

Measuring Police Effectiveness

By Anika Ludwig, Mike Norton and Iain McLean

Abstract

There are 43 territorial Policing areas in England and Wales, one in Northern Ireland, and one in Scotland. There is no serious political movement in favour of a single force to cover England and Wales. The multiplicity of forces makes it, in principle, possible to compare their performance. The UK government introduced elected Police & Crime Commissioners (PCCs) in 2012 with a view to improving the cost-effectiveness and accountability of local policing. It should be possible to compare performance figures from before and since this change to see whether the change itself is likely to have made a difference. As the PCC regime does not cover London or Scotland, it may be possible to make comparisons between forces that have a PCC and forces that do not. However, policing in Northern Ireland is so different in character from that in Great Britain that it cannot meaningfully be compared on the same criteria.

We examine previous attempts to measure police efficiency, which are beset by methodological difficulties. Inputs and outputs to measure police efficiency are difficult to measure because of the variety of work police are responsible for and carry out, and because crimes prevented are not measurable. As we can only measure the measurable, we present data on whether the introduction of PCCs have had any impact on public perception, feelings of safety and trust in the police forces.

The Labour government (1997-2010) had focused on community safety, and in its later years used levels of public confidence to measure police effectiveness. Effectiveness is, in principle, easier to assess than efficiency. There are two widely used sources: recorded crime statistics produced by the police, and the British Crime Survey. The former are subject to manipulation because of the degree of discretion in deciding what counts as a crime, and what counts as a clear-up. Recorded crime data are no longer certified as National Statistics by the UK Statistics Authority. Our results therefore use the British Crime Survey, but are subject to that survey's limitations.

We conclude that the introduction of PCCs has coincided with both a real-terms cut in police spending and an improvement in public satisfaction. But we are unable to show that the introduction of PCCs had any causal effect.

Measuring Police Effectiveness

Introduction: the policy context

Traditional policing occurs reactively, a response to whatever the current 'threat' may be. This means that resources are allocated in response to operational and political demands and public calls for service (DenHeyer 2014). In recent years, there have been efforts to direct resources to specific geographic areas of high crime or to specific crimes, and to apply intelligence-led targeted policing initiatives proactively (Innes 2011; Wilson & Weiss 2014). Demand for police services is rising but increased expenditure on resources is not feasible due to budgetary constraints, managing and allocating resources is crucial (Stockdale *et al.* 1999).

Researchers have traditionally found that the police utilised deterrence measures as their approach to crime control; via random foot patrols, emergency response, random stop-and-search processes, investigation and detection, etc., all of which are part of contemporary policing activity (Karn 2013). Increasingly, police forces are moving towards identifying and managing risk; shifting resources towards specific individuals (e.g. prolific offenders, repeat victims) or specific places (e.g. high crime areas or hot-spots) (Karn 2013).

As well as internal reforms and government budget cuts, socio-economic, demographic and technological changes also affect current patterns of crime which demand new responses from policing. The globalisation of markets for goods and services, the rapid expansion of new forms of communication, information technology and social media, the increase in personal mobility and migration, the growing income inequality and the fragmentation of families and communities are changing the patterns of crime globally that police officers face (Karn 2013). New threats create new forms of harm, particularly for the most vulnerable groups (e.g. children, migrants, the elderly, the poor). Identity theft, people trafficking and exploitation, investment scams and internet fraud and other emerging crimes present new challenges for the police, who are now required to work across local, regional and national boundaries to deal with criminal networks and changing modus operandi (Innes 2011; Karn 2013). The challenge facing police forces is to balance resources and service delivery levels with a decreasing level of funding and increasing expectations (Wilson & Weiss 2014).

In England & Wales, central government revenue grant to police authorities dropped by 22% between 2010 and 2015 (Ludwig & McLean 2016). At the same time, the UK (Coalition) government introduced wide-ranging reforms to police accountability, in particular the introduction of elected Police and Crime Commissioners (PCCs), and re-emphasised the police role as being first and foremost to fight crime. The previous (Labour) government had focused on community safety, and in its later years used levels of public confidence to measure police effectiveness.

Austerity means greater scrutiny of value for money, better evidence-based practice, and the reduction of long-term harm and demand through 'up-stream' intervention and prevention. The introduction of the College of Policing, which includes the assessment of police improvement through better use of research evidence and formal scrutiny of the

effectiveness of police forces has resulted since 2014 in Her Majesty's Inspectorate of Constabulary (HMIC)'s PEEL (police effectiveness, efficiency and legitimacy) inspection regime for all police forces in England and Wales. (Higgins & Hales 2016).

The policing mission is also being impacted by significant changes. As well as funding cuts police forces are dealing with new forms of governance and scrutiny, and there is a general shift away from volume crime reduction towards managing threat, risk, harm and vulnerability (Higgins & Hales 2016). Statistics show that many forms of recorded crime are falling (Farrell *et al.* 2010) but police workload is becoming more complex (College of Policing 2015); the internet has created new forms of crime and transformed old ones (McGuire & Dowling 2013) while growing international mobility, migration and more globalised markets have created new opportunities for criminals that manifest as harm in local communities. In some neighbourhoods global socio-economic factors have resulted in a number of factors which may lead to sectors of the population that are "*less visible to the police, more isolated, more difficult to engage and less capable of dealing with problems as a community*" (Higgins & Hales 2016).

Increasing decentralisation of police organisations means the manner in which resources are allocated between geographical areas for different services is of increasing importance. Previously, these allocations were negotiation or based on historical precedent. With increasing pressure for improved police accountability there is a need for police agencies to use justifiable methods to allocate resources (Schulenberg 2014), however limited amounts of information is available. How resources are allocated differs between jurisdictions (e.g. US forces allocate resources based on the number of calls for service, whereas in the UK it is based on the funding regime), making comparisons difficult (Loveday 2000; DenHeyer 2014). As a result of the changing environment, forces need to evidence transparent decisions, be able to evaluate outputs and outcomes, and demonstrate that resources are being used to generate the best returns for communities and society (DenHeyer 2009; 2014).

Previous attempts to measure efficiency

There have been numerous attempts to analyse crime using economic and econometric techniques. Dilulio (1996) argues that economists have not focused adequate attention on modelling crime or resource allocation through the use of sophisticated quantitative and modelling skills that are part of the economist's toolkit; instead remaining the domain of sociologists and criminologists who tend to use less sophisticated empirical analyses.

In a landmark, pioneer, contribution, Carr-Hill & Stern (1973; 1979) created a simultaneous equation model based on the premise that the demand for police services is partly determined by the crime rate, which in turn is affected by the level of police resources. This is a simultaneity issue in that the number of police officers can affect the level of crime and the level of crime can affect the number of police officers. Research which ignores this simultaneity problem falls foul of what we call the Carr-Hill-Stern test, and is therefore unreliable.

Benson and Rasmussen (1998) recommended the use of either time series or panel data as the foundation for a regression analysis model to develop a police resource allocation policy using an econometric approach. Official crime statistics (those reported and recorded by police) do not accurately reflect the actual number of crimes committed, but both cross sectional and time series analytical studies use this information as their basis. Inaccuracy can stem from the definition of crime, its interpretation and the administrative processes devised to record it (Weisburd & Eck 2004; DenHeyer 2014).

Proposed econometric models have been based on one dependent variable: the number of police officers, as a function of a number of different socio-economic and socio-demographic variables believed to be relevant to allocating resources. Developing an explanatory model through the construction of a regression equation will facilitate a better understanding of the situation under study and will allow experimentation with different combinations of inputs to examine and analyse their effects on the dependent variable (DenHeyer 2014).

The economic analysis of crime is concerned with the effect of incentives on criminal behaviour and the evaluation of alternative theoretical and operational strategies to reduce crime. Becker (1986) proposed that welfare maximising behaviour optimally allocates resources according to perceived returns, and links socio-economic conditions to an individual's expected return from legal and illegal activity. The economic literature focuses on the theoretical supply of offences in which crimes are related to the probability and the severity of punishment for the type of crime, the expected income from criminal activity, and perceived returns from alternative legal activities (DenHeyer 2014).

Stockdale *et al.* (1999) used criminological theory and economic techniques to assess the relative efficiency of police services. They concluded that as police services expand above a specific size, they typically encounter either diseconomies of scale, reduced technical efficiency, or a combination of the two. This was an extremely significant finding in terms of police organisational structure and resource allocation, and suggests that there is an optimum size of a police organisation, and this occurs at a relatively low minimum efficient scale or at a low number of police officers.

The use of the crime rate, for example, as a measure of the outcome of police activity, can be criticised on the grounds that it reflects only a small proportion of crime which actually takes place in a community (Mosher *et al.* 2002). Many studies (Coleman & Bottomley 1976; McCabe & Sutcliffe 1978; Weisburd & Eck 2004; Braga & Weisburd 2010) have indicated that recorded crime statistics are influenced by police discretion and their recording practices. The principal weakness in using aggregate data to analyse police effectiveness arises from the need to rely on gross data that is supplied by police (DenHeyer 2014). Difficulties in interpreting the research and the inconsistencies between the studies make it difficult to evaluate the aggregate approach in studying police effectiveness (DenHeyer 2014).

Current variables available for researching police efficacy have improved, due to increased desire to performance manage police force activities. Input variables are more diverse and more specific to policing than formerly, and variables relating to outcomes and outputs no longer rely on 'clear-up' rates. Crimes recorded and cleared have been used as indicators of

protection and crime prevention, however both are subject to measurement errors and manipulation (e.g. not all crimes are recorded) (Cameron 1989).

Official crime statistics are commonly used to judge police performance. However, crime is a social phenomenon and recording of criminal incidents is a cooperative venture between the police and the citizens. The level of cooperation varies from one area to the next and hence crime rates cannot be used to compare performance of police agencies in detecting and controlling crime (Verma & Gavirneni 2006). Criminal statistics are a function of the choices made by police (Carr-Hill & Stern 1979).

Police numbers and the level of crime

The relationship between police numbers and levels of crime have been debated for a number of years (see table 1). In 2008, an Australian study by Ogilvie and colleagues aimed to provide a comprehensive review of the literature on this topic. They focused on the impact of levels of police resources (mainly staff FTE) on crime activities (via recorded crime rates, rates of clearance, apprehension and arrest rates, and number of prosecutions) (Ogilvie *et al.* 2008). Their main conclusions were: police levels and crime rates are reciprocally related/ mutually interactive (a change in one may affect the other); an increase in the level of crime is correlated to an increase in police numbers; and there is no evidence which suggests that increasing police number effectively reduced crime (Ogilvie *et al.* 2008).

Many of these studies fail the Carr-Hill-Stern test. To avoid it, studies must assess policing before and after some exogenous shock, as in, for example, the introduction of a government initiative. Demonstrating an effect of police numbers on crime is also affected by the issue of causality – there are many issues that might affect both police numbers and crime including economic cycles or social change. Sherman and Eck (2002) did conclude that while there is consistent evidence that having *no* police (e.g. during strikes) significantly increases crime, the evidence of a marginal effect of increasing police numbers on crime is weak. Debate continues as to which social variables are associated with police staffing levels and which are associated with the level of crime (Den Heyer 2014).

The majority of studies which assess police performance have focussed on the impact of police activities on crime. Despite the multiple goals of policing and the complex nature of the relationships between police activities and crime rates, most of these efforts are postulated on a simple input-output relationship, which assumes a direct and simple relationship between policing and crime rates. In these studies, inputs typically include police budgets, number of personnel, and some type of-police strategy, e.g. patrol, criminal investigation or the use of technology. Typical output measures use official crime rates to measure the impact of police on crime (Murphy 1985).

Table 1: Previous research assessing the link between the number of police officers and levels of crime.

Author	Country	Study results
Cordner (1989)	USA	Investigated the relationship between police agency size and investigative effectiveness. Survey found no consistent variations in clearance rates by police agency size, it was reported that clearance rates decreased with increased investigative workloads.
Marvel & Moody (1996)	USA	Used police numbers per capita and crime rate to assess relationship. Found that causality worked in both directions, with a stronger effect of police numbers on crime. Also assessed relationship at city and state level and for a number of crime types. Statistical effects for homicide, robbery, burglary, auto crime and all crime were found to be significant.
Klinger (1997)	USA	Designed a police workload model to explain variations in police behaviour, used to explore the effects of increasing police numbers on crime levels. The level of police resources is a major determinant of how police respond to crime; as the workload increases there will be less time and resources available to respond to particular cases (and certain crimes may be prioritised). Main implications: workload effects could be reduced and more crimes could be cleared if police numbers are expanded.
Levitt (1998)	USA	The data used are a panel of 59 U.S. cities, with observations running from 1970-1992. Size of a police force systematically affects the willingness of crime victims to report crimes, and/or affects a police department's abilities to record crimes. It is assumed that an increase in police numbers increases the capacities of police to record crime, resulting in higher crime rates. This notion was not consistently supported. The size of variation is much more extreme in the USA.
Corman & Mocan (2000)	USA	Used two criminal-justice sanction variables: arrests for the specific crime (monthly data), and the number of police officers, from January 1970-December 1996 in New York City to provide new evidence on the relationship among crime, deterrence, and drug use. Used a regression analysis with lagged time effects and found significant deterrent effects of police numbers of robberies and burglaries, but not motor vehicle crime and homicide.
Kovandzic & Sloan (2002)	USA	Assessed impact of police numbers on crime - strength of association was small. Findings indicated that increased police levels led to lower rates of total crime - 10% increase in police levels reduced crime rates by 1.4% over time. Significant and substantial impacts of police levels of robbery, burglary and total crime. No effect on aggregate assault or murder.
Levitt (2002)	USA	Follow up to 1997 study which addresses criticisms by McRary (2002) over miscalculations. Used a two stage least square regression and found that original findings were still significant – there is a negative effect of police numbers on violent and property crimes. The impact of police on crime is estimated using two-stage least squares (2SLS) treating the police variables as endogenous and the other right-hand-side variables as exogenous

Zhao et al. (2003)	USA	Used police data from 4,482 cities and grant data was used to assess money spent on community schemes (including hiring – proxy measure of police numbers). Results suggest increasing police numbers was an effective method of increasing police arrests (particular offences), caution is advised (increased police numbers occurred within a framework of community oriented policing). Difficult to separate out the effects of increased police numbers from community policing activities.
DiTella & Schargrodsky (2004)	Argentina	The focus was to estimate the deterrent effect of police on car theft and to explore the internal validity of estimates. Utilised data on the location of car thefts before and after terror attacks. Found a large local deterrent effect of observable police on crime. No appreciable impact outside the narrow area in which the police were deployed was found.
Jackson & Boyd (2005)	USA	Examined the role of workload levels on police-recorded crime rates. Found that workload had an impact on police behaviour, police were more lenient when crime rates/workload increased. As workloads increases, crime prioritisation is likely to occur, therefore more police officers may produce increase crime clearance rates. Need to take into account that effectiveness of policing may vary between locations and crime types.
Klick & Tabarrok (2005)	USA	Looked at effect on crime of police mobilisation on ‘high alert’ days. The large declines in crime involving theft of and from cars support the idea that increased police presence reduces ‘street crimes’ during high-alert periods. Temporary increases in street police and CCTV are had less effect on homicide. An increase in police presence of 50% leads to a statistically and economically significant decrease in the level of crime on the order of 15%. Provide analyses that suggest this decrease is not due to changing tourism patterns induced by changes in the terror alert level.
Machin & Marie (2005)	UK	Used the introduction of the Street Crime Initiative in 2002 to assess the increased police presence and expenditure on robberies. Found it significantly reduced the number of robbers in the areas introduced.
Vollaard (2005)	Netherlands	Dutch study examining the effects of police numbers on crime reduction (via police and victimisation data between 1994 and 2003). In this time frame police personnel per capita grew by 20%, coinciding with the concurrent decline in crime rates. Increased police levels were associated with reductions in victimisation levels for violent crime, most types of property crime and nuisance (10%).
Chappell et al. (2006)	USA	Examined police organisational effects on officers’ arrests rates in the US. Utilised multivariate analysis to show that number of officers (per 100,000 residents) was negatively associated with arrest rates. Areas with fewer officers per population results in greater workload.
Paré et al. (2007)	Canada	Examined how crime clearance rates varied across communities in Canada. Found that crime workload had no effect on crime clearance rates. Found most criminal incidents remained unsolved (77.3%) and a significant degree of variation in the workloads of officers (17.5-85.1 crimes per officer). Community socio-economics found to influence workloads (poorer areas resulted in heavier workload). Police effectiveness varies according to the characteristics of places.

Draca et al. (2008)	UK	Used increased police presence post 7/7 in London to assess impact on crime rates. 'Susceptible' crimes (violence, sexual offences, theft and handling, robbery) fell significantly in the treatment areas.
Holmes et al. (2008)	USA	Examined the effects of demography in police resource allocations in the US. Results indicated that population/pop density were significantly/positively associated with the number of police officers and expenditure. Results highlight that demographic characteristics of communities must be considered to obtain a clear understanding of the effects of police numbers on crime.
Lin (2009)	USA	Used a two stage least squares (2SLS) and two-stage quantile (2SLAD) regression model and found significant effects between numbers of police and number of property crimes, murder, robbery, burglary and auto theft.
Vollaard & Koning (2009)	Netherlands	Found significant negative effects of higher police levels on property and violent crimes. Combined victimisation survey data and 'precaution taking' methods with data on police expenditure and numbers.

Cordner (1989) investigated the relationship between police agency size and investigative effectiveness (measure of police success). Looking at various sizes of police departments in the US, he assessed whether the consolidation of small departments into larger ones was beneficial to investigative effectiveness. Small agencies were presumed to lack expertise in special aspects of policing, have inadequate equipment, employ low quality personnel, and suffer from wasteful administrative duplication. Larger departments were thought to devote a smaller portion of their personnel to general patrol duties and employ more specialized operational strategies. Although the survey found no consistent variations in clearance rates by police agency size, it was reported that clearance rates decreased with increased investigative workloads. In addition, greater resources devoted to investigations was associated with more clearances per officer, while clearance rates decreased with increased investigative specialization (Cordner 1989).

Neither police agency size nor police officer workload was significantly related to investigative effectiveness in the state-wide multivariate analysis. Both variables had inverse bivariate relationships with clearance rates, but the relationships washed out in the multiple regression analysis. When separate analyses were conducted, agency size and officer workload were positively related to clearance rates for metropolitan area police agencies, but negatively associated with investigative effectiveness for nonmetropolitan agencies (Cordner 1989).

Corman and Mocan (2000) used monthly arrest data over a 25-year period from New York City to provide new evidence on the relationship between crime, deterrence and drug use. They consider that current arrests may be influenced by current criminal activity, creating a simultaneity bias "*if contemporaneous values of arrests are included in the crime equation*" (Corman & Mocan 2000). Overall, they found that murders, robberies, burglaries, and motor-vehicle thefts decline in response to increases in arrests; an increase in the size of the police force generates a decrease in robberies and burglaries (Corman & Mocan 2000). They noted that between 1970 and 1980, the police force of NYC decreased by about one third, but felony

arrests increased (approximately 5%). Simultaneously, arrests for misdemeanours decreased 40% and for violations also decreased over 80%. Police were able to relocate resources to combat the most serious crimes (Corman & Mocan 2000).

Vollaard and Koning (2009) argue that given the current reliance on police statistics as a source of crime data, evidence on the deterrent effect of police is limited mainly to crimes that are relatively well reported and well recorded (e.g. domestic burglary and theft of motor vehicles). To address endogeneity between police and crime in non-experimental data, they use the fact that the distribution of police resources across municipalities in the Netherlands is based on a specific funding formula which includes predictors of local police workload such as housing density and length of roadways (Vollaard & Koning 2009). Given the time needed to hire and train police personnel and the practice of smoothing year-to-year changes in local police resources, actual police levels differ from police levels prescribed by the funding formula. The difference between actual and prescribed police levels are used as a source of exogenous variation in police levels. They find significantly negative effects of higher police levels on property and violent crime, public disorder, and victim precaution (Vollaard & Koning 2009).

As noted, there is an endogeneity problem with many of these studies, which arises from the simultaneous determination of crime and police presence. Another problem with most previous studies is that they have examined the effects of police numbers on crime levels in isolation from other variables which affect policing and crime (e.g. organisational characteristics, community demographics and composition, deployment procedures and policies, demands of different types of crimes) (Ogilvie *et al.* 2008). It is not possible to fully comprehend the effect of police numbers on crime levels without considering many of these variables as well as the variety of activities police are involved in and engage in. Research has now moved away from this simple relationship and has begun to focus on *“how police resources may be used more effectively to reduce crime and how the effectiveness of officer behaviour varies according to policing activities”* as well as the social characteristics of specific areas (Ogilvie *et al.* 2008).

It has been suggested that due to the varying effects police have in different areas, characteristics of places (e.g. population demographics, socio-economic variables, and types of prevalent crimes) should be considered in models assessing police efficiency (Ogilvie *et al.* 2008). Variables such as the number of patrol cars, officers on the beat, the level of expenditure, etc. are not indicators of police effectiveness (Thanassoulis 1995), but indicators of budget priorities. Thanassoulis (1995) also states that *“merely spending more money on the police [...] does not [necessarily] translate into crime-control action”*, focus must be on what officers do in the field that ought to be measured. Ogilvie *et al.* (2008) suggest that efficient utility of police resources may have a significant impact on crime as some *“policing activities are more effective than others”*.

Some studies have also focused on the impact of one-off, large-scale changes in police deployment under unique circumstances (e.g. terror offences or police strikes) resulting in short, sharp increases or reductions in visible numbers. Evidence suggests that the large-scale

deployment of (temporary) additional officers following the bombings in London was associated with a reduction in crime (Draca *et al.* 2008). Given that such deployment patterns are short-term – and often related to other events – they provide limited evidence about the potential impact of smaller, marginal increases or decreases in police numbers (Braga & Weisburd 2010). There is little consensus amongst academics as to which social factors are related to criminal activity, how to model criminal activity or police resources appropriately, and which public policies serve to lessen criminal activity (Carr-Hill & Stern 1973; 1979; Benson *et al.* 1992; Weisburd & Eck 2004; Higgins & Hales 2016).

The number of arrests, response times, and reported levels of crime, are natural measures of effectiveness and tend to be the primary measures police utilise to appeal for expanded budgets (Carr-Hill and Stern 1973). Researchers almost always use either police expenditures or the number of police to measure police levels. The latter has a more direct relationship with the crime-reduction impact and because expenditures are sensitive to changes in budgeting classifications.

Data Envelopment Analysis (DEA) and attempts to measure Value for Money

The role of police in any given society is not defined clearly and police work covers a wide range of activities. Efficiency measurement also has to take into account that socio-economic and environmental factors influence the success of police work (PSPP 2000). Consequently, it is difficult to develop efficiency indicators, the main issue being how to quantify inputs and outputs of the police service, enabling Chief Constables to evaluate business decisions for resource allocation.

Furthermore, it is difficult to obtain data and to quantify the time spent by police on different activities (Cameron 1989). Traditional outputs related to police response (reactive policing) are often used in performance management studies due limited quantifiable data on non-accountable services. Many (Todd & Ramanathan 1994, Byrne *et al.* 1996, Drake & Simper 2002) have argued that even though much of the police's work cannot be measured, output and outcome measures can still be estimated. Stockdale *et al.* (1999) identified the *“growing need for the police to make resource allocation decisions transparent, to evaluate outputs and outcomes, and to demonstrate that resources are being used to generate the best returns”*.

In order to begin to assess efficiency, effectiveness and value for money in policing, academics have sought to utilise a number of analytical techniques previously used in other industries. Data envelopment analysis (DEA) can be used to measure the relative efficiency of decision making units (DMUs) within an organisation or industry utilising a range of inputs and outputs. DEA has been used since the late 1970s as a relative performance measure in public sector services such as education and health. Coined by Charnes *et al.* (1978), DEA is a linear programming non-parametric technique for constructing extremal piecewise frontiers as originally developed by Farrell (1957). DEA can be utilised in order to undertake a complete analysis of cost efficiency and its constituent components, allocative and technical efficiency

(Drake & Simper 2004). In traditional DEA models, these DMUs usually set their input and output targets in recognition of their autonomy separately (PSPP 2000, Fang 2013).

The research about resource allocation by DEA may be classified into two categories. One category assumes the efficiency of DMUs is constant (Yan *et al.* 2002, Korhonen & Syrjänen 2004, Amirteimoori & Shafiei 2006, Hadi-Vencheh *et al.* 2008) while the other assumes the efficiency of DMUs is changeable (Beasley 2003, Korhonen & Syrjänen 2004, Wu *et al.* 2008, Lozano *et al.* 2009).

When utilised in police efficiency studies it has also allowed variations to be applied to the weights on outcomes in a manner that was able to show police performance from its most favourable position. DEA is similar to ratio analysis, it uses paired data elements (input and output) and ranks the results in order of their relative performance (Nyhan & Martin 1999). DEA provides a single measure of efficiency based on the inclusion of a number of inputs and performance variables (or outputs) (Carrington *et al.* 1997). Output variables can be broadly defined to include measures of efficiency (output), quality and effectiveness (outcome).

DEA assigns optimal weights to all input and output variables, based on the analysis of the “*maximum weights [...] for which an individual DMU compares most favourably*” and minimum weights for those variables for which it compares least favourably (Nyhan & Martin 1999). This produces a single DEA score (or efficiency score), a scalar measure of performance for DMU included in the analysis between 0 and 1.00. A low score, close to 0 means that the DMU is inefficient compared to all the other DMUs in the analysis, and the opposite is true for a score close to 1.00. Frequently these variables are turned into percentages, where a score of 1.00 is 100% efficient (Nyhan & Martin 1999, Verma & Gavirneni 2006).

Estimates of efficiency have previously been divided into two major analytical groups: frontier (measures of absolute efficiency) and non-frontier models (measures of relative efficiency) (Aristovnik *et al.* 2013). Absolute efficiency is a measure of the maximum theoretically possible performance of a police force (e.g. resources are utilised in the best possible manner). Unfortunately, this is not measurable (PSPP 2000). Relative efficiency compares performance levels whilst recognising “*that even the best relative performers should not be standing still, but improving their performance over time*” (PSPP 2000). Despite limitations on evaluating police efficiency, a number of studies evaluating the efficiency of police performance employing non-parametric (non-frontier) methods such as DEA have been carried out.

Thanassoulis (1995) analysed 41 police forces in England and Wales, and was the first to apply DEA to measure performance of police efficiency at a regional level. Adopting an output-oriented model, he analysed three outputs (the number of clear-ups of violent crimes, burglaries, and other crimes) against four inputs (the police officers employed at each force, and the numbers of violent crimes, burglaries, and other crimes recorded). The research identified weaknesses in the accountability of performance due to omitted external factors, which may have led to dissimilarities in the identification of efficient performance units and their comparison with other peers (Aristovnik *et al.* 2014). Sun (2002), using the same inputs and outputs, measured the relative efficiencies of 14 police precincts in Taipei City, Taiwan.

Utilising DEA and a number of other analyses (window analysis, slack variable analysis), this study found no significant influence of the resident population and the location factor on police efficiency levels (Wu *et al.* 2010).

In Australia, Carrington *et al.* (1997) measured the technical efficiency of the New South Wales Police Service using a two-stage procedure: DEA to calculate the efficiency for all police precincts and regression analysis to analyse the external environment and operating factors. They found that on average better management could reduce input usage by 13.5%.

García-Sánchez (2007) evaluated the effectiveness of Spanish police forces using DEA, and divided overall effectiveness into two categories, according to the classification of police actions. These were investigative effectiveness (the actions the officers perform in order to determine those guilty of the offence committed or the disturbance organized) and coercive effectiveness (the capability of police organization to control the delinquents operating in their district, as well as the quality, accuracy and reserve of the investigative actions). Overall they found that the *“economic works on the functions of police production can mainly be found in the empirical area and can be classified into two categories: those that attempt to test the postulates of the economics of crime through non-frontier methods; and those that concentrate on evaluating efficiency by means of frontier techniques”* (García-Sánchez 2007).

In the UK, Drake and Simper have published a number of related articles covering a 10-year period (2000, 2001, 2002, 2003, 2004). Building on Thanassoulis' work, they mitigated the limitation bias on relative efficiency scores by including *“environmental, socioeconomic, and demographic variables in their comparative analysis”* (Drake & Simper 2005). With a two-stage procedure the authors verified that the exclusion of external factors may lead to inaccurate efficiency scores with respect to some police units of analysis. To quantify the role of a given police force in its society, Drake and Simper (2004) identified outputs covering a wide range of policing functions, which they divided into three groups: *managerial* efficiency outcomes, *response/reactive* outcomes, and *proactive/preventative* outcomes.

In their first study, Drake and Simper found an interesting dichotomy – the levels of *pure technical efficiency (PTE)* appear to decline with police force size, however there was evidence of an inverted U-shaped relationship with respect to *scale efficiency (SE)* (Drake & Simper 2000). This was particularly noticeable for the Metropolitan police force which had an SE score of 57.5, the lowest of all forces; but a PTE score of 100 in each of the years they studied suggesting that it is a highly efficient police force with no obvious inefficiencies in resource utilisation. They concluded that *“given that the Met is the largest force in the country, this result strongly suggests that there are significant diseconomies of scale at work with respect to large police force operations. As in other large organizations, this is probably attributable to the extra bureaucracy and layers of management structure that tend to accompany large scale operations”* (Drake & Simper 2000). Similar findings in their study the following year appear to confirm this finding. They suggest that *“in terms of police force structure and resource allocation there is an optimal scale for police forces and that this occurs at a relatively low minimum efficient scale. It also suggests that attempts to utilise resources and improve*

efficiency by merging police forces is likely to be counter-productive for all but the smallest police forces” (Drake & Simper 2001).

The 2001 study identified the West Midlands as the least efficient police force; it is 79% less efficient than *“its efficient reference set forces in terms of translating its available resources into the specified outputs”* (Drake & Simper 2001). They further identified that the bulk of the inefficiency is caused by a failure to operate under constant returns to scale (PTE = 64.3 and SE = 33.3). Comparing their DEA scores and the Audit Commission scores of efficiency, they conclude that indicators in which performance is measured by single, non-joint, survey data which exclude the economic transformation of inputs into outputs, can produce misleading results (Drake & Simper 2001).

The difficulty in utilising DEA in policing efficiency studies has been in determining accurate input and output variables. Distinct categories of inputs outlined by the Chartered Institute of Public Finance and Accountancy (CIPFA) Police Force Statistics include: employment costs, premises-related expenses, transport-related expenses and capital and other costs. The output variable sets that have been utilised also show that the modelling of police force efficiency is far from straightforward (see Table 2). Due to the large range of services police forces provide, it is difficult to identify an appropriate output variable.

DEA has not appeared since 2005 as a technique to measure police efficiency, and the main UK authors (Drake & Simper) have moved on to use this technique in other sectors. After thorough consideration, we have determined not to use DEA due to the high fluctuations of results on a yearly basis between different force data.

There is a plethora of indicators and information about police outputs and outcomes. But, to date, it has not been possible to draw this information together to build a comprehensive or systematic measure of relative police *efficiency* in meeting their ultimate objectives of promoting safety and reducing crime, disorder and the fear of crime (PSPP 2000).

Table 2: Inputs and Outputs in Policing Studies.

Study	Inputs	Outputs	Country
Law Enforcement Agencies as Multiproduct Firms (Darrough & Heineke 1979)	1. Weighted average of all police wages.	1. Burglary clearances. 2. Robbery clearances. 3. Motor vehicle theft clearances. 4. Larceny clearances. 5. Total number of crimes against the person. 6. Population.	USA
Assessing police forces in England and Wales using DEA (Thanassoulis 1995)	1. Number of violent crimes. 2. Number of burglaries. 3. Number of other crimes. 4. Number of officers.	1. Violent crime clear up rate. 2. Burglary clear up rate. 3. Other crime clear up rate.	England
Performance Measurement in Government Service Provision (Carrington et al. 1997)	1. Number of police officers. 2. Number of civilian employees. 3. Number of police cars.	1. km travelled by police cars. 2. Responding to offences recorded. 3. Number of summons served. 4. Number of major car accidents attended.	Australia
Assessing the Performance of Municipal Police Services Using DEA (Nyhan & Martin 1999)	1. Total department costs. 2. Total FTE.	1. Number of crime report clearances. 2. Response time to call out. 3. Crime rate.	USA
Productivity estimation and the size-efficiency relationship in English and Welsh police forces (Drake & Simper 2000)	1. Total employment costs. 2. Premises related costs. 3. Transport related costs. 4. Capital and other costs.	1. Total crime clear up rate. 2. Total number of traffic offences. 3. Total number of breathalyser tests.	England
The Economic Evaluation of Policing Activity (Drake & Simper 2001)	1. Total employment costs. 2. Premises related costs. 3. Transport related costs. 4. Capital and other costs.	1. % of time officers spend patrolling beat. 2. Violent crime clear up rate. 3. Burglary clear up rate. 4. % success rate in answering 999 call. 5. % of officers arriving at a scene within a specified response time.	England

X-efficiency and scale economies in policing (Drake & Simper 2002)	<ol style="list-style-type: none"> 1. Total employment costs. 2. Premises related costs. 3. Transport related costs. 4. Capital and other costs. 	<ol style="list-style-type: none"> 1. Clear up rate. 2. Total number of traffic offences. 	England
Measuring the relative efficiency of police precincts using data envelopment analysis (Sun 2002)	<ol style="list-style-type: none"> 1. Number of crimes recorded. 2. Number of FTE police officers. 3. Number of civilian employees. 4. Level of expenditure. 5. Capital equipment used. 6. Other inputs. 	<ol style="list-style-type: none"> 1. Number of crime clear ups. 2. Number of non-crime activities recorded. 3. Number of police activities to prevent crime and investigate criminal cases. 4. Other outputs. 	Taiwan
The efficiency of the Spanish police service (Diez-Ticio & Mancebon 2002)	<ol style="list-style-type: none"> 1. Capital (number of vehicles). 2. Labour (number of officers). 	<ol style="list-style-type: none"> 1. Property clear-up rate. 2. Violent clear-up rate. 	Spain
The measurement of English & Welsh police force efficiency (Drake & Simper 2003)	<ol style="list-style-type: none"> 1. Total employment costs. 2. Transport related costs. 3. Capital and other costs. 	<ol style="list-style-type: none"> 1. Total number of cleared up crimes. 2. Number of cleared up violent crimes. 3. Number of cleared up burglaries. 4. Total breathalyser tests. 	England
The Economics of Managerialism and the Drive for Efficiency in Policing (Drake & Simper 2004)	<ol style="list-style-type: none"> 1. Total employment costs. 2. Premises related costs. 3. Transport related costs. 4. Capital and other costs. 	<ol style="list-style-type: none"> 1. Number of complaints per officer. 2. Average number of days lost per officer. 3. Number of crimes solved. 4. No. of emergency calls answered in target time. 5. Number of breathalyser tests. 	England
The measurement of police force efficiency (Drake & Simper 2005)	<ol style="list-style-type: none"> 1. Number of offences. 2. Net budget revenue. 	<ol style="list-style-type: none"> 1. Total offences cleared. 	England
Police Efficiency in Offences Cleared (Drake & Simper 2005)	<ol style="list-style-type: none"> 1. Number of offences. 	<ol style="list-style-type: none"> 1. Total offences cleared. 	England
Measuring police efficiency in India (Verma & Gavirneni 2006)	<ol style="list-style-type: none"> 1. Total expenditure. 2. Number of police officers. 3. Number of cases investigated. 4. Total reported crime. 	<ol style="list-style-type: none"> 1. Number of persons arrested. 2. Number of persons charged. 3. Number of persons convicted. 4. Number of trials completed. 	India

Evaluating the effectiveness of the Spanish police force through DEA (García-Sánchez 2007)	1. Number of offences	1. Number of offences solved. 2. Number of individuals arrested and taken to court	Spain
Evaluating US state police performance using DEA (Gorman & Ruggiero 2008)	1. Number of sworn officers 2. Number of other employees 3. Number of vehicles.	1. Murder rate 2. Other violent crime rate. 3. Total property crime rate.	USA
Measuring the efficiency of local police force (García-Sánchez 2009)	1. Total number of police officers. 2. Capital assets (total number of vehicles.	1. Km travelled by police vehicles 2. No. of arrested taken before the court 3. Number of accusations formulated. 4. Number of breathalyser tests carried out. 5. Number of vehicles removed from roads. 6. Number of accident reports drawn up.	Spain
Measuring the performance of police forces in Taiwan using DEA (Wu <i>et al.</i> 2010)	1. Labour cost. 2. General operating costs. 3. Equipment purchasing costs.	1. Number of crimes cleared up. 2. No. of RTAs resulting in death or serious injury. 3. No. of emergency calls. 4. Satisfaction levels with public security.	Taiwan
Evaluation of the efficacy and effectiveness of the Spanish security forces (García-Sánchez <i>et al.</i> 2013)	1. Number of crimes committed. 2. The operational efficacy in the previous years. 3. The variation in the no. of crimes.	1. Number of arrests. 2. The inverse of the variation of crimes committed 3. Changes in the population, economic activity rates and tourism index over 3 years.	Spain
Yearly evolution of police efficiency in Spain & explanatory factors (García-Sánchez <i>et al.</i> 2013)	1. Total no. of police officers. 2. Total number of vehicles.	1. Percentage of solved crimes.	Spain
Relative efficiency of police directorates in Slovenia (Aristovnik <i>et al.</i> 2013)	1. Number of criminal offenses. 2. Population. 3. Violations of public order regulations. 4. Violations detected during road traffic controls.	1. Investigative and other measures taken while investigating criminal offenses. 2. Police measures against offenders. 3. Police measures applied during road traffic controls.	Slovenia

Regression results

Our review has shown that measuring police effectiveness, efficiency, and value for money is important, urgent, and difficult. It is important because as with any other public service, policy-makers want to know whether money is well spent when it could be spent on many other good things, or not spent at all and tax reduced instead. It is urgent in any context where spending on policing and/or crime are decreasing. It is difficult for multiple reasons:

- An important output of policing is crime and disorder which do not happen because of the police presence. How can we ever know what would have happened absent the police presence?
- Recorded crime statistics are unreliable because police officers and authorities have both the motive and the opportunity to manipulate the numbers;
- Carr-Hill and Stern's (1973, 1979; hereafter CHS) challenge is rarely dealt with: increasing police numbers may lead to an increase in reported crime, because more police record more crimes. A weaker version of the CHS critique is that police numbers and crime numbers may co-vary, both relating to some common underlying factor;
- Although sophisticated statistical techniques (notably DEA) have been applied to the problem, they produce very unstable results.

Therefore, we have restricted ourselves to some plain vanilla tests using standard techniques and exploiting the natural experiment resulting from 41 police force areas (PFAs) introducing PCC elections in 2012 while the Metropolitan and City of London police forces did not.

We employ difference-in-differences, fixed-effects, panel regressions using the Metropolitan and City of London police forces as the control groups. Fixed-effects regressions control for differences across geography and time. For example, Bedfordshire and Surrey in 2013 have different ratings, while Surrey in 2012 would differ from Surrey in 2013. We isolate any change in rating metrics to the introduction of PCC elections. We cluster the standard errors on the police force area level, given that treatment occurs by PFA. Moreover, controlling for geography encompasses additional place-based factors affecting crime such as poverty level, educational attainment, and ethnic/religious diversity. Time trends account for factors that affect the nation as a whole, such as the recession beginning in 2007.

Expanded models include the covariate most likely to predict police force ratings: police spending. We utilize net revenue expenditures per head of population within the police

force area, which take into account formula spending, specific funding, council tax, and reserves.¹

We first list the summary statistics for police spending and all ten separate dependent variables for reference in Table 3. A cursory view of the standard deviations provides evidence of fairly heterogenous distributions of the data, minus “Percent Agree Police Treat You with Respect.” Its high mean and median combined with unusually small standard deviation suggest a very narrow distribution, providing less variation and therefore weaker models for differences-in-differences regressions.

Table 3. Police Force Rating Descriptive Statistics²

Rating	Mean	Median	Standard Deviation	Source	Years
Net Revenue Per Head of Population	196.1	185.2	39.2	HMIC	2010-2016
Percent of Victims Satisfied	86.1	85.6	4.03	HMIC	2010-2016
Percent Agree Police Doing Excellent/Good Job	57.6	57.6	6.6	CSEW	2007-2015
Percent Agree with Police Dealing with Local Concerns	58.1	58.3	5.8	CSEW	2008-2015
Percent Agree Police Can Be Relied Upon When Needed	54.6	55.2	6.5	CSEW	2009-2015
Percent Agree Police Treat You with Respect	86.1	86.2	2.0	CSEW	2009-2015
Percent Agree Police Treat Everyone Fair	66.1	66.2	4.0	CSEW	2009-2015
Percent Agree Police Understand Issues Affecting Community	70.3	70.4	4.4	CSEW	2009-2015
Percent Agree Taking Everything Into Account Have Confidence in Police	73.0	73.4	5.2	CSEW	2009-2015
Percent Perceive Local Crime Has Not Gone Up	65.0	65.0	7.2	CSEW	2009-2015
Percent Confident Police Effective at Catching Criminals	65.3	65.6	4.8	CSEW	2009-2015

¹ We used models with both nominal and real values, adjusting for inflation using the OECD consumer price index for the United Kingdom. The spending coefficient was slightly muted with inflation-adjusted figures, but was largely insignificant in most models as well. Ultimately, we settled on nominal values, considering the time fixed-effects account for inflation over time and affect the nation as a whole.

² The observation count ranges from 294 to 378, depending on the metric and its years available, covering 42 police force areas (Metropolitan and City of London are combined in all but one metric).

For the regression results, it is worth noting the ratings were coded on a 100-point scale. That is, 62% of a police force area agreeing with said metric is coded as 62 in the dataset. Thus, a PCC coefficient of say +2.0 refers to a 2-percentage point increase of that metric with the introduction of PCC elections, or an increased rating of 64%. This makes the results easily interpretable.

Largely, the results, shown in Table 4, are insignificant, and the few statistically significant results provide a more negative narrative. We list a limited number of the metrics, but the remainder can be found in the data appendix. Victim satisfaction, overall confidence in the police force, and those agreeing their police force is good or excellent all appear to be unchanged by the introduction of PCC elections. The percentage agreeing that police treat everyone fairly experienced a slight uptick of 1%, yet this result dissipates once we control for police force spending.

Interestingly, we find statistically significant decreases in the percent of the population perceiving local crime has *not* gone up and percent agreeing that police are dealing with local concerns, which are robust to models with spending controls. That is, after controlling for spending, the introduction of PCC elections resulted in about a 6% decrease in those perceiving local crime has gone down or remained stable and a 3% decrease in those agreeing that police are dealing with local concerns. In other words, PCC elections may in fact be making crime a more salient topic among public discourse, resulting in higher perceptions of crime and more discontent with addressing said crime.

Additionally, the coefficients for police spending has an insignificant effect on all ten metrics, suggesting from the data that increased spending does not lead to increased citizen satisfaction in each police force area.

Taken altogether, the mixed results in Table 4 provide no statistical claim for PCC elections resulting in higher citizen satisfaction of their police forces, and in fact per the data currently available suggests PCC elections have made crime rise to the forefront of public consciousness. Larger claims may be made as PCC elections become a more regular pattern for voters.

The final two models investigate the democratic mechanisms of PCC reforms, asking if ratings affect turnout or vice versa.

It would be an interesting democratic argument if for instance higher (lower) ratings preceding an election led to lower (higher) turnout, or if higher (lower) turnout led to higher (lower) ratings after an election. Recent literature suggests the former, that higher citizen satisfaction results in lower turnout (Ezrow & Xezonakis 2016). Because PCC elections have only occurred twice, the observations are limited of course.

Table 4. Effect of PCC Reforms on Police Force Ratings

Rating	Percent of Victims Satisfied	Percent of Victims Satisfied	Percent Perceive Local Crime Has Not Gone Up	Percent Perceive Local Crime Has Not Gone Up	Percent Agree with Police Dealing with Local Concerns	Percent Agree with Police Dealing with Local Concerns
PCC Reform	0.672 (2.666)	0.762 (2.356)	-3.106*** (0.459)	-5.699** (1.995)	-0.791* (0.346)	-3.074* (1.444)
Net Revenue Per Head of Population	—	-0.00228 (0.0341)	—	0.0480 (0.0348)	—	0.0442 (0.0272)
Constant	84.59*** (0.298)	85.08*** (7.223)	55.35*** (0.493)	50.09*** (7.276)	51.22*** (0.380)	46.14*** (5.816)
Time/Place Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
Rating Source	HMIC	HMIC	CSEW	CSEW	CSEW	CSEW
Years	2010-2016	2010-2016	2009-2015	2010-2015	2008-2015	2010-2015
R ²	0.75	0.75	0.87	0.84	0.83	0.78
N	299	299	294	252	336	252

Table 4 (continued) Effect of PCC Reforms on Police Force Ratings

Rating	Percent Agree Police Doing Excellent/Good Job	Percent Agree Police Doing Excellent/Good Job	Percent Agree Police Treat Everyone Fair	Percent Agree Police Treat Everyone Fair	Percent Agree Taking Everything Into Account Have Confidence in Police	Percent Agree Taking Everything Into Account Have Confidence in Police
PCC Reform	0.742 (0.372)	-1.741 (2.058)	1.041* (0.398)	-0.407 (1.582)	0.633 (0.313)	-0.861 (1.395)
Net Revenue Per Head of Population	—	0.0420 (0.0388)	—	0.0335 (0.0287)	—	0.0232 (0.0276)
Constant	51.41*** (0.287)	46.81*** (8.225)	65.92*** (0.272)	58.71*** (6.159)	67.43*** (0.344)	64.31*** (5.886)
Time/Place Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
Rating Source	CSEW	CSEW	CSEW	CSEW	CSEW	CSEW
Years	2007-2015	2010-2015	2009-2015	2010-2015	2009-2015	2010-2015
R ²	0.89	0.85	0.71	0.72	0.88	0.87
N	378	252	294	252	294	252

For observing how ratings affected turnout, we averaged a PFA's ratings for all years before and including 2012 and also for 2013 to the most recent year available for the rating. Since Greater Manchester Police are phasing out the PCC role in 2017 and did not have 2016 elections, they were excluded, bringing the observation count to 40 for each election (80 total). The summary statistics are shown in Table 5 for reference.

Table 5. PCC Election Turnout Descriptive Statistics by Year

Year	Mean	Median	Standard Deviation
2012	15.1	14.8	1.8
2016	26.4	24.1	6.7

The models, listed in Table 6, had high R^2 values (all about 0.78), but the coefficients for ratings showed that they had a statistically insignificant effect on turnout. That is, using multiple metrics for how constituents viewed their police force, higher or lower ratings did not appear to affect their likelihood of actually voting. The strongest coefficient for all models was the dummy variable for 2016 observations, meaning the largest factor for increased turnout was simply the second iteration of PCC elections.

Table 6. Effect of Pre-2012/Pre-2016 PFA Ratings on 2012/2016 PCC Election Turnout

	Turnout Percentage	Turnout Percentage	Turnout Percentage	Turnout Percentage	Turnout Percentage
Rating	0.271 (0.502)	0.0216 (0.334)	0.0457 (0.387)	-0.0375 (0.288)	-0.333 (0.833)
Constant	-7.972 (42.64)	13.87 (18.63)	12.50 (21.78)	17.00 (14.82)	43.61 (71.60)
Time/Place Fixed Effects	Yes	Yes	Yes	Yes	Yes
Rating Source	HMIC Percent of Victims Satisfied	CSEW Percent Agree Police Doing Excellent/Good Job	CSEW Percent Agree with Police Dealing with Local Concerns	CSEW Percent Agree Police Can Be Relied Upon When Needed	CSEW Percent Agree Police Treat You with Respect
Years	2010-2016	2007-2015	2008-2015	2009-2015	2009-2015
R ²	0.78	0.78	0.78	0.78	0.78
N	80	80	80	80	80

Table 6 (continued) Effect of Pre-2012/Pre-2016 PFA Ratings on 2012/2016 PCC Election Turnout

	Turnout Percentage	Turnout Percentage	Turnout Percentage	Turnout Percentage	Turnout Percentage
Rating	-0.315 (0.380)	-0.0274 (0.475)	0.0758 (0.477)	-0.446 (0.336)	0.499 (0.510)
Constant	36.04 (25.32)	16.97 (33.01)	9.675 (33.89)	42.32* (20.38)	-16.63 (32.55)
Time/Place Fixed Effects	Yes	Yes	Yes	Yes	Yes
Rating Source	CSEW Percent Agree Police Treat Everyone Fair	CSEW Percent Agree Police Understand Issues Affecting Community	CSEW Percent Agree Taking Everything Into Account Have Confidence in Police	CSEW Percent Perceive Local Crime Has Not Gone Up	CSEW Percent Confident Police Effective at Catching Criminals
Years	2010-2016	2007-2015	2008-2015	2009-2015	2009-2015
R ²	0.78	0.78	0.78	0.79	0.78
N	80	80	80	80	80

The second model reverses the causal direction, asking if turnout affects how citizens rate their police force (e.g. higher involvement in PCC elections begets greater community buy-in and therefore higher ratings). It is shown in Table 7.

Because we only have one election with post-election ratings (2012, and average PFA ratings for years 2013 to the most recent year available), that limits the observations to 40. Without a time component, these models are basic, run of the mill OLS regressions with limited statistical power, which can be observed in their extremely small R^2 values.

The coefficients are again statistically insignificant, meaning that higher turnout in 2012 did not lead to higher PFA ratings in the years following.

Overall, both of these results, PFA's pre-election ratings' effect on election turnout or election turnout on PFA's post-election ratings, suggest a nearly non-existent link between how voters view their police force and the process of voting. Given we have only observed two cycles of PCC elections, this democratic mechanism may improve as voters acclimate themselves to electing police crime commissioners in charge of their local police forces and as more data becomes available.

Conclusion. Observable implications and further work.

In an earlier paper, written before we had full access to the CSEW data, we reported two relevant findings:

- we undertook textual analysis of all the successful 2012 manifestoes. We wished to see whether Conservative and Labour candidates stressed different themes, and whether political-party winners stressed different themes to successful independent candidates. The results were null. There was no measurable difference among the three classes of successful candidates. In political science jargon, the initial PCC elections were 'valence' not 'position' elections.
- Using the Metropolitan Police, which has no PCC, as a benchmark, we compared the HMIC ratings of forces for the six years (2007-2012) before the introduction of PCCs and the three years (2013-15) since their introduction for which data are currently available. We found a statistically significant increase in police force rating, about 6 percentage points increase, from the PCC introduction ($R^2=0.65$). (McLean *et al.* 2016, pp. 12-13).

Table 7 Effect of 2012 PCC Election Turnout on 2013-2016 PFA Ratings

	Percent of Victims Satisfied	Percent Agree Police Doing Excellent/Good Job	Percent Agree with Police Dealing with Local Concerns	Percent Agree Police Can Be Relied Upon When Needed	Percent Agree Police Treat You with Respect
Turnout	-0.302 (0.219)	-0.191 (0.434)	-0.201 (0.356)	-0.550 (0.346)	0.193 (0.138)
Constant	91.85*** (3.326)	64.71*** (6.586)	64.40*** (5.398)	66.98*** (5.247)	83.90*** (2.100)
Time/Place Fixed Effects	Yes	Yes	Yes	Yes	Yes
Rating Source	HMIC	CSEW	CSEW	CSEW	CSEW
Years	2012-2016	2012-2015	2012-2015	2012-2015	2012-2015
R ²	0.05	0.01	0.01	0.06	0.05
N	40	40	40	40	40

Table 7 (continued) Effect of 2012 PCC Election Turnout on 2013-2016 PFA Ratings

	Percent Agree Police Treat Everyone Fair	Percent Agree Police Understand Issues Affecting Community	Percent Agree Taking Everything Into Account Have Confidence in Police	Percent Perceive Local Crime Has Not Gone Up	Percent Confident Police Effective at Catching Criminals
Turnout	0.297 (0.319)	-0.0465 (0.294)	-0.154 (0.344)	0.442 (0.400)	0.0431 (0.267)
Constant	61.24*** (4.832)	72.71*** (4.458)	77.97*** (5.215)	64.12*** (6.067)	67.24*** (4.054)
Time/Place Fixed Effects	Yes	Yes	Yes	Yes	Yes
Rating Source	CSEW	CSEW	CSEW	CSEW	CSEW
Years	2012-2016	2012-2015	2012-2015	2012-2015	2012-2015
R ²	0.02	0.001	0.01	0.03	0.001
N	40	40	40	40	40

Some of the questions in CSEW potentially reveal sensitive information about individuals (although none of the ones that we wanted to use did). Therefore, although our only interest in more detailed data was to get reliable disaggregation to the level of each PFA, we had to undergo screening for access to sensitive data. This process took some months (and will do for follow-up researchers). The results we now present therefore complete the work we have been able to do in the lifetime of our project. Readers may be disappointed that most of our results are null. But statistical insignificance is not policy insignificance. We therefore conclude by underlining the policy significance of our work, and offering some pointers for future research.

1. The introduction of PCCs in 2012 had very modest results. Although HMIC ratings of forces with PCCs have been improving, compared to the ratings of the PCC-less Metropolitan Police, this result could just as easily be read as the relative deterioration of one force, rather than the relative improvement of 41. PCC candidates did not differ in their policy approaches – the election is a ‘valence’ one (who will do the best job?) rather than a position one (what is your policing priority?). We found no effect, however, from the performance of forces between 2012 and 2016, and the political outcome of the 2016 PCC election.
2. Measuring efficiency of police forces is very difficult, and may be a blind alley despite the very substantial resources put into it by many research teams in many jurisdictions. Fundamental problems include: the non-measurability of crime prevented; the Carr-Hill-Stern observation that more police leads to more reported crime, and the prevalent problem of showing direction of causation. Our own results tend to confirm this. Although satisfaction with police services is increasing, the imminence of PCC elections seems to have been associated with increased perception of local crime – as if, plausibly, the very fact of the election leads more people to talk about crimes.
3. Measuring effectiveness, on the other hand, is more promising. As conventionally defined, effectiveness is a measure of outputs achieved per input. One relevant output is the degree of public belief in their security. This may be secured by unmeasurable efficiency improvements (for instance, increasing the number of crimes prevented), but is in itself measurable. Here, the picture looks relatively rosy for UK policing. Inputs have sharply declined, see e.g., Ludwig and McLean 2016b, Table 1. But the output measures of public satisfaction, and of HMIC ratings, have not declined. It is thus tentatively possible to conclude that the effectiveness of policing in England & Wales has improved since 2010.

Further research is needed. It always is. Our attempt to extend the analysis to Scotland was defeated by data limitations. This is unfortunate for at least three reasons. Scotland has no PCC. It has a single force, recently, and controversially, amalgamated from eight. And the Scottish Government has put relatively less resources into policing than has the UK

Government. That is an obvious direction for further research, as are broader international comparisons. We do not pretend to have exhausted the subject.

References

- Amirteimoori, A. and M. Shafiei (2006). "Characterizing an equitable omission of shared resources: a DEA-based approach." Applied Mathematics and Computation **177**(1): 18-23.
- Andersen, P. and N. C. Petersen (1993). "A Procedure for Ranking Efficient Units in Data Envelopment Analysis." Management Science **39**(10): 1261-1264.
- Aristovnik, A., J. Seljak and J. Mencinger (2013). "Relative efficiency of police directorates in Slovenia: A non-parametric analysis." Expert Systems with Applications **40**: 820-827.
- Aristovnik, A., J. Seljak and J. Mencinger (2014). "Performance measurement of police forces at the local level: A non-parametric mathematical programming approach." Expert Systems with Applications **41**: 1653.
- Beasley, J. E. (2003). "Allocating fixed costs and resources via DEA." European Journal of Operational Research **147**(1): 198-216.
- Becker, G. S. (1968). "Crime and punishment: an economic approach." Journal of Political Economy **16**(2): 169-217.
- Benson, B. L., I. Kim and D. W. Rasmussen (1994). "Estimating Deterrence Effects: A Public Choice Perspective on the Economics of Crime Literature." Southern Economic Journal **61**(1): 161-168.
- Benson, B. L. and D. W. Rasmussen (1998). "Deterrence & public policy: Trade-offs in the allocation of police resources." International Review of Law and Economics **18**: 77-100.
- Benson, B. L., D. W. Rasmussen and D. L. Sollars (1995). "Police bureaucracies, their incentives, & the war on drugs." Public Choice **83**(1-2): 21-45.
- Bradford, B. (2011). *Police Numbers and Crime Rates – A Rapid Evidence Review*. London, HMIC.
- Braga, A. and D. Weisburd (2010). Policing problem places: Crime hot spots and effective prevention. New York, Oxford University Press.
- Byrne, D., H. Dezhbakhsh and R. King (1996). "Unions and Police Productivity." 35 **4**(556-584).
- Cameron, S. (1989). "Police cost function estimates for England and Wales." Applied Economics **21**: 1279-1289.
- Carr-Hill, R. A. and N. H. Stern (1973). "An econometric model of the supply and control of recorded offences in England & Wales." Journal of Public Economics **2**: 289-318.
- Carr-Hill, R. A. and N. H. Stern (1979). Crime, the police and criminal statistics: An Analysis of Official Statistics for England and Wales using Econometric Methods. London, Academic Press.

Carrington, R., N. Puthycheary, D. Rose and S. Yaisawarng (1997). "Performance Measurement in Government Service Provision: The Case of Police Services in New South Wales." Journal of Productivity Analysis **8**: 415-430.

Chappell, A. T., J. M. MacDonald and P. W. Manz (2006). "The Organizational Determinants of Police Arrest Decisions." Crime & Delinquency **52**(2): 237-306.

Charnes, A., W. W. Cooper and E. Rhodes (1978). "Measuring the efficiency of decision making units." European Journal of Operational Research **2**: 429-444.

Coleman, C. and A. Bottomley (1976). "Police conceptions of crime and no crime." Criminal Law Review: 344-360.

College of Policing (2015). College of Policing analysis: Estimating demand on the police service.

Cordner, G. W. (1989). "Police agency size and investigative effectiveness" Journal of Criminal Justice **17**: 145-155.

Corman, H. and H. N. Mocan (2000). "A Time-Series Analysis of Crime, Deterrence and Drug Abuse in New York City." The American Economic Review **90**(3): 584-604.

Darrough, M. N. and J. M. Heineke (1979). "Law Enforcement Agencies as Multiproduct Firms: An Econometric Investigation of Production Cost." Public Finance **34**: 176-195.

Den Heyer, G., M. Mitchell, S. Ganesh and C. Devery (2008). "An econometric method of allocating police resources." International Journal of Police Science & Management **10**(2): 192-213.

Den Heyer, G. (2009). Econometric modelling and analysis to support operational policing: A New Zealand police case study. Saarbrücken, Germany, VDM Verlag Dr. Müller.

Den Heyer, G. (2014). "Examining police strategic resource allocation in a time of austerity." Salus Journal, Special Issue on Policing and Security in Practice **2**(1): 63-79.

Diez-Ticio, A. and M.-J. Mancebon (2002). "The efficiency of the Spanish police service: an application of the multiactivity DEA model." Applied Economics **34**(3): 351-362.

Dilulio, J. J. (1996). "Help Wanted: Economists, Crime and Public Policy." Journal of Economic Perspectives **10**(1): 3-24.

DiTella, R. and E. Schargrotsky (2004). "Do police reduce crime? Estimates using the allocation of police forces after a terrorist attack." American Economic Review **94**(1): 115-133.

Draca, M., S. Machin and R. Witt (2008). Panic on the Streets of London: Police, Crime and the July 2005 Terror Attacks. CEP Discussion Paper No 852, LSE Centre for Economic Performance.

Drake, L. and R. Simper (2001). "The Economic Evaluation of Policing Activity: An Application of a Hybrid Methodology." European Journal of Law and Economics **12**: 173-192.

Drake, L. M. and R. Simper (2000). "Productivity estimation and the size-efficiency relationship in English and Welsh police forces: An application of data envelopment analysis and multiple discriminant analysis." International Review of Law and Economics **20**: 53-73.

Drake, L. M. and R. Simper (2002). "X-efficiency and scale economies in policing: a comparative study using the distribution free approach and DEA." Applied Economics **34**: 1859-1870.

Drake, L. M. and R. Simper (2003). "The measurement of English & Welsh police force efficiency: a comparison of distance function models." European Journal of Operational Research **147**: 165-186.

Drake, L. M. and R. Simper (2004). "The Economics of Managerialism and the Drive for Efficiency in Policing." Managerial and Decision Economics **25**: 509-523.

Drake, L. M. and R. Simper (2005). "The measurement of police force efficiency: An assessment of UK Home Office policy." Contemporary Economic Policy **23**(4): 465-482.

Drake, L. M. and R. Simper (2005). "Police Efficiency in Offences Cleared: An Analysis of English "Basic Command Units"." International Review of Law and Economics **25**: 186-208.

Ezrow, L. and G. Xezonakis (2016). "Satisfaction with democracy and voter turnout: A temporal perspective." Party Politics **22**(1): 3-14.

Fang, L. (2013). "A generalized DEA model for centralized resource allocation." European Journal of Operational Research **228**: 405-412.

Farrell, G., N. Tilley, A. Tseloni and J. Mailley (2010). "Explaining and sustaining the crime drop: Clarifying the roles of opportunity-related theories." Crime Prevention and Community Safety **12**: 24-41.

García-Sánchez, I.-M. (2007). "Evaluating the effectiveness of the Spanish police force through data envelopment analysis." European Journal of Law and Economics **23**: 43-57.

García-Sánchez, I.-M. (2009). "Measuring the efficiency of local police force." European Journal of Law and Economics **27**: 59-77.

García-Sánchez, I.-M., L. Rodríguez-Domínguez and J. Parra-Domínguez (2013). "Evaluation of the efficacy and effectiveness of the Spanish security forces." European Journal of Law and Economics **36**: 57-75.

García-Sánchez, I.-M., L. Rodríguez-Domínguez and J. Parra-Domínguez (2013). "Yearly evolution of police efficiency in Spain and explanatory factors." Central European Journal of Operations Research **21**(1): 31-62.

Gorman, M. F. and J. Ruggiero (2008). "Evaluating US state police performance using data envelopment analysis." International Journal of Production Economics **113**: 1031-1037.

Greenberg, D. F., R. C. Kessler and C. Loftin (1983). "The effect of police employment on crime." Criminology **21**(3): 375-394.

Hadi-Vencheh, A., A. F. Ali and M. Soleimani-damaneh (2008). "A DEA model for Resource Allocation." Economic Modelling **25**(5): 983-993.

Higgins, A. and G. Hales (2016). *Police effectiveness in a changing world: Cutting crime in the 21st Century: Informed proactivity in the midst of social and organisational change*. London, Police Foundation.

HMIC (1998). *What Price Policing*.

Holmes, M. D., B. W. Smith, A. B. Freng and E. A. Muñoz (2008). "Minority Threat, Crime Control, and Police Resource Allocation in the Southwestern United States." Crime & Delinquency **54**(1): 128-152.

Home Office (2001). *Demonstration Project*. London.

Innes, M. (2011). "Doing Less with More: The 'new' politics of policing." Public Policy Research **18**(2): 73-80.

Jackson, A. L. and L. M. Boyd (2005). "Minority-Threat Hypothesis and the Workload Hypothesis: A Community-Level Examination of Lenient Policing in High Crime Communities." Criminal Justice Studies **18**(1): 29-50.

Karn, J. (2013). "Policing and Crime Reduction: The evidence and its implications for practice." Police Effectiveness in a Changing World Project, The Police Foundation.

Klick, J. and A. Tabarrok (2005). "Using Terror Alert Levels to Estimate the Effect of Police on Crime." Journal of Law and Economics **48**(1): 267-279.

Klinger, D. A. (1997). "Negotiating Order in Patrol Work: An Ecological Theory of Police Response to Deviance." Criminology **35**(2): 277-306.

Korhonen, P. and M. Syrjänen (2004). "Resource allocation based on efficiency analysis." Management Science **50**(8): 1134-1144.

Kovandzic, T. V. and J. J. Sloan (2002). "Police levels and crime rates revisited: A county-level analysis from Florida (1980–1998)." Journal of Criminal Justice **30**: 65-76.

Langworthy, R. H. (1983). "Effects of police agency size on the use of police employees: A Re-examination of Ostrom, Parks, and Whitaker." Police Studies **5**: 11-19.

Levitt, S. D. (1998). "The Relationship Between Crime Reporting and Police: Implications for the Use of Uniform Crime Reports." Journal of Quantitative Criminology **14**(1): 61-81.

Levitt, S. D. (2002). "Using Electoral Cycles in Police Hiring to Estimate the Effect of Police on Crime: Reply." The American Economic Review **92**(4): 1244-1250.

Lin, M.-J. (2009). "More Police, Less Crime: Evidence from US State Data." International Review of Law and Economics **29**: 73-80.

Loveday, B. (2000). "Managing Crime: Police Use of Crime Data as an Indicator of Effectiveness." International Journal of the Sociology of Law **28**: 215-237.

Lozano, S., G. Villa and R. Brännlund (2009). "Centralized reallocation of emission permits using DEA." European Journal of Operational Research **193**(3): 752-760.

Ludwig, A. and I. McLean (2016a). Police Resource Allocation: What Do the Professionals Think? A Gwilym Gibbon Centre for Public Policy Working Paper.

Ludwig, A. and I. McLean (2016b). Principles of Police Funding, A Gwilym Gibbon Working Paper.

Machin, S. and O. Marie (2005). Crime and Police Resources: the Street Crime Initiative. CEP Discussion Paper No 680. LSE, Centre for Economic Performance.

Marvell, T. B. and C. E. Moody (1996). "Specification problems, police levels and crime rates." Criminology **34**(4): 609-646.

McCabe, S. and F. Sutcliffe (1978). Defining crime: A study of police decisions. Oxford, England, Blackwell Publishing.

McGuire, A., J. Henderson and G. Mooney (1998). The economics of health care: an introductory text. London, Routledge and Kegan Paul.

McLean, I., M. Norton and A. Ludwig (2016). What are PCCs for? The intended role of Police and Crime Commissioners and their performance since 2012. Gwilym Gibbon Centre Working Paper.

Mosher, C., T. Miethe and D. Phillips (2002). The mismeasure of crime. Thousand Oaks, California, Sage Publications.

Murphy, C. (1985). Assessing Police Performance: Issues, Problems and Alternatives. Programs Branch User Report.

Nyhan, R. C. and L. L. Martin (1999). "Assessing the Performance of Municipal Police Services Using Data Envelopment Analysis: An Exploratory Study." State and Local Government Review **31**(1): 18-30.

Ogilvie, J. M., T. J. Allard and A. L. Stewart (2008). Impact of Police Numbers on Crime. J. M. Griffiths, Griffiths University, Australia.

Paré, P.-P., R. B. Felson and M. Ouimet (2007). "Community Variation in Crime Clearance: A Multilevel Analysis with Comments on Assessing Police Performance." Journal of Quantitative Criminology **23**: 243-258.

Public Services Productivity Panel (2000). Improving police performance. London, UK, HM Treasury.

Redshaw, J., R. I. Mawby and P. Bunt (1997). "Evaluating core policing in Britain: the views of police and consumers." International Journal of the Sociology of Law **25**: 283-301.

Ross, A. and A. Walker (2014). Low Priority Laws and the Allocation of Police Resources. Department of Economics, West Virginia University Working Paper.

Schulenberg, J. L. (2014). "Systematic social observation of police decision-making: the process, logistics, and challenges in a Canadian context." Quality & Quantity **48**: 297-315.

Sherman, L. W. and J. E. Eck (2002). Policing for Prevention. Evidence Based Crime Prevention. L. W. Sherman, D. P. Farrington, B. C. Welsh and D. L. MacKenzie. New York, Routledge: 295-329.

Sherman, L. W., D. C. Gottfredson, D. L. MacKenzie, J. Eck, P. Reuter and S. D. Bushway (1997). Preventing crime: what works, what doesn't, what's promising. Report to the United States Congress. Washington, DC: U.S, Government Printing Service.

Smith, K., P. Taylor and M. Elkin (2013). Crimes detected in England and Wales 2012/13 - HOSB: 02/13. London, Home Office. **HOSB: 02/13**.

Stockdale, J. E., C. M. E. Whitehead and P. J. Gresham (1999). Applying economic evaluation to policing activity. Research, Development and Statistics Directorate. London, UK, Policing and Reducing Crime Unit. **Police Research Series Paper No. 103**.

Sun, S. (2002). "Measuring the relative efficiency of police precincts using data envelopment analysis." Socio-Economic Planning Sciences **36**: 51-71.

Thanassoulis, E. (1995). "Assessing police forces in England and Wales using Data Envelopment Analysis." European Journal of Operational Research **87**: 641-657.

Todd, R. and K. V. Ramanathan (1994). "Perceived social needs, outcomes, measurements, and budgetary responsiveness in a not-for-profit setting; some empirical evidence." The Accounting Review **69**(1): 122-137.

Verma, A. and S. Gavirneni (2006). "Measuring police efficiency in India: an application of data envelopment analysis." Policing: An International Journal of Police Strategies & Management **29**(1): 125-145.

Vollaard, B. (2005). Police Numbers Up, Crime Rates Down: The Effect of Police on Crime in the Netherlands, 1996-2003. The Hague, the Netherlands, CPB Netherlands Bureau for Economic Policy Analysis.

Vollaard, B. and P. Koning (2009). "The Effect of Police on Crime, Disorder and Victim Precaution. Evidence from a Dutch victimisation survey." International Review of Law and Economics **29**: 336-348.

Walker, M. A. (1992). "Do we need a clear-up rate?" Policing and Society: An International Journal of Research and Policy **2**(4): 293-306.

Weisburd, D. and J. E. Eck (2004). "What can police do to reduce crime, disorder and fear?" The Annals of the American Academy of Political and Social Science **593**: 42-65.

Wilson, J. M. and A. Weiss (2014). "Police Staffing Allocation and Managing Workload Demand: A Critical Assessment of Existing Practices." Policing **8**(2): 96-108.

Wu, J., Q. An, S. Ali and L. Liang (2008). "DEA based resource allocation considering environmental factors." Mathematical and Computer Modelling **58**: 1128-1137.

Wu, T.-H., M.-S. Chen and J.-Y. Yeh (2010). "Measuring the performance of police forces in Taiwan using data envelopment analysis." Evaluation and Program Planning **33**: 246-254.

Yan, H., Q. L. Wei and G. Hao (2002). "DEA models for resource reallocation and production input/output estimation." European Journal of Operational Research **36**(1): 19-31.

Zhao, J. S., M. C. Scheider and Q. Thurman (2003). "A National Evaluation of the Effect of COPS Grants on Police Productivity (Arrests) 1995-1999." Police Quarterly **6**(4): 387-409.