THE IMPACT OF THE FIRST PROFESSIONAL POLICE FORCES ON CRIME

Anna Bindler
University of Cologne and
University of Gothenburg

Randi Hjalmarsson
University of Gothenburg

Abstract
This paper evaluates the effect on crime of creating a fundamental modern-day institution: centralized professional police forces tasked with preventing crime. We study the 1829 formation of the London Metropolitan Police—the first professional force worldwide. Using newly digitized and geocoded crime and police data together with difference-in-differences and pre–post designs, we find evidence of a significant reduction in violent crimes (despite the possibility of offsetting increases in clearance and reporting rates). In contrast, a reduction in property crime is not visible. (JEL: K42, N93, H0)

Teaching Slides
A set of Teaching Slides to accompany this article are available online as Supplementary Data.

The test of police efficiency is the absence of crime and disorder, not the visible evidence of police action in dealing with them.
—Sir Robert Peel’s 9th Principle of Law Enforcement (1829)

Acknowledgments: This paper would not have been possible without the tremendous effort of our research assistants Srinidhi Srinivasan and Vu Tran, and the financial support of Vetenskapsrådet, The Swedish Research Council, Grants for Distinguished Young Researchers (VR 2014-01735). Dr Bindler is funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) under Germany’s Excellence Strategy—EXC 2126/1-390838866. We thank seminar and workshop participants at the University of Gothenburg, VATT Institute for Economics Research, Uppsala, University of Bologna, NBER Summer Institute, Transatlantic Workshop of Crime (Science Po Paris), IZA Bonn, CEP Workshop on Crime & Policing, Berlin’s Applied Micro Seminar, German Economists Abroad Christmas Meetings (Bonn), University of Cologne, University of Lund, RES Annual Conference (Warwick), SOLE Annual Meetings (Arlington), Brunel University (Crime Workshop), and the Warwick Applied Workshop for helpful comments and discussion. We are also very grateful for the many helpful comments from the editor and referees provided in the course of revising this paper.

E-mail: bindler@wiso.uni-koeln.de (Bindler); randi.hjalmarsson@economics.gu.se (Hjalmarsson)

Journal of the European Economic Association 2021 19(6):3063–3103 DOI: 10.1093/jeea/jvab011 © The Author(s) 2021. Published by Oxford University Press on behalf of European Economic Association. This is an Open Access article distributed under the terms of the Creative Commons Attribution License (http://creativecommons.org/licenses/by/4.0/), which permits unrestricted reuse, distribution, and reproduction in any medium, provided the original work is properly cited.
1. Introduction

Professional police forces are a fundamental component of criminal justice systems today. The origins of this modern-day institution date to 1829, when Sir Robert Peel founded the first professional force in the world—the London Metropolitan Police (the “Met”). Many features distinguished this new force from the decentralized and community organized policing of the past, including a substantial increase in officer numbers, full-time salaried employees, accountability to a centralized government authority, an emphasis on officer quality, and equal policing independent of community wealth. One of the most innovative features of the Met was a shift from reactionary policing focused on catching criminals to tasking officers—for the first time ever—with *deterring* crime by slowly walking a small beat and being visibly present (Emsley 2009). The Met subsequently served as a model for forces around the world.1

Yet, despite the historical (and contemporary) importance of this institution, the impact of creating the Met has never been empirically evaluated.2 Using newly digitized archival data sources together with quasi-experimental research designs, this paper provides the first empirical evidence on the causal effect of the formation of the Met on crime. We make two key contributions to the current understanding of the police–crime relationship. The first is to study old questions—the deterrence effect of more (visible) police—in a new context, that is, the creation of a force rather than its expansion. The second emphasizes little studied aspects of policing (force administration and quality).

Specifically, two broad conclusions have been reached by the many empirical papers studying the deterrence effects of police—a la Becker’s (1968) economic model of crime. The first, based largely on evaluations of (often temporary) expansions to existing police forces, is that increases in manpower reduce crime (Chalfin and McCrary 2017).3 The second, seen both in studies that rely on temporal variation in police deployment (e.g. Di Tella and Schargosky 2004; Draca et al. 2011) and in studies that rely on spatial variation generated by geographic boundaries (MacDonald et al. 2015; Heaton et al. 2016), is that visible police presence reduces crime.4 Both

---

1. Police forces of the United States were established in New York City (1845), New Orleans and Cincinnati (1852), Boston and Philadelphia (1854), Chicago and Milwaukee (1855), and Baltimore and Newark (1857). See Uchida (2015) and https://www.britannica.com/topic/police/Early-police-in-the-United-States (accessed October 22, 2018).

2. Historical crime in England has been studied by Wong (1995), who emphasizes opportunities for legal and illegal gains from 1857 to 1892, and Wolpin (1978), who looks at the role of clearance, conviction, imprisonment, and fines, as well as average sentences from 1894 to 1967. See Bignon et al. (2017), Mehlum et al. (2006), and Traxler and Burhop (2010) for studies of historical crime in France, Bavaria, and Prussia, respectively.

3. For specific examples, see, for instance, Evans and Owens (2007), Mello (2019), and Weisburd (2019).

4. One exception is Blanes I Vidal and Mastrobuoni’s (2018) study of non-terrorist attack related temporary patrol increases. Other deployment studies include Klick and Tabarrok (2005) and Blanes I Vidal and Kirchmaier (2018).
of these commonly evaluated features of policing today—manpower and visibility—were fundamental to the newly founded Met. Do the crime deterring effects of more, visible police generalize to this different margin of police force expansion: the first recruits? The answer can shed light on whether officer effectiveness depends on the institutional structures in place and whether officer visibility deters crime, even when forces are smaller and less experienced.

In contrast, “there is relatively scant evidence on the extent to which non-deployment related police policies reduce crime” (Owens 2020). Two dimensions particularly salient in the context of the Met are the public versus private provision of policing and monitoring officer quality. Cheng and Long’s (2018) study of the New Orleans French Quarter Task Force is one of the few papers studying this question; it finds that monitoring and performance incentives under private provision decreased crime. In this context, our paper makes an important contribution: It evaluates the effect of replacing decentralized, local community organized policing with a centralized, public force in which officers of poor quality (e.g. shirking, inappropriate on-the-job behavior, and corruption) were quickly dismissed.

Finally, this paper contributes to a literature beyond the economics of crime—namely that studying state capacity. While much of this literature takes a macro-perspective, our paper uses micro-data to study the institution of policing—an institution that is after all fundamental to a government’s ability to implement and enforce its intended policies and laws. Moreover, studying the extent to which such a major institutional reform was successful sheds light on the degree of state capacity in early-19th century England.

We face three fundamental challenges to identifying the effect of the newly created Met on crime. The first is specific to the historical context: It is not easy to measure crime before the institution designed to monitor and systematically record it is created. We overcome this by digitizing multiple “new” sources of archival crime and police data. The remaining challenges are not context-specific: Even if crime behavior does not change, crime statistics could change after the formation of the Met due to changes in reporting behavior and/or clearance rates. The next section discusses these issues in more depth. To the extent that one believes the Met likely to increase both reporting and clearance rates, these potential confounders would work against us finding a deterrence effect of the Met.

Our analysis relies on two sources of London crime data. The first is felony “trial” data for selected offenses—burglary, homicide, and robbery—from London’s Old

5. Other studies have considered the temporary destruction of a police force, including the extensive margin effects of police strikes (Pfuhl 1983) and slowdowns (Cann Chandrasekher 2016).

6. Other studies of private provision include business improvement districts (Brooks 2008) and university policing (MacDonald et al. 2015; Heaton et al. 2016). Banerjee et al. (2021) study incentives, community involvement, training, and managerial autonomy for police in India. Another dimension of policing that is increasingly being studied is the role of diversification; see, for instance, Linos (2018) and Miller and Segal (2019).

7. See e.g. Acemoglu et al. (2015), Besley and Persson (2010), Dencecco and Katz (2014), or Johnson and Koyama (2017).
Bailey Central Criminal Court; we include all cases, regardless of trial outcome. These data have three key advantages. First, we can geocode offenses and assign them to “treated” and “control” areas of London: With the exception of the “City” of London (a 1 square mile area in Central London’s business district, which still has its own force today), the initial catchment area of the Met was within an approximate 7-mile radius of Charing Cross. To implement a difference-in-differences design, we thus aggregate individual crimes into a panel data set of the number of crimes in each month and area from 1821 to 1837. A second advantage is that trial transcripts contain details about police witnesses, which allows us to document reform implementation: There is an instant shift in the type of police witness (“old”/e.g. pre-Met watchman to “new”/Met policeman) that is by far largest in the treated area. Finally, the selection of these most serious offenses limits concerns about biases in our estimates due to reporting changes.

Yet, given the historical felony classification of many offenses that are misdemeanors today (e.g. pickpocketing and simple assault), these three offenses (homicide, robbery, and burglary) comprise just 7% of felony trials. Concerns about generalizability to a broader set of property and violent crimes are addressed by our second data source, which consists of daily police reports for nine (magistrate-run) Police Offices. These offices were tasked—both pre- and post-Met—with processing crimes and apprehending offenders for all of London, including crimes outside the Met’s catchment area; constables in these offices were not tasked with deterrence in either period. The daily reports include three types of crime measures: stolen property incident reports, “information” (another type of incident report), and charges. Though all property, violent, and other offenses are included in these reports (and not just selected felonies), these data also come with limitations. Offenses cannot be geocoded into treated and control areas, necessitating pre–post designs that estimate the effect of creating the Met on crime in all of London (i.e. net of any offsetting effects in the control areas).

Analyses of both London data sources provide evidence consistent with a crime-reducing effect, especially for violent crimes. In the Old Bailey data, a significant and persistent reduction in trials is seen for the violent crime of robbery (more than 40%), but there are no (consistently) significant effects for homicide and burglary trials. These results are robust to specification choice [ordinary least-squares (OLS) versus Poisson], sample windows, and the units of temporal and geographic aggregation. We provide evidence in support of the underlying parallel trend assumption, and rule out biases due to a number of potential confounders (especially in a short window around the reform), including the 1832 cholera epidemic, contemporaneous criminal justice changes, population growth, police spillovers, and crime displacement. The lack of an observable crime-reducing effect for the property crime of burglary can be due to there actually being no deterrence effect or crime-reducing effects being offset by increases in clearance and/or reporting rates, leading to more trials. The latter is consistent with the findings of the daily police report analysis: A significant decrease is seen for all violent crime measures, but for property crimes, reductions in incidents are more than offset by increases in charges.

Finally, we assess whether the crime-reducing effects of new forces are specific to the Met by evaluating the 1839–1856 rollout of police forces to the counties of...
England and Wales. Taking advantage of variation in initial force characteristics, we use difference-in-differences models to estimate the effect of creating professional forces overall and by force size. Though there is no overall crime-reducing effect of creating just any professional police force, we find that the largest forces per capita significantly reduced both violent and property crimes; event study specifications show these effects are not immediate and increase over time. These results highlight the potential importance of force quality, and are in fact consistent with an additional finding of the London pre-post analyses—namely that effects increase over time.

The remainder of the paper is organized as follows. Section 2 highlights the differences between the pre- and post-Met “police” and discusses the various channels through which the formation of the Met can impact crime and crime statistics. Sections 3 and 4 estimate the effect of the Met using the Old Bailey geocoded data and the daily report data, respectively. Section 5 evaluates the rollout of county forces throughout England and Wales. Section 6 concludes.

2. Institutional Background and Framework

2.1. “Policing” before the Met

This section describes the forms of public and private “policing” that existed in London before the 1829 Metropolitan Police Act. Local watchmen, tasked with guarding the streets at night, date to the 1285 Statute of Winchester, which required all towns to form a watch manned by the local residents (Rawlings 2008). The unpopularity of this community policing system—neighbors watching out for (but also arresting) neighbors—led to a series of Watch Acts in the 1700s, “in which households exchanged the duty to serve in the watch for the duty to pay a watch rate” to hire substitutes (Rawlings 2008). Watchmen initially stood guard in a fixed location, but by 1800, the better watches aimed to prevent crime by patrolling the streets at set intervals and yelling out the time on the hour (Reynolds 1998). Because they were locally funded, watch quality varied substantially with parish wealth. \(^8\) Just three wealthy parishes (St Marylebone, St James, and St George) had superior watches, with larger numbers of relatively well paid watchmen patrolling in multiple night shifts (Reynolds 1998). By the early 1800s, the many criticisms of the local watches included incompetence, improper and corrupt conduct, responsibility to only their local parish, and an inability to cope with major disturbances, such as riots and large crowds, without military support (Rawlings 2008). \(^9\)

---

8. There are also examples of prosperous neighbors paying private watchmen to guard gates (Rawlings 2008).

9. The five days of rioting in the 1780 Gordon Riots are a prime example. Calling in the army was delayed given the fear it would not arrive in time and would further incite unrest (Rawlings 2008). The Gordon Riots led to some calls for police reform and more unified policing, but these failed and were eventually forgotten (Lyman 1964).
Another form of (public) policing was London’s Bow Street Runners, who were sworn constables of Westminster and date to around 1750 (Emsley 2009). As there were initially only eight Runners, they did not have a physical presence and were not meant to deter crime but rather to clear it by locating and arresting serious offenders. Initially, these Runners were similar to the 18th century thief-takers, that is, men who earned livings from rewards upon the convictions of serious criminals. By the end of the 1700s, however, the Runners were essentially full-time policemen located at the Bow Street house, which became a centralized collection point of crime incidents. Other Bow Street forces were created: The Horse Patrol (1805) on the highways leading to London, the Foot Patrol (1790) in central London, and the Dismounted Horse Patrol beyond these jurisdictions. Both horse patrols remained following the creation of the Met, but the Foot Patrol was disbanded. Rawlings (2008) estimates that these patrols comprised about 400 persons when the Met was created.

Following the Bow Street template, seven additional Police Offices were established in the Middlesex Justices Act of 1792: Queen’s Square, Great Marlborough Street, Worship Street, Lambeth Street, Shadwell (replaced by Marylebone High Street by 1816), Union Hall, and Hatton Garden. A Thames River police was established in 1798. Each office was staffed by three magistrates and up to twelve constables, whose primary purpose was to follow up on crime reports (Emsley 2009). According to Reynolds (1998), “these men did not patrol the streets on any regular basis”. In contrast to the local watches, the jurisdictions of the Police Offices were not confined by parish boundaries (Rawlings 2008). The magistrates’ main responsibility was to process criminal cases. They retained this judiciary role upon the formation of the Met (Davis 1984). That is, even post-Met, a known offender would be processed through these same offices (which existed until 1839) and recorded in the offices’ daily reports.

The pre-Met police were poorly paid. Office constables were paid 21 shillings per week (55 pounds per year) in 1811 (Radzinowicz 1953), while 1820s watchman salaries ranged from 12 to 16 shillings per week (Reynolds 1998). These low salaries led to a system with additional sources of income: systematic schedules of charges for all “extras” (e.g. witness testimony), rewards for catching criminals, and corrupt behavior. Such a system inherently does not emphasize deterrence: No one can pay a reward for a crime that does not occur. A telling summary of the perception of the pre-Met Police Offices is provided in 1850 by Charles Dickens (Radzinowicz 1953): “We are not by any means devout believers in the Old Bow Street Police. […] as a Preventive Police they were utterly ineffective, and as a Detective Police were very loose and uncertain in their operations, … .”

The weaknesses of the decentralized, parish-wealth-dependent, non-cooperative, poorly paid, too small for mass crowd control, and oftentimes-corrupt policing agents were not unrecognized before the creation of the Met. But, all earlier calls and attempts to reform (unify) policing encountered strong resistance by rich parishes—and

---

10. With respect to corruption, Radzinowicz (1953) quotes witnesses examined during evaluations of pre-Met policing as saying an annual salary of 100–120 pounds is needed “to be above temptation and to do nothing mean.”
especially the City of London—who wanted to control their own watch and worried that any replacement would reduce local coverage to provide for the poorer locales (Rawlings 2008).

2.2. The Introduction of the Met

The Metropolitan Police Act (10 Geo.4, c.44) created the London Metropolitan Police (the “Met”) on September 29, 1829.11 Though the Act states that property crime was rising, it also highlights the many previously recognized deficiencies in the existing “police”:

[... ] the local establishments of nightly watch and nightly police have been found inadequate to the prevention and detection of crime, by reason of the frequent unfitness of the individuals employed, the insufficiency of their number, the limited sphere of their authority, and their want of connection and co-operation with each other [...].

Why did this Act pass, while other reform attempts failed? This perhaps comes down to Sir Robert Peel, who became Secretary of the Home Office in 1822. Peel emphasized legal reforms, including (i) the consolidation of laws (1823), (ii) abolishing capital punishment (offense by offense), and (iii) improving prison conditions (1823). Peel learned how to negotiate the political landscape to get reforms through the Parliament; importantly, he convinced parishes that the new police would not cost more than localities paid before (Reynolds 1998).12

The new Met initially consisted of 1,000 men, but there were more than 3,000 officers by May 1830. Panel (a) of Figure 1 documents the weekly number of hires from 1829 until 1856. It shows that (i) a first wave of hiring occurred for six inner divisions in September 1829 and a second wave for eleven outer divisions in February 1830 (see also Online Appendix Table A.1) and (ii) the Met grew continuously to about 6,000 men in 1856. Each parish was notified of the date that the Met would take over, and the existing watch disbanded, by a notice posted on the parish’s church door two Sundays prior to that date (Reynolds 1998).

The initial catchment area was within an approximately 7-mile radius from Charing Cross in Central London and extended to 15 miles in 1839.13 Excluded were the City of London (which established its own force in 1832, which is still distinct today) and,

---

11. The bill to create the Met was introduced in April 1829, amended in May, and passed by both houses in June.
12. The Watch Rate was replaced with a Police Rate (equal to the parish Poor Rate, a property tax used to provide poor relief) on property holders, with a maximum of 8 pence per pound of annual property value (Lyman 1964).
13. While all descriptions of the formation of the Met describe this 7-mile radius, no explicit distance was written in the original Act. Rather, the Act includes a “List of the parishes, townships, precincts, and places constituting ‘The Metropolitan Police District’”. That list includes 88 parishes or places for which we geocoded the main point of interest (e.g. parish church); 85 lie within 7 miles from Charing Cross and all are within 8 miles. Moreover, 75% of the locations are within 4 miles. We test the robustness of our results to an 8-mile radius.
FIGURE 1. The Met—weekly hires and dismissals. Panel (a) shows the weekly number of total police, appointments as well as removals from the Met between 1829 (1830 for appointments and removals) and 1857. This figure is based on manually transcribed data from the Weekly State of the Metropolitan Police 1829–1857 available at the London National Archives (MEPO 4/1). Panel (b) shows the weekly number of leavers from the Met among those officers who were recruited between September 1829 and March 1831. The figure is based on transcribed data from the Register of recruits into the Met sourced from the London National Archives (MEPO 4/31 und 4/32). Panel (c) shows the weekly number of dismissals split up by detailed reason (drunkenness, neglect or misconduct, criminal behavior, and others). This figure is based on manually transcribed data from the Home Office: Police Entry Books, Series I. Metropolitan Police sourced from the London National Archives (HO 65/11, 65/12, and 65/13).
until 1839, the Thames River Police. Panel (a) of Figure 2 presents a historical map of the original jurisdiction of the Met. Panel (b) shows that the preexisting Police Offices were centrally located within the 7-mile radius (and even a smaller 4-mile radius). An equal number of police were hired into each division, regardless of the geographic size (see Online Appendix Table A.1). As beats in the outer divisions were often larger, it is plausible that police in these divisions were less visible on their beats (Emsley 2009). There is thus a potentially more intense treatment in a shorter radius around Charing Cross.

In addition to the sharp increase in the number of “police” and the centralized structure, many other changes characterized the new police, including their tasks, uniforms, incentive structure, the ability to police across parishes, and an emphasis on quality and proper behavior. First, the primary task of an officer was deterrence. Met officers walked a regular route at a pace of 2.5 miles per hour; the beat was intentionally small to increase visibility. The deterrence goal was also salient in the provision of uniforms that “caught the eye of criminals and the law-abiding public so both would know the police were on hand” (Reynolds 1998). There was little expectation that the new police would solve crimes; the Bow Street Runners remained until 1839, and were replaced with a Detective Division in 1842 (Reynolds 1998).

Second, while the new police received little formal training, there were clear behavioral guidelines, including Peel’s nine principles of law enforcement. The higher quality achieved via enforcing these rules is seen in historical records. Panel (b) of Figure 1 shows the weekly number of leavers among those recruited before March 1831, while Panel (c) shows the weekly number of post-1833 dismissals by reason (drunk, neglect/misconduct, criminal behavior, and others). There is high turnover, especially at the beginnings of the Met, including annual firing spikes for being drunk on duty around Christmas. Moreover, id numbers and division letters on uniform collars made it easier to identify officers behaving improperly (Reynolds 1998).

Who were the new police? Selection criteria included being at least 170 centimeters, under 35, physically and mentally fit, and able to read and write (Reynolds 1998). Were the new officers just the old “police” with a new job title? Though this could have happened to some extent, it would have been far from a one-to-one replacement: Peel’s new force consisted of more than 3,000 men, while Lyman (1964) estimates that there were 1,000 men in total (constables and watchmen) in 1800. While some of

---

14. Unfortunately, while we can measure the initial number hired per division, we cannot track officers per division over time.

15. On October 6, 1829, the Met opened its first station on a street called Great Scotland Yard (near 4 Whitehall Place and Charing Cross). This became the home station of the Met, including the two police commissioners. As the force expanded, other buildings in this area were taken over, but eventually the Met started opening police stations throughout London. To the best of our knowledge, few if any stations were opened in the 1830s.
FIGURE 2. The Met’s jurisdiction (1829). Panel (a) presents a map of the original Met district. Shaded in red is the City of London Police area, outside of the Met’s jurisdiction. The large letters indicate the various districts of the Met. The map is available from the British Library’s online map collection: http://www.bl.uk/onlinegallery/onlineex/crace/j/007000000000019u00055000.html. Panel (b) shows a map of London centered on Charing Cross, with the preexisting Police Offices indicated by blue squares and 4- (dashed), 7-, and 15-mile radii around Charing Cross in red. The borders represent modern-day postcode areas; the shapefiles were obtained from Maproom’s UK Postcodes Shapefiles and contain OS, Royal Mail, and National Statistics data.
the old became the new force, we can demonstrate that many did not last long in the job.\textsuperscript{16} Anecdotally, many did not even qualify.\textsuperscript{17}

The Parliament was responsible for the finances of the new police. Online Appendix Table A.2 lists the annual pay for constables (47 pounds), sergeants (55 pounds), inspectors (100 pounds), and superintendents (200 pounds). Thus, there was now a system with ranked officers, providing promotion incentives. But, constable and sergeant salaries were similar to that of pre-Met constables; continued reports of external fee schedules (e.g. witness testimony) are thus unsurprising (Radzinowicz 1953). Did Peel keep his budgetary promises? Many parishes were actually surprised by the size of the initial fees, and had to rely on loans from the Treasury; continued struggles led to an 1833 statute that allowed parishes to pay 75\% of the full amount, with the remainder covered by the government (Reynolds 1998).

2.3. Framework: The New Police, Crime, and Crime Statistics

The new police could have affected the actual number of crimes via two channels. As highlighted above, Peel’s ultimate goal was the absence of crime via deterrence: This was to be accomplished with more police, who were highly visible during the day and night, with uniforms, and constant patrols. The second crime-reducing channel is incapacitation: If the new police apprehended more offenders who were sentenced to prison, death, or transportation to Australia, then crimes were prevented by removing these persons from London’s society. Though incapacitation was not an intention of the new Met, we cannot rule out such effects. Despite the new police not being tasked with apprehending offenders (a task that remained in the hands of the Police Offices), officers could be more likely to clear crimes by catching offenders in the act on the streets. Our empirical analysis cannot disentangle such deterrence and incapacitation effects, which both move crime in the same direction. Both effects may increase in size over time as the new police become more experienced and disciplined.

We now turn to the potential confounders of measuring this crime reduction in administrative data. Estimates of the effect of any contemporary or historical policy on crime can be biased if the policy under consideration affects crime reporting and/or

\textsuperscript{16} We compare registers of the first 3,000 officers hired by the Met (Source: MEPO 4/31, National Archives London) to the 156 men hired into the Bow Street Foot Parole between 1823 and the Met’s 1829 creation (Source: MEPO 4/508, National Archives London). We can observe (using names) that a number of them were hired by the Met in the initial hiring wave. For instance, 24 of the last 34 Bow Street hires pre-Met subsequently joined the Met, but 9 were already dismissed by May 1830 and a number of others soon after. The Met, however, maintained a steady force size in these months, quickly replacing dismissed officers.

\textsuperscript{17} Some of the pre-Met police clearly did meet the age criteria. Though our digitalization of those hired into the Bow Street Foot Patrol (Source: MEPO 4/508, National Archives London) indicates all men were between ages 20 and 35, this is not the case for watchmen. Reynolds (1998) documents that just one-third of St Luke’s watch met all Met eligibility requirements and that, in the early 1800s, many watchmen were older than 50.
clearance rates (i.e. the share of incidents that have an identified/known offender associated).

First, did the introduction of the Met increase clearance rates? Even if crime remained constant, higher clearance rates could imply more arrests, charges, and/or trials. Though this type of bias would not be relevant for incident data (i.e. reported crimes), it could be a potential confounder in outcomes measured later in the justice system, like charges or trials. As noted above, since the task of following up on crime reports remained with the pre-Met Police Office constables, one may not expect clearance rates to go up that much (except perhaps for street crimes, like pickpocketing, in which the new police could be present to immediately apprehend offenders). Moreover, if the new police impacted clearance rates, one would expect this to increase administrative crime records: that is, it works against us finding a crime-reducing effect.

A second potential confounder is that the new police affected crime reporting. In particular, crime reporting may have increased if people felt that there was a new trusted authority to report to, who would take the report seriously regardless of the potential rewards. That is, creating the Met might have lowered the victim’s opportunity cost of reporting a crime. This would increase observed crime (in all data types, ranging from reports to trials). This may be especially relevant for relatively minor property crimes, for which the effort of reporting pre-Met may not have been worth it. At the same time, there is anecdotal evidence, at least initially, of anti-police sentiments towards the Met. Thus, if changes in reporting biases crime measures upward, one would expect this to increase over time with growing trust in the new police—as Peel’s emphasis on a non-corrupt high-quality force becomes recognized—and to be more prominent for minor offenses. For those most serious violent offenses, for example, robbery and homicide, one would not expect reporting biases to play a significant role.

3. The Impact of the Met on Crime: Evidence from the Old Bailey Proceedings

3.1. Old Bailey Data Description

Our first data source is the Proceedings of the Old Bailey. The Old Bailey was the Central Criminal Court for felony crimes in both London and the surrounding county of Middlesex. In other words, offenses both inside and outside the Met’s catchment area were trialed at the Old Bailey; the only difference is that there were different juries for offenses in London as opposed to Middlesex. Trials took place in (approximately) monthly sessions. At this time, victims often acted as prosecutors and a charge would

---

18. One magistrate is quoted as stating that “a strong feeling existed against the new police” upon the Met’s formation in an October 1, 1829 issue of The Morning Journal. See the Open University collection (accessed May 7, 2019): https://www.open.ac.uk/Arts/history-from-police-archives/MphcR1/Scrapbooks/sbIntro.html.
go to trial if a grand jury decided that there was sufficient evidence. The verdict and sentences were decided by juries and judges, respectively.

The Proceedings were published after each court session, and include records of more than 200,000 trials from 1700 to 1913. The Old Bailey Proceedings Online (Hitchcock et al. 2013) has digitized and tagged many variables (e.g. offense type, verdict, and sentence). We complement these readily available data with manually extracted (i) crime locations, geocoded into treatment and control areas of London, (ii) offense dates to identify treatment exposure, and (iii) police witness characteristics (number, type, and crime scene presence) to evaluate the “first-stage” implementation of the reform. Given the time-intensive nature of this work, we only transcribed and geocoded trials for homicide (murder/manslaughter), robbery, and burglary/housebreaking from 1821 to 1837, whose felony status (and thus representation at the Old Bailey rather than a lesser court for misdemeanors) did not change during this period.19

A potential limitation of the Old Bailey data is that it only contains those serious felony charges that resulted in a trial. Importantly, we use charges brought to trial (i.e. the occurrence of a trial and not its outcome) as a proxy for crime. It is of course possible though that crimes did not result in a trial—either because the victim did not report/prosecute or because the report was not cleared (i.e. did not result in an arrest/charge of a suspect). However, our focus on the most serious offenses of murder, robbery, and burglary limits concerns about the former; as highlighted in Section 2.3, such serious crimes are likely to always be reported. We further note that these historical trial data are fairly encompassing: Cases that result in pleas and acquittals are reported in the Proceedings and included in our data. Acquittal rates were around 25% in the 1820s and 1830s, while pleading did not play a prominent role at this time. See Bindler and Hjalmarsson (2020) for more details about trials at this time.

We geocoded the data using the most detailed address available in the Proceedings (e.g. an intersection, parish/district name, or street end/mid points) and mapped these locations into modern-day London maps to obtain postcodes and geo-coordinates.20 We measure the distance of each offense location to Charing Cross, and classify them as being in one of four areas. Only parishes within a 7-mile radius of Charing Cross (but outside the “City” of London) are treated, that is, within the Met’s original jurisdiction.

19. We also geocoded robbery, burglary, and murder/manslaughter for 1820–1850, but focus on pre-1837 years prior to avoid the abolition of the death penalty for robbery and burglary as a confounder. In addition, after the initial data coding, we noted that hardly any offenses were labeled burglary from the mid-1820s to the mid-1830s, while there was a sharp increase in offenses labeled housebreaking. We thus geocoded housebreaking offenses for 1820–1837 and treat housebreaking and burglary as one combined offense category for the entire period.

20. Whenever locations changed names (e.g. street names), we identify the current address using historical maps (roughly 40% of our sample). When the most detailed address is a long street (about 11% of our sample), we geocode the nearest street endpoint as the location (i.e. assign potentially untreated observations to the treatment area). Results are qualitatively robust to excluding either of those “fuzzy” locations (see Online Appendix Table A.4).
We subdivide this 7-mile circle into two treatment areas: the more intensely treated area within a 4-mile radius and the less intensely treated area within a 4–7 mile radius of Charing Cross (see Section 2.2). The two control areas are those outside the 7-mile radius of Charing Cross and within the City of London. Using the coded date of offense to classify incidents as before and after the Met’s introduction, we can thus estimate both simple pre–post and difference-in-differences specifications.

Online Appendix Table A.3 overviews the data, and shows the number of trials by crime type and police witness details within and outside a 7-mile radius of Charing Cross and in the City of London. One statistic that stands out is the relatively low number of homicide trials (just 258 in all areas over 1821–1837). This does not appear to be driven by the quality of Old Bailey trial data: Online Appendix Figure A.1 shows that trials track well with an alternative series of (non-trial) London homicides. A more likely explanation is the limited ability of coroners at this time to identify potential murders due to the limitations of forensic technology, for instance, to identify murder by poison (Emmerichs 2001). These measurement issues may, of course, limit our ability to say a lot about the impact of the new police on homicides.

Finally, for our main analysis, we aggregate the data to an area by month panel—that is, a data set of the number of crimes (by crime type) in a given area and month. The baseline specification uses the four areas (two treatment and two control) described above. We study two windows around the reform: January 1821–December 1837 (4 areas × 17 years × 12 months = 816 observations) and January 1828–December 1832 (240 observations).

The raw data (i.e. the number of crimes in month by area cells) are presented in time-series plots in Online Appendix Figure A.2 and summarized in Table 1, which presents, for each area and period (pre- and post-Met), the mean number of crimes per month. Panels A and B use the short and long windows, respectively. These statistics emphasize that the pre-reform number of crimes in the treated area is greater than that in the uncertain and control areas. In a preview of the results in Section 3.4, Table 1 shows that there is a significant reduction in robbery in the 4 miles around Charing Cross, which is seen in both the short and longer windows. No significant changes in burglary or homicide are observed.

3.2. The First Stage: Evidence of the Introduction of Met Police as Trial Witnesses

We use the Old Bailey trial reports to assess whether there is evidence of the introduction of the Met. Police witnesses were called constables (both before and after the creation of the Met), policeman (a post-Met label), watchman (a pre-Met label), and a handful of other labels that were either predominantly pre- or post-Met. Do we see an increase in the number and/or different types of police witnesses at trial after the Met was created?

We estimate the effect of the Metropolitan Police Act on witness type for each (of the four) treatment and control areas \( a \) in simple pre-post regressions depicted in equation (1). Specifically, we regress each measure of police presence at trial \( i \) for offense \( o \) at date \( t \) on a dummy variable indicating whether the offense occurred after
TABLE 1. Differences in means in the Old Bailey Proceedings.

|          | Burglary |          | Robbery |          | Homicide |          |
|----------|----------|----------|---------|----------|----------|----------|
|          | Before   | After    | Δ       | Before   | After    | Δ        |
| Treated  | 4.70     | 4.40     | −0.30   | 2.80     | 1.33     | −1.48*** |
| Uncertain| 0.40     | 0.40     | 0.00    | 0.05     | 0.10     | 0.05     |
| Control  | 1.15     | 1.03     | −0.13   | 0.30     | 0.28     | −0.03    |
| City of London | 0.95 | 1.05 | 0.10 | 0.50 | 0.35 | −0.15 |
| All      | 1.80     | 1.72     | −0.08   | 0.91     | 0.51     | −0.40**  |

Panel A: 1828–1832, \(Y = \text{number of crimes per month/area}\)

|          | Before   | After    | Δ       | Before   | After    | Δ        |
|----------|----------|----------|---------|----------|----------|----------|
| Treated  | 4.81     | 4.90     | 0.09    | 2.59     | 1.46     | −1.13*** |
| Uncertain| 0.36     | 0.57     | 0.21**  | 0.11     | 0.14     | 0.03     |
| Control  | 0.68     | 0.78     | 0.10    | 0.18     | 0.35     | 0.17**   |
| City of London | 0.79 | 1.14 | 0.35** | 0.57 | 0.25 | −0.32*** |
| All      | 1.66     | 1.85     | 0.19    | 0.86     | 0.55     | −0.31*** |

Panel B: 1821–1837, \(Y = \text{number of crimes per month/area}\)

Notes: The table shows the average number of monthly trials for crimes that took place before and after the introduction of the Met (and their difference), for each offense as well as by area (separately and all areas together). Trials are allocated to areas according to the following definitions: Treated (within 4 miles from Charing Cross), Uncertain (between 4 and 7 miles from Charing Cross), Control (more than 7 miles from Charing Cross), and City of London (in the City of London). Panel A shows the results for 1828–1832 and panel B for 1821–1837. The numbers are based on data from the Old Bailey Proceedings Online and own transcriptions/calculations; the sample includes trials for robbery, burglary, and homicide. See the text for details. Statistical significance of the difference is based on corresponding before–after regressions. *\(P < 0.1\); **\(P < 0.05\); ***\(P < 0.01\).

Offense-type fixed effects (\( \alpha_o \)) capture inherent differences across crime types in police presence; for example, police might be more likely to be present for street crimes like robbery. In contrast to the main analysis, which aggregates crimes to the month by area unit, these first-stage specifications keep the data at the individual trial level. Results for two windows—1821–1837 and 1828–1832—are presented in Table 2.

There is little evidence that the creation of the Met affected the extensive margin of having any police witness at a trial (see columns (1) and (2)). But, columns (3)–(6) show that it significantly changed the type of witness: The likelihood of a trial having a “new” police witness increased by 57 and 46 percentage points in the 4 and 4–7 miles radius areas (using the 1821–1837 window), while the presence of “old” police decreased by 49 and 25 percentage points, respectively. In the control area, there was an increase (16 percentage points) in “new” but no change in “old” police, while in the City of London, potentially treated after April 1832, a shift from old to new police is seen in the larger window (including post-1832) but is much smaller and/or insignificant in the shorter window. Finally, as the Met officers were constantly walking a short beat, it is possible that they were increasingly present at the crime scene itself,
TABLE 2. Evidence of the introduction of the Met: Police witnesses at the Old Bailey.

| Outcome:                      | (1) Any police witness | (2) Any "new" police witness | (3) Any "old" police witness | (4) Police at crime scene |
|-------------------------------|------------------------|------------------------------|-----------------------------|---------------------------|
| Sample:                       | 1821–1837              | 1821–1837                    | 1821–1837                   | 1821–1837                 |
| Treated (≤4 miles)            |                        |                              |                             |                           |
| Post-Met                      | 0.025                  | –0.018                       | 0.574***                    | –0.487***                 |
|                              | (0.018)                | (0.032)                      | (0.019)                     | (0.021)                   |
| Observations                  | 1.586                  | 427                          | 1.586                       | 427                       |
| Uncertain (4–7 miles)         |                        |                              |                             |                           |
| Post-Met                      | 0.034                  | –0.071                       | 0.455***                    | –0.254***                 |
|                              | (0.064)                | (0.145)                      | (0.057)                     | (0.082)                   |
| Observations                  | 135                    | 31                           | 135                         | 31                        |
| Control (≥7 miles)            |                        |                              |                             |                           |
| Post-Met                      | 0.014                  | –0.007                       | 0.155***                    | –0.056                    |
|                              | (0.052)                | (0.083)                      | (0.033)                     | (0.063)                   |
| Observations                  | 230                    | 89                           | 230                         | 89                        |
| City of London                |                        |                              |                             |                           |
| Post-Met                      | –0.062                 | –0.096                       | 0.268***                    | –0.259***                 |
|                              | (0.039)                | (0.059)                      | (0.042)                     | (0.056)                   |
| Observations                  | 314                    | 100                          | 314                         | 100                       |

Notes: The table shows the results of estimating equation (1). Specifically, it regresses the first-stage outcomes of police presence (dummy variables for any police witness at the trial, any "new" police witness, any "old" police witness, and whether any police was at the crime scene) on a dummy indicating whether the offense occurred after the creation of the Met (Post-Met) and offense fixed effects. These results are estimated separately for each treatment and control area (see each panel) and for the long and short windows (odd and even columns, respectively). As highlighted in the text, these regressions are based on a data set of individual trials, where we consider the police presence at each trial. The data are sourced from the Old Bailey Proceedings Online and own transcriptions/calculations; the sample includes trials for robbery, burglary, and homicide. See Section 3.1 for details. Robust standard errors are shown in parentheses below the coefficient. *P < 0.1; **P < 0.05; ***P < 0.01.
either by witnessing the crime or being close enough to be called for assistance, that is, a shorter response time. This may be especially relevant for street crimes. However, columns (7) and (8) of Table 2 only find support for this in the longer window for the treated area.

Table 2 shows indeed that the reform was implemented in both treated areas, but with a stronger intensity in the inner (4-mile) circle. Consistent with the immediate elimination of the parish watch described in the Act, the old police—that is, watchmen—are observed to be substituted out. An unexpected finding, however, is the increase in new police in the control areas, albeit substantially smaller than that for the treated areas. This could be explained by (i) geocoding errors or (ii) an increased use of the term “police” in the Proceedings by court reporters, regardless of the actual type. (The same reporter is responsible for the entire Proceedings.) Alternatively, (iii) there could be spillovers of the Met into the control areas. This could occur because the 7-mile radius/City of London is not a perfect boundary and some officers actually patrol this area or some crimes committed outside the 7-mile radius or in the City led to arrests within the 7-mile radius. If such spillovers existed and the control group was partially treated, we would underestimate the treatment effect in a difference-in-differences analysis of the effect of the Met on crime. To assess the plausibility of such spillovers, Figure 3 presents kernel densities of post-Met trials with and without police at the crime scene (panel (a)) and with at least one new or one old type of police witness (panel (b)), both by distance from Charing Cross. If there were policing spillovers beyond the Met’s catchment area, then one would expect to see a non-zero density in police presence at distances from Charing Cross beyond 7 miles, with potentially larger densities at distances closer to the Met boundaries. The figures do not suggest a substantial spillover of Met policing since the densities are close to zero around and outside the 7-mile mark.

3.3. Old Bailey Empirical Strategy

Having established that the creation of the Met affected policing in London, we turn to its effect on crime. Equation (2) presents a difference-in-differences model, which uses the area outside the 7-mile radius and the City of London as the best possible control groups. Again, we split the potentially treated areas into a 4-mile radius area with a more intense treatment and a 4–7 mile radius area with potentially less intense (i.e. uncertain intensity) treatment; these are labeled Treatment and Uncertain, respectively. The outcome is the number of trials for offense o in area a during month m of year y. The baseline analysis aggregates the data at the month (m) and area (a) level, and estimates equation (2) separately by offense category o.

\[
Trials^{o}_{amy} = \gamma_1 (Treatment \times PostMet)_{amy} + \gamma_2 (Uncertain \times PostMet)_{amy} + \alpha_y + \alpha_m + \alpha_a + \epsilon_{amy}.
\]

(2)

Year fixed effects (\(\alpha_y\)) flexibly control for criminal justice and societal trends over our sample period that are common to both treated and control areas, while
FIGURE 3. Spillovers in policing and crime displacement. Panel (a) shows the kernel density of homicide, robbery, and burglary trials at the Old Bailey from 1829 (after the introduction of the Met) to 1837 with/without police at the crime scene by distance from Charing Cross (in miles). Panel (b) shows the respective kernel density for trials with at least one “new” or “old” type of police witness at the trial (see the text for details on the types of police). Panels (c)–(e) show the kernel densities of crime locations for trials at the Old Bailey for 1821–1829 (before the introduction of the Met) and 1829–1837 (after the introduction of the Met), for burglary, robbery, and homicide. All figures exclude the City of London. The dashed vertical lines mark the thresholds in terms of distance from Charing Cross for the treated, uncertain, and control areas, respectively. The figures are based on data from the Old Bailey Proceedings Online and own transcriptions/calculations.
month of year fixed effects ($\alpha_m$) capture seasonality. Area fixed effects ($\alpha_a$) control for preexisting differences between areas, such as population density. $\gamma_1$ and $\gamma_2$, the coefficients on the interaction between a treatment (or uncertain) area dummy and whether the month of offense is October 1829 or later, capture the treatment effect for the inner and outer circles, respectively. This specification allows for causal interpretations of the estimated treatment effects ($\gamma_1$ and $\gamma_2$) if the usual parallel trend assumption holds, that is, nothing else changed differentially in the treatment and control areas that could have affected crime rates in the estimation window. We test for pre-reform differences between the treatment and control areas in an event-study design and discuss a number of potential confounders, including the 1832 cholera epidemic, contemporaneous criminal justice changes, population growth, police spillovers, and crime displacement.

We highlight four aspects of the baseline estimation. First, robust standard errors are used. Online Appendix Table A.5 assesses the sensitivity to a wild cluster bootstrap procedure (Cameron et al. 2008), clustering by the four baseline areas: Our findings are generally robust. Second, the baseline classifies the City of London as a control area. But, as the City Day Police (which was not part of the Met) became operational in April 1832, it is possible that a treatment similar to the Met was introduced in the City of London. We thus also estimate specifications that demonstrate robustness to which treatment or control group the City of London is assigned to: the treatment group after April 1832, the uncertainty group, or dropped completely. Third, we present results for two windows (1821–1837 and 1828–1832); the shorter window alleviates concerns about potential confounders discussed in Section 3.5.

Fourth, we estimate the baseline using OLS. Arguably, however, the correct model depends on a number of considerations. For example, should the impact of the Met be proportional to the population? Another example is a potential concern about differences in the moments (mean and variance) of the distribution in the outcome variable across the treatment and control groups; for instance, as seen in Table 1, mean crime levels are significantly higher in the treated than in the control areas. Such differences would make the control group potentially unsuitable for a difference-in-differences setup. These arguments would speak in favor of a Poisson model: In the absence of information on the number of officers per treatment/control area (note that we cannot reliably map divisions of 150 officers each into these areas), a Poisson model allows for a proportional interpretation of the introduction of the Met. Moreover, the Poisson model transforms the outcome variable in ways that can mitigate the second concern. (Note that a simple log-transformation is not suitable here given the zeros in the outcome variable.) However, given that there are reasons that speak against such a model (e.g. it is not obvious that a proportional effect should be expected, given differential population density and officer density per policed area), we remain more agnostic and show the results for a Poisson model alongside the results for the OLS model.21 Furthermore, we estimate OLS specifications with a binary (extensive

---

21. We use QML Poisson to estimate the fixed-effects models and adjust standard errors (Wooldridge 1999).
margin) outcome: whether there is any crime in a given area and period (using smaller areas or time periods than the baseline). These specifications are less dependent on the assumptions discussed above.

3.4. Main Results and Robustness Checks: Old Bailey Data

Table 3 presents these results. Columns (1) and (2) correspond to the baseline specification, estimated via OLS with the City as a control area, for the 1821–1837 and 1828–1832 windows, respectively. Panels A–C present the results for burglary, robbery, and homicide. The introduction of the Met significantly decreases the average number of monthly robbery trials in the treatment area (<4 miles from Charing Cross) relative to the control areas by 1.0 (long window) and 1.3 (short window), respectively, or 12–15 per year. There are no significant effects for burglary and homicide. Relative to the average number of pre-Met robberies in the treatment group, robberies decrease by 40% and 46% in the long and short windows, respectively. These findings are consistent with Draca et al. (2011), who, using modern-day London data, also find no effect on burglary but a decrease in robbery. They argue that the nature of the offense makes burglary less susceptible to a “public deterrence technology” than robbery (which can be classified as a street crime). An alternative explanation is that offsetting biases due to increased clearance and/or reporting rates are larger for burglary than for robbery. Though at least partially treated, we find no significant effects of the Met on crime in the uncertainty area; in fact, most of the coefficients are positive. This could imply that crime levels did not change in this area (maybe due to a smaller deterrence effect, as police were less visible in larger divisions) or that the crime-reducing effect was offset by increased apprehensions or reporting. Columns (3)–(5) of Table 3 show that these results are robust to how the City of London is classified.

The rest of Table 3 presents three alternative estimation strategies for both sample windows. For ease of interpretation, columns (6) and (7) show incident rate ratios from estimating a quasi-maximum likelihood (QML) Poisson model. These estimates generally have the same sign as baseline OLS (i.e. the Poisson estimates are less than 1 when the OLS estimates are negative), but differ in precision. Specifically, we find an approximately 30% reduction in robberies in the treatment area, which is only significant (at the 10% level) in the short window. While the OLS found no significant effects on burglary and homicide in the main treatment area, the Poisson results find a significant reduction in the long window. In the area of uncertain treatment, the Poisson specification finds significant increases in burglary in the long window and in robbery in the short window. Again, these results suggest that if there is any reduction in robberies or burglaries in the uncertainty area, it is (more than) offset by increases in either reporting or clearance rates.

Columns (8)–(11) present the results when we instead move to OLS specifications with a binary (extensive margin) outcome: whether there is any crime in a given area and period. To create enough variation in this new outcome variable (and, at the same time, test the robustness of our results to different aggregation levels), we aggregate the data to the week by area level in columns (8) and (9) (i.e. a smaller temporal
| Sample:          | (1) 1821–1837 | (2) 1828–1832 | (3) 1828–1832 | (4) 1828–1832 | (5) 1828–1832 | (6) 1821–1837 | (7) 1828–1832 | (8) 1821–1837 | (9) 1828–1832 | (10) 1821–1837 | (11) 1828–1832 |
|-----------------|----------------|----------------|----------------|----------------|----------------|----------------|----------------|----------------|----------------|----------------|----------------|
| Specification:  | Intensive margin: by month/area | Intensive margin: by month/area | Intensive margin: by week/area | Intensive margin: by month/area | Intensive margin: by week/area |
| City of London: | Treated from 2 April 1832 | Uncertain | Excluded | Control | Control | Control | Control | Control | Control | Control | Control |
|                 | OLS | QML | Poisson | OLS | OLS | OLS | OLS | OLS | OLS | OLS | OLS |
| Panel A: burglary | | | | | | | | | | | |
| Post-Met × treatment area | -0.106 | -0.294 | -0.285 | -0.234 | -0.120 | 0.821* | 0.951 | -0.069* | -0.062 | -0.032 | -0.064 |
|                            | (0.370) | (0.660) | (0.474) | (0.686) | (0.701) | (0.091) | (0.124) | (0.036) | (0.069) | (0.036) | (0.072) |
| Post-Met × uncertainty area | 0.016 | 0.006 | -0.013 | 0.116 | 0.180 | 1.290** | 1.015 | -0.066 | -0.029 | 0.031 | -0.028 |
|                            | (0.149) | (0.288) | (0.298) | (0.293) | (0.377) | (0.144) | (0.133) | (0.026) | (0.049) | (0.031) | (0.059) |
| Panel B: robbery | | | | | | | | | | | |
| Post-Met × treatment area | -1.032*** | -1.297**** | -0.832*** | -1.281*** | -1.336*** | 0.727 | 0.693** | -0.115*** | -0.127* | -0.116*** | -0.168** |
|                            | (0.219) | (0.428) | (0.299) | (0.433) | (0.438) | (0.341) | (0.141) | (0.033) | (0.065) | (0.034) | (0.069) |
| Post-Met × uncertainty area | 0.129 | 0.228 | 0.288* | 0.144 | 0.189 | 1.705 | 2.928** ** | 0.020 | -0.002 | 0.018 | 0.027 |
|                            | (0.088) | (0.162) | (0.170) | (0.189) | (0.208) | (0.800) | (0.597) | (0.016) | (0.031) | (0.018) | (0.030) |
| Panel C: homicide | | | | | | | | | | | |
| Post-Met × treatment area | -0.008 | 0.120 | 0.072 | 0.188 | 0.222 | 0.618*** | 1.053 | 0.026 | 0.064 | 0.000 | 0.044 |
|                            | (0.139) | (0.251) | (0.181) | (0.262) | (0.266) | (0.094) | (0.482) | (0.027) | (0.047) | (0.028) | (0.051) |
| Post-Met × uncertainty area | -0.049 | -0.055 | -0.062 | 0.100 | 0.047 | 0.975 | 0.433* | -0.005 | -0.007 | 0.002 | -0.008 |
|                            | (0.057) | (0.115) | (0.119) | (0.133) | (0.143) | (0.148) | (0.198) | (0.012) | (0.024) | (0.015) | (0.024) |
| Sample:  | (1) 1821–1837 | (2) 1828–1832 | (3) 1828–1832 | (4) 1828–1832 | (5) 1828–1832 | (6) 1821–1837 | (7) 1828–1832 | (8) 1821–1837 | (9) 1828–1832 | (10) 1821–1837 | (11) 1828–1832 |
|----------|---------------|---------------|---------------|---------------|---------------|---------------|---------------|---------------|---------------|---------------|---------------|
| Specification: | Intensive margin: by month/area | Intensive margin: by month/area | Extensive margin (1/0): by week/area | Extensive margin (1/0): by month/distance band |
| City of London: | Control | Control | Treated from | Uncertain | Excluded | Control | Control | Control |
| Treated from | 2 April 1832 | Uncertain | Excluded |
| Observations | 816 | 240 | 240 | 240 | 180 | 816 | 240 | 3,604 | 1,060 | 3,672 | 1,080 |
| Year fixed effects | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Month fixed effects | Yes | Yes | Yes | Yes | Yes | Yes | Yes | No | No | Yes | Yes |
| Week fixed effects | No | No | No | No | No | No | No | Yes | Yes | No | No |
| Area fixed effects | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |

Notes: The table shows the regression results corresponding to the baseline difference-in-differences model in equation (2). Panel A shows the results for burglary, Panel B for robbery, and Panel C for homicide. The dependent variable is the number of crimes (that are brought to trial) per month and area in columns (1)–(7), a dummy variable indicating whether there is any crime in a given week and area in columns (8) and (9), and a dummy variable indicating whether there is any crime in a given month and distance band from Charing Cross in columns (10) and (11). For all columns, the treated area includes crimes located within 4 miles from Charing Cross, the uncertain area those located between 4 and 7 miles from Charing Cross, the control area those located more than 7 miles from Charing Cross and City of London those located in the “City of London”. In columns (2)–(5), we vary whether we treat the City of London as part of the control area, as treated from April 1832 when they introduced their own police force, as part of the uncertainty area, or exclude the City of London from the analysis. Columns (6) and (7) report incident rate ratios from QML Poisson regressions (using the xtpqml Stata program); the other columns report the estimated coefficients from OLS regressions. Distance bands in columns (10) and (11) are circles around Charing Cross: less than 1 mile, 1–2 miles, 2–3 miles, . . . , 13–14 miles, and more than 14 miles. Robust standard errors are shown in parentheses below the coefficient. Regressions are based on manually geocoded data from the Old Bailey Proceedings Online and own transcriptions/calculations; see the text for details. *P < 0.1; **P < 0.05; ***P < 0.01.
period than in the baseline) and to the month by 1-mile distance band level in columns (10) and (11) (i.e. smaller geographic areas). We see the same pattern of results as in the baseline. There is a significant reduction in robbery trials (ranging from 11.5 to 16.8 percentage points) after introducing the Met, regardless of window or aggregation level. In contrast to the baseline (but comparable to the Poisson estimation), we do observe a marginally significant reduction in the chance of burglary (6.9 percentage points) in column (8). As in the OLS specifications, no significant effects are seen in the uncertainty area.

Figure 4 shows the results from event-study estimations by crime type (burglary—panel (a), robbery—panel (b), and homicide—panel (c)) for both the treated and uncertain areas. We estimate a more flexible specification that interacts the treatment and uncertainty indicators in equation (2) with a dummy for each 1-year interval before and after the introduction of the Met. In other words, we estimate a series of lead and lag effects that average all months within a given year, relative to the reform. To account for the mid-year creation of the Met, we define a year from September to August. These specifications allow us to both test the plausibility of the parallel trends assumption and study the dynamic effects of creating the Met. The results are supportive of parallel trends for all crime categories, and for both treatment and uncertainty areas: The coefficients are not significantly different from zero in the years leading up to the reform. (The only exception is the third lead for burglary in the treatment area.) There is no evidence of a short- or long-run effect for homicide or burglary. For robbery, the change in the point estimates in the treatment area is immediate and persistent (though not quite significant), and is close to the overall difference-in-differences estimate in every post reform year.22

Finally, Online Appendix Table A.4 demonstrates the robustness of the results to (i) area-specific time trends, (ii) excluding crimes reported “somewhere” on a long street, which could lead to miss-classified treated offenses, (iii) including only crimes for which we could identify the coordinates without referring to historical maps, and (iv) excluding offenses with missing crime dates (rather than assigning trial dates, as in the baseline). The main result—a significant reduction in robberies in the treated area with no observable effect for burglary—does not change. Controlling for area-specific trends does yield marginally significant increases in robbery trials in the uncertainty area.

3.5. Spillovers, Crime Displacement, and Other Potential Confounders

This section discusses a number of potential confounders to interpreting our results causally. The first is whether the 1832 cholera epidemic, which resulted in the deaths of almost 7,000 Londoners, differentially affected treatment and control areas.23 The epidemic could have affected crime through public riots (Tynkynen 1995) or by

22. These estimates are significant when using (less-demanding) 2-year intervals.
23. See, for instance, https://www.choleraandthethames.co.uk/ (accessed April 29, 2019).
FIGURE 4. Event study: the effect of the Met on crime (Old Bailey Proceedings). The figures show the estimated coefficients and 95% confidence intervals corresponding to the event-study specifications described in Section 3.4. The figures are based on expanding equation (2) by allowing interactions between the treatment and uncertainty dummies with each year leading up to/following the introduction of the Met:

\[
\text{Trials}_{amy}^0 = \sum_{t=-8}^{8} \gamma_1(Treatment \times PostMet)_{amy}(t = y - 1829) \\
+ \sum_{t=-8}^{8} \gamma_2(Uncertain \times PostMet)_{amy}(t = y - 1829) + \alpha_y + \alpha_m + \alpha_a + \varepsilon_{amy}
\]

The figures to the left represent the estimates for the treated area (\( \gamma_1 \)) and the figures to the right those for the uncertain area (\( \gamma_2 \)). Panels (a.1) and (a.2) show the results for burglary, panels (b.1) and (b.2) for robbery, and panels (c.1) and (c.2) for homicide. A year is defined as September to August. The vertical line represents the year before the introduction of the Met (September 1829), which is the omitted category. The dashed horizontal line represents the (average) difference-in-differences estimate. The figures are based on data from the Old Bailey Proceedings Online and own transcriptions/calculations.
affecting police resources (directly through ill/dying officers or indirectly as officer responsibilities are shifted away from crime prevention) or criminals (who may be incapacitated by the disease or driven to commit crimes). To explicitly look at the geographical and temporal distributions of cholera in London, we use a new data source (the *Returns to Death* of Met officers from 1829 to 1889), which includes the officer’s police division and date and (often) reason of death. Online Appendix Figure A.3 demonstrates that cholera arrived and peaked in 1832, diminished by 1833, and almost disappeared by 1834. A total of 19 officers died; all but one death were in July–September (peaking in August). Despite the equal number of officers across divisions, deaths were not equally distributed, with more deaths in inner London. However, our shorter estimation window mostly avoids this concern and the results are robust to dropping all trials after May 1832.

A second concern is whether other criminal justice reforms differentially affected treatment and control areas. The main reform is the abolition of capital punishment for burglary and robbery in 1837, which we avoid by ending our sample before 1837. The shorter time window further limits the possibility of omitting such shocks. Our previous research (Bindler and Hjalmarsson 2018) documents that there are no anticipation effects related to the timing of the offense by offense capital punishment reforms. Other legal reforms pushed by Peel (highlighted in Section 2) occurred in 1823, well outside the short window.

A third potential confounder is dynamic population growth. Online Appendix Figure A.4 shows decadal population estimates and growth for Inner London, Outer London, and the City of London. In the first half of the 19th century, Inner London grows at the highest rate (almost 25%) between each census, with Outer London not far off, while the City of London does not grow at all. To the extent that population growth implies more potential criminals and increases in crime, this biases us against finding a crime-reducing effect in the pre–post analysis. Likewise, faster growth in the treated areas would yield a similar bias in a difference-in-differences analysis. Using a short window around the reform mitigates these concerns.

Finally, to determine whether the introduction of professional police decreased crime overall or just locally (where police were introduced), one needs to understand whether police spillover or crime displacement effects exist. Section 3.2 concluded that there was little evidence of a substantial police spillover; however, if it did exist, it would attenuate any crime-reducing effects of the police. With regard to crime displacement, if criminals chose to commit crime in less policed areas than the newly treated Met jurisdiction, this would bias the difference-in-differences estimates in the direction of a crime-reducing effect. To assess the extent to which displacement is a concern, we return to Table 1, which compares the average number of pre- and post-Met offenses. In the shorter window, there is a significant reduction of 53% (2.88 to 1.33) in the average number of monthly robberies in the treated area. A comparable decrease

---

24. The classification of Inner and Outer London areas is done by the historical census, and cannot be matched to the “areas” in our data as it does not necessarily correspond to a 7-mile boundary.
is not seen in the less intensively treated uncertainty area, neither in the control area or City of London. There were no significant changes to burglary or homicide, neither any evidence of an increase in crime in the control areas. In other words, there is no evidence of a substitution of offenses from the treated to control areas in the short window.

However, such displacement cannot be ruled out in the longer window, as although it is still an order of magnitude smaller than in the treated area, the average number of monthly robberies outside the 7-mile radius actually doubles. While this increase could reflect displacement, it could be due to other potential confounders in the longer window. To take a closer look at potential displacement, Figure 3 plots kernel densities of crime locations (relative to Charing Cross) for pre- and post-Met periods. If there is displacement, one would expect an increase in the post-reform density just outside the 7-mile radius, where the Met was not introduced. While one can in fact see such a blip for each crime category around this distance, we highlight that (i) it is negligible relative to the amount of crime in the treated area and (ii) similar blips are seen in the pre-Met period, suggesting it is not completely driven by displacement.25 Online Appendix Figure A.5 reinforces this conclusion by plotting the pre–post Met differences in the mean number of trials (by crime type) separately for each 1-mile radius from Charing Cross. The only evidence of a crime reduction is for robbery in the treated area.

4. The Impact of the Met on Crime: Evidence from Daily Police Office Reports

A limitation of the Old Bailey data used in Section 3 is its restriction to three offenses—robbery, homicide, and burglary. Though the severity of these offenses mitigates concerns about changes in reporting, they make up a small fraction of Old Bailey trials: Taken together, they represent about 7% of all trials from 1820 to 1830.26 Thus, a natural question is whether the Old Bailey findings—that the Met led to an observable reduction in robbery but not burglary—generalize to other types of violent crimes (e.g. simple assault) and property crimes (e.g. shoplifting, pickpocketing, and simple/petty larceny). Many of these offenses were also classified as felonies at this time and trialed at the Old Bailey, but not transcribed and geocoded by us, as described in Section 3. This section thus complements the Old Bailey trial analysis with an alternative data source (the Report or Account of the Proceedings at the Several Police Offices) that is all encompassing in terms of offense categories and that measures crime at an earlier stage of the judicial process than trials.

25. In this period, criminals would likely be travelling on foot. Horse drawn stage coaches could be hired, and from 1829, the first omnibuses were introduced in central London (horse-drawn buses), but these alternatives were expensive. See https://www.oldbaileyonline.org/static/Transport.jsp, last accessed June 19, 2018, and Heblich et al. (2020). In this context, the control area with a radius of 7 to 15 miles from Charing Cross is not small.

26. This is based on tabulations at the Old Bailey website: https://www.oldbaileyonline.org/forms/formStats.jsp.
4.1. Daily Police Reports: Data Description

Online Appendix Figure C.1 presents an excerpt from this data source, which contains the reports by the nine Police Offices (described in Section 2) that were run by the pre-1829 magistrates (and continued until 1839). We sourced the reports from the National Archives and manually transcribed the data from January to April of 1828 (the year pre-reform), 1830 (the year post-reform), 1831, and 1832. Unfortunately, these daily police reports did not exist before 1828 and those for the second half of 1828 and 1829 are missing from the Archives. For each office and day (except Sundays), a detailed description of persons charged, informations, and property stolen are reported. As can be seen from the Online Appendix excerpt, informations are varied in nature, and typically provide details about a specific incident, including descriptions of missing property, suspect names, and potential awards.

For each day and office, we create three crime measures: (i) number of property stolen entries, (ii) number of property, violent, and other informations, and (iii) number of charges by crime type (property, violent, and other). The first two measure crime incidence (i.e. they report a crime incident, and while the suspect may be known, he is not caught), while charges are most similar to modern-day arrest data; there is a known suspect who is caught. One possible interpretation of these measures is to distinguish between uncleared and cleared crimes, where the informations and stolen property reports would be classified as the former and charges as the latter. The main offenses included in violent crime are robbery, assault, stabbing, wounding, and murder/manslaughter, while the main property offenses are burglary, theft, animal theft, and pickpocketing. Thus, while the violent and property crime categories used here include the Old Bailey offenses of homicide, robbery, and burglary, they encompass many other offenses. To address the possibility that the introduction of police only shifts uncleared crimes into the cleared category, we also create a measure that aggregates all types of incidents (informations, stolen property, and charges) together; this is our best proxy of the total number of crimes.

Figure 2 shows that all of the Police Offices are physically located in the 4-mile inner circle around Charing Cross. These Police Offices cannot be uniquely matched to the treatment and control areas used in the first part of the paper. And, in fact, the jurisdiction of these offices extends beyond the treatment areas to include crimes in the City of London as well as beyond the 7-mile radius. We thus, instead, use a simple pre–post design to estimate the net effect on crime in London (i.e. accounting for displacement to control areas) in a narrow window around the reform. We also note that an implication of the wider range of crime categories included in the Daily

---

27. Specifically, the 1832 Daily Police Reports (National Archives MEPO 4/17) often (but not always) include handwritten notes indicating the Metropolitan Police division in which the offense is located. There are two important takeaways. First, this is not a one-to-one match, with some divisions corresponding to multiple Police Offices and vice versa. Second, crimes from the City of London and beyond the 7-mile radius are included in these documents, that is, the catchment area for the Police Offices includes both treated and control areas of London.
Reports analysis is a potentially increased susceptibility to concerns about reporting biases than for those offenses (murder, robbery, burglary) studied in the Old Bailey.

4.2. Daily Police Reports: Summary Statistics

These historical records are digitized into a panel data set of crime measures at the office by date level for the nine offices and each Monday–Saturday in January through April of 1828, 1830, 1831, and 1832. We exclude the Thames Police Office as the Thames River Police is not in the jurisdiction of the Met and the composition and nature of crimes in this jurisdiction (docks and water) are likely to differ from other offices.28 The full analysis sample includes 3,232 office \times day observations, while the 1828–1830 sample window includes 1,616 observations. Table 4 presents summary statistics for the entire sample and by period: pre-reform (1828), 1 year post (1830), and 3 year post (1830–1832). For the entire period, there are on average 0.5 informations, 6.4 charges, and 0.4 reports of stolen property per day and office. Aggregating all crime measures together (top panel of Table 4), we find on average 7.3 crime reports per day and station. For all measures, property offenses comprise the largest crime category. Looking at the total number of crimes across years (1828 versus 1830), we see a reduction in the average number of violent crimes (per day and office) of 40%, which persists until 1832, but a small increase in property crimes (8%) and other crimes (4%). These increases are due to increases in the number of charges, while decreases are seen in both stolen property reports and property and other informations.

4.3. Daily Police Reports: Empirical Specification

Equation (3) presents the pre–post model used to estimate the effect of the introduction of the Met on daily crime reported by the different Police Offices:

\[ Y_{it} = \beta \text{PostMet}_t + \alpha_w + \alpha_d + \alpha_i + \varepsilon_{it}, \] (3)

where \( Y \) is the daily measure of crime reported in office \( i \) on date \( t \). Our main variable of interest, PostMet, equals 1 on all dates \( t \) after the introduction of the Met (i.e. 1830–1832 in our analysis sample) and zero in the year before (i.e. 1828). Our baseline specification includes Police Office fixed effects (\( \alpha_i \)) to control for unobservable office characteristics, such as differences across magistrates, as well as fixed effects for calendar weeks (\( \alpha_w \)) and day of the week (\( \alpha_d \)) to control for seasonality and variation in crime rates over the days of a week.

These data and simple pre–post specification clearly have limitations. One is an inability to control for confounding factors that changed at the same time as the

28. We also exclude the “Metropolitan Police Office” as this tenth office was created only in April 1831; this restriction is only relevant for the larger estimation window. It is not the case that crimes in the Met jurisdiction were displaced to the Thames office; a pre–post analysis of the Thames office indicates a reduction in all crime measures in the first year after the Met and the baseline results are robust to including the Thames in the analysis.
Table 4. Summary statistics: daily crime reports.

| Sample period                  | N  | All       | Pre-Met: 1828 | Post-Met: 1830 | Post-Met: 1830–1832 |
|-------------------------------|----|-----------|---------------|-----------------|---------------------|
|                               |    | Mean      | Mean          | Mean            | Mean                |
|                               |    | SD        | SD            | SD              | SD                  |
| **All incidents: informations ± charges ± property stolen** |    |           |               |                 |                     |
| Number of incidents: all      |    | 7.299     | 3.801         | 6.685           | 3.569               | 7.025               | 3.672               | 7.501               | 3.854               |
| Number of incidents: property |    | 5.748     | 3.264         | 5.189           | 3.139               | 5.592               | 3.138               | 5.932               | 3.283               |
| Number of incidents: violent  |    | 0.203     | 0.492         | 0.295           | 0.599               | 0.178               | 0.445               | 0.173               | 0.447               |
| Number of incidents: others   |    | 1.348     | 1.373         | 1.201           | 1.320               | 1.255               | 1.322               | 1.396               | 1.386               |
| **Informations**              |    |           |               |                 |                     |                     |                     |
| Number of informations: all   |    | 0.513     | 0.982         | 0.791           | 1.154               | 0.489               | 0.875               | 0.421               | 0.899               |
| Number of informations: property |    | 0.401     | 0.806         | 0.566           | 0.922               | 0.396               | 0.743               | 0.346               | 0.757               |
| Number of informations: violent |    | 0.048     | 0.242         | 0.101           | 0.362               | 0.036               | 0.192               | 0.030               | 0.182               |
| Number of informations: others |    | 0.065     | 0.282         | 0.124           | 0.386               | 0.059               | 0.274               | 0.045               | 0.236               |
| Any informations: all         |    | 0.311     | 0.463         | 0.465           | 0.499               | 0.316               | 0.465               | 0.260               | 0.439               |
| Any informations: property    |    | 0.269     | 0.444         | 0.371           | 0.483               | 0.279               | 0.449               | 0.256               | 0.424               |
| Any informations: violent     |    | 0.043     | 0.204         | 0.086           | 0.281               | 0.037               | 0.188               | 0.029               | 0.168               |
| Any informations: others      |    | 0.057     | 0.232         | 0.107           | 0.310               | 0.053               | 0.224               | 0.040               | 0.197               |
| **Charges**                   |    |           |               |                 |                     |                     |                     |
| Number of charges: all        |    | 6.382     | 3.590         | 5.281           | 3.154               | 6.161               | 3.419               | 6.744               | 3.651               |
| Number of charges: property   |    | 4.946     | 3.064         | 4.010           | 2.746               | 4.834               | 2.878               | 5.254               | 3.101               |
| Number of charges: violent    |    | 0.155     | 0.421         | 0.194           | 0.479               | 0.143               | 0.402               | 0.143               | 0.399               |
| Number of charges: others     |    | 1.284     | 1.355         | 1.077           | 1.271               | 1.199               | 1.306               | 1.352               | 1.375               |
| Any charges: all              |    | 0.991     | 0.0943        | 0.983           | 0.131               | 0.990               | 0.099               | 0.994               | 0.078               |
| Any charges: property         |    | 0.976     | 0.153         | 0.949           | 0.221               | 0.979               | 0.143               | 0.985               | 0.121               |
| Any charges: violent          |    | 0.136     | 0.343         | 0.164           | 0.370               | 0.127               | 0.334               | 0.127               | 0.333               |
| Any charges: others           |    | 0.660     | 0.474         | 0.598           | 0.491               | 0.627               | 0.484               | 0.681               | 0.466               |
| **Property stolen**           |    |           |               |                 |                     |                     |                     |
| Number of incidents           |    | 0.405     | 0.750         | 0.613           | 0.951               | 0.376               | 0.655               | 0.337               | 0.656               |
| Any incident                  |    | 0.292     | 0.455         | 0.394           | 0.489               | 0.295               | 0.456               | 0.258               | 0.438               |

Notes: The table shows summary statistics for the analysis sample based on the daily crime reports described in more detail in Section 4.1. The first two columns show the mean and standard deviations for the different crime measures for the complete sample, and the remaining columns separately for 1828 (one year pre-reform), 1830 (one year post-reform), and the years 1830–1832 (three years post-reform). The number of observations is shown at the top of each column. The data were manually transcribed from the Report or Account of the Proceedings of the Several Police Offices sourced from the National Archives (MEPO 4/12, 4/13, 4/15, and 4/17).
Based on the Old Bailey discussion, however, we believe these concerns are alleviated by (i) the short window emphasized for most of this analysis, (ii) the lack of contemporaneous reforms in these years, and (iii) comparability between the Old Bailey pre–post and difference-in-differences estimates. Second, having only one pre-period of data (January–April of 1828) limits our ability to test for preexisting crime trends. But, as one argument for the new police was rising crime rates, it is hard to imagine deterrence being confounded by a downward trend in crime.

How is the parameter, $\beta$, interpreted? Compared to the Old Bailey analysis, where we could geocode offenses into two treated and controls areas, the catchment area for the pre-Met Police Offices includes both the 0–4 mile treatment and 4–7 mile uncertainty areas, as well as the control areas. Thus, this specification estimates the net effect of introducing the Met on crime in all of London; a perhaps unintended advantage is that we do not need to worry about crime displacement effects. As highlighted above, the other main difference from the Old Bailey analysis is the crime measures used: The daily police report data include all offenses, regardless of offense type or severity. While this increases the generalizability of the results beyond the selected offenses of robbery, burglary, and homicide, it does increase the possibility that the results are confounded by changes in reporting behavior.

4.4. Daily Police Reports: Main Results

Table 5 presents the results of estimating equation (3) using the daily crime reports for each outcome: any number of informations (panels B and C), any stolen property reports (panel D), and number of charges (panel E). Panel A presents the total number of incidents, that is, the aggregated number of informations, stolen property offenses, and charges. Column (1) shows the raw pre–post difference with the short sample window (i.e. 1828 and 1830 only) for all offenses. There is a significant reduction in the chance of observing any informations by 14.9 percentage points (32% relative to the 1828 mean), the number of informations by 0.302 (38%), and the likelihood of any stolen property incidents by 9.8 percentage points (25%). In contrast, there is an increase in the total number of charges by 0.88 (16.6%). Panel A shows that the combined effect on the total number of incidents (aggregated across categories) is positive (0.34% or 5.1%) and marginally significant. Adding Police Office, calendar week, and day of week fixed effects in column (2) has little impact on the raw estimates. Expanding the post-period to include 1831 and 1832 results in larger estimates, with unchanged sign and precision. We study possible reasons for this pattern in Table 6. Finally, consistent with the Old Bailey analysis, we present results from QML Poisson model estimations for the non-binary outcomes; column (7) combines all offense

29. Online Appendix Table A.6 presents a number of robustness checks, including the following estimates: (i) excluding one office at a time, (ii) at the weekly instead of daily level, (iii) excluding incomplete weeks of data, e.g. due to holidays, and (iv) based on alternative specifications, including logarithms of the dependent variable (where appropriate). Online Appendix Table A.7 shows the robustness to conventional clustering (by Police Office) as well as a wild cluster bootstrap procedure.
Table 5. Daily crime reports: baseline results.

| Sample:          | (1)       | (2)       | (3)       | (4)       | (5)       | (6)       | (7)       | (8)       | (9)       | (10)      |
|------------------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|
| Crime type:      | 1828–1830 | 1828–1830 | 1828–1832 | 1828–1830 | 1828–1830 | 1828–1830 | 1828–1830 | 1828–1830 | 1828–1830 | 1828–1830 |
|                  | Total     | Total     | Total     | Property  | Violent   | Other     | Total     | Property  | Violent   | Other     |
| Estimator        | OLS       | QML Poisson|
| Panel A: number of all incidents per day/station |          |           |           |           |           |           |           |           |           |           |
| Post-Met police  | 0.340**   | 0.347**   | 0.821***  | 0.406***  | –0.115*** | 0.056     | 1.052     | 1.078     | 0.609**   | 1.046     |
|                  | (0.180)   | (0.154)   | (0.134)   | (0.135)   | (0.026)   | (0.063)   | (0.047)   | (0.060)   | (0.119)   | (0.144)   |
| Panel B: any informations per day/station |          |           |           |           |           |           |           |           |           |           |
| Post-Met police  | –0.149*** | –0.148*** | –0.206*** | –0.090*** | –0.049*** | –0.055*** | –          | –          | –          | –          |
|                  | (0.024)   | (0.022)   | (0.019)   | (0.022)   | (0.012)   | (0.013)   |           |           |           |           |
| Panel C: informations per day/station |          |           |           |           |           |           |           |           |           |           |
| Post-Met police  | –0.302*** | –0.301*** | –0.371*** | –0.170*** | –0.064*** | –0.065*** | 0.619**   | 0.700     | 0.359**   | 0.473**   |
|                  | (0.051)   | (0.046)   | (0.041)   | (0.039)   | (0.014)   | (0.016)   | (0.149)   | (0.171)   | (0.129)   | (0.140)   |
| Panel D: any stolen property per day/station |          |           |           |           |           |           |           |           |           |           |
| Post-Met police  | –0.098*** | –0.099*** | –0.137*** | –          | –          | –          | –          | –          | –          | –          |
|                  | (0.024)   | (0.023)   | (0.019)   |           |           |           |           |           |           |           |
| Panel E: number of charges per day/station |          |           |           |           |           |           |           |           |           |           |
| Post-Met police  | 0.879***  | 0.890***  | 1.471***  | 0.827***  | –0.050**  | 0.126**   | 1.168**   | 1.206**   | 0.742     | 1.117     |
|                  | (0.164)   | (0.140)   | (0.120)   | (0.121)   | (0.022)   | (0.061)   | (0.068)   | (0.092)   | (0.146)   | (0.157)   |
| Observations     | 1,616     | 1,616     | 3,232     | 1,616     | 1,616     | 1,616     | 1,616     | 1,616     | 1,616     | 1,616     |
| Office fixed effects | No | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Calendar week fixed effects | No | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Day of week fixed effects | No | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |

Notes: The table shows the regression results corresponding to equation (3), discussed in Section 4.3. For a description of the underlying data, see Section 4.1. The dependent variable in panel A is the number of all incidents (i.e. the sum of informations, stolen property incidents, and charges), in panel B a dummy variable indicating whether there are any informations, in panel C the number of informations, in panel D a dummy variable indicating whether there are any stolen property reports, and in panel E the number of charges. The independent variable of interest is a dummy variable Post-Met police equal to 1 for any date after the introduction of the Met police. The top of each column indicates the years included in the sample and where appropriate the crime category, as well as the estimation details (OLS, QML Poisson). Robust standard errors are shown in parentheses below the estimated coefficients. *P < 0.1; **P < 0.05; ***P < 0.01.
| Sample: | (1) 1828–1832 | (2) 1828–1832 | (3) 1828–1832 | (4) 1828–1832 | (5) 1828–1832 | (6) 1828–1832 | (7) 1828–1832 | (8) 1828–1832 | (9) 1828–1832 | (10) 1828–1832 |
|---------|---------------|---------------|---------------|---------------|---------------|---------------|---------------|---------------|---------------|---------------|
| Y:      | Number of all incidents | Number of charges | Any informations | Any stolen property |
| Crime type: | Total | Property | Violent | Total | Property | Violent | Total | Property | Violent | Property |
| Post-Met: 1830, January | −0.496* | −0.158 | −0.130*** | 0.030 | 0.214 | −0.056 | −0.114*** | −0.045 | −0.055*** | −0.114*** |
| | (0.268) | (0.233) | (0.043) | (0.246) | (0.209) | (0.037) | (0.038) | (0.038) | (0.019) | (0.040) |
| Post-Met: 1830, >January | 0.632*** | 0.606*** | −0.111*** | 1.177*** | 1.040*** | −0.049** | −0.160*** | −0.106*** | −0.046*** | −0.094*** |
| | (0.176) | (0.153) | (0.027) | (0.163) | (0.140) | (0.023) | (0.024) | (0.023) | (0.012) | (0.025) |
| Post-Met: 1831 | 0.772*** | 0.688*** | −0.131*** | 1.382*** | 1.146** | −0.065*** | −0.220*** | −0.141*** | −0.054*** | −0.125*** |
| | (0.165) | (0.141) | (0.026) | (0.151) | (0.129) | (0.021) | (0.021) | (0.021) | (0.011) | (0.023) |
| Post-Met: 1832 | 1.350*** | 1.147*** | −0.117*** | 2.157*** | 1.783*** | −0.034 | −0.250*** | −0.174*** | −0.068*** | −0.187*** |
| | (0.175) | (0.151) | (0.026) | (0.160) | (0.137) | (0.022) | (0.021) | (0.021) | (0.011) | (0.022) |
| P-value | 0.000 | 0.000 | 0.853 | 0.000 | 0.000 | 0.483 | 0.000 | 0.000 | 0.079 | 0.000 |
| Observations | 3,232 | 3,232 | 3,232 | 3,232 | 3,232 | 3,232 | 3,232 | 3,232 | 3,232 | 3,232 |
| Office fixed effects | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Calendar week fixed effects | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Day of week fixed effects | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |

Notes: The table shows the regression results corresponding to an expansion of equation (3), where we allow for separate coefficients by time after the introduction of the Met. (Note that the second wave of hiring, mainly in the outer divisions, occurred in February 1830.) In other words, we allow the Met to have differential effects in four post-Met periods (January 1830, February–April 1830, January–April 1831, January–April 1832). For a description of the underlying data, see Section 4.1. The dependent variable in columns (1)–(3) is the total number of incidents (charges + informations + property stolen incidents), in columns (4)–(6) the number of charges, in columns (7)–(9) a dummy variable indicating whether there are any informations, and in column (10) a dummy variable indicating whether there are any stolen properties. The top of each column indicates the years included in the sample and the crime category. The P-value corresponds to the test of equality of all four shown coefficients. Robust standard errors are shown in parentheses below the estimated coefficients. *P < 0.1; **P < 0.05; ***P < 0.01.
categories. Similar effect sizes are seen for the number of informations (−38%), charges (16.8%), and total incidents (5.2%). Except for the latter, these estimates are statistically significant.

Columns (4)–(6) of Table 5 look at property, violent, and other crimes with OLS. The respective results using a Poisson model are shown in columns (8)–(10). The OLS results are consistent with the comparison of pre- and post-means. For informations, we see negative effects for all crime categories, with a reduction of any property and violent informations of 24% and 57%, respectively. (The corresponding Poisson estimates are −30% and −64.1%, though only the latter is statistically significant.) For charges, there is a more heterogeneous pattern: Charges increase (decrease) by about 21% (−26%) for property (violent) crime. Aggregating informations and charges, panel A shows that the overall effect is a significant increase (8%) and decrease (−39%) in property and violent crime, respectively, but no effect on other crimes. The corresponding Poisson estimates are almost exactly the same in magnitude for both property (7.8%) and violent (−39.1%) crimes, but less precisely estimated (not significant for property crime). For violent crimes, the overall number of incidents decreases on average by 0.115 per day and office: This corresponds to a reduction throughout London of 287 violent crimes per year (0.115 × 8 offices × 6 days per week × 52 weeks). A comparison to the Old Bailey analysis suggests that about 5% (15 of 287 violent crimes averted per year) of these offenses are robberies. Of course, to the extent that the estimates are biased by offsetting increases in clearances or reporting, these numbers of violent crimes averted are a lower bound.

The combined measure—informations plus stolen property plus charges—is our best proxy for all criminal incidents, regardless of whether there are known suspects and associated charges. To summarize, the above analysis found a reduction in total violent crime that was driven by reductions in both violent informations (uncleared offenses) and charges (cleared offenses). This in fact parallels the reduction seen for robbery, a violent offense, in the Old Bailey analysis. In contrast, there is an increase in observed property crime incidents overall, which is driven by an increase in charges, which outweighs decreases in informations and stolen property reports. Increases in clearance rates or reporting are plausible drivers of the increase in charges: Increased clearance is possible as the physical presence of the Met officers walking the streets may have allowed them to apprehend property offenders, such as pickpocketers, as crimes were being committed. And given that the data include more (and less serious) offenses than just burglary (as in the Old Bailey analysis), reporting changes are also feasible. In this respect, the results are consistent with the Old Bailey analysis, where there was no burglary effect—that is, deterrence and clearance/reporting channels offset each other.

4.5. Daily Police Report Extension: Short- and Medium-Term Dynamics

What are the dynamic effects of creating the Met? We take advantage of the two-stage initial hiring described in Section 2.2 (i.e. inner divisions in September 1829 and outer
divisions in February 1830) to estimate an extended version of equation (3) that allows for different coefficients on the treatment variable in (i) January 1830 (post-Met but before the second hiring wave), (ii) all other months in 1830 (after the second wave), (iii) 1831, and (iv) 1832. This allows us to study the immediate effect of the large February 1830 hiring wave and whether the impact of the Met changes over time. Table 6 shows the results for the combination of all incidents in columns (1)–(3), the number of charges in columns (4)–(6), any informations in columns (7)–(9), and stolen property in column (10). There are two key takeaways. First, the point estimates generally increase over time, which may not be surprising given the increasing quality of police after the initial introduction of the Met and the continued hiring. Second, while some of the reduction in “uncleared” crimes is immediate (for violent informations and stolen property reports), the significant increase in property crime charges does not kick in until the second wave. This could be because reporting did not increase until the new police began to gain the trust of the people or because the new police needed to gain experience and manpower to sufficiently increase apprehension/clearance rates, which was at least initially not their explicit task.

5. Extension: The Introduction of Professional County Police

The results thus far show that the creation of the Met significantly reduced violent crimes: Reductions in crime due to deterrence or incapacitation dominate increases in clearance and reporting. For property crimes, the latter effects are at least as large (in the Old Bailey) or larger (in the daily reports), such that there is less visible evidence of crime deterrence. Are these patterns specific to the creation of a professional force in London in the 1830s?

We answer this question by studying the rollout of forces to the rural counties of England and Wales via the County Police Act of 1839 and the County and Borough Police Act of 1856. The former allowed for the creation of county forces, while the latter made it mandatory. We digitized the year of county force formation and initial force size (see Online Appendix Table B.1) from the Police History Society (Stallion and Wall 1999). Panel (a) of Figure 5 maps the rollout, and demonstrates no obvious clustering in the years of force creation by neighboring counties. Panel (b) depicts the evolution over time: 16 forces were created in 1840 (when permitted), 23 in 1857 (when mandatory), and 9 in the intervening years. The 1856 Act also established a national Inspectorate to annually certify force “efficiency”. One important input was whether the officer-to-population ratio was sufficiently large; the 1839 Act recommended 1 officer per 1000 people. We study the effect of having any county force as well as whether higher quality forces—measured by initial force size per capita—have differential effects.

30. See Bindler and Hjalmarsson (2019) for a more in-depth discussion of the county force rollout.
Figure 5. The rollout of police forces for English and Welsh counties. The map illustrates the different start years of police forces across counties in England and Wales. Each color represents a different start year. The counties of York, Sussex, and Suffolk are excluded (left blank) because of multiple start dates for the same county. This map is based on 1851 county registration districts, from the Great Britain Historical GIS Project (2012) “Great Britain Historical GIS”. Boundary data were downloaded from https://vision.port.ac.uk/downloads/download_free/boundaries.jsp. Panel (b) shows the number of county police forces in each year for our analysis sample of 48 counties, that is, excluding Middlesex, York, Suffolk, and Sussex. The red vertical line marks 1857, the year when the creation of a county police force became mandatory. See Section 5 for details on the data and sample.
To measure crime, we digitized the annual number of persons committed or bailed for trial for each fiscal year ending September 29, as reported in the *Judicial Statistics* (1857–1892), for six crime classes (see Online Appendix Table B.2). We focus on the total number of charges and the sub-categories of violent, property, and other property crimes.\(^3\) We also use 1851 and 1861 census records to create county-level controls.\(^2\) Of the 52 counties, we drop Middlesex, York, Sussex, and Suffolk, as they represent regions where multiple forces were created (at different times) but crime data are only available for the aggregate (entire) county. We use 1832–1865 as the baseline sample years, resulting in a county by year panel of 1,632 observations (48 counties × 34 years). Online Appendix Table B.3 shows the summary statistics.

We first estimate the overall effect of having any professional force (regardless of its characteristics) on crime with the following difference-in-differences model.

\[
\text{Crime}_{ct} = \beta \text{Force}_{ct} + \alpha_c + \alpha_t + X_{ct} \theta + \varepsilon_{ct}. \tag{4}
\]

*Crime* is the log number of trials in county *c* and fiscal year *t*, while *Force* equals 1 for county–year combinations with a professional force for all of year *t*.\(^3\) County fixed effects (\(\alpha_c\)) control for unobservable but constant differences across counties, for example, preexisting crime levels that may be related to force creation decisions as well as any fixed characteristics of neighboring counties. Year fixed effects (\(\alpha_t\)) capture national shocks, such as other criminal justice reforms (e.g. abolition of the death penalty and transportation and an 1855 Act that shifted non-violent larceny cases out of the courtroom). Column (1) of Table 7 presents the results for all charges in panel A, and violent, property, and other property charges in panels B–D, respectively. The creation of a county force, on average, does not significantly affect trials in any crime category.

Yet, this non-effect could mask differential effects for forces of varying quality. We observe one dimension of quality—force size upon creation relative to the population. We thus evaluate the differential effect of introducing sufficiently or insufficiently large forces in an expanded specification, where superscript *j* indicates the threshold used to measure sufficiency.

\[
\text{Crime}_{ct} = \beta_s \text{SuffForce}^j_{ct} + \beta_{ins} \text{InsuffForce}^j_{ct} + \alpha_c + \alpha_t + X_{ct} \theta + \varepsilon_{ct} \tag{5}
\]

Columns (2)–(4) of Table 7 present the results for selected thresholds of 1,500, 2,000, and 2,500 people per policeman. Under the strictest and weakest criteria, there are 10 and 30 sufficiently large forces, respectively. Force size matters: Creating a sufficiently

---

\(^3\) Violent includes crimes against person and violent property classes. Property includes the class of non-violent property and other property includes the classes of malicious property and forgery. Online Appendix Figure B.1 shows that alternative crime measures (available after 1857) move in lockstep with our measure—charges.

\(^2\) Census data are from the *North Atlantic Population Project, UK Censuses*, https://www.nappdata.org/napp/; the controls are share male, married, native, in various age groups, unemployed or out of the labor force, and farmers.

\(^3\) Results are robust to using log trials *per capita*; this is not our preferred choice given imperfect population data.
### Table 7. The effect of county police forces on crime.

| Panel | Dependent variable: log (number of charges) | Police force size defined according to the below thresholds in the number of people per policeman (upon police force creation) | (1) | (2) | (3) | (4) |
|-------|--------------------------------------------|---------------------------------------------------------------------------------------------------------------------------------|-----|-----|-----|-----|
|       |                                             |                                                                                                                               | Any force | 1,500 | 2,000 | 2,500 |
| Panel A: all charges | Force | –0.024 | 0.062 | 0.055 | 0.048 |
|       | Sufficiently large force | –0.190*** | 0.022 | 0.053 | 0.066 |
|       | Insufficiently large force | 0.022 | 0.053 | 0.066 |
|       | P-value | 0.011 | 0.048 | 0.151 |
| Panel B: violent charges | Force | –0.031 | 0.104 | 0.070 | 0.062 |
|       | Sufficiently large force | –0.183* | 0.002 | 0.034 | 0.048 |
|       | Insufficiently large force | 0.002 | 0.034 | 0.048 |
|       | P-value | 0.129 | 0.128 | 0.267 |
| Panel C: property charges | Force | 0.017 | 0.050 | 0.057 | 0.083 |
|       | Sufficiently large force | –0.143** | 0.064 | 0.090 | 0.101 |
|       | Insufficiently large force | 0.064 | 0.090 | 0.101 |
|       | P-value | 0.010 | 0.090 | 0.2351 |
| Panel D: other charges | Force | 0.002 | 0.120 | 0.097 | 0.089 |
|       | Sufficiently large force | –0.181 | 0.027 | 0.009 | 0.030 |
|       | Insufficiently large force | 0.027 | 0.009 | 0.030 |
|       | P-value | 0.082 | 0.549 | 0.489 |

Notes: The table presents the results of the county-level difference-in-differences model [see equation (4) for column (1) and equation (5) for columns (2)–(4)]. The dependent variable in all columns is the log of number of charges overall (panel A), for violent crime (panel B), for property crime (panel C), and for other crimes (panel D). In column (1), the variable of interest “Force” is equal to 1 for a county c in any year t after which a county police force has been in place for the entire year. In columns (2)–(4), the variables of interest “Sufficiently large force” and “Insufficiently large force” are equal to 1 for a county c in any year t after which a sufficiently large or insufficiently large county police force has been created. Sufficiency is defined according to the number of people per officer, with varying thresholds as indicated at the top of each column. The P-value (in italics) corresponds to a hypothesis test of equality of the two parameters for “Sufficiently large force” and “Insufficiently large force”. All specifications include county and year fixed effects. The baseline sample includes 48 counties for the years 1832–1865. For four years (1832, 1833, 1840, and 1852), we do not have charges by each crime type. In the few instances when there are zero charges in a given year/county, we drop this observation when taking the logarithm of the outcome variable. This results in sample sizes of 1,632 observations in panel A, 1,431 observations in panel B, 1,440 observations in panel C, and 1,267 observations in panel D. Standard errors are clustered by county and shown in brackets below the coefficient. *P < 0.1; **P < 0.05; ***P < 0.01.
large force with less than 1,500 people per policeman decreases the overall number of crimes by about 19%; comparable effects are seen for violent and property crimes (18% and 14%, respectively) but not for other offenses. In contrast, creating an insufficiently large force (insignificantly) increases the number of property and other crimes. Moving across the table, a significant crime-reducing effect of a “sufficiently large” force is also seen using the weaker threshold of 2,000. The effects of the sufficiently and insufficiently large forces on total charges are significantly different from each other for the 1,500 and 2,000 thresholds. ($P$-values are shown in the table.)

A causal interpretation of these findings relies on the assumptions of parallel trends, random timing of force formation, and conditional randomness of having a sufficiently large force. Online Appendix B summarizes and presents a number of robustness tests addressing these assumptions. Specifically, event-study specifications in Online Appendix Figure B.2 support the difference-in-differences assumptions and show that the crime-reducing effect of relatively large forces is not immediate but starts around 3 years post-reform and increases over time. Online Appendix Table B.4 further shows that the timing of force adoption is unaffected by lagged crime rates in the county or neighboring counties, the size of preexisting local forces, or neighboring county force characteristics. Online Appendix Table B.5 finds no evidence of crime displacement from London to neighboring counties upon the creation of the Met. Finally, few observables predict the force’s initial type/size (Online Appendix Table B.6).

What do we learn from these findings—that is, a crime-reducing effect of sufficiently large forces but a null effect of insufficient forces? Similar to what we argue in the London case, the net negative effect for sufficiently large forces suggests that deterrence and incapacitation (decreasing crime) can outweigh reporting and apprehension channels (increasing reported crime). But, these results also highlight that this will only happen when a force’s quality is high enough (at least in the dimension measured here). Moreover, the increasing size of the crime-reducing effect over time further highlights the potential role of quality: Force quality likely improved over time as people per officer ratios decreased (more officers were hired), more supervisors were hired, and experience was gained. These results reinforce our findings in the London daily police reports analysis (Table 6) that the estimates increase over time.

6. Conclusion

This paper estimates the effect on crime of the introduction of the Met—the first professional force in the world—and provides evidence that this new force significantly decreased violent crimes. Specifically, we estimate an annual reduction in violent

34. Online Appendix Table B.7 presents a number of additional robustness checks, including controlling for whether neighboring counties had insufficiently or sufficiently large forces.
crimes of at least 15 fewer robberies and 300 violent crimes more generally; given the potential biases due to increased clearance rates and (especially in the daily report data) reporting, these estimates are likely to be lower bounds. For property crimes, these potential biases may offset any crime-reducing effects, such that a crime reduction is not visible (though it may exist) in administrative charge or trial data.

The relative size of the violent crime reduction (about 40%) is of similar magnitude to contemporary studies. Using terror-related shocks to deployment, Draca et al. (2011) and DiTella and Schargrodsky (2004) find a decrease in crime of approximately 0.3% with a 1% increase in police (elasticity of –0.3). MacDonald et al. (2015) find a 45%–85% difference in the number of crimes across boundaries of police catchment areas. But, comparisons of this unique historical context to policing studies today are not perfect. Thus, one can also think about the size of these effects relative to the size of the reform or its cost. About 3,000 officers were hired by May of 1830, with annual salary costs estimated in Online Appendix Table A.2 to be about 157,000 pounds. Using these numbers and our results (i.e. a lower bound of about 300 violent crimes reduced), one could infer that one violent crime per year is deterred per 10 officers on the force or per approximately 500 pounds (measured in 1829). These ratios are clearly conservative given both the potential biases discussed above, especially for property crimes, and the fact that we over-estimate the cost of the reform (since we do not measure the marginal cost relative to what was spent on pre-Met “policing”). Finally, given the inherent difficulties, even in a contemporary context, in estimating the (averted) cost of crime, we hesitate to benchmark our findings with such comparisons.

Our study demonstrates that the creation of this first professional police force was successful: crime decreased. Even an increase in clearance and/or reporting can be thought of as an achievement for an institution aiming to protect society. The successful implementation of a reform of this scale is evidence in and of itself of state capacity in early-19th century England. The consequences that this modern-day institution had on societal dimensions besides crime, ranging from local economic activity and the enforcement of property rights to inequality and even political outcomes, is an important question for future research.

References

Acemoglu, Daron, Camilo Garcia-Jimeno, and James A. Robinson (2015). “State Capacity and Economic Development: A Network Approach.” American Economic Review, 105(8), 2364–2409.

Banerjee, Abhijit, Raghabendra Chattopadhyay, Esther Duflo, Daniel Keniston, and Nina Singh (2021). “Improving Police Performance in Rajasthan, India: Experimental Evidence on Incentives, Managerial Autonomy and Training.” American Economic Journal: Economic Policy, 13, 36–66.

Becker, Gary (1968). “Crime and Punishment: An Economic Approach.” Journal of Political Economy, 76, 169–217.

Besley, Timothy and Torsten Persson (2010). “State Capacity, Conflict, and Development.” Econometrica, 78, 1–34.

Bignon, Vincent, Eve Caroli, and Roberto Galbiati (2017). “Stealing to Survive? Crime and Income Shocks in Nineteenth Century France.” Economic Journal, 127, 19–49.
Lyman, J. L. (1964). “The Metropolitan Police Act of 1829: An Analysis of Certain Events Influencing the Passage and Character of the Metropolitan Police Act in England.” *Journal of Criminal Law, Criminology, and Police Science*, 55, 141–154.

MacDonald, John, Jonathan Klick, and Ben Grunwald (2015). “The Effect of Private Police on Crime: Evidence from a Geographic Regression Discontinuity Design.” *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, 179, 831–846.

Mehlum, Halvor, Edward Miguel, and Ragnar Torvik (2006). “Poverty and Crime in 19th Century Germany.” *Journal of Urban Economics*, 59, 370–388.

Mello, Steven (2019). “More COPS, Less Crime.” *Journal of Public Economics*, 172, 174–200.

Miller, Amalia and Carmit Segal (2019). “Do Female Officers Improve Law Enforcement Quality? Effects on Crime Reporting and Domestic Violence.” *Review of Economic Studies*, 86, 2220–2247.

Owens, Emily (2020). “The Economics of Policing.” In *Handbook of Labor, Human Resources and Population Economics*, edited by Zimmermann, K. Springer, Cham. https://doi.org/10.1007/978-3-319-57365-6_146-1.

Pfuhl, Erdwin (1983). “Police Strikes and Conventional Crime. A Look at the Data.” *Criminology*, 21, 489–504.

Radzinowicz, Leon (1953). “Trading in Police Services: An Aspect of the Early 19th Century Police in England.” *University of Pennsylvania Law Review*, 102, 1–30.

Rawlings, Philip (2008). “Policing before the Police.” In *Handbook of Policing*, edited by T. Newburn. Routledge.

Reynolds, Elaine (1998). *Before the Bobbies: The Night Watch and Police Reform in Metropolitan London, 1720–1830*. MacMillan Press.

Stallion, Martin and David Wall (1999). *The British Police: Police Forces and Chief Officers 1829–2000*. Police History Society, Bramshill.

Traxler, Christian and Carsten Burhop (2010). “Poverty and Crime in 19th Century Germany: A Reassessment.” Working paper, Max Planck Institute for Research on Collective Goods.

Tynkkynen, K. (1995). “Four Cholera Epidemics in Nineteenth-Century London.” *Hippokrates*, 12, 62–88.

Uchida, Craig (2015). “The Development of the American Police: An Historical Overview.” In *Critical Issues in Policing*, 7th ed., edited by R. Dunham and G. Alpert. Waveland Press, Long Grove, IL, pp. 11–30.

Weisburst, Emily (2019). “Safety in Police Numbers: Evidence of Police Effectiveness from Federal COPS Grant Applications.” *American Law and Economics Review*, 21, 81–109.

Wolpin, Kenneth I. (1978). “An Economic Analysis of Crime and Punishment in England and Wales, 1894–1967.” *Journal of Political Economy*, 86, 815–840.

Wong, Yue-Chim Richard (1995). “An Economic Analysis of the Crime Rate in England and Wales, 1857–92.” *Economica*, 62, 235–246.

Wooldridge, Jeffrey M. (1999). “Distribution-Free Estimation of Some Nonlinear Panel Data Models.” *Journal of Econometrics*, 90, 77–97.

**Supplementary Data**

Supplementary data are available at [JEEA online](https://academic.oup.com/jeea/article/19/6/3063/6174678).