariate data is not available. As such, these fixed effects will be relied on to capture any time-invariant between unit variation and unit-invariant time period effects that could have been picked up by the inclusion of other covariates into the models. Census data, such as the ACS American Community Survey, which would otherwise be used to provide demographic controls are not available at this finite block level. Data for larger groupings are unreliable and offered at lengthy intervals.

The crux of my empirical strategy rests upon network link analysis, estimating average changes that occur in one node when the link's connected node is treated. To meet this purpose, a k-nearest neighbor adjacency matrix was first constructed, tying each block to its 24 closest neighbors. This square matrix has a length and a width equal to the total number of city blocks, and each row and column indexes a particular block. Every cell in the matrix has a value of zero, aside from those blocks that are among the set of k-nearest neighbors to one another. Their intersections will have a value of one (while the main diagonal, representing a block's intersection with itself will have a value of zero as well). For example, if block three and block four are neighbors, cell (4,3) and cell (3,4) will have values of one while cells (4,4) and (3,3) will have values of zero.

To generate the directional linkage dyads, each coordinate pairing with a value of one is taken as its own observation for each year. With this strategy, the average treatment effect on linked blocks is able to be efficiently estimated.

1.3.3 Measuring Gentrification

There exists no definitive method by which gentrification should be identified quantitatively (Barton, 2014). Glaeser et al. (2020) notes that there are nearly as many different measures of gentrification as there are papers written on the subject. Each definition is at its core a product of the inherent data limitations of each individual project. Some of these different definitions and measurements are discussed in Appendix A.

Because of the narrow resolution at which I am attempting to measure gentrification and its effects, along with the peculiarities of Philadelphia itself, none of the established gentrification measures are operable or desirable. I adopt an approach to measurement similiar to that of Holms and Schulz (2017) and Glaeser et al. (2020). The "GentriMap" model of Holms and Schulz (2017) is a two pronged approach to identifying gentrification, utilizing both real estate factors (upgrades to properties and value changes) and social demographic factors as identifying criteria. The approach of Glaeser et al. (2020) is similar. The authors first identify neighborhoods that are capable of gentrifying, based on poverty rates from the first five-year American Community Survey (ACS) to occur in the window of their study. Next, to identify gentrification, the authors track growth in home rental prices; with those that have above median rental price growth rates being labeled as gentrifying.

The approach I utilize builds upon the methods described above, but instead identifies gentrification at its finest resolution at its earliest stage, at the level of an individual city block at the very beginning of redevelopment. As in the above mentioned approaches, the feasible set of blocks that are capable of gentrification are identified by demographic characteristics. Here, I make use of household income levels from the 2011-15 ACS for this purpose. The finest resolution that this data is reliably available for is the Census Block Group. [5] As such, each city block is mapped to the Census Block Group that encompasses it. Each block that is located in a block group having an average household income level below the city-wide median is labeled as having the potential to gentrify.

Of those blocks identified as having the potential to gentrify, the gentrification treatment is defined as beginning based upon real estate trends. Since home sale and rental prices (or home valuations) are not easily available for the area of study, I make use of the quantity of home sales and the number of building/renovation permits to approximate the real estate component of gentrification. Spikes in home sales are likely to be accompanied by rising prices. Helms (2003) provides justification for the use of permitting data[6], finding that trends in building permit issuance do in fact match the types and locations of properties that are predicted to experience real estate interest brought on by gentrification. Permitting data is provided at the property level before being mapped to a city-block. I use the income and housing data to create a conservative identification of gentrification.

To be labeled as gentrifying, the block in question must either have a change from the previous year in the number of new construction or building alteration permits issued that is more than three standard deviations above the city-wide mean for that year, have a change in the number of home sales that is greater than three standard deviations above the city-wide mean, or have both a change in the number of permits issued and a change in the number of home sales that each exceed two standard deviations above the city-wide mean. Limiting the gentrification label to the far-right tail of the permitting and home sale distributions has the benefit of preventing the definition from being arbitrarily applied and ensures that only relatively unprecedented shocks are able to turn this treatment indicator on. Further, as stated, for a block to be labeled as gentrifying it must also be in a Census Block Group with below-median household income as measured in the 2010 Census. The numerical values of these classification criteria are summarized in the table below.[7]

[5]Census Blocks mentioned here are not the same as the city blocks utilized as the basic level of observation in this study. Philadelphia has 384 Census Tracts, 1336 Census Block Groups, 18872 Census Blocks, and 25,113 actual city blocks

[6]Using permitting data as a justification for gentrification implies a lagged effect on the treatment, as there is undoubtedly a delay between a permit being issued and any sort of residential demographic shifts. This is addressed in the treatment dynamics subsection of the robustness checks portion of this paper. Further, the primary results are replicated with a lagged treatment definition in the supplemental appendix.

[7]The supplemental appendix replicates this paper's main analysis with a separate gentrification criteria.

Year	Conditions to be met in addition to below median household income condition	# blocks newly meeting criteria
2011	ΔSales $\geq$ 2 or ΔPermits $\geq$ 12 or (ΔSales $\geq$ 1 & ΔPermits $\geq$ 8)	265
2012	ΔSales $\geq$ 2 or ΔPermits $\geq$ 12 or (ΔSales $\geq$ 2 & ΔPermits $\geq$ 8)	292
2013	ΔSales $\geq$ 2 or ΔPermits $\geq$ 13 or (ΔSales $\geq$ 2 & ΔPermits $\geq$ 9)	268
2014	ΔSales $\geq$ 2 or ΔPermits $\geq$ 14 or (ΔSales $\geq$ 2 & ΔPermits $\geq$ 9)	241
2015	ΔSales $\geq$ 3 or ΔPermits $\geq$ 15 or (ΔSales $\geq$ 2 & ΔPermits $\geq$ 10)	131
2016	ΔSales $\geq$ 3 or ΔPermits $\geq$ 16 or (ΔSales $\geq$ 2 & ΔPermits $\geq$ 11)	152
2017	ΔSales $\geq$ 3 or ΔPermits $\geq$ 17 or (ΔSales $\geq$ 2 & ΔPermits $\geq$ 10)	189
2018	ΔSales $\geq$ 3 or ΔPermits $\geq$ 18 or (ΔSales $\geq$ 2 & ΔPermits $\geq$ 11)	152
2019	ΔSales $\geq$ 3 or ΔPermits $\geq$ 19 or (ΔSales $\geq$ 2 & ΔPermits $\geq$ 13)	115
2020	ΔSales $\geq$ 3 or ΔPermits $\geq$ 18 or (ΔSales $\geq$ 2 & ΔPermits $\geq$ 12)	35
Total		1840

Table 1.2: Criteria for Gentrification Classification

It is evident from the above table that the necessary conditions to be labeled as gentrifying grow more stringent over time. This reflects the overall rising tide of development across the city. As more areas experience revitalization, a more significant amount of development is needed to represent a development shock to a neighborhood.

Utilizing changes in permitting and sales, rather than absolute numbers, ensures that areas of longer term sustained real estate interest do not get mislabeled as newly gentrifying. The primary results of the study are replicated with varying definitions of gentrification in the supplemental appendix. In the upcoming econometric evidence, I will suppose that once a block is labeled as having begun the gentrification treatment, it maintains the gentrification label for the remainder of the study. The map below displays the locations of gentrifying blocks in the city.

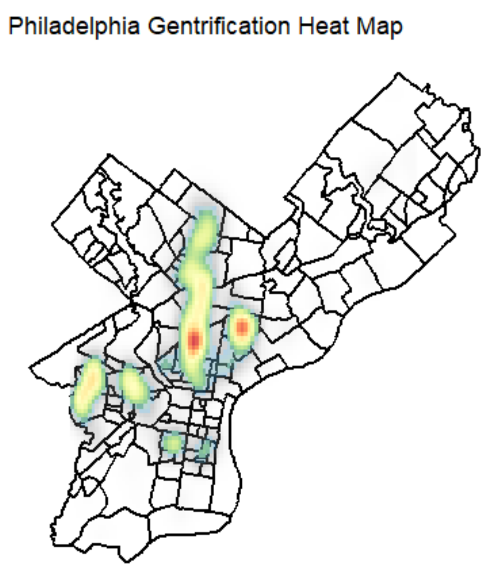


Figure 1.4: Heat map depicting density of blocks that are labeled as gentrifying during the period of the study

Appendix B details an alternative gentrification measure where gentrification is treated as a continuous variable. Appendix D presents evidence of the validity of the treatment definition utilized, by comparing the spatial distribution of blocks labeled as gentrifying by this strategy and of locations identified as gentrified by media reports.

1.4 Estimation Strategy and Results

1.4.1 Primary Specification

A two-way fixed effect differences-in-differences model is specified to identify the relationship between gentrification and gun violence. [8] City blocks gentrify independently of one another across time. Exploiting this variation provides an ideal setting from which to estimate gentrification's impact on violence.

This primary difference-in-differences specification is reported below.

$$Shootings_{it} = \beta_i + \psi_t + \alpha_1 Gentrified_{it} + \epsilon_{it} \tag{1.1}$$

Where, i indexes dyads, t indexes years, β_i is a vector of dyad fixed effects, and ψ_t is a vector of year fixed effects. $Gentrified_{it}$ is a dummy variable equal to 1 if block i has been identified as having begun gentrifying by period t.

The α_1 parameter will capture the estimated causal effect that gentrification has on each treated dyad. The two fixed effect vectors partial out the time and dyad invariant components of gun violence, allowing for the treatment parameter to capture the estimated treatment effect. A positive value for this coefficient indicates that the gentrification treatment has led to an increase in shootings on that block.

The results of fitting this primary empirical model to the data using both the total number of shootings and an Inverse Hyperbolic Sine transformation are reported below.[9] The IHS transformation of the dependent variable is introduced so as to be able to interpret the treatment effect as an elasticity and to deal with the large right tale of shootings data (Bellemare and Wichman, 2019).

[8] Additional specifications and robustness tests are included in the supplemental appendix.

[9] Estimation methods aimed at addressing the large amount of zeros in the data are reported in the appendix.

	Dependent variable:	
	Neighbor Block Shootings	
	Total	Inverse Hyperbolic Sine
Gentrified	0.021***	0.012***
	(0.003)	(.002)
Dyad Fixed Effects	✓	✓
Year Fixed Effects	✓	✓
Mean	0.11	0.085
Observations	6,027,120	6,027,120
Adjusted R^2	0.13	0.16
AIC	7512901	2248264
Note: All Standard Errors clustered at the sending block level.		
		*p<0.1; **p<0.05; ***p<0.01

Table 1.3: Output from primary differences-in-differences specification from various estimation methods

Gentrification has a positive effect on levels of gun violence observed in the neighborhood surrounding the gentrified block. This effect represents an 18% increase in gun violence on blocks linked to gentrified blocks (in the primary OLS specification). This result is analyzed in detail in the robustness section of this paper.

1.4.2 Common Trends Assumption and Treatment Dynamics

Given that the empirical model makes use of a staggered treatment timing difference-in-differences estimator, this estimator's unbiasedness as a causal estimator rests upon the assumption of parallel trends in outcome between both control and treatment groups (prior to those treatment blocks receiving their gentrification treatment). Following the bulk of recent literature, an event-study plot is used to visually demonstrate the average treatment effect at a number of periods before or after treatment.

This plot is made by specifying the following regression model, adapted from He and Wang (2017).

$$Shootings_{it} = \sum_{k=-9,k\neq-1}^{k=9} D_{it}^{k} \cdot \delta_k + \gamma_t + \psi_i + \epsilon_{it} \tag{1.2}$$

Here, the coefficients of interest are δ_k. D_{it}^k represents a vector of dummy variables equal to one, if the sending block of dyad i in period t is k periods away from beginning gentrification. $k = 0$ in the initial treatment period. Further, as in He and Wang (2017), $k = -1$ is omitted so that post-treatment event study estimators are relative to the period immediately before treatment.

The results are plotted in the graphic below.

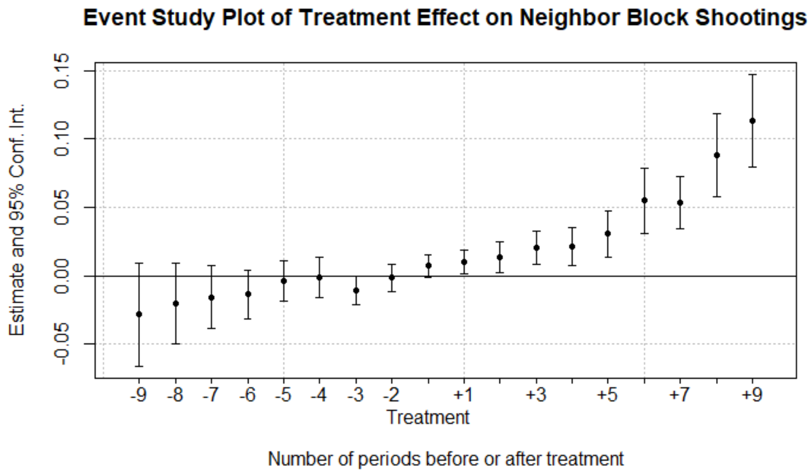


Figure 1.5: Observed treatment effect by number of years pre or post treatment

Of the eight pre-treatment coefficients, only one is significant at the 95% confidence level. This result is not dissimilar from those surveyed by Roth (2020). However, the primary criteria by which the common trends assumption will be tested on is that of a joint-significance χ^2 test. All of the pre-period coefficients will be tested together against a null hypothesis of joint insignificance.

This test returns a Wald statistic of 1.41, which has a p-value of 0.18. As such, the null hypothesis cannot be rejected. There is not evidence that the assumption of common trends is violated. Gentrification's effect on gun violence is increasing over time. While this may initially seem counter-intuitive (as it should be expected that the market will re-stabilize quickly), this feature is likely due to the fact that I have measured gentrification at the point of permit *issuance*. This represents the beginning point of neighborhood change, the effects of which are likely to grow more pronounced across time. Further, it should be noted that since units are treated in all years of the study's window, less observations are available on the tails of the event study. As such, confidence intervals for these cohorts grow quite large and the precision of these estimates is questionable. Inference based off of these event study estimates should be done cautiously. However, the event study specification here is sufficient for the purpose of evaluating the parallel trends assumption.

1.4.3 Addressing the Potential Non-Random Assignment of the Gentrification Treatment

There is certainly cause to question whether the patterns of real estate investment and development that are used here to proxy for gentrification can in fact be treated as exogenous or random. Despite the above evidence of parallel trends for both treated and untreated blocks, arguments can easily be made that the blocks that gentrify are not selected at random. This would obviously impact the ability of the findings reported here to be interpreted as plausibly causal in nature. Fortunately, the synthetic difference-in-differences estimator of Arkhangelsky et al. (2021) allows for causal identification without reliance on the assumption of parallel trends and in settings in which treatment assignment is not independent of unobserved time invariant unit-specific characteristics or unobserved time-varying global shocks. These shocks may have differential impacts on individual units, removing the threat of bias introduced by the omission of variables with potentially non-uniform impacts across units. Synthetic difference-in-differences estimation solely requires that the error term be uncorrelated with treatment status. This movement away from the standard additive fixed effects framework to that of interactive fixed effects allows for identification when regressors are functions of unit fixed effects or systematic time-varying unobserved characteristics (Bai 2009).

The above event study analysis (Figure 5) provides evidence of parallel trends between treated and control units. This suggests that eventual treatment status is not correlated with the outcome variable in pre-treatment periods. Further, the correlation between the binary treatment variable and the error term from the two-way fixed effects specification is near zero and statistically insignificant when tested using the point-biserial correlation test of Tate (1954). This test specifically addresses correlation between a continuous and binary variable. As such, the only cause for concern regarding possible non-random treatment assignment rests on the potential correlation between treatment assignment and some unobserved characteristics, ether time-varying or time-invariant. Estimation under the interactive fixed effects framework underpinning the synthetic difference-in-differences model allows for identification when treatment status is correlated with unit-specific time invariant characteristics and time-varying unobservables. For example, if a spate of development across the city is viewed as an unobserved global shock, with differential impacts on different city blocks, an interactive fixed effects framework allows for identification to still be possible even if these differential impacts are predictive of treatment status.

This estimator, as extended to staggered treatment adoption settings in Porreca (2022), allows for a localized causal estimator in which treated units are compared to a never treated control group that is weighted on both the time and unit dimensions. This produces a control group for identification that is composed of those units most similar to those that are treated and that follows a time path similar to the treated units. The synthetic counterfactual's construction differs from that of the traditional synthetic control method, in which the counterfactual is built from the weighting of observed covariates. The synthetic difference-in-differences model instead builds a counterfactual from the weighting of the outcome variable. Unlike the more standard two-way fixed effects estimator used in the primary specification, there is no earlier treated versus later treated component to the estimator (Goodman-Bacon 2021). Treated units are solely compared to a weighted composite of never treated units.

In the context of this study, treated dyads are compared to never treated dyads that are most similar in both pre-treatment periods and in the time trend those units follow in the post-treatment period. This local estimator provides a comparison between the outcomes of gentrified blocks and those blocks that themselves would be seen as "potentially gentrifiable". The results of replicating the primary specification with this estimator are reported in the table below.

	Dependent variable:
	Neighbor Block Shootings
Gentrified	0.905***
	(0.004)
Dyad Fixed Effects	✓
Year Fixed Effects	✓
Observations	6,027,120

Note: Standard errors are calculated with jackknife procedure. This is a local estimator, with only those similar dyads of the control group being used a basis for comparison.

**p<0.1; **p<0.05; ***p<0.01*

Table 1.4: Output replicating the primary specification with the synthetic difference-in-differences estimator.

As is evident from the table above, gentrification still has a positive and statistically significant impact on neighbor block shootings. This estimate, however, does need to be interpreted with some caution as it is solely a local estimate of this effect. Its magnitude reflects the estimated difference in shootings across the entire time window that the neighbors to a gentrified block will experience relative to the neighbors of an extremely similar, yet ungentrified, block. While I will avoid drawing any conclusions about this estimator's magnitude, it does sufficiently demonstrate that even with the possible endogeneity of gentrification, the positive plausibly causal relationship found in the primary specification is sound.

1.4.4 SUTVA and Gentrification's Impact on Other Gentrified Blocks

The stable unit treatment value assumption (SUTVA) of Rubin (1980) is essential for identification in difference-in-differences. In this setting, in which I am seeking to measure treatment spillovers onto neighbor blocks, this assumption requires that the gentrification of one block does not have an effect that extends beyond its 24 block neighborhood. If this assumption is violated here, and gentrification has the same effect in inducing increases in gun violence in an even greater surrounding neighborhood, this simply would introduce a downward bias into the treatment effect estimator. As such, if there is cause to suspect that SUTVA has been violated, then the positive effect of gentrification spurring neighborhood violence that is estimated here is in fact *understating* gentrification's true impact.

It is similarly worth examining the effect that this treatment has on neighboring blocks that have themselves already been treated, Gentrification does not occur disparately across the city. As evidenced in the map displayed earlier, gentrification tends to occur in clusters. The table below shows, for the final year of this study (2020), the degree of clustering among gentrified blocks. At the end of the study, there were 1,840 blocks recorded as having begun gentrifying. The table displays the frequency of a gentrified block having a number of gentrified neighbors in its 2 block radius (k=24).

Number of treated neighbors	Number of treated blocks in each category
0 to 5	246
6 to 10	378
11 to 15	535
16 to 20	518
21 to 24	163

Table 1.5: Frequency of gentrified observations with a number of gentrified neighbors falling into a specific range

It is evident from the above table that gentrification occurs in clusters, but very few gentrified blocks are entirely surrounded by gentrified neighbors. The majority of treated blocks, have less than half of their neighbors eventually treated.

To examine the heterogeneous impact that gentrification has on other gentrified blocks versus non-gentrified blocks, the data is subset into dyads that connect to gentrified blocks and dyads that connect to non-gentrified blocks. The primary model specification was replicated on these two subsets. Here the coefficient is disaggragated into the treatment effect on those blocks that are also already gentrified and those that are not.

	Dependent variable:	
	Neighbor Blocks Shootings	
	(Gentrified Response Block Subset)	(Non-Gentrified Response Block Subset)
Gentrified	0.011	0.016**
	(0.008)	(0.005)
Dyad Fixed Effects	✓	✓
Year Fixed Effects	✓	✓
Mean	0.245	0.103
Observations	292,231	5,734,889
Adjusted R^2	0.13	0.13
AIC	601560	6799962

Note: All Standard Errors clustered at the sending block level. *p<0.1; **p<0.05; ***p<0.01

Table 1.6: Dissagregation of treatment effect between neighbor blocks that have themselves previously been treated and those that have not.

There is no effect observed on the gentrified response blocks. The insignificance of this estimated effect suggests that the increases in neighborhood gun violence that stems from gentrification are limited to those neighbor blocks that are themselves yet to begin the redevelopment associated with gentrification. This provides evidence that the effect being captured in this paper's primary regression specification is largely occurring on non-gentrified blocks.

1.5 Mechanism

1.5.1 Resident Turnover

Implicit in the earlier discussed hypothesized mechanism is an assumption that the redevelopment of urban neighborhoods brings with it a turnover of the resident population. To validate this assumption, providing a basis for the hypothesis of gentrification changing the geographic constraints of illegal drug markets, I employ data at the census block group level of geographic aggregation from the 2006-2010, 2011-2015, and 2016-2020 American Community Surveys. I then map permit issuance and home sale data its constituent census block group and ACS period. A constructed index of ethnic fractionalization is regressed on a continuous gentrification index (similar to the index used to replicate the main result of this paper in Appendix B). A positive relationship between gentrification and ethnic fractionalization implies a turnover of neighborhood residents and a growing degree of racial/ethic diversity in the neighborhood.

Criminal organizations have traditionally been able to operate in areas of ethnic homogeneity (Skaperdas 2001). Some empirical evidence of this proposition is provided in Trawick and Howsen (2006), who document that crime rates across Kentucky are positively correlated with a county's degree of ethnic homogeneity. Criminal organizations rely on tacit community support to function. Changes to the structure of that community, including a loss of in-group identity for criminal actors, has the potential to lessen the degree of support and community deference that allowed the organization to previously flourish. In the context of the present study, increased ethnic fractionalization would presumably limit the ability of the criminal organization to entrench itself in a neighborhood to operate illegal drug markets. Subsequently, the criminal organization wishing to continue as a competitor in the illegal drug market would need to relocate to some other territory.

I follow Alesina et al. (1999) in constructing an index of ethnic fractionalization. This index measures the probability that two randomly drawn people from a particular geography belong to different ethnic groups. Higher values of this index represent higher degrees of ethnic and racial diversity in a particular geography. The index values are calculated as follows:

$$\begin{gathered} Ethnic\ Index = 1 - \sum_{h}^{H} (Race_h)^2 \\ \forall\, h \;\in\; \{Asian, Black, Multiracial, Pacific\ Islander, White, Other\} \end{gathered} \tag{1.3}$$

Here, h indexes racial groups represented in US Census ACS surveys.

The gentrification index utilized captures both of the operational components discussed earlier in constructing the binary gentrification indicator: the potential to gentrify, based on poverty rates, and the level of development as captured by permit issuance and home sales. More real estate interest in more impoverished census block groups leads to higher values on this spectrum, while less interest in wealthier census block groups leads to lower values.

The real estate development component of this index is calculated as follows.

$$Real\ Estate\ Index_{it} = \frac{(\frac{sales_{it}}{\bar{sales_t}} + \frac{permits_{it}}{\bar{permits_t}})}{\frac{1}{N} \cdot \sum_{i=1}^{N} (\frac{sales_{it}}{\bar{sales_t}} + \frac{permits_{it}}{\bar{permits_t}})} \tag{1.4}$$

This index is then multiplied by its constituent census block group's poverty rate to provide a simple continuous measure of gentrification.

$$Gentrification\ Index_{it} = Real\ Estate\ Index_{it} \cdot Poverty\ Rate_{it} \tag{1.5}$$

Using these two indices, a a fixed effects OLS specification allows for estimation of the relationship between gentrification and ethnic fractionalization.

$$Ethnic\ Index_{it} = \beta_i + \psi_t + \alpha_1 Gentrification\ Index_{it} + \epsilon_{it} \tag{1.6}$$

In this specification, i indexes census block groups and t indexes 5 year ACS survey periods. β_i and ψ_t are vectors of unit and survey period fixed effects. The results from this specification are displayed in the table below.

	Dependent variable:
	Ethnic Fractionalization
Gentrification Index	0.056***
	(0.016)
Census Block Group Fixed Effects	✓
ACS Year Fixed Effects	✓
Mean	0.36
Observations	3,795
Adjusted R^2	0.69
AIC	3888.6
Note: Standard errors are clustered at the census tract level.	*p<0.1; **p<0.05; ***p<0.01

Table 1.7: Output from regressing ethnic fractionalization index on gentrification index.

These results demonstrate a positive relationship between the degree of gentrification in a particular census block group and its level of ethnic fractionalization. This provides suggestive evidence that as lower income areas are redeveloped there is a turnover of residents which would in turn likely make these areas less viable for sustaining illicit drug markets and organized criminal activity.

1.5.2 Drug Markets

The driving presumption behind this paper's empirical strategy is that a block's gentrification makes that block no longer suitable for a drug market. There are multiple potential rationales that would support this assumption. As described above, gentrification includes a turnover of residents and as such likely reduces the original resident base that supported the illicit activity. Further, gentrification is associated with increased policing that is likely to have a deterrent effect. Therefore, my hypothesis is that the drug trade is disrupted at the gentrified signal block and displaced to the response blocks.

To empirically test the accuracy of this hypothetical mechanism, the following difference-in-differences estimation model is specified and tested. This specification seeks the plausibly causal impact of gentrification on the levels of drug crime observed on a block following that block's gentrification.[10] For evidence of my hypothesized mechanism's validity, the post-gentrification treatment coefficient should be negative and significant.

$$Drugs_{it} = \beta_i + \psi_t + \alpha_1 Gentrified_{it} + \epsilon_{it} \tag{1.7}$$

In this case, i indexes individual blocks rather than dyads, t indexes time periods, β_i is a vector of block fixed effects, ψ_t is a vector of year fixed effects, ϵ_{it} is an error term, and α_1 is the parameter

[10] This specification mirrors the primary estimation strategy, for gentrification's effect on shootings.

of interest. The results of this specification are reported below. As the result demonstrates, there is sufficient evidence to support the validity of this mechanism. Newly gentrified blocks do not support the same level of drug trade post-gentrification as they did pre-gentrification.

	Dependent variable:
	Own Block Drug Crimes
Gentrified	-0.218***
	(0.068)
Dyad Fixed Effects	✓
Year Fixed Effects	✓
Mean	0.39
Observations	251,130
Adjusted R^2	0.56
AIC	22504829
Note: Standard errors are clustered at the sending block level.	*p<0.1; **p<0.05; ***p<0.01

Table 1.8: Output from differences-in-differences estimation of gentrification's negative effect on immediate block drug crime.

Theory predicts that the removal of viable territory from an illegal market will increase the relative value of the remaining territory to be competed for. I have provided evidence that gentrification reduces drug crime on the immediate gentrified block, implying that this block has been removed from the set of viable territory to be competed for. This is not enough to explain how gentrification is able to instigate increases in gun violence. It needs to be seen that in the short run the gentrification of one block leads to increases in drug crime on linked neighbor blocks. As such, I will limit the treatment variable to it's effect in the immediate period of treatment[11] and replicate the same specification with i now indexing dyads and drug crime at this level as the response variable.

[11] This is done by replacing the Gentrification Indicator with a new indicator that is equal to one if a given block is newly gentrified in that particular period.

	Dependent variable:
	Neighbor Block Drug Crimes
Immediate Year Gentrification Indicator	0.06**
	(0.023)
Dyad Fixed Effects	✓
Year Fixed Effects	✓
Mean	0.38
Observations	6,027,120
Adjusted R^2	0.56
AIC	22408438
Note: Standard errors are clustered at the sending block level.	*p<0.1; **p<0.05; ***p<0.01

Table 1.9: Output from differences-in-differences estimation of gentrification's short run positive effect on neighboring block drug crime.

In the short run there is a clear increase in neighbor block drug crime following a block's gentrification. When taken in conjunction with gentrification's effect on drug crime on the immediate block, there is a clear suggestion that gentrification pushes drug crime away from the immediate block and into the surrounding neighborhood. This provides evidence of a likely mechanism behind gentrification's positive relationship with gun violence.

1.5.3 Displacement of Drug Activity Spurring Violence

An unfortunate attribute of crime data is the inability to directly observe the motives and context behind criminal incidents. Here, this means that I am unable to directly test the assumption that it is the spatial displacement of drug activity, discussed above, that creates the increases in gun violence observed in the neighborhood of newly gentrifying blocks. However, Sobrino (2019) derives a theoretic basis for the assumption that more valuable drug market territory will see more criminal organizations seeking to enter that territory and that levels of violence will increase as more criminal organizations enter the same territory. Further, she demonstrates empirically that once a second cartel enters a municipality homicide rates in that municipality increase, continue to increase with the entry of subsequent cartels, and do not decline until all but one cartel exit the municipality (Sobrino 2019).

Similarly, Taniguchi et al. (2011), Brantigham et al. (2012), and Bruhn (2021) all show that significant outliers in levels of violence are most likely to be observed at the borders of gang territories. Taken together, this empirical and theoretic evidence provides support to my proposition that the displacement of drug activity is a driving factor behind the uptick in gun violence observed following a block's gentrification. This displacement pushes rival criminal organizations into closer proximity to one another, bringing their informal territorial boundaries into conflict. Consequently, this in turn causes the same sort of violent response observed in other studies. Gang competition at territorial (or market) boundaries fosters increasing levels of violence at those boundaries.

1.5.4 Heterogeneous Treatment Effects: The gentrification of high drug crime blocks

Here, I analyze the heterogeneous treatment effects, providing evidence that the treatment has a significantly stronger effect on those blocks that house the illegal drug trade. It was argued earlier that gentrification makes a block less able to sustain an illegal drug market and that the loss of physical territory from such an illicit market would be a destabilizing force. That this can lead to violence should be no surprise. The territory that remains gains in value, incentivizing fiercer competition for its control. Because of this, it is expected that the treatment effect will be more pronounced when it is blocks that have been observed to have high levels of illegal drug activity that gentrify.

To begin, as mentioned previously, drug incidents from the same Philadelphia Police Department crime incidents database are mapped to their nearest city block. These incidents are listed in the city's records as "narcotic/ drug law violations". Those blocks that have a number of incidents in the previous year that exceeded three standard deviations above the city-wide mean are labeled as high drug crime areas. Once labeled a drug block that block will maintain that label for the remainder of the study. Allowing for blocks to gain this label at later time periods is meant to capture some of the dynamic nature of the movement of the drug trade across the city. A map of the location of these blocks is shown below.

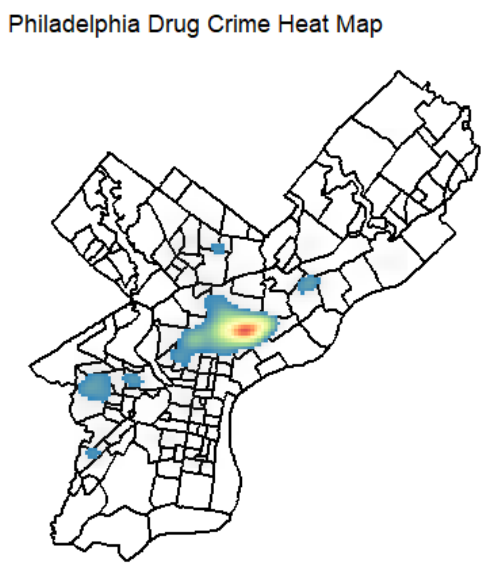


Figure 1.6: Heat map depicting density of blocks designated as high drug crime areas during the period of study

The table below charts the number of drug blocks both initially identified and gentrified in the city in each year. It is evident that drug blocks are both a relative rarity in the city and that their gentrification is an even rarer occurrence. Across the ten year window examined, the threshold for this classification is on average at least seven drug incidents occurring on that block in a given year.

Year	Total New Blocks Gentrified	Total New Blocks Labeled High Drug crime	Number of Drug Blocks Gentrified
2011	265	282	23
2012	292	0	24
2013	268	71	21
2014	241	62	23
2015	131	34	5
2016	152	35	13
2017	189	42	7
2018	152	49	13
2019	115	29	10
2020	35	50	4
Total	$1,840$	654	143

Table 1.10: Counts of unique blocks labeled as gentrified or as drug blocks by year

With this drug block indicator, the primary difference in difference specification is able to be disaggregated into the treatment effect on non-drug blocks and the effect on drug blocks.

$$\begin{aligned} Shootings_{it} = \beta_i + \psi_t + \alpha_1 NonDrugBlockGentrified_{it} + \\ \alpha_2 DrugBlockGentrified_{it} + \epsilon_{it} \end{aligned} \tag{1.8}$$

This model is essentially the same as the primary specification. However, here the treatment effect parameter is split into its constituent effects on the the two different types of treated units.

	Dependent variable:	
	Neighbor Block Shootings	
	Total	Inverse Hyperbolic Sine
Drug Block Gentrification Indicator	0.116***	0.068***
	(0.016)	(0.009)
Non-Drug Block Gentrification Indicator	0.014***	0.008***
	(0.003)	(0.002)
Dyad Fixed Effects	✓	✓
Year Fixed Effects	✓	✓
Mean	0.11	0.085
Observations	6,027,120	6,027,120
Adjusted R^2	0.13	0.16
AIC	7512571	2247992

Note: Standard errors are clustered at the sending block level. *p<0.1; **p<0.05; ***p<0.01

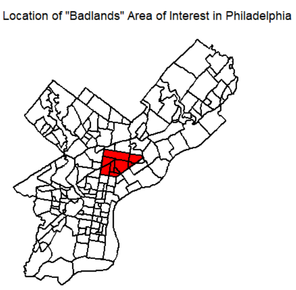
Table 1.11: Output from disaggregating difference and difference coefficient with various estimation methods

As is evident in both specifications, the gentrification of a drug block has a highly significant positive effect on the level of gun violence exhibited in the surrounding neighborhood. Further, it is of note that the effect on drug blocks is consistently of a higher magnitude than that on non-drug blocks. Focusing on the totals results, the treatment effect is more than ten times as large when it is a drug block that gentrifies. Testing the equality of these two coefficients returns a Wald Chi Square of 12.38, with a corresponding p-value below 0.01. These two coefficients are significantly different than one another.

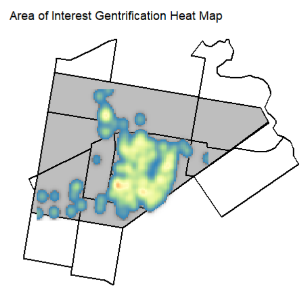
It is clear that the gentrification of a high drug crime block has a greater violent spillover into the surrounding neighborhood than does the gentrification of a more typical city block. This is consistent with the hypothesized mechanism behind the story of this paper. In the next section, I focus on one particular Philadelphia neighborhood, known locally as the “Badlands", that has a reputation as housing an open-air heroin market. This neighborhood has the largest concentration of high drug crime blocks of the city, and as such based on this result is expected to demonstrate an even more pronounced increase in neighborhood violence following gentrification.

1.6 Case Study of the “Badlands"

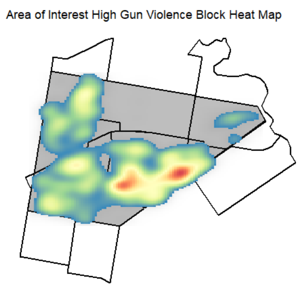
Within the city of Philadelphia, one particular section stands at the cross roads of the trends here-in described: the so-called North Philadelphia “Badlands" (which has shown up in red in all of the above heat maps). This section, located in the lower northeast of the city, is known locally as a hotbed of illegal drug activity and has for some time been plagued by rampant gun violence. However, despite this, the rapid gentrification of surrounding neighborhoods has led to an influx of development within this area. The maps below depict this section’s location within Philadelphia, the distribution of gentrifying blocks within the area, and the distribution of high drug crime and high gun violence blocks in the area.



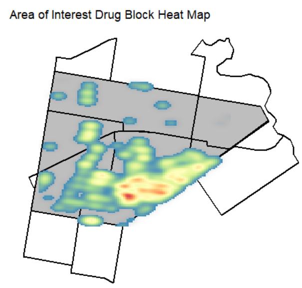
(a) Area location



(b) Gentrification heat map



(c) Gun violence heat map



(d) Drug block heat map

Figure 1.7: Location and heat maps for Badlands area of interest

Graphic depicting the Badlands are of interest’s location within the city (in red) along with heat maps of gentrification, high gun violence blocks, and high drug crime blocks. For the heat maps, the area of interest is in gray.

It is of note that there are no formal boundaries to this area. As such, I have adopted a commonly held definition of its borders as described in conversations with locals.[12] This definition has the Badlands encompassing the area of north Philadelphia bounded by Kensington Avenue to the east, Broad Street to the west, York Street to the south, and Hunting Park Avenue to the north.

This area is widely known for its quite visible open air drug markets, where illegal narcotics (primarily heroin and fentanyl) are sold openly on gang controlled street corners.[13] This unregulated market is operated by crews from around the city and attracts drug users from across the greater-Philadelphia region. Newer development and an influx of young professionals into the surrounding areas of Fishtown, Olde Kensington, and Northern Liberties has created a rise in housing prices and and spate of development within the traditional Badlands area. Because of these trends, and its uniquely large scale and unregulated drug market, the Badlands provides an opportunity to replicate this paper’s city-wide finding in a more geographically concentrated area.

The results from replicating the primary difference-in-differences specification and the disaggregation specification for heterogeneous treatment effects on this geographic subset are displayed in the tables

[12]This same definition is also mirrored at https://en.wikipedia.org/wiki/Philadelphia_Badlands

[13]See https://www.nytimes.com/2018/10/10/magazine/kensington-heroin-opioid-philadelphia.html for a New York Times piece examining the Badlands’ heroin market

below. This is simply replicating the primary analysis on a truncated data set limited to this particular section of the city.

	Dependent variable: Neighbor Block Shootings	
	Total	Inverse Hyperbolic Sine
Gentrified Indicator	0.09***	0.05***
	(0.03)	(.015)
Dyad Fixed Effects	✓	✓
Year Fixed Effects	✓	✓
Mean	0.30	0.22
Observations	308,880	308,880
Adjusted R^2	0.14	0.16
AIC	687461	399057
Note: Standard errors are clustered at the sending block level.		*p<0.1; **p<0.05; ***p<0.01

Table 1.12: Output from primary differences-in-differences specification from various estimation methods

	Dependent variable: Neighbor Block Shootings	
	Total	Inverse Hyperbolic Sine
Drug Block Gentrification Indicator	0.16***	0.09***
	(0.05)	(0.02)
Non-Drug Block Gentrification Indicator	0.05**	0.03**
	(0.018)	(0.011)
Dyad Fixed Effects	✓	✓
Year Fixed Effects	✓	✓
Mean	0.30	0.22
Observations	308,880	308,880
Adjusted R^2	0.14	0.16
AIC	687413	399021
Note: Standard errors are clustered at the sending block level.		*p<0.1; **p<0.05; ***p<0.01

Table 1.13: Output from disaggregating difference and difference coefficient with various estimation methods

As is evident from these results, the same direction of the effects seen across the city are mirrored in this area; albeit with much larger magnitudes. The impact of gentrification on gun violence is felt more acutely in this neighborhood. This is consistent with the result demonstrated earlier. In a neighborhood like this, with a greater clustering of high drug blocks, gentrification's destabilizing affect is of a greater magnitude.

1.7 Discussion

I have demonstrated a robust positive relationship between gentrification and gun violence. Following a block's gentrification, there is a substantial increase in shootings in the surrounding neighborhood. Economic intuition provides a potential mechanism behind some of these increases. I have argued that due to a lack of enforceable territorial claims, competition in illegal markets grows more violent with shocks to the total viable territory available and that gentrification serves as such a shock. However, this is not the only mechanism behind gentrification's role in spurring gun violence.

It is likely that gentrification forces intra-city migration. Residents are displaced from their long term homes, and forced into the remaining viable tracts of affordable housing. These sorts of situations, where disaffected low-income residents are forced to live in unfamiliar neighborhoods surrounded by similarly disaffected and displaced neighbors, have the potential to cause excessive tension. That this can give rise to explosions in gun violence is not surprising.

Despite these computational and data limitations, I have presented a readily replicable measure of gentrification at an extremely narrow spatial resolution. Further, through leveraging a massive networked data set, I have been able to generate robust significant estimates of the effect that gentrification has on levels of gun violence.

All of the coefficients estimated are at the dyad level and as such are of seemingly small magnitudes. However, when considering that each of these estimates is of the average effect on the single linkage between a newly gentrified block and each of its neighbors, these numbers add up quickly to a startling total. Over the ten year window of the study, back of the envelope calculations suggest that the city experienced some 5,800 shootings that can be attributed to gentrification. This means, that of the 27,000 shootings that occurred across the city during this time period, around 21% may have been spillover effects of gentrification.

Exploring the mechanism posited by this study's economic intuition, the estimator was significantly higher. The gentrification of drug blocks accounted for roughy 2,400 additional shootings during the ten year time span. This attributes some 8% of the city's gun violence across the decade to gentrification giving rise to instability in the city's illicit drug markets.

These estimates are striking. Each episode of gun violence has the potential to forever impact the lives of victims, perpetrators, families, and community members at large. As such, in light of this unintended impact, it is crucial that urban development occur responsibly and intentionally. Forced displacement, through pricing out residents, has very real effects on the surrounding neighborhood.

Future research would do well to examine the natural limitations to where illegal drug markets can feasibly exist. Models that are able to predict the spatial movements of these markets across time can help to prevent these high levels of violence from continuing. Further, this result suggests that police resources could be better utilized in the neighborhoods surrounding newly developed blocks. City policy may benefit from efforts that seek to stave off this sort of violent spillover effect through deployment of officers and social workers into areas that are likely experiencing significant population displacements. It is likely that displacement will give rise to volatility and violence. It is reasonable that resources be deployed proactively alongside the forces of development in an attempt to prevent this sort of community violence.

1.8 References

[1] Alesina, A., Baqir, R., & Easterly, W. (1999). Public goods and ethnic divisions. The Quarterly journal of economics, 114(4), 1243-1284.

[2] Arkhangelsky, D., Athey, S., Hirshberg, D. A., Imbens, G. W., & Wager, S. (2021). Synthetic difference-in-differences. American Economic Review, 111(12), 4088-4118.

[3] Autor, D., Palmer, C., & Pathak, P. (2017). Gentrification and the amenity value of Crime reductions: Evidence from rent deregulation. https://doi.org/10.3386/w23914

[4] Bai, J. (2009). Panel data models with interactive fixed effects. Econometrica, 77(4), 1229-1279.

[5] Baker, A., Larcker, D. F., & Wang, C. C. (2021). How much should we trust staggered difference-in-differences estimates? SSRN Electronic Journal. https://doi.org/10.2139/ssrn.3794018

[6] Barnum, J. D., Campbell, W. L., Trocchio, S., Caplan, J. M., & Kennedy, L. W. (2016). Examining the environmental characteristics of drug Dealing Locations. Crime & Delinquency, 63(13), 1731–1756. https://doi.org/10.1177/0011128716649735

[7] Barton, M. (2014). An exploration of the importance of the strategy used to identify gentrification. Urban Studies, 53(1), 92–111. https://doi.org/10.1177/0042098014561723

[8] Becker, G. S. (1968). Crime and punishment: An economic approach. In The economic dimensions of crime (pp. 13-68). Palgrave Macmillan, London.

[9] Bellemare, M. F., & Wichman, C. J. (2020). Elasticities and the inverse hyperbolic sine transformation. Oxford Bulletin of Economics and Statistics, 82(1), 50-61.

[10] Brantingham, P. J., Tita, G. E., Short, M. B., & Reid, S. E. (2012). The ecology of gang territorial boundaries. Criminology, 50(3), 851-885.

[11] Bruhn, J. (2021). Competition in the Black Market: Estimating the Causal Effect of Gangs in Chicago. SSRN Electronic Journal. https://doi.org/10.2139/ssrn.3837695

[12] Callaway, B., & Sant'Anna, P. H. C. (2020). Difference-in-differences with multiple time periods. Journal of Econometrics. https://doi.org/10.1016/j.jeconom.2020.12.001

[13] Castillo, J. C., Mejía, D., & Restrepo, P. (2020). Scarcity without Leviathan: The Violent Effects of Cocaine Supply Shortages in the Mexican Drug War. The Review of Economics and Statistics, 102(2), 269–286. https://doi.org/10.1162/rest_a_00801

[14] Chalfin, A., & McCrary, J. (2017). Criminal Deterrence: A Review of the Literature. Journal of Economic Literature, 55(1), 5–48. https://doi.org/10.1257/jel.20141147

[15] Clogg, C. C., Petkova, E., & Haritou, A. (1995). Statistical methods for comparing regression coefficients between models. American Journal of Sociology, 100(5), 1261–1293. https://doi.org/10.1086/230638

[16] DeAngelo, G. (2012). Making space for crime: A spatial analysis of criminal competition. Regional Science and Urban Economics, 42(1-2), 42–51. https://doi.org/10.1016/j.regsciurbeco.2011.04.008

[17] de Chaisemartin, C., & d'Haultfoeuille, X. (2020). Difference-in-differences estimators of intertemporal treatment effects. SSRN Electronic Journal. https://doi.org/10.2139/ssrn.3731856

[18] Dewey, M., Míguez, D. P., & Saín, M. F. (2016). The strength of collusion: A conceptual framework for interpreting hybrid social orders. Current Sociology, 65(3), 395–410. https://doi.org/10.1177/0011392116661226

[19] Dixon, P. M. (2001). Ripley's K function. Encyclopedia of environmetrics, 3, 1796.

[20] Dragan, K., Ellen, I. G., & Glied, S. (2020). Does gentrification displace poor children and their Families? New evidence from Medicaid data in New York City. Regional Science and Urban Economics, 83, 103481. https://doi.org/10.1016/j.regsciurbeco.2019.103481

[21] Easton, S., Lees, L., Hubbard, P., & Tate, N. (2019). Measuring and mapping displacement: The problem of quantification in the battle against gentrification. Urban Studies, 57(2), 286–306. https://doi.org/10.1177/0042098019851953

[22] Eck, J. E. (1994). Drug markets and drug places: A case-control study of the spatial structure of illicit drug dealing.

[23] Ellen, I. G., Horn, K. M., & Reed, D. (2019). Has falling crime invited gentrification? Journal of Housing Economics, 46, 101636. https://doi.org/10.1016/j.jhe.2019.101636

[24] Glaeser, E. L., Luca, M., & Moszkowski, E. (2020). Gentrification and neighborhood change: Evidence from yelp (No. w28271). National Bureau of Economic Research.

[25] Goodman-Bacon, A. (2021). Difference-in-differences with variation in treatment timing. Journal of Econometrics. https://doi.org/10.1016/j.jeconom.2021.03.014

[26] Grossman, H. I., & Kim, M. (1995). Swords or Plowshares? A Theory of the Security of Claims to Property. Journal of Political Economy, 103(6), 1275–1288. https://doi.org/10.1086/601453

[27] Harbring, C. (2006). The effect of communication in incentive systems—an experimental study. Managerial and Decision Economics, 27(5), 333–353. https://doi.org/10.1002/mde.1266

[28] He, G., & Wang, S. (2017). Do college graduates serving as village officials help rural china? American Economic Journal: Applied Economics, 9(4), 186–215. https://doi.org/10.1257/app.20160079

[29] Helms, A. C. (2003). Understanding gentrification: an empirical analysis of the determinants of urban housing renovation. Journal of Urban Economics, 54(3), 474–498. https://doi.org/10.1016/s0094-1190(03)00081-0

[30] Henningsen, A. (2020, August 4). Estimating Censored Regression Models in R using the censReg Package. Retrieved September 23, 2021, from https://cran.r-project.org/web/packages/censReg/vignettes/censReg.pdf.

[31] Hirshleifer, J. (1989). Conflict and rent-seeking success functions: Ratio vs. difference models of relative success. Public Choice, 63(2), 101–112.

[32] Holm, A., & Schulz, G. (2017). GentriMap: A Model for Measuring Gentrification and Displacement. Gentrification and Resistance, 251–277. https://doi.org/10.1007/978-3-658-20388-7_10

[33] Laniyonu, A. (2017). Coffee Shops and Street Stops: Policing Practices in Gentrifying Neighborhoods. Urban Affairs Review, 54(5), 898–930. https://doi.org/10.1177/1078087416689728

[34] Manson, A., Schroeder, J., Van Riper, D., Kugler, T., & Ruggles, S. IPUMS National Historical Geographic Information System: Version 16.0 [dataset]. Minneapolis, MN: IPUMS. 2021. http://doi.org/10.18128/D050.V16.0

[35] O'Sullivan, A. (2005). Gentrification and crime. Journal of Urban Economics, 57(1), 73–85. https://doi.org/10.1016/j.jue.2004.08.004

[36] Papachristos, A. V., Smith, C. M., Scherer, M. L., & Fugiero, M. A. (2011). More Coffee, Less Crime? The Relationship between Gentrification and Neighborhood Crime Rates in Chicago, 1991 to 2005. City & Community, 10(3), 215–240. https://doi.org/10.1111/j.1540-6040.2011.01371.x

[37] Polo, M. (2005). Internal cohesion and competition among criminal organizations. In The economics of organized crime (pp. 87–109). essay, Cambridge University Press.

[38] Porreca, Z. (2022). Synthetic Difference-In-Differences Estimation With Staggered Treatment Timing. Economics Letter, 220, 110874

[39] Roth, J. (2019). Pre-test with Caution: Event-study Estimates After Testing for Parallel Trends. Department of Economics, Harvard University, Unpublished Manuscript.

[40] Rossiter, D. G. (2021, April 6). Exercise: Spatial Point Pattern Analysis. Retrieved October 20, 2021, from http://www.css.cornell.edu/faculty/dgr2/_static/files/R_PDF/exPPA.pdf.

[41] Rubin, D. B. (1980). Randomization analysis of experimental data: The Fisher randomization test comment. Journal of the American Statistical Association, 75(371), 591-593.

[42] Skaperdas, S. (1992). Cooperation, Conflict, and Power in the Absence of Property Rights. The American Economic Review, 82, 720–739.

[43] Skaperdas, S. (1996). Contest success functions. Economic Theory, 7(2), 283–290. https://doi.org/10.1007/bf01213906

[44] Skaperdas, S. (2001). The political economy of organized crime: providing protection when the state does not. Economics of Governance, 2(3), 173–202. https://doi.org/10.1007/pl00011026

[45] Skarbek, D. (2016). Covenants without the Sword? Comparing Prison Self-Governance Globally. American Political Science Review, 110(4), 845–862. https://doi.org/10.1017/s0003055416000563

[46] Smith, C. M. (2012). The Influence of Gentrification on Gang Homicides in Chicago Neighborhoods, 1994 to 2005. Crime & Delinquency, 60(4), 569–591. https://doi.org/10.1177/0011128712446052

[47] Sobrino, F. (2019). Mexican Cartel Wars: Fighting for the U.S. Opioid Market.

[48] Strumpf, E. (2011). Medicaid's effect on single women's labor supply: Evidence from the introduction of Medicaid. Journal of health economics, 30(3), 531-548.

[49] Taniguchi, T. A., Ratcliffe, J. H., & Taylor, R. B. (2011). Gang set space, drug markets, and crime around drug corners in Camden. Journal of research in crime and delinquency, 48(3), 327-363.

[50] Tate, R. F. (1954). Correlation between a discrete and a continuous variable. Point-biserial correlation. The Annals of mathematical statistics, 25(3), 603-607.

[51] Trawick, M. W., & Howsen, R. M. (2006). Crime and community heterogeneity: Race, ethnicity, and religion. Applied Economics Letters, 13(6), 341-345.

[52] Varese, F. (2010). What is organized crime?, In Organised Crime: Critical Concepts in Criminology (Vol. 1, pp. 11–33). introduction, Routledge.

1.9 Appendix

1.9.1 Appendix A: Discussion of Other Gentrification Definitions

First and foremost, there exists no overarching consensus regarding the method by which gentrification should be identified quantitatively (Barton, 2014). Authors have employed novel identification strategies, such as utilizing the number of coffee shops in a given neighborhood [14] (Papachristos et al., 2011). Easton et al. (2019) have stated that the difficulties in developing a quantitative identification strategy for gentrification is primarily a product of insufficient data. They posit that gentrification is the *displacement* of lower income residents, and that neighborhoods undergoing gentrification can be identified by demographic changes, real estate trends, and survey data. However, the authors concede that this sort of data is limited in its usefulness and may not necessarily capture all of this sort of resident displacement.

Helms (2003) provides a thorough model of real estate dynamics as a product of gentrification. His model's framework provides much of foundation on which this study's quantitative definition of gentrification rests. Using data from Chicago to test his theoretical model, Helms (2003) finds that trends in building permit issuing do in fact match the types and locations of properties that are predicted to experience real estate interest brought on by gentrification. This finding supports this study reliance on building/ renovation permit and home sale data to serve as a marker of gentrification.

Holm and Schulz (2017) provide a robust statistical model for identifying gentrification. As is typical of this literature, the authors concede that data availability is the largest limitation to the quantitative identification of gentrification. Their "GentriMap" model includes both real estate factors (upgrades to properties and value changes) and social demographic factors as identifying criteria.

This study embodies a similar approach to that of the "GentriMap" model, subject to the data limitations. Sales counts are used in place of value changes. This proxy is a reasonable substitution, given data limitations. An increase in sales reflects an increase in demand which in turn should imply an increase in property valuations. Permit data is used in the same manner as in other gentrification models. Median household incomes are utilized as a demographic factor indicative of a block's ability to become gentrified -as most definitions agree that low income household displacement is a necessary condition.

Several other recent papers have put forth measurements of gentrification. Dragan et al. (2019) made use of 5-year ACS data at the Census Tract level (the third smallest level of Census data aggregation) to identify potentially gentrifying blocks based in part on changes in demographic characteristics captured by these surveys. This data was augmented by administrative data identifying the addresses of individuals, to track their movement across the city over the study period. Ellen et al. (2019) used confidential census data to identify household demographic characteristics and residency, to track high-income movements into the city that are associated with gentrification. This analysis takes place at the Census Tract level.

[14] Such a strategy would not be useful for a city such as Philadelphia which embodies a striking divide between residential and commercial neighborhoods. As such, this method is not replicated in this study

Autor et al. (2017) use spatial exposure to the buildings that end their rent control status as a measure of gentrification. This is done on the basis of some literature the authors cite, that establish the abatement of rent control as being associated with numerous changes in neighbor housing stock quality and pricing. Glaeser et al. (2020) puts forth both a continuous and a binary gentrification measure, built on 5 year ACS survey data and rental price data. That analysis is conducted at the zip-code level.

As is evident, this study is not alone in offering forth an empirical measure of gentrification. Nor is the measurement strategy utilized herein disparate from that utilized in the bulk of the literature. However, this study differs in that it measures gentrification at the finest resolution available, while using solely open access freely available data.

1.9.2 Appendix B: Continuous Definition of Gentrification

It is worth briefly examining a potential continuous measure of gentrification, in which gentrification exists on a spectrum. More real estate interest on more impoverished blocks leads to higher values on this spectrum, while less interest of wealthier blocks leads to lower values.

The real estate component is represented by the following index.

$$Real\ Estate\ Index = \frac{(\frac{\Delta sales_{it}}{\Delta sales_t} + \frac{\Delta permits_{it}}{\Delta permits_t})}{\frac{1}{N} \cdot \sum_{i=1}^{N}(\frac{\Delta sales_{it}}{\Delta sales_t} + \frac{\Delta permits_{it}}{\Delta permits_t})} \tag{1.9}$$

This index value is block specific. It is then multiplied by its constituent Census Block Group's poverty rate to provide a simple continuous measure of gentrification.

$$Gentrification\ Index = Real\ Estate\ Index \cdot Poverty\ Rate \tag{1.10}$$

This measure has a mean of 0.032, a standard deviation of 1.05, and ranges from -10.07 to 9.3. Higher values are associated with higher degrees of gentrification.

Replacing the binary gentrification treatment indicator from the initial specification with this continuous measure creates a simple OLS fixed effects model which yields the following results.

$$shootings_{it} = \beta_i + \psi_t + \alpha_1 Gentrification\ Index_{it} + \epsilon_{it} \tag{1.11}$$

	Dependent variable:
	Neighbor Block Shootings
Continuous Gentrification Measure	0.0007**
	(0.0003)
Dyad Fixed Effects	✓
Year Fixed Effects	✓
Mean	0.11
Observations	6,027,120
Adjusted R^2	0.13
AIC	7512447
Note: Standard errors are clustered at the sending block level.	*p<0.1; **p<0.05; ***p<0.01

Table 1.14: Continuous model of gentrification's effect

1.9.3 Appendix C: Addressing Non-standard Distribution of Shootings

As is to be expected, the distribution of shootings is heavily clustered at zero. The vast majority of blocks and linkages do not experience any shootings. Typically, this is addressed with with a Tobit model. However, due to the large number of fixed effects used in the empirical model, this is extremely computationally costly and not feasible. This sort of estimation would rely on specialty atypical maximization routines that would still provide inconsistent estimations of the variance (Henningsen, 2020).

As such, a random effects Tobit model is specified here, the results of which are reported in the table below.

	Dependent variable:
	Number of Shootings
Gentrification Indicator	0.11***
	(0.0008)
Log Sigma Mu	-0.62***
	(0.0007)
Log Sigma Nu	-0.78***
	(0.00002)
Dyad Random Effects	✓
Year Random Effects	✓
Mean	0.11
Observations	6,027,120
Log Liklihood	9419719
Note: Standard errors are clustered at the sending block level.	*p<0.1; **p<0.05; ***p<0.01

Table 1.15: Random Effects Tobit Model

The disaggregation of this coefficient, as shown earlier with OLS estimators in the heterogeneous treatment effects section, is displayed below.

	Dependent variable:
	Number of Shootings
Drug Block Gentrification	0.40***
	(0.003)
Non-Drug Block Gentrification	0.08***
	(0.0009)
Log Sigma Mu	-0.62***
	(0.0007)
Log Sigma Nu	-0.78***
	(0.00002)
Dyad Random Effects	✓
Year Random Effects	✓
Mean	0.11
Observations	6,027,120
Log Liklihood	9419083
Standard Errors	Clustered at Census Block Group Level
Note:	*p<0.1; **p<0.05; ***p<0.01

Table 1.16: Random Effects Tobit Model- disaggregation of primary treatment estimator

Now, addressing the fixed effects issue, a hurdle model is approximated. This will provide two separate estimators: one, a logit estimator of the probability of the shooting outcome variable not being equal to zero and the other a differences-in-differences OLS estimator ran solely on the subset of the data in which the number of shootings is greater than zero.

As above, the disaggregated coefficient is reported as well.

	Dependent variable:	
	Pr(Number of Shootings>0)	Number Shootings When Number of Shootings>0
	Logit	OLS
Gentrification Indicator	0.017	0.06***
	(0.028)	(0.022)
Dyad Fixed Effects	✓	✓
Year Fixed Effects	✓	✓
Mean	0.08	1.37
Observations	2,391,250	483,363
Adjusted R^2	0.12	0.14
AIC	2627291	1548094
All Standard Errors Clustered at Census Block Group Level		
Note: Logit Adjusted R^2 is Squared Correlation	*p<0.1; **p<0.05; ***p<0.01	

Table 1.17: Output from replicating primary difference and difference specification with hurdle approximation

	Dependent variable:	
	Pr(Number of Shootings>0)	Number Shootings When Number of Shootings>0
	Logit	OLS
Drug Block Gentrification Indicator	0.18***	0.19**
	(0.068)	(0.060)
Non-Drug Block Gentrification Indicator	-0.001	0.04*
	(0.027)	(0.022)
Dyad Fixed Effects	✓	✓
Year Fixed Effects	✓	✓
Mean	0.08	1.37
Observations	2,391,250	483,363
Adjusted R^2	0.12	0.14
AIC	2627269	1548063
All Standard Errors Clustered at Census Block Group Level		

Note: Logit Adjusted R^2 is Squared Correlation *p<0.1; **p<0.05; ***p<0.01

Table 1.18: Output from disaggregating difference and difference coefficient with hurdle approximation

1.9.4 Appendix D: Validating the Gentrification Definition with Newspaper Text Analysis

To validate the measure of gentrification utilized in this study, I began by downloading the text of every newspaper article referencing gentrification from the *Philadelphia Inquirer* and the *Philadelphia Daily News* published between 2010 and 2020. This yielded 663 articles. These articles were then parsed by a script to extract any complete address or intersection. Limiting this to those location's within the city limits left 68 specific locations. These were then geocoded using the Google Maps API, and assigned to their corresponding city blocks, as was done with the bulk of this paper's data.

It should be noted that the points extracted from these newspapers typically came from restaurant reviews or articles relating to well established previously gentrified neighborhoods. The newspaper references stem more from the long term effects of gentrification, not the first wave of initial investment that is used as the treatment in this study. Gentrification in Philadelphia is largely an outward moving process, expanding away from Center City. As is evident in the map below (displaying the 68 newspaper generated points superimposed over the previously shown gentrification heat map) these newspaper points are clustered in the areas between my early stage gentrification treatment locations and the city's central business district.

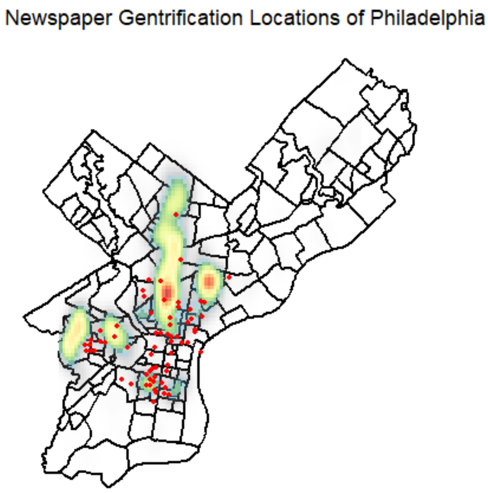


Figure 1.8: Comparison of empirical definition of gentrification to newspaper scraped examples

Blocks identified as gentrified from newspapers superimposed over the treatment gentrification heat map used in this study.

To evaluate whether it is likely that these two spatial distribution's arose from the same underlying data generating process, a bivariate Ripley's K function is used to analyze the two spatial point processes. The procedure for evaluating the independence of two spatial distributions, as detailed by Dixon (2002) and Rossiter (2021), entails generating the K measure of how many observations of one type are observed at increasing radii surrounding an observation of the other type. This line is then plotted against a theoretical Ripley's K generated by a homogeneous Poisson process, which represents complete spatial randomness (CSR) or independence of the two distributions. This result is plotted below.

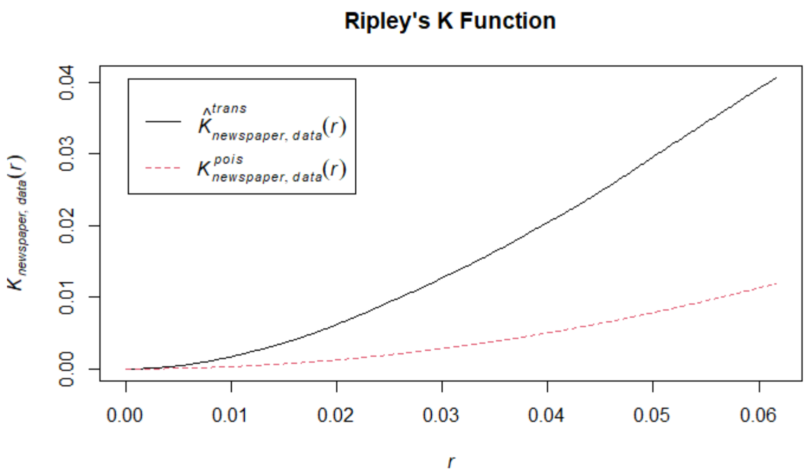


Figure 1.9: Ripley's K function for newspaper and empirically defined gentrified blocks

Plotting of Ripley's K function, the Poisson (CSR) process is in red and the empirically observed K function is in black

As is evident from this figure, the empirically observed K function is consistently above the Poisson process line. This means that within every radius[15] surrounding a given observation of a newspaper generated gentrified block there are more blocks identified by my empirical gentrification definition than would be at random. This is indicative of the two spatial distributions not being independent of one another, and provides some measure of validation for the empirical definition used herein.

[15]The units of radius here are in hundredths of a longitude-latitude square. As such, 0.06 is around one kilometer in size.

2 Synthetic Difference-In-Differences Estimation With Staggered Treatment Timing

The published version of this chapter is available in Economics Letters, Volume 220.

2.1 Introduction

Causal inference has become the dominant aim of empirical micro-economic research, making unbiased causal estimation an issue at the forefront of the econometric literature. Difference-in-differences (DiD) estimation is now applied in settings with staggered treatment adoption, rather than solely in the more restrictive simultaneous treatment adoption framework. Recent research has focused on best practices in these settings (de Chaisemartin and D'Haultfœuille 2020; Sun and Abraham 2020; Roth 2020; Goodman-Bacon 2020; Baker et al. 2021; Callaway and Sant'anna 2021).

Simultaneously, an alternative path to estimating causal inference, the Synthetic Control Method (SCM) of Abadie and Gardeazabal (2003), has been developed. This method synthesizes an optimal control group for settings with only one treated unit, through the application of weights to pre-treatment period control units. SCM allows researchers to make use of empirical causal inference in case study settings. This method has been extended to settings with multiple treated units. Ben-Michael, Felller, and Rothstein (2021) explored these multi-unit synthetic control estimators and began the formalization of pooled synthetic control estimators.

Arkhangelsky et al. (2021) introduced a new causal inference estimator that can be employed in both traditional DiD settings and in those settings typically in the domain of SCM, such as those with a single treated unit and no clear control group. Their method, synthetic difference-in-differences estimation (SynthDiD), does not rely on parallel trends assumptions or assumptions of exogeneity of the treatment as in DiD. SynthDiD does this by introducing the weighting of both pre-treatment time periods and cross-sectional units in the construction of a synthetic counterfactual for causal estimation. This method allows more heterogeneity of outcomes and has been suggested to improve the precision of the estimator (Arkhangelsky et al. 2021). Details of this method are discussed in the next section. The application of this estimator to staggered treatment timing settings has yet to be thoroughly explored.

The process by which a SynthDiD estimator is obtained by weighting of individual cohort estimators is briefly described in an appendix to Arkhangelsky et al. (2021). However, the actual functional form of the estimator has not been formalized until the present note. Here, I offer a formalization of this estimator, present an estimator of its variance, and offer a practical example demonstrating its implementation. To illustrate this estimator's importance and distinction, I employ data from Abrams (2012) to demonstrate the differences in estimators obtained through SynthDiD, a standard two-way fixed effects DiD estimator, the group-time average treatment effect estimator of Callaway and Sant'Anna (2021), and the partially pooled synthetic control method estimator of Ben-Michael et al. (2021). This paper is meant to serve as an introduction to the use of SynthDiD estimation in the staggered treatment timing settings often encountered in empirical research.

First, I will present a brief overview of the single treatment period SynthDiD estimator as described by Arkhangelsky et al. (2021). I then proceed to formalize the staggered treatment timing SynthDiD estimator and describe the practicalities of its estimation. Last, I will demonstrate the outcomes obtained by implementing this estimator using open source software.

2.2 Synthetic DiD

In depth detail regarding the performance, precision, and theoretical basis of the synthetic difference-in-differences estimator can be found in Arkhangelsky et al. (2021). In short, this estimator is novel in its inclusion of both time and unit weights to create a reliable counterfactual, its inclusion of a unit fixed effect (which is missing in SCM), and its reliance on weighting control observations to achieve parallel trends with treated observations rather than weighting control observations to match with treated observations directly.

The synthetic difference-in-differences estimator assumes that the data generating process follows a latent factor model:

$$\boldsymbol{Y} = \boldsymbol{L} + (\boldsymbol{W\tau})_{i,t} + \boldsymbol{E} \quad where \quad \boldsymbol{L} = \boldsymbol{\Gamma\Upsilon^T} \quad and \quad (\boldsymbol{W\tau})_{i,t} = \boldsymbol{W_{i,t}\tau_{i,t}} \tag{2.1}$$

In the above data generating process, $W_{i,t}$ represents a binary treatment indicator that is equal to 1 if unit i is treated in period t and 0 otherwise. Γ is a vector of latent time factors and Υ is a vector of latent unit factors. τ is the average treatment effect and E is an error matrix.

The average treatment effect in this data generating process is defined as:

$$\tau = \frac{1}{N_{tr}T_{post}} \sum_{i=N_{co}+1}^{N} \sum_{t=T_{pre}+1}^{T} \tau_{it} \tag{2.2}$$

N_{tr} and N_{co} represent the number of treatment and control group observations respectively. T_{pre} and T_{post} represent the number of time periods before and after treatment.

The estimator of the average treatment effect, $\hat{\tau}$, is formally defined as follows:

$$\hat{\tau} = \left(\frac{1}{N_{tr}} \sum_{i=N_{co}+1}^{N} \hat{\delta}_i\right) - \left(\sum_{i=1}^{N_{co}} \hat{\omega}_i \hat{\delta}_i\right) \tag{2.3}$$

$$\hat{\delta}_i = \left(\frac{1}{T_{post}} \sum_{t=T_{pre}+1}^{T} Y_{it} \right) - \left(\sum_{t=1}^{T_{pre}} \hat{\lambda}_t Y_{it} \right) \tag{2.4}$$

Y_{it} is the observed outcome variable of interest. $\hat{\lambda}_t$ and $\hat{\omega}_i$, the time period and unit weights, are described below. Together, the overall estimator can be expressed as follows:

$$\hat{\tau} = \left[\frac{1}{N_{tr}} \sum_{i=N_{co}+1}^{N} \left(\frac{1}{T_{post}} \sum_{t=T_{pre}+1}^{T} Y_{it} - \sum_{t=1}^{T_{pre}} \hat{\lambda}_t Y_{it} \right) \right] - \left[\sum_{i=1}^{N_{co}} \hat{\omega}_i \left(\frac{1}{T_{post}} \sum_{t=T_{pre}+1}^{T} Y_{it} - \sum_{t=1}^{T_{pre}} \hat{\lambda}_t Y_{it} \right) \right] \tag{2.5}$$

The procedure by which the weights $\hat{\lambda}_t$ and $\hat{\omega}_i$ are chosen is described in Arkhangelsky et al. (2021). The time period weights are defined as follows:

$$\begin{aligned} (\hat{\lambda}_0, \hat{\lambda}^{sdid}) &=_{\lambda_0 \in \mathbb{R}, \lambda \in \Lambda} \ \ell_{time}(\lambda_0, \lambda) \ \text{ where} \\ \ell_{time}(\lambda_0, \lambda) &= \sum_{i=1}^{N_{co}} (\lambda_0 + \sum_{t=1}^{T_{pre}} \lambda_t Y_{it} - \frac{1}{T_{post}} \sum_{t=T_{pre}+1}^{T} Y_{it})^2, \\ \Lambda = \{ \lambda \in \mathbb{R}_+^T &: \sum_{t=1}^{T_{pre}} \lambda_t = 1, \lambda_t = T_{post}^{-1} \ \text{ for all } \ t = T_{pre}+1, ..., T \} \end{aligned} \tag{2.6}$$

Similarly, the unit weights are defined as follows:

$$\begin{aligned} (\hat{\omega}_0, \hat{\omega}^{sdid}) &=_{\omega_0 \in \mathbb{R}, \omega \in \Omega} \ \ell_{unit}(\omega_0, \omega) \ \text{ where} \\ \ell_{unit}(\omega_0, \omega) &= \sum_{t=1}^{T_{pre}} (\omega_0 + \sum_{i=1}^{N_{co}} \omega_i Y_{it} - \frac{1}{N_{tr}} \sum_{i=N_{co}+1}^{N} Y_{it})^2 + \zeta T_{pre} ||\omega||_2^2, \\ \Omega = \{ \omega \in \mathbb{R}_+^N &: \sum_{i=1}^{N_{co}} \omega_i = 1, \omega_i = N_{tr}^{-1} \ \text{ for all } \ i = Nco+1, ..., N \}, \end{aligned} \tag{2.7}$$

The regularization parameter, ζ, is defined as:

$$\begin{aligned} \zeta = (N_{tr} T_{post})^{1/4} \hat{\sigma} \ \text{ with } \ \hat{\sigma}^2 &= \frac{1}{N_{co}(T_{pre} - 1} \sum_{i=1}^{N_{co}} \sum_{t=1}^{T_{pre}-1} (\Delta_{it} - \bar{\Delta})^2, \\ \text{where } \ \Delta_{it} = Y_{i(t+1)} - Y_{it}, \ \text{ and } \ \bar{\Delta} &= \frac{1}{N_{co}(T_{pre} - 1)} \sum_{i=1}^{N_{co}} \sum_{t=1}^{T_{pre}-1} \Delta_{it} \end{aligned} \tag{2.8}$$

2.3 Staggered Treatment Timing

Arkhangelsky et al. (2021) describe in their appendix a procedure for obtaining a weighted average of each treatment cohort's individual $\hat{\tau}$ to estimate the aggregate group average treatment effect. This proposed estimator is obtained by iterating through the SynthDiD algorithm on subsets of the data limited to the entire never-treated subset of units and each treatment cohort individually. This method is conceptually similar to that of Ben-Michael et al. (2021) and Callaway and Sant'Anna (2021).

Assuming the same data generating process described in Arkhangelsky et al. (2021) (and shown in Equation 1), the average treatment effect in the staggered treatment setting is defined as follows:

$$\tau = \frac{1}{N_{tr} T_{post} L} \sum_{\ell \in L | \ell > 0}^{L} \sum_{i=N_{co}+1}^{N} \sum_{t=T_{pre}+1}^{T} \tau_{it} \tag{2.9}$$

The subscript $\ell \in L : \{0, 1, ...\ell\}$ indexes a cohort membership. $\ell = 0$ is the cohort of never-treated control group observations.

This parameter can be estimated by allowing for a distinct $\hat{\tau}_\ell$ for each of L treatment cohorts.

$$\hat{\tau}_\ell = \left[\frac{1}{N_{tr}} \sum_{i=N_{co}+1}^{N} \left(\frac{1}{T_{post}} \sum_{t=T_{pre}+1}^{T} Y_{it\ell} - \sum_{t=1}^{T_{pre}} \hat{\lambda}_t Y_{it\ell} \right) \right] - \left[\sum_{i=1}^{N_{co}} \hat{\omega}_i \left(\frac{1}{T_{post}} \sum_{t=T_{pre}+1}^{T} Y_{it} - \sum_{t=1}^{T_{pre}} \hat{\lambda}_t Y_{it} \right) \right]$$

(2.10)

$\forall \ell \in L$. These individual $\hat{\tau}_\ell$ estimators are combined together by weighted averaging. The weights, μ_ℓ, are equal to the proportion of treated units that belong to each cohort, ℓ (Arkhangelsky et al. 2021). As such, the the staggered treatment timing SynthDiD estimator can be formally expressed as:

$$\hat{\tau} = \sum_{\ell \in L | \ell > 0}^{L} \left(\mu_\ell \cdot \hat{\tau}_\ell \right) \tag{2.11}$$

The weight applied to the ℓth $\hat{\tau}$ cohort estimator is μ_ℓ. This weight is defined as:

$$\mu_\ell = \frac{N_\ell}{\sum_{\ell \in L | \ell > 0} N_\ell} \tag{2.12}$$

This weight is equivalent to the proportion of the of non-zero row-sum rows in the treatment assignment matrix $\mathbf{W}$ for which the first non-zero column of that row is column ℓ.[16]

Applying these weights, the overall $\hat{\tau}$ estimator can be expressed as:

$$\hat{\tau} = \sum_{\ell \in L|\ell>0}^{L} \left(\left[\frac{N_\ell}{\sum_{\ell \in L|\ell>0} N_\ell} \right] \left[\frac{1}{N_{tr}} \sum_{i=N_{co}+1}^{N} \left(\frac{1}{T_{post}} \sum_{t=T_{pre}+1}^{T} Y_{it\ell} - \sum_{t=1}^{T_{pre}} \hat{\lambda}_t Y_{it\ell} \right) \right] - \left[\sum_{i=1}^{N_{co}} \hat{\omega}_i \left(\frac{1}{T_{post}} \sum_{t=T_{pre}+1}^{T} Y_{it} - \sum_{t=1}^{T_{pre}} \hat{\lambda}_t Y_{it} \right) \right] \right) \tag{2.13}$$

Practically, this estimator is simply a weighted average of cohort-specific estimated average treatment effects, where the weight applied to any individual cohort's specific estimator is equal to the proportion of treatment group observations that originate in a specific cohort Arkhangelsky et al. (2021).

Following Kahn (2015), it is a property of the influence functions of such estimators of summary parameters that the following holds (with $N = \sum_{\ell \in L|\ell>0}^{L} N_\ell$):

$$\frac{1}{\sqrt{N}}(\hat{\tau} - \tau) = \sum_{i=1}^{N} \psi_\tau(x_i) + o_p(1) \tag{2.14}$$

Where $\psi_\tau(x_i)$ is the influence function for the n-th observation and the summary parameter, τ. As such, the variance of the summary parameter $\hat{\tau}$ can be computed through the procedure described in Erickson and Whited (2002) and Kahn (2015). Paraphrasing Kahn (2015), this procedure entails calculating the empirical equivalents of the influence functions for each estimator and stacking them into a single matrix, Ψ, in which the rows correspond to each estimator and the columns to each observation. This procedure is similar to that used to calculate the variance of the summary parameter in Callaway and Sant'Anna (2021). The variance-covariance matrix for the individual cohort estimators can be calculated as follows:

$$\hat{V} = \frac{1}{N^2}(\Psi^T \Psi) \tag{2.15}$$

Thus, using the $\ell \times 1$ vector of weights, μ, described in Equation 12, an estimator for the variance of the aggregated summary parameter can be computed as follows:

$$\hat{V}_{\hat{\tau}} = \mu^T \hat{V} \mu \tag{2.16}$$

[16]These weights correspond to those described in the appendix to Arkangelsky et al. (2021).

2.4 Application

To demonstrate the use of the staggered-treatment timing synthetic difference-in-differences estimator, replication data from Abrams (2012) are used to estimate the causal impact of a law adopted by different states in different years. This estimation is done with a typical two-way fixed effects OLS estimator (TWFE), the synthetic difference-in-differences estimator presented here, the partially-pooled synthetic control estimator of Ben-Michael, et al. (2021), and an aggregated group-time average treatment effect estimator (CS) obtained through the estimation procedure detailed in Callaway and Sant'Anna (2021).[17]

Abrams (2012) estimates the deterrent effect of firearm sentencing enhancements. Rather than systematically replicating the entirety of Abrams (2012), replication data aggregated at the state level (used in an appendix to that paper) is used here for demonstrative purposes. Further, for simplicity of computation and demonstration, data is limited to a balanced panel of 45 states measured in each of 38 years.

$$y_{it} = \tau \cdot W_{it} + \gamma_i \upsilon_t^T + \epsilon_{it} \tag{2.17}$$

This equation is analogous to the latent factor model described as the data generating process in Equation 1. [18]For the analysis herein, $\hat{\tau}$ is the parameter of interest: the average treatment effect. W_{it} is an indicator equal to one if a state i is treated in year t. Once a unit is treated, it remains treated for all subsequent time periods. The outcome variable of interest is the per-capita armed robbery rate in state i in year t. The treatment is the enactment of a state-wide sentencing enhancement for crimes that involve firearms.

First, I evaluate whether there is a violation of the parallel trends assumption that the traditional difference-in-differences estimation relies on for identification. An event study is estimated following the example of He and Wang (2017). Periods greater than 5 periods from treatment are aggregated (to ± 6 on either tail). Further, $t = -1$ is omitted to serve as the point of reference.

$$y_{it} = \sum_{k>-5, k\neq -1}^{k>6} W_{it}^k \cdot \delta_k + \beta_t + \alpha_i + \epsilon_{it} \tag{2.18}$$

W_{it}^k is an indicator equal to one if state i in year t is k years away from initially enacting a sentencing enhancement. δ_k is the k period specific event-study estimator. The other variables and parameters remain unchanged. Equation 18 assumes the additive form of the latent factor model in the data generating process discussed in Footnote 4. By making the assumption that $\gamma_i \upsilon_t^T = \lambda_i + \omega_t$, this event

[17]The individual cohort estimators for the group-time average treatment effect estimator (CS) are estimated using the doubly robust estimation procedure as detailed in Callaway and Sant'Anna (2021).

[18]As mentioned in Arkhangelsky et al. (2021), if the latent factor component of the data generating process takes an additive form ($L_{it} = \alpha_i + \beta_t$) this is analogous to the standard two-way fixed effects estimator utilized in most difference-in-differences frameworks.

study equation is reconciled with the assumed data generating process.

The results of this estimation are displayed below. Given that some non-zero trend is evident in the pre-treatment periods, inference relying on the parallel trends assumption of difference-in-differences may be unreliable. A Wald χ^2 test of the joint significance of the pre-treatment estimators can serve as a parametric test of that parallel trends assumption (Roth, 2020). This test yields a χ^2 value of 2.01 with p-value of 0.075. As such, the null-hypothesis of joint significance fails to be rejected at the 5% level. Further analysis of this assumption is beyond the scope of this paper. However, there is some evidence that this assumption may be violated here, introducing unwarranted bias into the fixed-effects estimator.

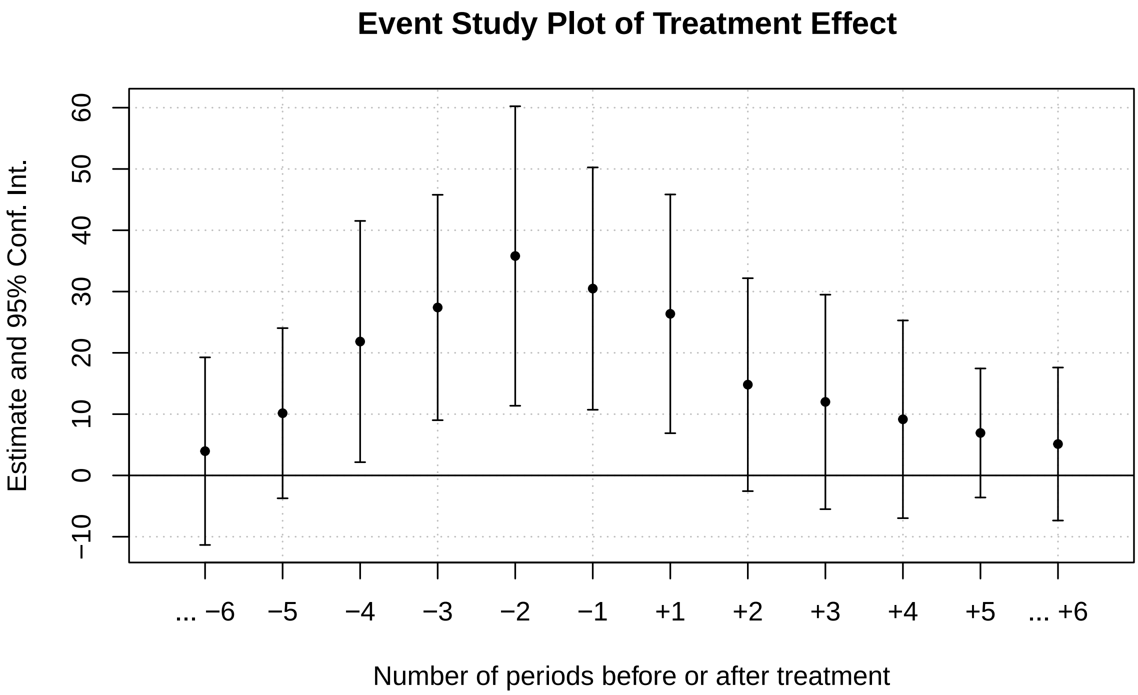


Figure 2.1: Observed treatment effect by number of years pre or post treatment

Estimating $\hat{\tau}$ under the aforementioned specifications without covariates yields the following:

	Dependent variable:			
	Per-capita armed robbery rate			
	SynthDiD	TWFE	CS	Partially-pooled SCM
$\hat{\tau}$	-16.697^{***}	-20.182^{***}	-23.724^{**}	-14.095
	(0.364)	(7.430)	(5.243)	(10.177)
Observations		1,710		1520

Note: *p<0.1; **p<0.05; ***p<0.01
Standard errors are in parentheses. Standard errors for the TWFE estimator are clustered at the state level. Standard errors for the SynthDiD estimator are calculated as described in Equations 14-16. Standard errors for the Callaway and Sant'Anna estimator are bootstrapped. Standard errors for the partially-pooled SCM estimator were constructed using a jackknife procedure. For the Partially-pooled SCM estimator, observations for units that were already treated in the initial period are dropped, as the estimator's usage requires.

Table 2.1: Estimation of $\hat{\tau}$ without covariates

SynthDiD produces an estimate of a smaller magnitude than both TWFE and CS. The variance of this estimator is smaller than all alternatives. This evidences the precision and reliability of the SynthDiD estimator discussed in Arkhangelsky et al. (2021).

In this applied context: while the impact of the sentencing enhancement is consistently estimated here to reduce armed robbery rates, SynthDiD estimates the effect (in the absence of covariates [19]) to be markedly more conservative than the alternative approaches. A brief discussion of what may be causing the discrepancy between these point estimates is included with the simulation results.

2.5 Simulation Results

A simple simulation is useful in evaluating the performance of this SynthDiD estimator and provides some insight into the differences in point estimates from the empirical results. 1000 simulated datasets were generated with their parameters calibrated to those of the Abrams (2012) replication data. Treatment is randomly assigned, with the true value of τ for each treated unit being normally distributed around

[19] SynthDiD estimation (including in the staggered treatment timing setting) allows for the potential inclusion of covariates. As detailed in a footnote to Arkhangelsky et al. (2021), an adjusted y_{it}^R outcome variable is the residual from regressing y_{it} on the matrix of covariates:

$$y_{it}^R = y_{it} - X_{it}\hat{\beta} \tag{2.19}$$

The systematic effect of unit-specific time-varying X_{it} covariates on y are partialled out as in OLS, providing an adjusted outcome variable: y_{it}^R. This is the component of y that remains unexplained after controlling for covariates.

-20 with a standard deviation of 2. The performance of the four estimators considered previously are compared in the table below. The first row (with $\hat{\tau}_{additiveFE}$) estimates the $\hat{\tau}$ average treatment effect parameter on data simulated to follow an additive fixed effects data generating process, while the second row (with $\hat{\tau}_{Factor}$) estimates the same parameter on data simulated to follow a latent factor model data generating process.

	Dependent variable:			
	Simulated per-capita armed robbery rate			
	SynthDiD	TWFE	CS	Partially-pooled SCM
$\hat{\tau}_{additiveFE}$	−20.011	−19.984	−20.013	−19.990
	(0.223)	(0.440)	(0.903)	(1.429)
$\hat{\tau}_{Factor}$	−19.926	−19.962	−19.924	−19.916
	(0.179)	(0.286)	(0.329)	(3.015)
Number of Simulations			1,000	
Note:				*p<0.1; **p<0.05; ***p<0.01
Standard errors are in parentheses.				

Table 2.2: Mean estimated $\hat{\tau}$

This simple simulation exercise yields similar point estimates and standard errors for both SynthDiD and TWFE.[20] This is unsurprising, as the data generating process of the first series of simulations relies on the assumption of an additive form for its latent factor component (see Footnote 4). When this is the case, typical two-way fixed effects DiD will consistently recover the parameter τ (Arkhangelsky et al. 2021). SynthDiD has performed almost identically to TWFE in this context. This provides some explanation for the difference in point estimates seen in the above empirical section. It implies that the data generating process underlying the replication data utilized follows an interactive fixed effects model rather than the additive form TWFE relies on.[21] The second specification, in which the data generating process follows that of a latent factor model, provides evidence that SynthDiD can consistently provide more precise estimates.[22]

[20] However, while these estimates are similar in magnitude and variance, the p value from a test of their equality is 0.52

[21] This implies that the use of SynthDiD as a preemptive correction may be analogous to the use of heteroscedasticity-robust standard errors. The estimator performs nearly the same as the standard methodology when the correction is not needed, but removes bias when the correction is merited.

[22] The p value from a test of equality of the SynthDiD and TWFE estimates from this simulation exercise is 0.46

2.6 Conclusion

In this note, I have formalized the functional form of the SynthDiD estimator in settings with staggered treatment timing and have demonstrated its practical use. While, the brief empirical demonstration is in no way meant to be a robust replication of Abrams (2012), it does demonstrate that different aggregated average treatment effect estimates can be obtained using this new and arguably more precise estimator. The simulation exercise demonstrates that in such settings SynthDiD performs as well as traditional DiD methods. However, in more realistic settings in which the underlying data generating process does not include simple additive fixed effects SynthDiD may be able to better estimate the average treatment effect.

SynthDiD may be an appropriate estimator in settings in which the underlying data generating process follows that of a latent factor model or additive fixed effects. However, due to the added computational burden of SynthDiD estimation, this estimator's usage may be inappropriate in big data settings in which the underlying data generating process follows an additive fixed effects model since both methods would be able to consistently estimate the average treatment effect, but with TWFE being much less computationally costly. This brief note provides a base for the implementation of SynthDiD in staggered treatment adoption settings, and hopefully a foundation for this estimator's further development.

2.7 References

[1] Abadie, A., & Gardeazabal, J. (2003). The economic costs of Conflict: A Case Study of the basque country. American Economic Review, 93(1), 113–132. https://doi.org/10.1257/000282803321455188

[2] Abadie, A., Diamond, A., & Hainmueller, J. (2010). Synthetic control methods for comparative case studies: Estimating the effect of California's tobacco control program. Journal of the American statistical Association, 105(490), 493-505.

[3] Abrams, D. S. (2012). Estimating the deterrent effect of incarceration using sentencing enhancements. American Economic Journal: Applied Economics, 4(4), 32–56. https://doi.org/10.1257/app.4.4.32

[4] Abrams, D. S. (2012, July 1). Replication data for: Estimating the deterrent effect of incarceration using sentencing enhancements. openICPSR. Retrieved January 18, 2022, from https://www.openicpsr.org/openicpsr/project/113838/version/V1/view

[5] Abrams, D. S. (n.d.). Web Appendix A. Robustness Checks . Estimating the Deterrent Effect of Incarceration using Sentencing Enhancements. Retrieved January 17, 2022, from https://assets.aeaweb.org/asset-server/articles-attachments/aej/app/app/2011-0005_app.pdf

[6] Arkhangelsky, D., Athey, S., Hirshberg, D. A., Imbens, G. W., & Wager, S. (2021). Synthetic difference-in-differences. American Economic Review, 111(12), 4088–4118. https://doi.org/10.1257/aer.20190159

[7] Baker, A., Larcker, D. F., & Wang, C. C. (2021). How much should we trust staggered difference-in-differences estimates? SSRN Electronic Journal. https://doi.org/10.2139/ssrn.3794018

[8] Ben-Michael, E., Feller, A., & Rothstein, J. (2021). Synthetic controls with staggered adoption (No. w28886). National Bureau of Economic Research.

[9] Callaway, B., & Sant'Anna, P. H. C. (2021). Difference-in-differences with multiple time periods. Journal of Econometrics, 225(2), 200–230. https://doi.org/10.1016/j.jeconom.2020.12.001

[10] De Chaisemartin, C., & D'Haultfœuille, X. (2020). Difference-in-differences estimators of intertemporal treatment effects. Available at SSRN 3731856.

[11] Erickson, T., & Whited, T. M. (2002). Two-step GMM estimation of the errors-in-variables model using high-order moments. Econometric Theory, 18(3), 776-799.

[12] Goodman-Bacon, A. (2021). Difference-in-differences with variation in treatment timing. Journal of Econometrics, 225(2), 254–277. https://doi.org/10.1016/j.jeconom.2021.03.014

[13] He, G., & Wang, S. (2017). Do college graduates serving as village officials help rural china? American Economic Journal: Applied Economics, 9(4), 186–215. https://doi.org/10.1257/app.20160079

[14] Kahn, J. (2015). Influence Functions for Fun and Profit. Ross School of Business, University of Michigan. Available from http://j-kahn. com/files/influencefunctions. pdf (version: July 10, 2015).

[15] Porreca, Z. (2022, January 11). Zachporreca/staggered_adoption_synthdid: Code to incorporate staggered treatment adoption (based on appendix from Arkhangelsky et al. 2021) into synthdid package. GitHub. Retrieved January 18, 2022, from https://github.com/zachporreca/staggered_adoption_synthdid

[16] Roth, J. (2022). Pre-test with Caution: Event-study Estimates After Testing for Parallel Trends. Forthcoming, American Economic Review: Insights.

[17] Sun, L., & Abraham, S. (2021). Estimating dynamic treatment effects in event studies with heterogeneous treatment effects. Journal of Econometrics, 225(2), 175–199. https://doi.org/10.1016/j.jeconom.2020.09.006

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

This chapter is coauthored with Bryan McCannon.

3.1 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 Wainright* (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.[23] A growing literature questions the quality of the publicly-provided defense in the U.S.[24] Since all have access 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.

[23]It 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.

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

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 Hjalmarsson (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.[25]

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, Anderson and Heaton (2012), Cohen (2012), and Roach (2014) identify 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 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.[26]

An important, closely-related work is that of Ater, Givati, and Rigbi (2017). They consider a reform in Israel that extended the right to counsel to suspects of crimes where it had previously only been guaranteed for defendants charged with a crime. They ask whether extending this right affects the decision to prosecute a case, and whether there is a spillover onto crime. They show that prosecutions and arrest duration decreased. Further, they show that police engaged in fewer arrests and that crime

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

[26] It is also worth pointing out that our work connects to research evaluating the impacts of legal representation in non-criminal settings. For example, Greiner and Pattanayak (2012) randomize access to legal aid services of a law school clinic to those appealing unemployment benefit eligibility rulings. Farmer and Tiefenthaler (2001) and Greiner *et al.* (2021) have investigated the impact of attorney representation in divorce disputes.

increased. Thus, defense counsel is effective at their jobs. Conditional on the willingness to pursue a conviction, we contribute to the understanding of how the right to counsel affects this prosecution.

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. Bjerk (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.

3.2 Historical Background

3.2.1 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.[27] 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).

3.2.2 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.[28] 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.

3.2.3 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

[27]Juries 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.

[28]See oldbaileyonline.org and Shoemaker (2008) for historical accounts.

harmed 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.[29] 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.[30]

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

[30] As 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).

3.2.4 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 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.[31] 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).[32] 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).[33]

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

[32] This 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.

[33] Stated 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)).

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 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).[34]

3.3 Identification

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

[34]Other legal reforms occurred around this time period as well. It was a period in English civil law procedures where the country moved from a system of "forms of action" to one with formal civil procedures. Parliamentary acts such as the Uniformity of Process Act of 1832, Real Property Limitation Act of 1833, Common Law Procedure Act of 1852, and Judicature Act of 1873 transformed the system (Maitland, 1909). These apply to civil disputes, such as property and contract, and would not affect the functioning of criminal cases.

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 nine crime categories that make up our control group are: assault, bigamy, indecent assault, libel, manslaughter, perjury, perverting justice, stealing from master, and sodomy. The 23 more-serious crimes become the treated ones. They, then, are those previously subject to the Bloody Code and were not those formally defined as misdemeanors.

Regarding the time period to study, data from the Old Bailey covers almost 200 years. 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.

3.4 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.*[35] 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

[35] www.oldbaileyonline.org

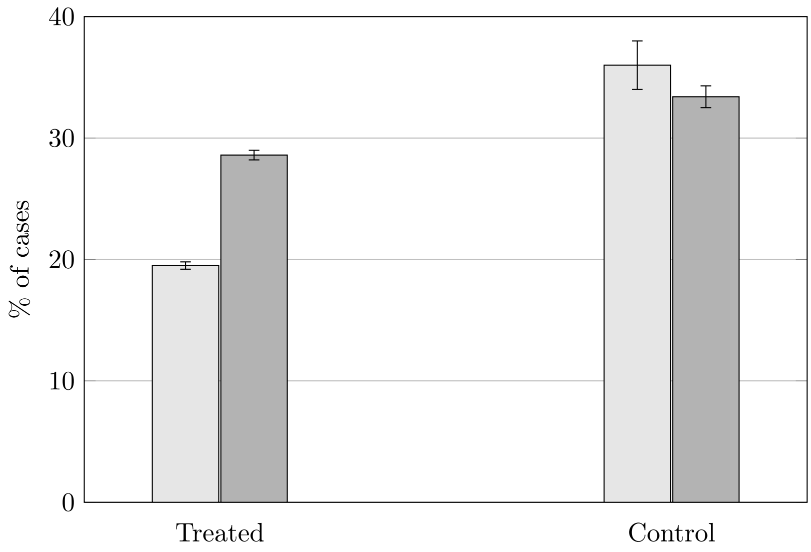


Figure 3.1: Use of Defense Counsel Before and After the 1836 Act

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.

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.[36] A question may not be assigned an author or a transcript can be incomplete.[37] Thus, our measure should be thought of as a lower bound to the actual prevalence of a defense attorney.[38] 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 3.1 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

[36] Recognizing 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.

[37] To 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).

[38] It 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.

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 3.1 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 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.

A final note regarding Figure 3.1 is that one may be concerned that the prevalence of defense counsel's use in felony cases after the institutional change is still rather modest. It is worth emphasizing that this measurement comes from a bag-of-words search in the published trial transcripts. We expect that this suffers from substantial under-reporting as a transcript can neglect to report who asked the questions or even fail to report all questions asked. It is best to view the approximately 30% rate of defense counsel presence as the proportion of cases with clearly-marked involvement. Both cases without defense counsel and cases with defense counsel but incomplete records make up the omitted category. What is important is that the growth rate in the use of defense counsel in felony cases is substantial.

3.5 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).[39] 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 3.1.

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 time frame, 69.9% of cases have only one defendant and the bulk of the observations

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

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

Table 3.1: Descriptive Statistics

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.

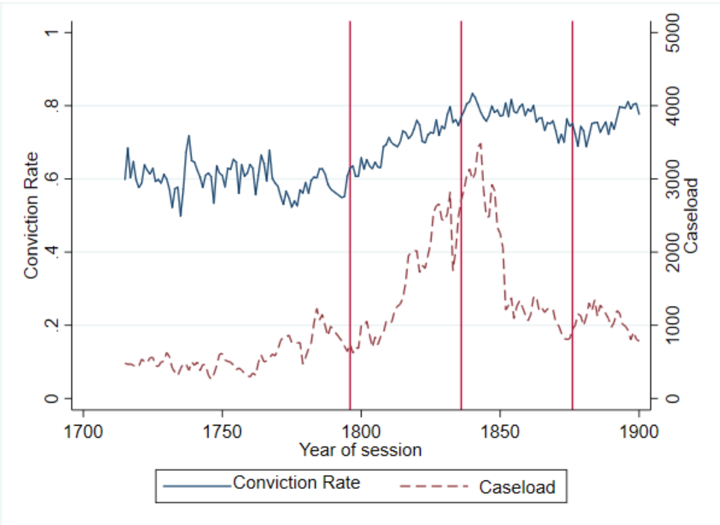


Figure 3.2: Conviction Rate over Time

The dashed line is the proportion of cases in each year that resulted in a conviction (measured on the left axis). The solid line is the number of total number of cases at the Old Bailey in each year (measured on the right axis). The three vertical lines depict the year of the treatment (middle line) and the endpoints of the baseline time period used in the analysis.

do not identify the jury's identify ($\approx 75\%$). As previously stated, our baseline analysis considers criminal cases at the Old Bailey between 1796 and 1876 (± 40 years around the Act's implementation). There are 135,365 observations.

Regarding the outcome variable of interest, for the baseline period, 75.2% of the observations had a conviction. The crime prosecuted falls within our treated crime types for 89.4% of the observations. Figure 3.2 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 3.2 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.

3.6 Results

The primary econometric model to estimate is

$$Conviction_{icy} = \beta_1 Post_{iy} \times Treated_{ic} + \tau_{iy} + \kappa_{ic} + X_{icy}\beta + \epsilon_{icy}. \tag{3.1}$$

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 defendant-specific control variables (those provided in Table 3.1). 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).

We proceed by first presenting the results from the difference-in-difference estimation. We follow this by a thorough investigation of potential threats to identification.

3.6.1 Difference-in-Difference 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 3.3 depicts them.

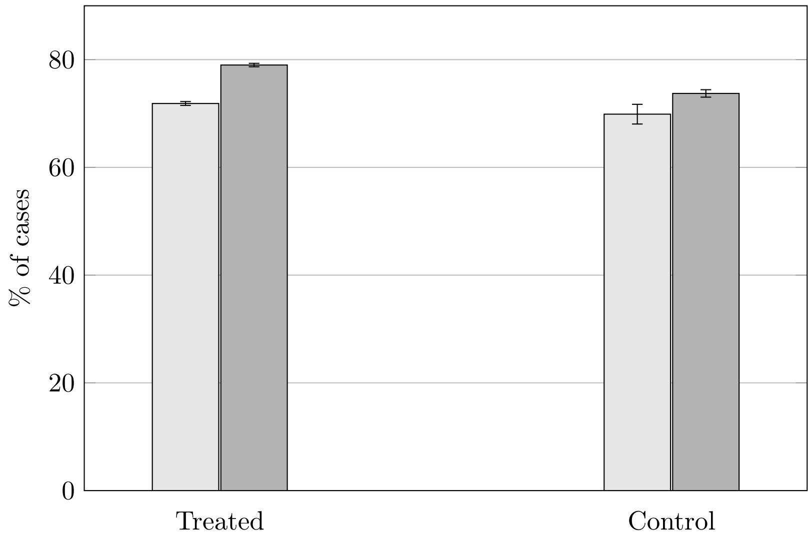


Figure 3.3: Convictions Before and After the 1836 Act

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.

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 3.3 is robust. Table 3.2 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 3.3.

Table 3.2 considers whether a conviction arises in each case. One concern is whether this effect is really arising from jury convictions as modern criminal justice, in the United States especially, rely heavily on plea bargaining. At the Old Bailey, prior to the Act's implementation, less than 3% of cases have the defendant pleading guilty. Thus, this is a rather rare phenomenon in early 19th century London. Further, while not presented here, if only jury convictions are used as the dependent variable, the difference-in-difference coefficient is larger in magnitude and still highly statistically significant. A related concern is whether the jury is convicting at the original charges, or for lesser offenses. Again, if the jury convicted at the original charges is used as the dependent variable, then the difference-in-difference

	Baseline	*Alternative Time Windows*		
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 ***	0.0239 ***	0.0190 **	0.0263 ***
	(0.0064)	(0.0063)	(0.0080)	(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

Table 3.2: Main Result

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.

coefficient is larger and highly significant.

3.6.2 Identification Threats

3.6.2.1 Parallel Trends & Dynamic Effects

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

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

$$\begin{aligned} Conviction_{icy} = {} & \alpha_1 Year_{iy} + \alpha_2 Year_{iy} \times Treated_{ic} \\ & + \alpha_3 Post_{iy} \times Year_{iy} + \alpha_4 Post_{iy} \times Treated_{ic} \times Year_{it} \\ & + \alpha_5 Post_{iy} + \alpha_6 Post_{iy} \times Treated_{ic} + \kappa_{ic} + \epsilon_{icy}. \end{aligned} \tag{3.2}$$

The coefficient α_1 captures the linear time trend in convictions for the non-treated crime types and α_2

	Baseline	*Alternative Time Windows*		
coverage:	±40 years	±30 years	±50 years	BH window
years:	[1796-1876]	[1806-1866]	[1786-1886]	[1803-1871]
	[1]	[2]	[3]	[4]
Year ($\widehat{\alpha}_1$)	0.0048 ***	0.0044 *	0.0040	0.0044 **
	(0.0013)	(0.0024)	(0.0009)	(0.0019)
Year x Treated	-0.0005	-0.00003	0.0004	0.0002
	(0.0012)	(0.00236)	(0.0010)	(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

Table 3.3: Test of (Linear) Pre-Act Parallel Trends

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 $\widehat{\alpha}_1$ and $\widehat{\alpha}_2$ 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.

measures whether the treated crimes follow a different time trend in the pre-Act period. Testing whether $\widehat{\alpha}_2 = 0$ 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 3.3 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.[40]

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. To do this, we evaluate 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* x *Treated* and *Post* x *Year* x *Treated* in equation (2).

Figure 3.4 presents the point estimates and the 95% confidence intervals of the six pentads following

[40] While not presented in Table 3.3 (since the focus in on the pre-Act time period), the time trend diverges after the Act's implementation. A level shift up in the time trend occurs ($\widehat{\alpha}_5 = 8.03$; $p < 0.01$) and the time trend for the treated crimes after the treatment does not differ from the non-treated crimes ($\widehat{\alpha}_6$ has $p > 0.6$).

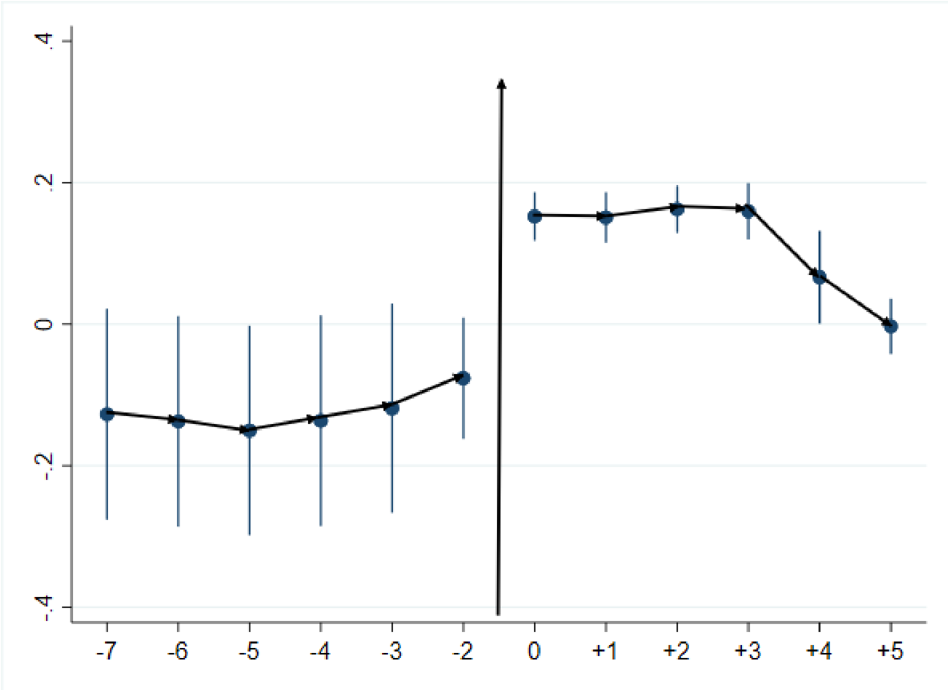


Figure 3.4: Dynamic Effects

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.

implementation (periods +1 through +6) and the six pentads prior to implementation (periods labeled -7 through -2).

First, 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.

This provides additional confirmation of common trends in the pre-Act periods. Thus, non-parallel trends does not pose a threat to our identification strategy's ability to produce causal identification.

Additionally, 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.

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

3.6.2.2 SUTVA

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 types of cases arising at the Old Bailey and shows that the main result is not driven by case mix.

3.6.2.3 Other Relevant Institutional Changes: Police & Capital Punishment

Without staggered adoption, a difference-in-difference identification strategy is subject to influences from other related changes occurring at approximately the same time. While we argue that it is the introduction of the right to counsel that is driving our results, if there is another policy change around 1836, then it is possible that we are capturing its effects instead.

Previous research has identified two other important changes to criminal justice in London – the introduction of a police force and changes in capital punishment. We seek evidence that it is the right to counsel and not these changes that we are identifying.

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

coverage:	±40 years	±40 years
years:	[1796-1876]	[1796-1876]
	[1]	[2]
Police x Treated	0.0131	-0.0150
	(0.0098)	(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
N	135,363	135,363

Table 3.4: Metro Police

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 (81 clusters); *** 1%, ** 5%, * 10% level of significance.

under 1000 officers of varying rank, divided into divisions and tasked with patrolling the city's seventeen newly defined police territories (Emsley, 1991).[41] 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 indicator variable for the treated cases, and add it to equation (1). Table 3.4 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

[41] See Davies (2002) for a further discussion of private policing in London.

	Baseline	*Alternative Time Windows*		
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	-0.1250	-0.0409	-0.1217
	(0.1328)	(0.1411)	(0.1522)	(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

Table 3.5: Triple Difference

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.

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.

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

$$\begin{aligned} Conviction_{icy} = {} & \gamma_1 Post_{iy} \times Treated_{icy} \times NotDeathEligible_{icy} + \gamma_2 Post_{iy} \times Treated_{ic} \\ & + \gamma_3 Post_{iy} \times NotDeathEligible_{icy} + \gamma_4 Treated_{ic} \times NotDeathEligible_{icy} \\ & + \gamma_5 NotDeathEligible_{icy} + \tau_{iy} + \kappa_{ic} + X_{icy}\gamma + \epsilon_{icy}. \end{aligned} \tag{3.3}$$

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 3.5 presents the estimation results.

Consistently, the triple-difference coefficient is statistically insignificant. Alternatively, if the main specification presented previously includes a control variable for whether the specific crime currently is eligible for the death penalty, the size of the difference-in-difference coefficient is relatively unaffected and stays highly statistically significant (and the coefficient on the new control is statistically indistinguishable from zero).[42]

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. Consequently, the introduction of a police force and mitigation of capital punishment's use, while consequential, cannot explain away our finding that the introduction of the right to counsel corresponds to more guilty convictions.

3.6.3 Mechanism

It is instructive to explore, to the degree possible, how the institutional change leads to outcome changes; i.e, the mechanism creating the result. We first consider the argument that it is working through changes in jury behavior by looking at sentencing recommendations they provide to the judge. Second, we differentiate cases with a defense counsel presence to those without to explore whether it is the destructive actions of the solicitors or truly an institutional change.

3.6.3.1 Jury Mitigation

Our argument is that the increased conviction works through the jury responding to the lopsided nature of the proceedings prior to the change by avoiding convictions. Along with acquitting an individual, the English juries also had the freedom to convict an individual for lesser offenses and, additionally, were permitted to recommend mercy asking the judge to reduce the sanction's severity. If juries were, in fact, pulling back on their desire to offset the lopsided nature to trials, then these actions should also reduce in prevalence.

Here, we consider only those cases with a conviction and ask whether the rate at which the jury convicts for lesser offenses and the rate at which it suggests mercy change. For our baseline time period, conditional on conviction, lesser offenses arise in 7.3% of the cases and mercy is recommended in 11.1% of the convictions. Table 3.6 presents the results re-estimating equation (1), but limiting attention to convictions and considering these alternative dependent variables.

The introduction of the right to counsel corresponds to fewer convictions for lesser offenses and a reduction in the rate at which the jury recommends mercy. Similar to the main result that convictions are more likely, juries cease to search for ways to reduce the punishment the defendant experiences when they are given the right to have defense counsel.

[42] The difference-in-difference coefficient is 0.0275 ($p < 0.001$) and the coefficient on death eligibility is 0.0036 ($p = 0.600$). Additionally, if 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.

	Convicted for a lesser offense [1]	Mercy recommended [2]
Post x Treated	-0.0653 ***	-0.0493 ***
	(0.0099)	(0.0152)
Crime Fixed Effects?	Yes	Yes
Year Fixed Effects?	Yes	Yes
Controls?	Yes	Yes
R^2	0.262	0.076
AIC	-12,718	36,958
DV μ	0.0734	0.1114

Table 3.6: Jury Mitigation

Only the sample of cases that resulted in a conviction considered; $N = 82,549$. 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 (81 clusters); *** 1%, ** 5%, * 10% level of significance.

3.6.3.2 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 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 3.1.

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:

$$\begin{aligned} Conviction_{icy} = {} & \delta_1 Post_{iy} \times Treated_{icy} \times Counsel_{icy} + \delta_2 Post_{iy} \times Treated_{ic} \\ & + \delta_3 Post_{iy} \times Counsel_{icy} + \delta_4 Treated_{ic} \times Counsel_{icy} \\ & + \delta_5 Counsel_{icy} + \tau_{iy} + \kappa_{ic} + X_{icy}\delta + \epsilon_{icy}. \end{aligned} \tag{3.4}$$

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 3.7 presents the estimation results.

	[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 ***	0.0137
	(0.0066)	(0.0114)
Counsel	-0.0849 ***	-0.0354 *
	(0.0051)	(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

Table 3.7: Use of Defense Counsel
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.

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 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.

3.7 Text Analysis

The analysis of the mechanism relies on easily quantifiable, but coarse, measurements of the Old Bailey trials. We do have access to the wealth of text from these trials. We now turn to an exploratory analysis

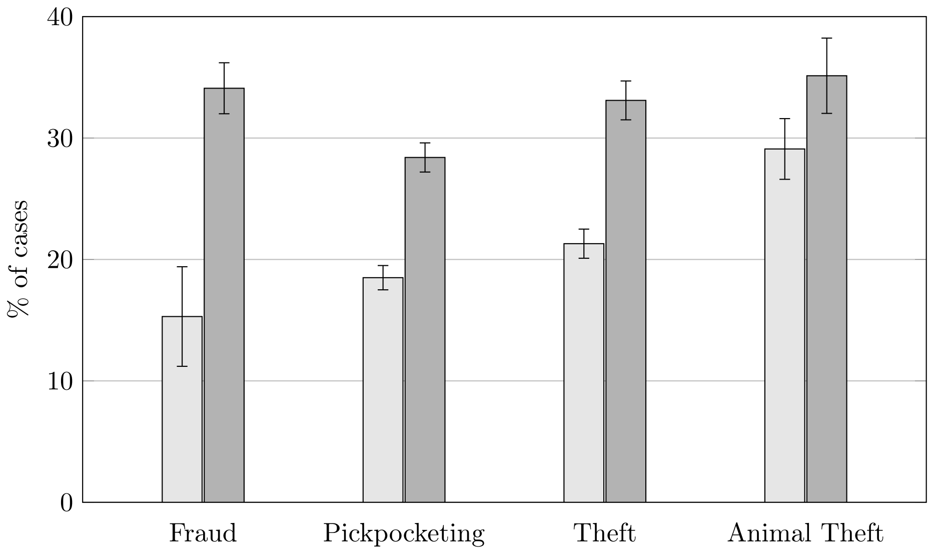


Figure 3.5: Before and After the 1836 Act

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.

of these transcripts.

Our objective is to search for changes in the text of the trial proceedings that correspond to the institutional change. We expect this will provide additional information on the mechanism. We do not *ex ante* have a hypothesis about which words, phrases, questions, etc. should be more or less likely to arise. We expect that the formalization of the criminal courtroom, provided by the Prisoner's Counsel Act, should create alterations in what information is solicited from witnesses, the defendant, and the accuser. This should be measurable in a text analysis. To do this, we first want to zone in on particular crime categories where the Act has had the greatest impact.

As previously described, we record the prevalence of defense counsel by identifying the existence of related keywords in the text depicted in Figure 3.1 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, theft[43], and animal theft. Figure 3.5 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 3.5 to exhibit large changes in the presence of defense counsel and investigate the trial transcript's text for each subset.

[43]Specifically, the crime "theft from place", which is distinct from petty larceny, simple larceny, grand larceny, and "stealing from master".

To achieve our goal of using text to identify mechanisms driving our main result, we use a popular topic modeling method in text analysis. What is important in this context is that we use a method which does not require *a priori* researcher knowledge of which words, phrases/n-grams, etc. to identify. Our objective is to uncover the "themes"/"topics" covered within each trial. While topic modeling methods are subject to interpretation, we use it to provide depth to the causal analysis already conducted.

3.7.1 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,

$$\theta = \begin{pmatrix} \theta_1 \\ \vdots \\ \theta_D \end{pmatrix} = \begin{pmatrix} P(t=1|d=1) & \cdots & P(t=T|d=1) \\ \vdots & \ddots & \vdots \\ P(t=1|d=D) & \ldots & P(t=T|d=D) \end{pmatrix}. \tag{3.5}$$

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,

$$\phi = \begin{pmatrix} \phi_1 & \cdots & \phi_T \end{pmatrix} = \begin{pmatrix} P(w=1|t=1) & \cdots & P(w=1|t=T) \\ \vdots & \ddots & \vdots \\ P(w=V|t=1) & \ldots & P(w=V|t=T) \end{pmatrix}. \tag{3.6}$$

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 its 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

$$\prod_{d=1}^{D} P(\theta_d|\alpha) \left(\prod_{n=1}^{N_d} \sum_{z_{d,n}} P(z_{d,n}|\theta_d) P(w_{d,n}|z_{d,n}, \phi) \right). \tag{3.7}$$

Gibbs sampling is used to estimate the conditional probabilities that best explain the corpus of documents.[44]

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.

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.[45] 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. Bandiera *et al.* (2019) use it in their analysis of CEO behavior. They use LDA to classify the types of activities CEOs engage in from diaries chronicling their work week. 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 regresssion.[46]

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.

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

[45] Edison and Carcel (2021) also analyzes FOMC texts using LDA.

[46] See 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.

Figure 3.6: *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.

3.7.2 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 3.6 represents the time series of Topic 1's prevalence. In Figure 3.6, 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 3.8 lists selected words that have a large probability of arising in a document that involves Topic 1.[47] In Table 3.8 we rank the words according to their probabilities derived from the LDA estimation.

[47]We do not report common words such as "said" and "man" which arise with a large probability in most topics.

	Fraud: Mail	*Other 9 Topics*	
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

Table 3.8: Differentiating Words in *Fraud: Mail*

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.

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*. 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 become 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.[48] 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 3.9 provides the selected words that dominate this topic for each of the four crimes.

[48] In 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.

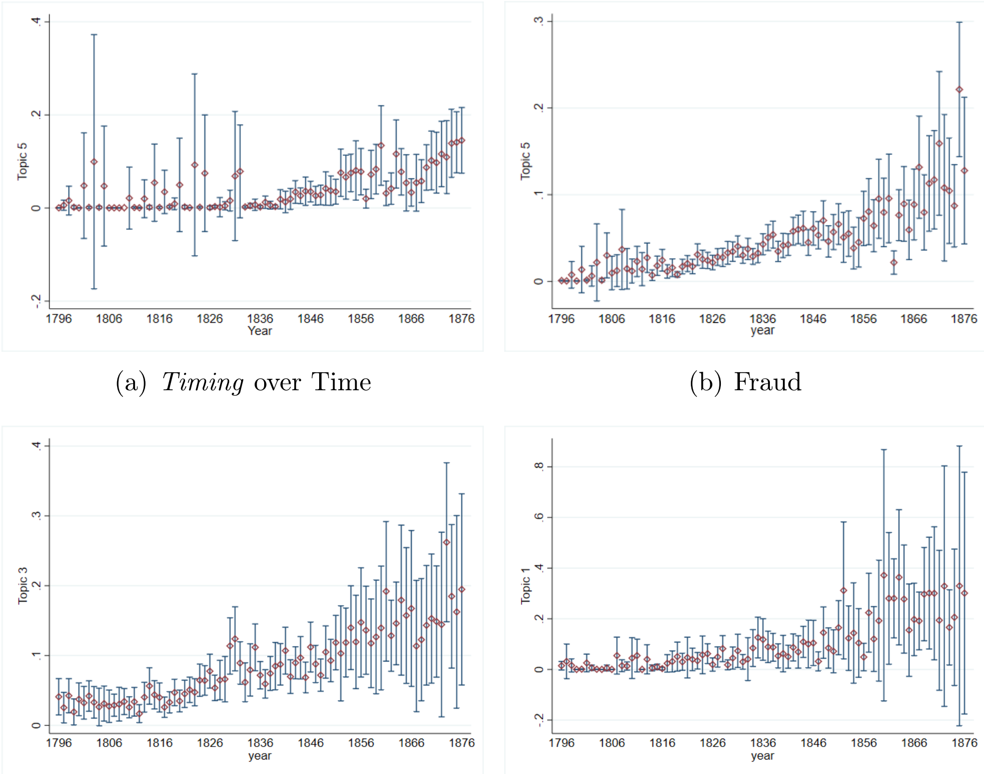
Fraud			*Pickpocketing*		
word	rank	avg rank (other 9)	word	rank	avg rank (other 9)
time	2	11.9	o'clock	5	17.2
day	17	23.4	night	15	102.0
never	18	15.3	time	16	81.2
last	41	39.78	morning	17	133.3
afterwards	50	44.6	never	23	69.4
date	51	342.8	day	30	115.0
months	52	56.1	minutes	31	61.2
since	53	101.9	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

Theft			*Animal Theft*		
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

Table 3.9: *Timing* Across Crimes

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.

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 3.7 depicts the time series of each.



(a) *Timing* over Time

(b) Fraud

(c) Pickpocketing

(d) Animal Theft

Figure 3.7: *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.

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

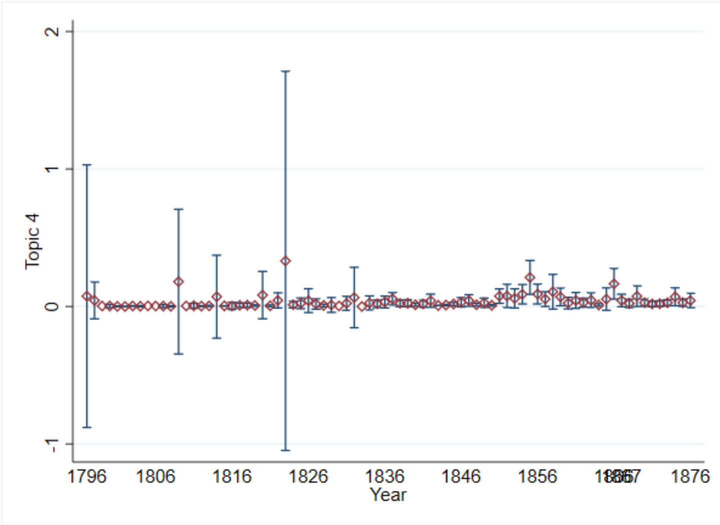


Figure 3.8: *Timing* in Misdemeanor Cases

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.

related to the timing of events.[49] Figure 3.8 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 the time trend changes course after 1836. To that end, for each topic we estimate the following model:

$$Topic_{iy} = \zeta_0 + \zeta_1 Post_{iy} + \zeta_2 Year_{iy} + \zeta_3 Post_{iy} \times Year_{iy} + \epsilon_{iy}. \tag{3.8}$$

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 $\widehat{\zeta}_1$ will be positive and large but $\widehat{\zeta}_3$ will be expected to be equal to zero. If, on the other hand, the effect of the Act grows gradually over time then $\widehat{\zeta}_3$ will be positive. Table 3.10 presents the results for each of the four crime category subsamples.

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

crime:	Fraud [1]	Pickpocketing [2]	Theft [3]	Animal Theft [4]	Control Crimes [5]
constant ($\widehat{\zeta_0}$)	0.00777	-1.67518 *** ***	-3.09613 ***	-3.37149 ***	-0.04080
	(1.30434)	(0.31958)	(0.39133)	(0.84928)	(1.55630)
Post ($\widehat{\zeta_1}$)	-5.28525 ***	-1.58490 ***	-2.26303 ***	-7.47237 ***	-1.25927
	(1.41390)	(0.43390)	(0.67407)	(1.40535)	(1.65438)
Year ($\widehat{\zeta_2}$)	0.00001	0.00093 ***	0.00173 ***	0.00188 ***	0.00004
	(0.00072)	(0.00081)	(0.67407)	(1.40535)	(0.00085)
Post x Year ($\widehat{\zeta_3}$)	0.00287 ***	0.01107 ***	0.00086 ***	0.00406 ***	0.00068
	(0.00077)	(0.00024)	(0.00037)	(0.00076)	(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

Table 3.10: Break in Time Trends for *Timing*

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

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.[50] 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 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.

3.8 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

[50]The 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.

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.[51]

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 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.

3.9 References

[51] This 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).

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

[2] 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.

[3] 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.

[4] 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.

[5] Anderson, James M. and Paul Heaton (2012), How Much Difference Does the Lawyer Make? The Effect of Defense Counsel on Murder Case Outcomes, *Yale Law Journal* 122: 154-217.

[6] Ater, Itai, Yehonatan Givati, and Oren Rigbi (2017), The Economics of Rights: Does the Right to Counsel Increase Crime?, *American Economic Journal: Economic Policy* 9(2): 1-37.

[7] Bandiera, Oriana, Andrea Prat, Stephen Hansen, Raffaella Sadun (2020), CEO Behavior and Firm Performance, *Journal of Political Economy* 128(4): 13325-1369.

[8] 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.

[9] 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.

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

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

[12] 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.

[13] 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.

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

[15] 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.

[16] 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.

[17] Blei, David, Ng, Andrew Y., and Jordan, Michael I. (2003), Latent Dirichlet Allocation, *Journal of Machine Learning Research* 3(1): 993-1022.

[18] 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.

[19] Cohen, Thomas H. (2012), Who is Better at Defending Criminals? Does Type of Defense Attorney Matter in Terms of Producing Favorable Case Outcomes?, *Criminal Justice Policy Review* 25(1): 29-58.

[20] 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.

[21] Dickenson, William and Talfourd, Thomas Noon (1829), *Practical Guide to the Quarter Sessions and Other Sessions of the Peace*, London: S. Sweet.

[22] Duggan, John and Martinelli, Cesar (2001), A Bayesian Model of Voting in Juries, *Games and Economic Behavior* 37(2): 259-294.

[23] 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.

[24] Edison, Hali and Carcel, Hector (2021), Text Data Analysis Using Latent Dirichlet Allocation: Application to FOMC Transcripts, *Applied Economics Letters* 28(1): 38-42.

[25] Emsley, Clive (1991), *The English Police: A Political and Social History*, London: Harvester Wheatsheaf.

[26] 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.

[27] Farmer, Amy and Jill Tiefenthaler (2001), Conflict in Divorce Disputes: The Determinants of Pretrial Settlement, *International Review of Law and Economics* 21(2): 157-180.

[28] 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.

[29] Flanagan, Francis X. (2018), Race, Gender, and Juries: Evidence from North Carolina, *Journal of Law and Economics* 61(2): 189-214.

[30] Friedman, Ezra and Wickelgren, Abraham L. (2006), Bayesian Juries and the Limits to Deterrence, *Journal of Law, Economics, and Organization* 22(1): 70-86.

[31] Gallanis, T.P. (2006), The Mystery of Old Bailey Counsel, *Cambridge Law Journal* 65(1): 159-173.

[32] 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.

[33] Greiner, D. James, Ellen Lee Degnan, Thomas Ferriss, and Roseanna Sommers (2021), Using Random Assignment to Measure Court Accessibility for Low-Income Divorce Seekers, *Proceedings of the National Academy of Science* 118(14): e2009086.

[34] Greiner, D. James and Cassandra Wolos Pattanayak (2012), Randomized Evaluation in Legal Assistance: What Difference Does Representation (Offer and Actual Use) Make?, *Yale Law Journal* 121(8): 2118-2214.

[35] Griffiths, Cerian Charlotte (2014), The Prisoners' Counsel Act 1836: Doctrine, Advocacy and the Criminal Trial, *Law, Crime and History* 2: 28-47.

[36] Grimmer, Justin (2010), A Bayesian Hierarchical Topic Model for Political Texts: Measuring Expressed Agendas in Senate Press Releases, *Political Analysis* 18(1): 1-35.

[37] 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.

[38] Halliday, Stephen (2007), *Newgate: London's Prototype of Hell*, The History Press.

[39] Hanly, Conor (2021), Jury Selection in Victorian England, *American Journal of Legal History* 61(1): 36-55.

[40] 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.

[41] Hoekstra, Mark and Street, Brittany (2021), The Effect of Own-Gender Jurors on Conviction Rates, *Journal of Law and Economics*, forthcoming.

[42] Jones, Mark and Johnstone, Peter (2012), *History of Criminal Justice*, New York: Routledge.

[43] Koyama, Mark (2012), Prosecution Associations in Industrial Revolution England: Private Providers of Public Goods?, *Journal of Law and Economics* 41: 95-130.

[44] Langbein, John H. (1978), The Criminal Trial before the Lawyers, *University of Chicago Law Review* 45(2): 263-316.

[45] 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.

[46] 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.

[47] Larsen, Vegard H. and Thorsrud, Leif A. (2019), The Value of News for Economic Development, *Journal of Econometrics* 210(1): 203-218.

[48] Maitland, F.W. (1909), The Forms of Action in Common Law, 1909, in Paul Haulsall, ed. (2021), *Internet Medieval Source Book*, Fordham University Center for Medieval Studies, https://sourcebooks.fordham.edu/basis/maitland-formsofaction.asp.

[49] McCannon, Bryan C. (2020), Wine Descriptions Provide Information, *Journal of Wine Economics* 15(1): 203-218.

[50] Nowak, Adam and Smith, Patrick (2017), Textual Analysis in Real Estate, *Journal of Applied Econometrics* 32(4): 896-918.

[51] 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.

[52] Roach, Michael A. (2014), Indigent Defense Counsel, Attorney Quality, and Defense Outcomes, *American Law and Economics Review* 16(2): 577-619.

[53] Roach, Michael A. (2017), Does Raising Indigent Defender Pay Rates Improve Defendant Outcomes?, *Applied Economics Letters* 23(14): 1025-1030.

[54] Rubin, Donald B. (1978), Bayesian Inference for Causal Effects: The Role of Randomization, *Annals of Statistics* 6(1): 34-58.

[55] 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.

[56] Shem-Tov, Yotam (2021), Make-or-Buy? The Provision of Indigent Defense Services in the U.S., *Review of Economics and Statistics*, forthcoming.

[57] 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/

[58] Starr, Sonja (2015), Estimating Gender Disparities in Federal Criminal Cases, *American Law and Economics Review* 17(1): 127-159.

[59] Tirunillai, Seshadri and Tellis, Gerard J. (2014), Mining Marketing Meaning from Online Chatter:

Strategic Brand Analysis of Big Date Using Latent Dirichlet Allocation, *Journal of Marketing Research* 51(4): 463-479.

[60] Voth, Hans-Joachim (1997), Time and Work in Eighteenth-Century London, *Journal of Economic History* 58(1): 29-58.

[61] Wontner, Thomas (1833), *Old Bailey Experience*, reprint (2018), Forgotten Books.

[62] 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.

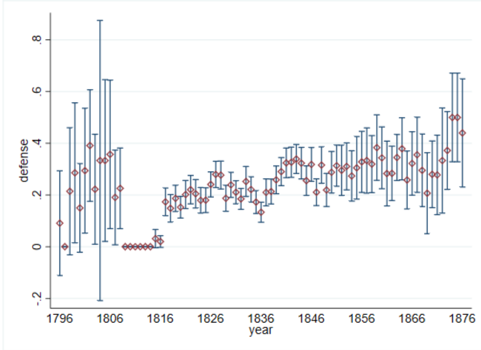
[63] Wade, Stephen (2009), *Britain's Most Notorious Hangmen*, South Yorkshire: Wharncliffe.

[64] 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.

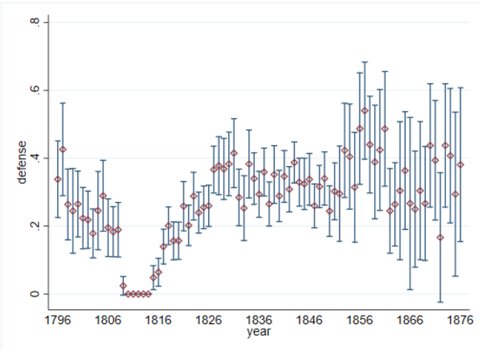
3.10 Appendix

The time series for the presence of defense counsel in the four selected crimes are presented in Figure 3.9. The time series for the other topics not presented in the main body, for each of the four crimes, are provided in Figures 3.9(a), 3.13(c), 3.9(c), 3.9(d).

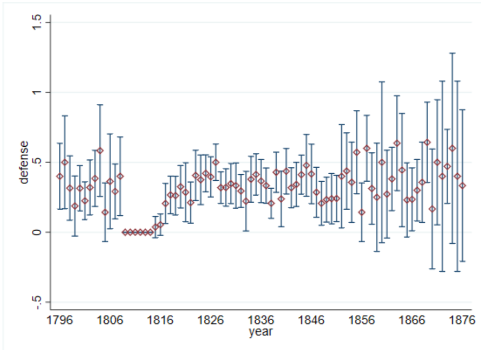
(a) Fraud



(b) Pickpocketing

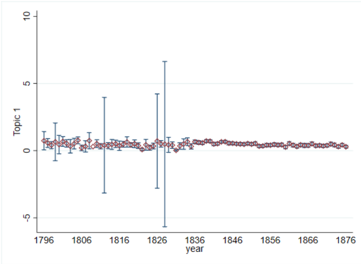


(c) Theft

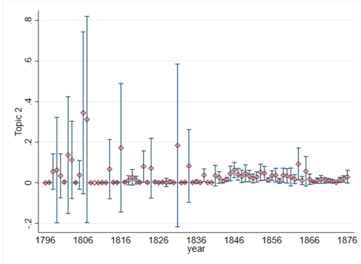


(d) Animal Theft

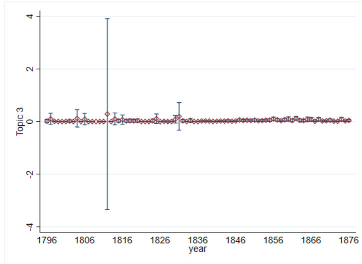
Figure 3.9: Defense Counsel Presence



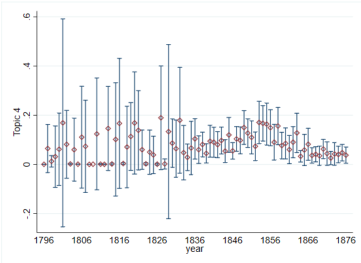
(a) Topic 1



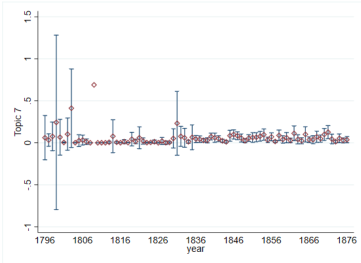
(b) Topic 2



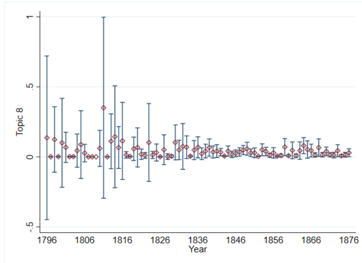
(c) Topic 3



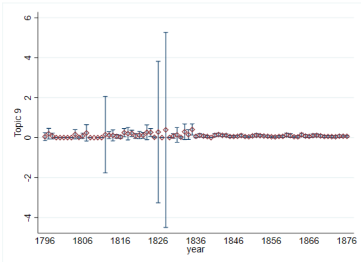
(d) Topic 4



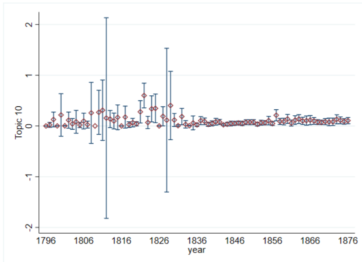
(e) Topic 7



(f) Topic 8



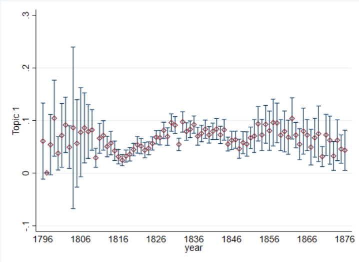
(g) Topic 9



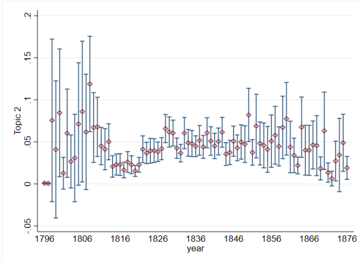
(h) Topic 10

Figure 3.10: Fraud Topics Over Time

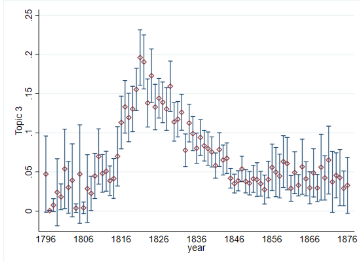
The topic numbered "Topic 5" is *Timing* and "Topic 6" is *Mail Fraud*. They are presented in the text.



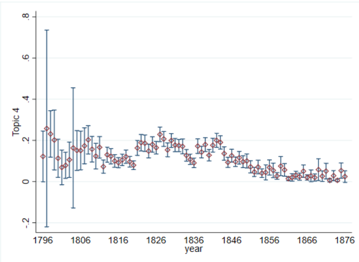
(a) Topic 1



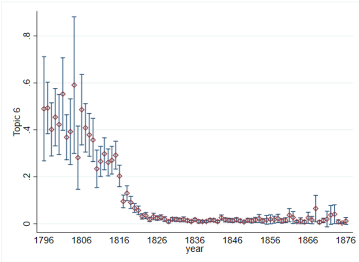
(b) Topic 2



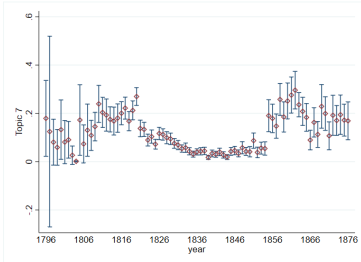
(c) Topic 3



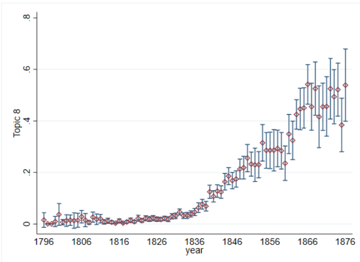
(d) Topic 4



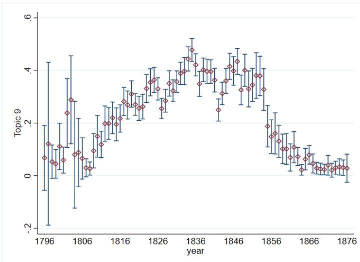
(e) Topic 6



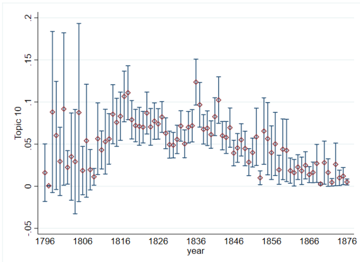
(f) Topic 7



(g) Topic 8



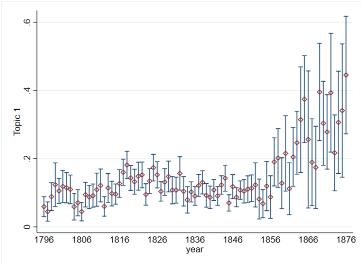
(h) Topic 9



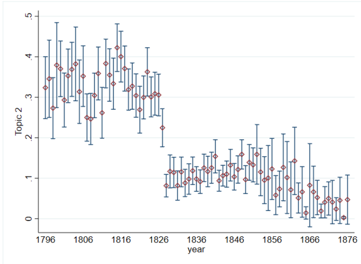
(i) Topic 10

Figure 3.11: Pickpocketing Topics Over Time

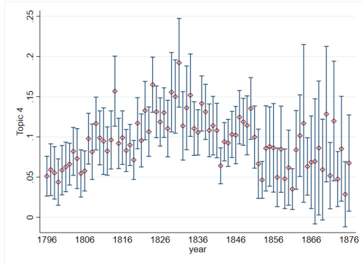
The topic numbered "Topic 5", *Timing*, is presented in the text.



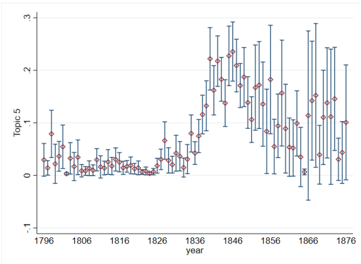
(a) Topic 1



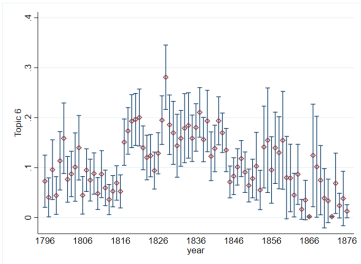
(b) Topic 2



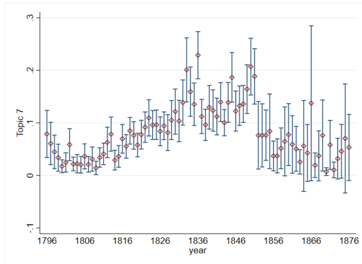
(c) Topic 4



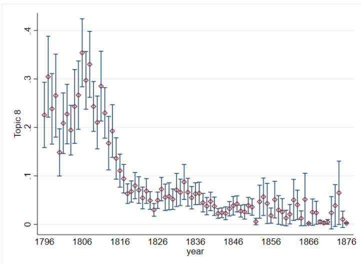
(d) Topic 5



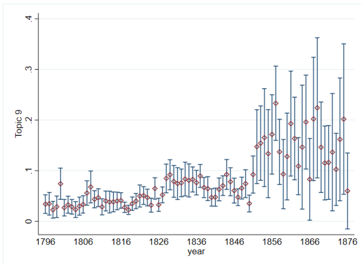
(e) Topic 6



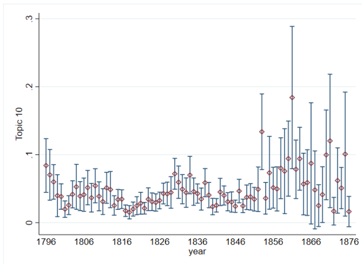
(f) Topic 7



(g) Topic 8



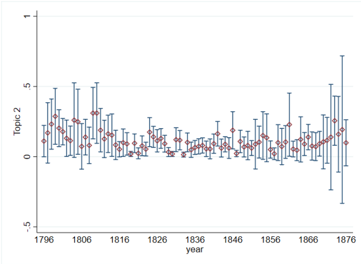
(h) Topic 9



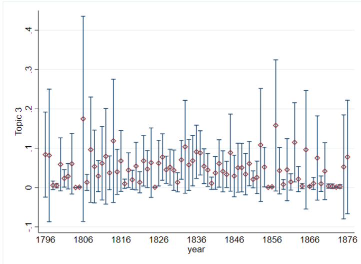
(i) Topic 10

Figure 3.12: Theft Topics Over Time

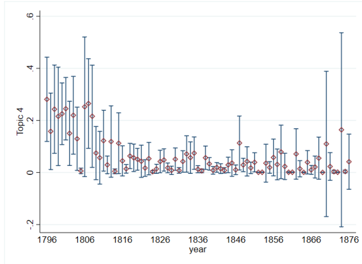
The topic numbered "Topic 3", *Timing*, is presented in the text.



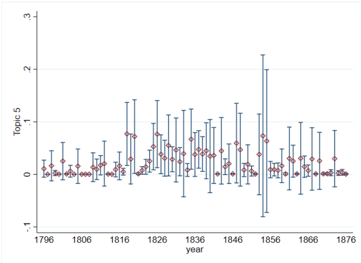
(a) Topic 2



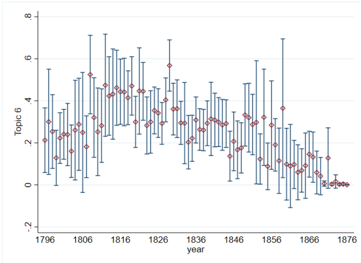
(b) Topic 3



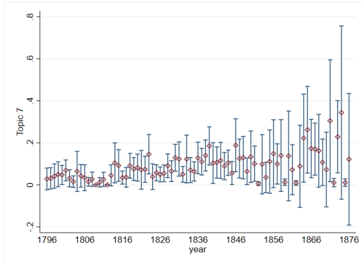
(c) Topic 4



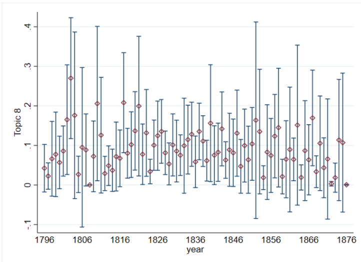
(d) Topic 5



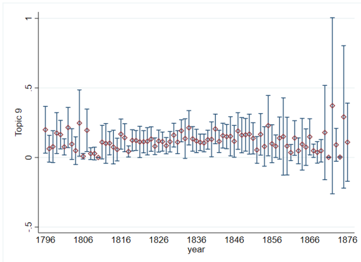
(e) Topic 6



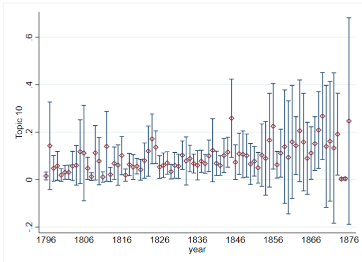
(f) Topic 7



(g) Topic 8



(h) Topic 9



(i) Topic 10

Figure 3.13: Animal Theft Topics Over Time

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