

Has the financial regulatory environment improved in the UK? Capture-Recapture approach to estimate detection and deterrence

John Ashton
Bangor University

Tim Burnett
Warwick University

Ivan Diaz Rainey
University of Otago

Peter L. Ormosi
Centre for Competition Policy
Norwich Business School
University of East Anglia

CCP Working Paper 18-3

Abstract

Have regulators improved their capacity to detect or deter financial misconduct since the global financial crisis? This study addresses this question through an analysis that differentiates between detection and deterrence of financial misconduct over the period 2002-2016. To make this determination a Capture-Recapture, trap response model within a difference-in-differences framework is used to examine regulatory notices and supplementary information issued by UK financial regulators with corresponding Australian data employed as an unbiased control. Examining pre- and post-financial crisis periods, we offer evidence, whilst there was no statistically significant change in detection rates for breaches of financial regulation, there was a significant increase in the rate of deterrence occurred after 2010. We interpret this result as an increase in the efficacy of regulation of financial misconduct in recent years.

Contact Details:

John Ashton

j.ashton@bangor.ac.uk

Has the financial regulatory environment improved in the UK? Capture-Recapture approach to estimate detection and deterrence

John Ashton, Bangor University

Tim Burnett, Warwick University

Ivan Diaz Rainey, University of Otago

Peter L. Ormosi, University of East Anglia

Abstract

Have regulators improved their capacity to detect or deter financial misconduct since the global financial crisis? This study addresses this question through an analysis that differentiates between detection and deterrence of financial misconduct over the period 2002-2016. To make this determination a Capture-Recapture, trap response model within a difference-in-differences framework is used to examine regulatory notices and supplementary information issued by UK financial regulators with corresponding Australian data employed as an unbiased control. Examining pre- and post-financial crisis periods, we offer evidence, whilst there was no statistically significant change in detection rates for breaches of financial regulation, there was a significant increase in the rate of deterrence occurred after 2010. We interpret this result as an increase in the efficacy of regulation of financial misconduct in recent years.

Keywords: misconduct risk, regulatory punishments, partial observability, Capture-Recapture, deterrence.

Acknowledgements

We would like to thank Yang Wang and participants from a seminar at Cardiff University (2017) and a presentation at EARIE (2017) for their helpful comments. Faults remain our own.

1. Introduction

How successful are financial regulators at contending with and constraining financial misconduct? Did the events of the financial crisis lead regulators to enhance their capacity to detect or deter financial misconduct? This study addresses these significant public policy questions using the population of Final Notices and Enforcement Undertakings issued by UK and Australian financial regulators between 2002 and 2016. Using a Capture-Recapture method to address partial observability concerns, a trap response model is employed within a difference-in-differences framework (hereafter DiD) to assess whether detection and deterrence rates changed in the UK over this period. Assuming Australia to be an unbiased control group, we offer evidence that whilst detected breaches of UK financial regulation fell after 2010, the corresponding level of deterrence actually rose in this period. The causes of these changes are uncertain yet could include highly-publicised changes in regulatory structures, the effectiveness of punishments, or even cultural change in the banking industry.

This paper contributes to the literature in two regards. We apply a new method of quantifying detection and deterrence effect of financial regulation on financial misconduct using a statistical approach drawn from biology, ecology and demography – previously applied to assess other areas of criminal and corporate offending (see Ormosi 2014). This paper therefore contributes to an emergent literature examining the partial observability of offending in financial markets (e.g. Dyck *et al* 2013, Wang *et al* 2010, Wang 2011).

Second, we revisit the long-standing discussion as to the efficacy of regulation and optimal levels of regulation and punishment (e.g. Becker 1968, Stigler 1970, Polinsky and Shavell, 1979, Ehrlich 1996, Gneezy and Rustichini, 2000, La Porta *et al* 2006). By employing a new empirical method to examine the detection and deterrence of financial misconduct, this paper makes a significant contribution to this long-standing discourse.

This study is important for three reasons. One, since the financial crisis, there has been greater awareness of the risks posed by the conduct of financial institutions and their

employees (Skinner 2016). While financial misconduct alone is rarely the cause of financial crises, poor conduct by financial firms amplifies macro-prudential risks (e.g. through mortgage fraud, mis-selling of financial services, or even market abuse). Thus, while it is important not only to limit the dangers posed by financial misconduct through appropriate assessment of the risk this poses and creation of resilient financial institutions, it is also necessary to enhance methods to quantify the effectiveness of regulatory detection and deterrence in order to best minimise misconduct by firms and individuals.

Secondly, in recent years regulatory enforcement has been applied more frequently with increasingly severe punishments imposed; for example, in the UK, fines and remediation totalled £38.7 billion (\$56 billion) between 2011 and 2014, accounting for 60% of bank's profits (The Economist 2016)¹. A similar process is witnessed in the USA where financial institutions paid around \$139 billion in fines between 2012-14 (Zingales 2015). Despite the scale of these punishments, it remains unclear if these regulatory efforts have acted as an effective deterrent reducing the proportion of firms and individuals breaching financial regulations. This is troublesome, as the frequency and scale of regulatory sanctions has also heightened uncertainty in financial industries, reduced valuations of the incumbent firms sector and reinforced public scepticism as to whether this sector can ever be reformed (Group of 30, 2015). Moreover, levying large fines on firms may be viewed as unfair, placing an inordinate cost on the shareholders of financial firms rather than those persons and corporate entities responsible for offending (Goodhart 2017). Thus, in light of the growing body of literature identifying the often negative outcomes from regulatory enforcement actions (Delis *et al* 2017, Danisewicz *et al*, forthcoming), it is important to assess if these increasingly harsh punishments have been effective in reducing misconduct or otherwise,

¹ The level of disgorgement or redress arising from financial offending has also been a considerable cost for financial firms. In the UK this cost has run into many billions. For example, over £22 billion has been repaid to customers by February 2016 for payment protection insurance alone (House of Commons Committee of Public Accounts 2016).

Lastly, this work also contributes to the important and voluminous discussion on reform of financial regulation. Regulation, compliance and enforcement activities come at considerable cost for both firms and regulators.² Equally, a large number of nations have reactively invested heavily in changing their regulatory architecture in recent years; not least the UK where the Financial Conduct Authority (FCA) was established after the break-up of the Financial Services Authority in 2013 (hereafter the FSA). Through developing further metrics to assess regulatory outcomes, including often overlooked yet important deterrence effect (Gordon and Squires 2008) this study hopes to provoke further work on the public accountability of financial regulators and the efficacy of regulatory arrangements.

The paper is organised as follows. Following this introduction, we review past studies addressing the partial observability problem and outline the context of the study, namely, UK and Australian financial regulation. In section 3 we sketch the theoretical basis of the assessment. The methods employed and data used are detailed in section four. In section five, the results are reported and conclusions, policy implications and further avenues for research are presented in section six.

2. Literature review.

The scope of financial regulation defies precise definition (Allen and Carletti 2010) and optimal regulatory outcomes are hard to gauge (McCraw 1975). This uncertainty has resulted in financial laws and regulation being viewed to have both a positive (e.g. La Porta *et al* 2006) and negative (e.g. Stigler 1964) influence on the operation of financial markets. Similarly, increased enforcement of financial regulation has been interpreted as evidence both of more active and successful regulators (Stigler 1970) and failure for ‘allowing’ regulatory transgression (Becker and Stigler 1974). Clearly, this conflagration of criticism

² The Financial Conduct Authority (hereafter FCA), the UK’s principal financial regulator since 2013 for matters other than prudential regulation, has an annual turnover of over £½bn (House of Commons Committee of Public Accounts 2016)

faced by regulators, as either undertaking too little or too much regulatory action simultaneously requires further investigation.

To engage with these issues, this discussion initially considers the importance of, and past attempts to address the partial observability problem in financial and economic markets. We then consider the context of the work; the UK and Australian financial systems.

2.1 Quantifying partial observability in financial wrongdoing

An impediment to measuring the success of regulatory policy in detecting and deterring aberrant behaviour is the inability to observe those cases where a misconduct exists yet is not detected. While we are aware how many firms and individuals have been caught for breaching regulations, due to the illicit nature of financial misconduct it is unclear how many firms' transgress regulations and are not caught. This is problematic because the number of cases detected does not lead to an unambiguous assessment of regulatory performance.³ Therefore, a major challenge confronting any assessment of regulation is partial observability; non-detection of misconduct is likely to lead to underestimation of the true level of misconduct, overestimate the effectiveness of regulation and base assessment on a biased sample.

Sample selection issues such as partial observability arise when analysis is limited to a non-random subsample of interest. Observations of firms caught for regulatory failings are selected through a process that is not independent of the outcome of interest (i.e. whether the firm has breached a certain regulation and is affected by a diversity of non-random influences). This non-random selection can arise from explicit and incidental sources, is biased explicitly on the outcome of interest (i.e. the sample maybe truncated or censored) or arises when other endogenous variables determine the selection process. Reflecting the

³ In an example where the number of cases detected increases, this could imply that the regulator has become better at detecting misconduct, but it could also be a sign of weakening deterrence.

general nature of this statistical challenge, a variety of methods have arisen in different disciplines to address partial observability:

An early approach to accommodate partial observability involved the analysis of duration. This approach (see Bryant and Echard 1991) used Markov chain approaches to quantify the population of cartels. It was assumed the probability of detection was higher when the duration of cartel conspiracies was lower and vice versa; the theoretical basis of such duration models was subsequently developed by Harrington and Change (2009) to test the efficacy of different competition policies.

Sample selection issues have also been addressed using the Heckman selection model (Heckman 1979). This requires the estimation of a Probit model for selection, followed by the insertion of a correction factor (an inverse Mills ratio). This technique, previously used to correct for partial observability in criminology and fraud studies (e.g. Tan *et al* 2017), can be problematic. Concerns arise with the use of dichotomous dependent variables, misestimated standard errors and the inappropriate application of the Heckman model for such sample selection circumstances (see Bushway *et al* 2007).

Allied approaches to address the partial observability problem have emerged in accounting and finance, building on the work of Poirer (1980) and Feinstein (1990). Wang *et al* (2010) and Wang (2011) examined the incidence of corporate fraud and how the attributes of captured firms can be used to estimate the characteristics of firms likely to undertake similar transgressions. These logistic regression models consider the latent processes underlying fraud commissioning and detection distinctly to estimate the characteristics of the population of potential offenders. Using this approach Wang *et al* (2010) reported firms are more likely to commit fraud when business conditions are good, yet less so when investor confidence becomes very high. Wang (2011) further broadened the range of factors linked to corporate fraud. This type of logistic regression model has subsequently been employed to assess the influence of social links between directors on fraud (Kuang and Lee 2017) and accounting

mis-statements (Zakulyunina 2018), and has been developed to address partial observability directly (Lancaster and Imbens 1996).

Quantifying partial observability using these logistic regression techniques also has drawbacks. These approaches require far more data than just the frequency of offending, demanding data as to the characteristics of firms concerned. Furthermore, the models forwarded by Wang *et al* (2010) and Lancaster and Imbens (1996) have been reassessed and shown to be sensitive to the model assumptions (see Hahn *et al* 2016; Phillips and Elith 2013 respectively).

Studies of accounting fraud and mis-statements have also employed a range of methods to quantify undetected offending. Financial accounts and reports, the subject of many accounting frauds have been used to predict overall levels of detected and non-detected offending. Descriptive statistical methods (Dechow 2011), machine learning (Cecchini *et al* 2010) and deviations from Benfords law (Amiram *et al* 2015) have all also been used to predict unseen offending. These approaches whilst promising, are data intensive requiring data on the subject of offending, in addition to the occurrence and frequency of offending; Amiram *et al* (2017) provides a review of these techniques.

The total level of fraud in a market has also been estimated using natural experiments. Dyck *et al* (2013) proposed the failure of Arthur Andersen was one such experiment. In the early 1990s and following the collapse of this accounting firm, a large number of firms suddenly required a new auditor. These firms were assumed to be closely examined by their new auditors, enabling estimates of fraud throughout corporate America to be made. From this sample of closely scrutinised firms, it was estimated that 14.5% of large US publically limited firms engaged in accounting fraud.

Lastly, Capture-Recapture methods have been used to address partial observability in a number of settings. These techniques, drawn from ecology, epidemiology and biology have been developed to accommodate situations where populations change over time, when heterogeneity within the sample exists, and if time dependence influences recapture. In its

most simple form, Capture-Recapture models estimate a population through examining repeated random samples taken from the population of interest. In this process, samples are marked and replaced, with common observations recorded (see Chai *et al* 2001). The proportion of recaptured animals is then used to infer population parameters such as population size, capture and survival rate (see Chao *et al* 2001; Huggins and Hwang 2011). Capture-Recapture methods have been used in their more conventional setting to estimate human and animal populations, but it has also been applied to the analysis of the frequency of economic crimes and other illicit behaviours. For example, applying these techniques Ormosi (2014) estimated that 13-17% of European cartels were caught in any given year by competition law regulators between 1985 and 2009. Other crimes such as prostitution (Rossmo and Routledge 1990), marijuana cultivation (Bouchard 2007), car theft (Collins and Wilson 1990) and criminal desistance (Brame *et al* 2003) have also been examined with these methods.

In summary, addressing partial observability is an emergent subject and, as such, applying new techniques is important to address measurement concerns (Amiram *et al* 2017). The application of Capture-Recapture (or mark and recapture) techniques in this study of financial misconduct is motivated by a number of factors. This assessment, examining the frequency of offending to evaluate regulatory detection, has much in common with seemingly unrelated fields where the effectiveness of population monitoring activities need to be estimated. The Capture-Recapture method is well suited to the statistical concerns raised in this study as it focusses on the estimation of population characteristics from incomplete data, where repeated captures of subjects (financial companies) are not independent. Moreover, previous work has shown it to be effective in the study of ‘white-collar’ crime (Ormosi, 2014). Moreover, alternative approaches and methods, with their different assumptions and requirements are either a concern owing to misapplication (Heckman selection models), or inappropriate due to the particularities of the parameters they estimate, or alternatively difficult to implement owing to onerous data requirements.

2.2 Financial Regulation with the UK and Australia

The basis of the UK financial regulatory regime during the sample period (2002-2016) is encapsulated by the Financial Services and Markets Act (2000). This Act introduced the Financial Standards Authority (FSA) as regulator and affected virtually all aspects of financial services operations, training and competence, and consumer compensation in the UK. Under this regime, a common set of rules and regulatory principles, manifested in the regulators Handbooks of Rules and Guidance, were developed. The new regulator also operated a particular discursive, self-regulating form of enforcement, which enabled negotiation between the regulator and the firm such that Final Notices were only issued in cases where firms or individuals persistently engaged in a behaviour or actions judged to be unacceptable.

Furthermore, UK financial regulation has evolved over time (see Georgosouli 2010). Perhaps most significantly, the structure of financial regulation shifted from a single super regulator (the FSA) to a twin peaks structure in 2013, where financial misconduct was considered by the FCA Financial Conduct Authority (FCA) (see Taylor 1995) and prudential concerns were addressed by the Prudential Regulatory Authority. Similarly the scope of the regulator has changed from its original mandate(s), for example becoming additionally responsible for regulating the mortgage business in October 2004, commodity derivatives in November 2007, and consumer credit in April 2014.

Australia, which is used as an unbiased control in this study, has had a far more consistent regulatory approach and structure over the sample period. Since the Wallis Inquiry (Hanratty, 1997) Australia has possessed a ‘twin peaks’ structure of regulation (see Bakir, 2009, Godwin *et al* 2014, Thompson and Abbot, 2000) similar to the current arrangements in the UK.⁴ Through its shift to a twin peaks structure, Australia allocated prudential

⁴ Prior to this date, Australia pursued an institutional form of regulation where banking, building societies, insurance, unit trusts and consumer protection were regulated separately by distinct bodies.

regulation to the Australian Prudential Regulatory Agency and oversight of financial conduct in multiple financial services markets to the Australian Securities and Investments Commission (hereafter ASIC). The ASIC has applied a similar discursive approach to discourage offending and used a comparable array of punishments to the UK to deter transgressions (Australian Securities and Investment Commission, 2014).

These two financial regulatory systems while formed through different processes adopt similar approaches to regulation and enforcement; Australia following a twin peaks structure throughout, and the UK using, first, a single super-regulator and a twin peaks structure since 2013. While these regulatory systems share many features, they developed through independent and distinct processes – the UK changes occurring after financial turmoil as reactive political responses (Ferran 2011), and Australian regulation developing through long-term political initiatives to ensure an effective regulatory community (Bakir 2009) and competitive neutrality between different financial institutions (Thompson and Abbot 2000)⁵. Moreover, while some differences in the scope of regulation exist, both nations adopt a common law system, share a common language and many regulatory principles and procedures (see Australian Securities and Investment Commission, 2014).

3. Theory

In this section we outline the theoretical relationships envisaged between regulatory punishment and levels of offending. We then present the framework used in our analysis.

3.1. Link between punishment and offending

Central to the literature as to how regulation and enforcement operates is the link between levels of punishment and offending, and associated questions such as the optimal level of fines. Since Becker (1968) the choice to engage in crime has been framed as an economic

⁵ Andersson (2016) provides a discussion of the role of financial crises in developing regulatory arrangements in financial industries.

decision involving cost and benefit trade-offs, where offending is associated with how much an economic actor can earn from the offence balanced against the costs and (crucially) the probability of being caught (Stigler 1970).

In this cost-benefit framework, firms and individuals only undertake criminal acts if their private benefits from these actions exceed the external costs (Polinsky and Shavell 1979) – suggesting that the efficacy of detection and severity of the punishment influences the likelihood of engaging in crime. Deterrence is, therefore, central to this reasoning. An enforcement system that successfully deters misbehaviour avoids the social costs that arise from lack of compliance and the considerable cost of policing of markets. Optimal punishment therefore has a central role of deterring future offending.

As higher expected punishments are assumed to reduce crime levels (Gneezy and Rusichini 2000) any fines should reflect both the damage caused to victims, and must to be adjusted upwards to account for the imperfect probability of apprehension (Becker and Stigler 1974). As such, determining appropriate levels of punishment requires a calculation to both deter poor behaviours, whilst also limiting the costs of regulation (see Becker 1968).

A trade-off between fines and probability of detection can therefore arise. For instance, implausibly large fines arising when the capture of offenders becomes improbable. Fining an individual in excess of the cost of the offending appears unfair and may only provide a marginal deterrence or even redirect and amplify offending activity (Stigler 1970). These relationships are further complicated in the presence of risk aversion (Polinsky and Shavell 1979), the wealth of the offending economic actor (Polinsky and Shavell 1984) and the scale and form of enforcement costs (Polinsky and Shavell 1992).

The level of enforcement and probability of being apprehended is also influenced by societies' perception of the offence (Becker and Stigler 1974), and any required levels of compliance must be compatible with socially acceptable behaviours (Stigler 1970). Within this market model of offending (Ehrlich 1996), we speculate that the perceived social cost

of financial offending has risen since the financial turmoil of 2007, in tandem with the scale of fines levied on offending firms (see Table 2, below).

Cognisant of the incentives to breach regulations, firms and individuals also invest effort and resources in various strategies to evade capture and to diminish the efficacy of regulatory enforcement. This effect is compounded by the relative or perceived infrequency of regulatory breaches and the adversarial setting in which these occur. Examples of such deceptive behaviours often include masking, dazzling, decoying, repackaging, mimicking and double play – all of which are observed in actions between firms' and their regulatory monitors with the aim of limiting detection (see Johnson *et al* 2001).

The actions of firm lobbyists may also constrain the incentives for regulators to intervene, encouraging limited regulatory actions and ensuring a sympathetic handling of firms (Yu and Yu 2011). These outcomes arise through many mechanisms from seducing naive politicians (Blau *et al* 2013; Gropper *et al* 2013), 'revolving doors' in the selection of regulators (Dal Bo 2006) and regulators socially identifying with the industry they regulate (Veltrop and de Haan 2014).

In summary, regulators will only capture a proportion of offenders. Subsequently the level and frequency of punishment has a significant, if imprecise influence over the decision to offend, and, critically, the level of deterrence. This degree of imprecision is influenced by many factors, not least the offenders' ability to evade discovery. Thus, a missing element when comprehending the effectiveness of detecting and deterring of misconduct, is the ability to gauge not only the firms guilty of transgressions, but also ascertain the proportion of offending firms evading capture.

3.2 Addressing the partial observability problem.

To address the problem of partial observability we formulate a simple model of detection and deterrence. Denote the population size of all registered financial sector firms by N , the probability of deterring a regulated business from committing an infringement by ω , and the

probability of detecting an infringement by ρ . The number of cases detected (n) is then given by $n = (1 - \omega)\rho N$. From this, the probability of deterrence is defined as:

$$\omega = \frac{1}{\rho} \frac{n}{N} \quad (1)$$

We denote the proportion of firms under financial regulation that engage in regulated misconduct by $\eta = \frac{n}{N}$. From equation (1), we conclude that deterrence increases if $\frac{\Delta\eta}{\eta} > \frac{\Delta\rho}{\rho}$ where, in the analysis of Section 5, $\Delta\eta$ and $\Delta\rho$ denote changes in, respectively, the average values of η and ρ between the period up to and including 2010 and the average for the period 2011 onward – where 2010 is identified as a structural break point in the dataset (see section 4.1, page 19) which also corresponds, in accordance with the objectives of the paper, to a point immediately post-financial crisis. Therefore we can establish the following proposition.

Proposition 1. A sufficient set of conditions to establish we have observed increased deterrence after 2010 is: $\Delta\eta < 0$ and $\Delta\rho \geq 0$, or $\Delta\eta \leq 0$ and $\Delta\rho > 0$.

Deterrence increases if the pre- and post-average proportion of firms engaging in misconduct should decline (or remain static) after 2010, coupled with the average probability of detecting an offending firm should be equal to or greater than 0 (or should increase) over the same interval. Accordingly, to test this proposition, our empirical strategy consists of two main elements: First, we estimate the impact of our structural break (2010) on the relative number of detected cases (to elicit $\Delta\eta$). Second, we then look to how the detection probability changed ($\Delta\rho$). Finally, we use these estimated effects to infer how regulatory deterrence has changed in the UK since 2010.

In the assessment of Section 5, Australia is used as a control group. Using Australia as counterfactual provides unbiased estimates if two assumptions are satisfied, (1) that Australia is unaffected by the regulatory reform in the UK (independence), and (2) in the absence of the treatment the outcome variables would behave similarly in Australia as in the

UK (similarity). In terms of the independence assumption, we have no reason to believe that UK reforms would have affected financial firms in Australia, as Australian regulation has developed through a distinct political process (see Section 2.2). Regarding similarity, the regulatory processes in both nations share many features as we have previously outlined (again, see Section 2.2).

4. Methods and Data

In this section we outline the sources of the data, its format and the processes employed to code and transform this data into firm-level data usable for the study. We also introduce the descriptive and inferential techniques employed. These methods include the Capture-Recapture (CR) methods used to address partial observability and the DiD techniques employed to handle the identification of the effect on detection and deterrence rates.

4.1 Data sources and variables

The study employs data from a number of UK and Australian sources. The primary data sources are the ‘Final Notices’ issued by the UK financial regulators, the FSA (in operation between 2001-13) and the FCA (operating since 2013), and the ‘Enforceable Undertakings’ issued by the Australian financial regulator, the ASIC. These documents are publicly available and provide details of misconduct and breaches of regulations, punishments and the characteristics of the offence.

In the UK, a sample of 1,869 Final Notices were collected, varying in document length from one to ninety pages and issued to firms and individuals between 2002-2015. A sample of 412 Enforceable Undertakings issued to individuals and firms over the same time-period were collated for Australia. These undertakings vary from three to 140 pages. The data included the date of the offence, the duration of offending, the date of the regulatory intervention (i.e. date of the ‘Final Notice’ for the UK or ‘Enforceable Undertaking’ for Australia from which

we create yearly and quarterly measures of offending), firm characteristics, punishments and the nature of the offence.

This hand-collected data was supplemented and manually cross-checked for both nations. For the UK, Financial Regulator Annual Reports and the Financial Services Register⁶ were used to cross-check the data. Furthermore, Supervisory, Warning and Decision notices and press releases issued by the FSA and FCA, as well as appeals to the Financial Services and Markets Tribunal (359 documents in total) were also consulted to augment and confirm Final Notice details. This process was also undertaken for the Australian Enforcement Undertakings using both associated press releases from the ASIC and the frequent and detailed reports on enforcement activity provided the ASIC.

The data was initially collected and coded at the level of individual offences according to classifications previously applied to Final Notices within annual reports issued by the FSA and to comply with existing forms of coding used within the Financial Services Register. In Australia, a similar coding strategy was adopted drawing on the classification of Enforcement Undertakings adopted by the ASIC and the summaries of enforcement activity and press releases issued by this regulator. Although the data was gathered and categorised according the different types of offences recognised by regulators, in the final analysis these were all aggregated into a single offender category.

To transform the Final Notices and Enforcement Undertakings into a firm-level data set, a number of assumptions have been made, resulting in the exclusion of some observations. For the UK, in 22 cases a Final Notice refers to a rejected application to extend a regulatory function, and in 32 Final Notices the judgement concerned a form of market abuse such as insider dealing and/or an accounting or listing reporting irregularity involving a non-financial firm. Furthermore, in 135 Final Notices a person or firm has provided financial services whilst not being regulated. These cases fall outside our frame of reference (focusing

⁶ The Financial Services Register records all firms and individuals licenced to operate in UK financial services markets.

on financial firms or employees, and breaches of financial regulation) and are, therefore, excluded from the analysis.

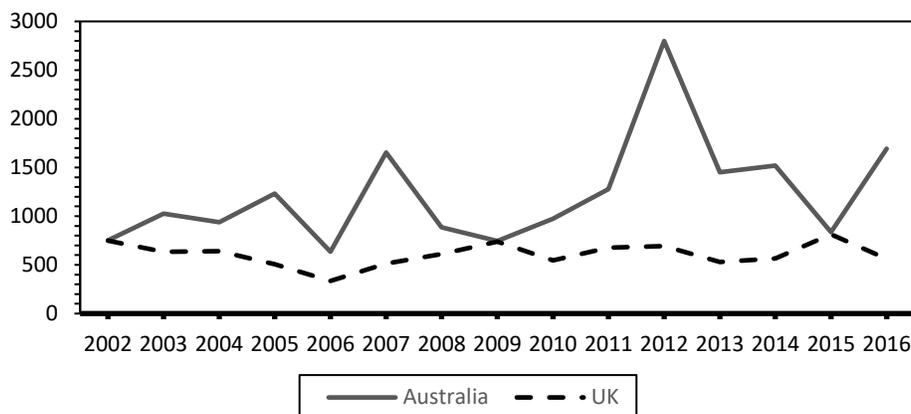
The remaining UK observations relate to 1,389 firms, including situations where firms, or their employees, have been issued multiple Final Notices.⁷ For 68 Final Notices multiple firms were involved; in these cases, each firm involved was considered distinctly. Finally, the data was annualised, such that we considered whether a firm or its employee(s) had offended within a given year (multiple offences within a single year were only considered once). Overall, for the UK, 1,295 firms only offended within one year and nearly 100 committed offences in two or more different years. For Australia, 412 Enforcement Undertakings were similarly coded. These cases pertain to 266 firms associated with financial misconduct. Of these firms, 21 offended in more than one year.

Lastly, data was collected from regulators' annual reports and accounts and other sources to determine the population of regulated firms and if variation in detection arises from regulatory resources or wider macro-economic concerns. We summarise the data and variables used in four tables in the Appendix. The variables employed as co-variables in the DiD assessment are described in Appendix

Table 4. Table 5 outlines the descriptive variables for the enforcement cases considered at an offence-level. Table 6 provides descriptive statistics for firm level data over time and descriptive statistics of the co-variables are reported in Table 7. Of particular interest in both nations is the significant rise in regulatory resources such as employees and operating costs of regulators between 2002 and 2016.

⁷ The maximum observed was the 11 Final Notices received by UBS AG.

Figure 1 - The Duration of Offences (# of days)



The average fines levied on firms and individuals are significantly higher in the UK than in Australia, though in both nations the average fine displays an upward trend. The average duration of offences shows increasing variability for the UK and an approximately steady rate for Australia (Figure 1). Regarding the type of observed financial misconduct and the punishments applied in the UK and Australia, reporting and compliance offences are the most frequently observed in both countries – an offence which incorporates many actions from non-payment of regulatory fees to failures submitting transactions data. Mis-selling of financial services, the sale of a financial service which is not needed by a customer, accounts for a far larger proportion of misconduct cases in Australia than the UK. In the UK, more cases of fraud and theft are reported, many of which are associated with corresponding criminal proceedings. Turning to punishments, in Australia there is far greater focus on disgorgement of the gains of offending and compensating victims of misconduct. In the UK non-financial punishments such as prohibition of individuals from working in the financial sector, or cancellation of regulatory permissions to trade as a financial services firm are used more frequently.

Figure 2 - Proportion of offenders in the UK and Australia

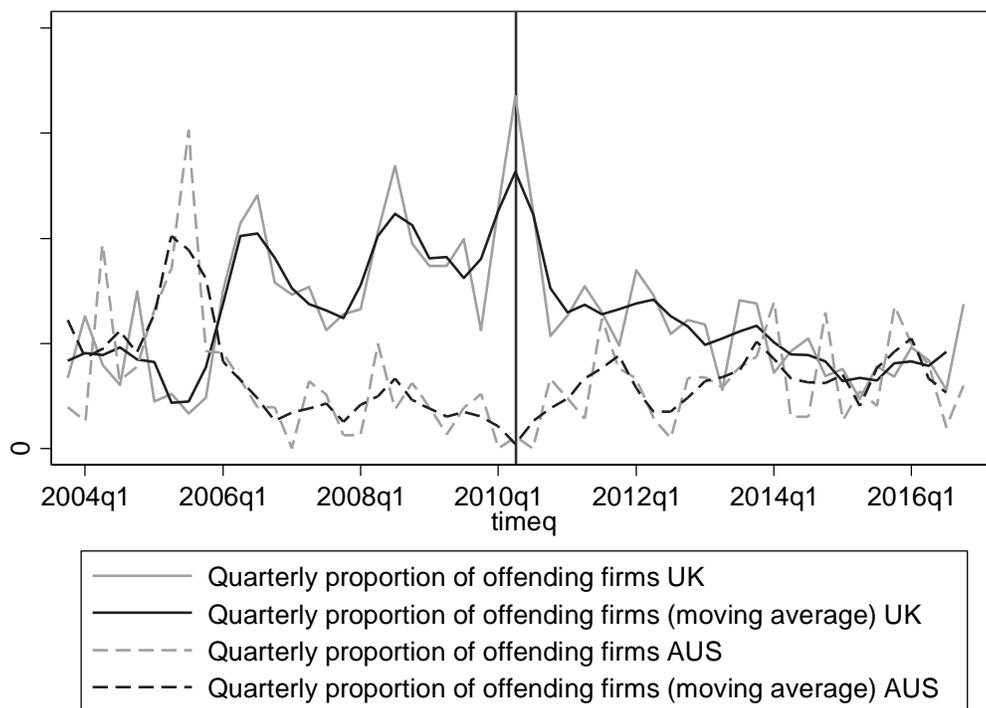
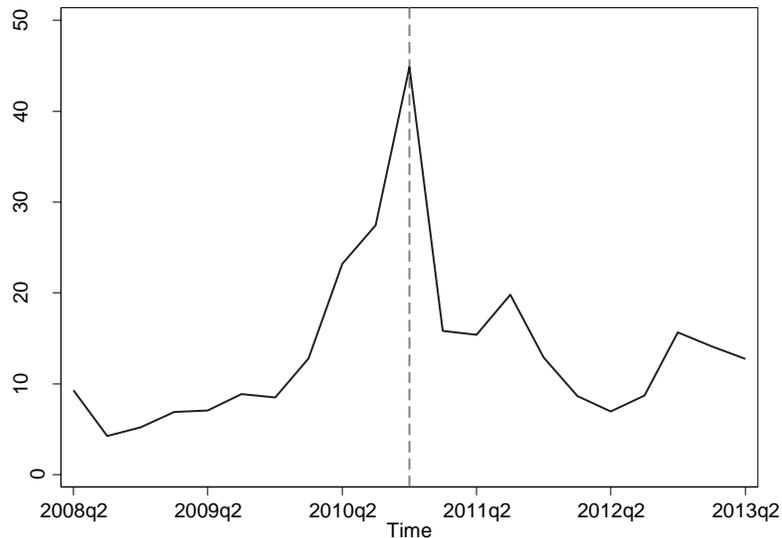


Figure 2 presents the percentage of ‘captured’ offending firms (those which were caught). To smooth the two curves and focus on the longer run trends rather than short-term variation, we also report 3-quarter moving averages. The ratio of captured offending firms to the number of regulated firms (n/N) varies over the sample period; in particular in the UK this increases until 2010 before declining (firm level data on observed offending and reoffending, the total number of regulated firms operating in UK and Australian financial services industries are reported in Table 5 in the Appendix). Australia did not seem to experience a similar change. To formally confirm this structural break in 2010, we run a set of Wald tests of whether the coefficients in a time-series regression (on UK data alone) vary over the periods defined by possible break dates. Figure 3 shows the test statistics for these break dates, which shows a peak at Q4 2010, implying the highest probability of a structural break

at this point in time. Moreover, as the average duration of offending is generally over 600 days, 2010 can also be seen as a lagged effect of regulatory changes occurring in response to 2008.

Figure 3 - Wald test statistics on a set of potential break dates



What also stands out from Figure 2 is that in our latter difference-in-differences model we need to exclude pre-2006 data, where the parallel trend assumption is clearly violated.

In the context of partial observability, and considering the information from Figure 2 in isolation, one could jump – potentially mistakenly – to the conclusion that enforcement has become less effective in the UK in the post-2010 period. Similarly, when considering Figure 1, which outlines the duration of offending we see a sharp rise and then a fall in Australia in the average length of offending and a static duration of offending in the UK suggesting the detection of financial misconduct in Australia and the UK is variable.

It is proposed that both Figure 1 and Figure 2 mask key information: A drop in the number of detected offences could be a sign of a decline in detection rates, but it could also be

interpreted as an improved regulatory environment with improved deterrence and fewer offences to detect. Whereas the former would be an undesired change, the latter is clearly a positive development. Similarly, the duration of offending might reflect biased samples of captured firms relative to not-captured firms; for instance, the Australian regulator may either be slower at addressing offending or could be relatively successful in discovering a broader scope of financial misconduct.

The need to unpick these contorted interpretations and accurately assess regulatory performance motivates the use of the CR framework to distinguish between these possible explanations and allow identification of these different effects.

4.2 *Methods*

To estimate the impact of the financial crisis on the proportion of detected offences, we use a simple DiD model. Denote the two countries by $j \in \{UK, Australia\}$ and a $(K \times 1)$ vector of time-dependent country/regulatory characteristics by χ_{jt} . These include the number of cases where fines were imposed, the mean fine, stock market index, total assets less current liabilities of regulated firms, net operational costs of regulated firms, number and cost of regulatory employees. μ_{jt} are unobserved firm and time fixed effects and ε_{jt} are idiosyncratic shocks with zero mean. This is defined as:

$$\eta = \beta_0 + \beta_1 I_{jt} + \gamma \chi_{jt} + \mu_t + \varepsilon_{jt} \quad (2)$$

where $I_{jt} = D_j \times T_t$ (D_j indicates whether the business operates in the UK or Australia and T_t is a dummy variable denoting whether period t was before or after 2010). Within the framework of Proposition 1, the DiD estimator $\beta_1 = \Delta\eta$ is our coefficient of interest. Provided Australia is an unbiased control, β_1 should give us an estimate of the effect of the 2010 structural break on the relative number of regulatory offences.

To estimate the other component in Proposition 1 we turn to Capture-Recapture (CR) methods. Ormosi (2014) offers a detailed explanation of the terminology, and how to apply CR methods to business data, thus below we restrict our methodological exposition to the extent we deviate from that work.

The two key parameters in a CR setting are detection and survival probability. ρ_{tm} is the probability of exposure (detection) of a financial misconduct of firm m at time t . The survival rate (ϕ_{tm}) in this application is an apparent survival estimate. It is apparent because it is not known whether the firm 'dies' because it does not exist anymore, because it refrains from future financial misconduct, or because it joins the subpopulation of those offending firms that are never re-captured (for example because the firm developed techniques to evade regulatory detection). Thus, in this setting, survival only means survival of the possibility of being captured again, such that the offending firm is still in the subset of firms which could be captured in the future. This could also be thought of as the survival of detectable evidence related to the offense, which is generated when the offense is committed and this evidence remains alive until capture (even if capture happens later than the end of the offense).

To record data for a CR analysis, the information on the timing of detection is organised into an $n \times K$ matrix \mathbf{X} (n is the number of offending firms detected, K is the number of years in our sample, $m \in n$, and $t \in K$), where $x_{mt} = 1$ if firm m was captured at sampling occasion t and $x_{mt} = 0$ otherwise. Let ϕ_{tm} denote the probability of an offending firm m surviving time $t = 1, 2, \dots$ which is the conditional apparent survival from year t to year $t + 1$, given that the same firm is 'alive' at the beginning of year t . Denote the probability of firm m being captured at sampling occasion $t = 1, 2, \dots$ by ρ_{tm} .

We compose a capture history (a $K \times 1$ matrix of 1s and 0s denoting, for each firm in the sample, whether a given firm was detected in a given year) and express the probability of

observing these capture histories. If we denote the time of the first capture of firm m as f_m , the last capture as l_m and the departure ('death' or migration) from the sample as $d_m (> l_m)$ the probability of observing any capture history can be written in a general form (summing up for all possible departure times d_m , which is necessary as d_m is typically not observed). Finally, let $x_{mt} = 1$ if firm m was captured at time t , and $x_{mt} = 0$ otherwise. Note that instead of a general parametric form, which assumes that both capture and survival rates are time-dependent, we are only interested in the effect of the 2010 break on the rate of detection (capture) and survival, therefore ϕ_t and ρ_t can assume only two values (pre, and post-2010):

$$Pr\{CH_m|f_m\} = \sum_{d_m=l_m}^K \left\{ \left(\prod_{t=f_m}^{d_m-1} \phi_t \right) (1 - \phi_{d_m}) \times \left(\prod_{t=f_m+1}^{d_m} \rho_t^{x_{mt}} (1 - \rho_t)^{(1-x_{mt})} \right) \right\} \quad (3)$$

Because we want to estimate how 2010 has affected ϕ_t and ρ_t , we use the following logit link functions for each parameter respectively:

$$\phi_t \text{ and } \rho_t = \frac{\exp(\beta_0^{\rho,\phi} + \beta_1^{\rho,\phi} DT_t + \beta_2^{\rho,\phi} D + \beta_3^{\rho,\phi} T_t)}{[1 + \exp(\beta_0^{\rho,\phi} + \beta_1^{\rho,\phi} DT_t + \beta_2^{\rho,\phi} D + \beta_3^{\rho,\phi} T_t)]} \quad (4)$$

where D denotes the country (UK or Australia), T_t denotes whether a given year t was before or after 2010, and $\beta_1^{\rho,\phi}$ are the effect of 2010 on the probability of detection and survival respectively. Using the individual capture history likelihoods and provided that all individuals are independent, the likelihood of observing all capture histories is therefore a product of the individual probabilities:

$$L = \prod_{m=1}^{2^K-1} Pr(CH_m|f_m) \quad (5)$$

Once capture histories are recorded for all captured individuals, the log of L can be used to find the parameters, including β_1 , that maximise the likelihood of observing the recorded capture histories.

5. Results

Next we formally test whether the drop in the proportion (and number) of detected offending firms in the UK is significant, using Australia as a control group. Because some of the variables vary significantly in their magnitude, we use standardised values for all but the dummy variables. Subsequently the coefficients should be interpreted as the standard deviation change in the dependent variable associated with a 1 standard deviation change in the independent variable. To remove the effect of size, we divided the stock index, the net operational costs, the employee number, and the employee costs over total assets. Table 1 displays the results of estimating the model in Eq.(2).

We use four different model specifications. The first column shows the estimates where the dependent variable is the proportion of detected offending firms and is estimated using a number of time-dependent co-variates as previously specified. The second column is the same as the first column but without co-variates. Columns 3 and 4 estimate the same models but now using the number of detected offenders as dependent variable.

The first row of Table 1 shows the difference-in-differences estimator (DiD), which is significant and negative for all model specifications. This is unsurprising, given visually apparent drop in the number of detected offenders after 2010, as observed in Figure 1. This is evidence that one of our sufficient conditions to establish an increased deterrence rate (Proposition 1) holds as $\Delta\eta < 0$ (i.e. the change in detected firms has declined since 2010). Beyond noting the significance of the number of fines, we refrain from the interpretation of the effect of the co-variates to maintain focus on the effect of the post-crisis effect indicated by the 2010 structural break.

Table 1 - Regression results on the proportion and number of offending firms (η)

	(1)	(2)	(3)	(4)
	Proportion	Proportion	Number	Number
DiD	-1.296***	-1.099***	-1.941***	-1.609***
	(0.273)	(0.233)	(0.328)	(0.266)
Number of cases with fines (1 year lag)	0.756***		0.645**	
	(0.231)		(0.277)	
Stock index	-0.991		-1.723*	
	(0.808)		(0.969)	
Total assets	-0.196		-0.509*	
	(0.225)		(0.270)	
Net operational costs	-1.842		-3.400	
	(2.355)		(2.825)	
Employees	2.332*		2.806*	
	(1.307)		(1.568)	
Employee costs	0.751		3.013	
	(2.049)		(2.458)	
Observations	90	92	88	88

Standard errors in parentheses

=* p<0.10 ** p<0.05 *** p<0.01

One issue with difference-in-difference estimators is that serial correlation in the data could lead to biased standard errors. Here we test for serial correlation in the number of detected offenders data using Wooldridge's (2002) test and reject the null that there is serial correlation.

As a next step, we examine the change in detection rate, $\Delta\rho$. In Eq.(3) we chose a parametrisation where both the detection rate (ρ) and survival rate (ϕ) can assume two values, one before and one after 2010 for the UK and Australia. This reflects our interest in estimating a difference-in-difference model (i.e. the impact of 2010). However, different assumptions about these parameters might be equally viable, for example where ϕ or ρ (or both) are time dependent and assumes a different value each year. Our job was to find which parametric assumption fits our data best.

Table 2 contains all the models (reflecting all possible parametric assumptions) estimated to acquire detection rate estimates before and after 2010. For the model names we use the following rules: (.) implies a model where the given parameter is assumed to be time-constant; (*t*) indicates time-dependent parameters; (*j*) implies a model with different parameters for our two countries $j \in (UK, Australia)$; ($j * t$) indicates a model with separate time-dependent parameters for each country; and (*./*) refers to a model where the given parameter is time-constant, but the values are assumed to be different in the year of detection than any other subsequent year. For example $\phi(./)$ refers to a model where we assume that firms' survival rate in the year of the detection is different from all subsequent years (the methodological literature knows this as a trap-response model). $\phi(./)(./)$ implies a trap response model where the parameters are different for pre-, and post- 2010. Finally, $\rho(did)$ refers to a model where we examine the effect of 2010.

Table 2 - Estimated capture recapture models and the corresponding DiD estimates⁸

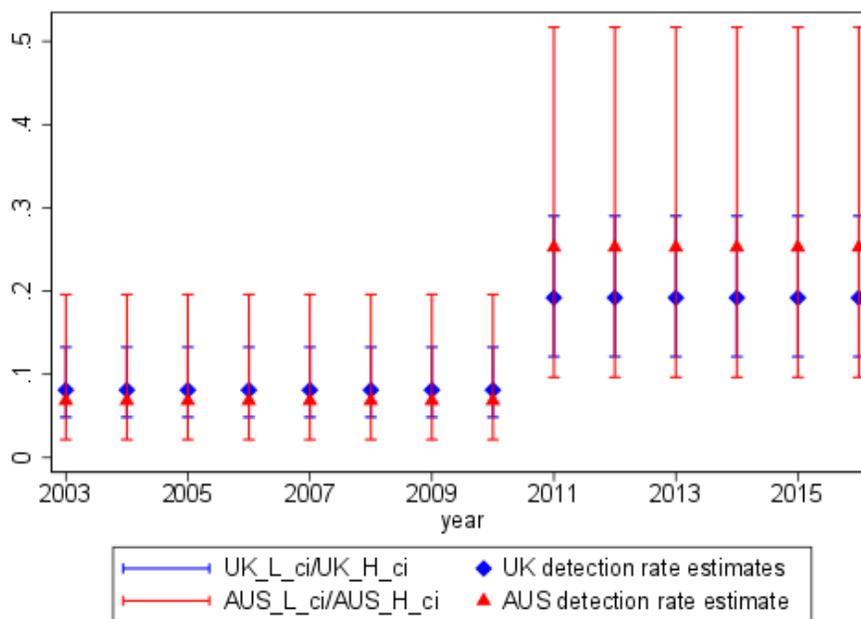
Model	AICc	AICc Weights	Model Likelihood	Num. Par	DiD	S.E.
$\phi(./)(./) \rho(did)$	844.4221	0.46229	1	11	-0.5392	0.8191
$\phi(./)(./)\rho(did)$ fines	845.8099	0.23097	0.4996	12	-0.2418	0.9205
$\phi(./)(./) \rho(did)$ fine amount	846.6009	0.15552	0.3364	13	-0.1991	0.9179
$\phi(./)(./)\rho(did)$ all covariates	846.7647	0.14329	0.31	18	1.233	1.0843
$\phi(./)\rho(did)$	852.5534	0.00793	0.0172	8	-1.1858	0.7099
$\phi(j*t) \rho(did)$	909.8406	0	0	18	-0.7136	0.914
$\phi(did) \rho(did)$	917.5198	0	0	8	-0.4239	0.8119

⁸ Here we are only showing the best performing models. The full table is given in the Appendix.

The AICc weights in Table 2 indicate that given the dataset, model $\phi(./.)(./.)\rho(\text{did})$ is approximately twice as likely to have generated the observed data than the second most efficient model. The *DiD* models have the best fit over models where the detection rate ρ is constant or time-dependent (see Table 8 in the Appendix). Table 2 also reports the *DiD* coefficients and their standard errors for each model, where ρ was modelled accordingly (as shown in Eq.(4)). We found no evidence that detection rate significantly changed after 2010 (assuming Australia is an unbiased control group). This is confirmed visually on Figure 4 where we plot the detection parameters (and their 95% confidence interval) as estimated using our difference-in-difference model $\phi(./.)(./.)\rho(\text{did})$.

Using this results together with our framework in Proposition 1, this would imply that $\Delta\rho = 0$, i.e. the second sufficient condition of showing that deterrence increased also holds.

Figure 4 - DiD estimates of the probability of detecting an offender



Regarding survival rates, the models with trap-response (with different trap response pre- and post-2010) performed best. The survival rate estimates for the best fitting model are shown in Table 3.

Table 3 - Estimated survival rates, pre-, and post-2010 for the UK and AUS

	UK	AUS
Before 2010 Year of detection	0.133	0.288
s.e.	(0.027)	(0.142)
Subsequent years	-	0.705
s.e.	-	(0.087)
After 2010 Year of detection	0.043	0.204
s.e.	(0.018)	(0.132)
Subsequent years	0.812	0.869
s.e.	(0.059)	(0.175)

These results include two interesting findings: In general, the chance of survival (a business remaining in operation following a detection) is very low in the year of the detection. This is particularly the case for the UK. However, businesses that survive the critical first year after detection, have a very good survival probability. This is in line with intuition and a simple observation of data. The effect of the 2010 structural break is seen in that the probability of survival in the year of detection became even smaller after this point, and those firms who survived past capture are more likely to persist.

To conclude this section, in Proposition 1 we formulated a sufficient pair of alternative conditions needed to establish that financial regulations were more deterring of misconduct after 2010. One of these conditions required that the proportion of detected offenders' decreased ($\Delta\eta < 0$) and the rate of detection did not decrease ($\Delta\rho > 0$). Evidence supporting both of these conditions in the UK is provided, where detection rates have

remained constant and the number of detected cases dropped significantly. We believe this is strong evidence that the UK regulatory environment improved after 2010 as the rate of deterrence has risen.

6. Conclusions

Since the financial crisis there has been much reflection as to the effectiveness of financial regulation. The UK financial regulator (the FSA) in particular was candid as to its failings surrounding this crisis and areas where improvement could be effected (Ferran 2011). Despite the importance of critically assessing regulatory performance, too much analysis has focused on deconstructing causes of past crisis events and often politically reactive regulatory developments. This study forwards and applies a new method for assessing the efficacy of financial regulation, through assessing regulatory detection and deterrence rates to aid this assessment of misconduct regulation.

Specifically, this study examines if detection and deterrence for financial misconduct in the UK has altered since the end of the financial crisis in 2010, a year which both satisfies the criteria for being immediately post-crisis, and which also exhibited a structural break in our dataset. The levels of financial misconduct are examined within a DiD framework using a Cormack-Jolly-Seber method (Jolly, 1965) with a trap-response model. These estimates include both firms and individuals caught and sanctioned by financial authorities and the proportion of economic agents estimated to be breaching regulation and evading capture. Results indicate that while detection rates appear not to have altered, deterrence of financial offending has enhanced. This finding raises a host of further questions as what might be driving this process. Clearly increased punishments, changing regulatory structures or cultural change in the industry as a whole could have all been influential; these explanations are beyond the scope of this study yet remain important and pressing areas for future investigation.

Whilst there has been a growing literature addressing partial observability in both criminal and fraud areas deploying a diversity of methods (as Hahn *et al* (2017) and Amiran *et al* (2017) respectively report), it is imperative to move beyond repetition of existing methods and to develop new techniques to refine estimations and disentangle continuing issues with causality. This study forwards a technique which is less data demanding and emerges from a developed statistical tradition with ecology and biology. Thus, this assessment is important both for the reasons outlined in the introduction, but also in its contribution to further refining the current methods of assessing financial.

References

- Allen F. and Carletti, E. (2010). "An overview of the Crisis: Causes, Consequences and Solutions", *International Review of Finance*, 10(1): 1-26.
- Amiram, D. Bozanic, Z. and Rouen E. (2015). "Financial Statement Errors: Evidence from the Distributional Properties of Financial Statement Numbers", *Review of Accounting Studies*, 20(4): 1540-1593.
- Amiram, D, Bozanic, Z. Cox, J. D. , Dupont, Q. Karpoff, J. M. and Sloan, R. (2017). "Financial reporting fraud and other forms of misconduct: A multidisciplinary review of the literature", *Review of Accounting Studies*, Forthcoming. <https://doi.org/10.1007/s11142-017-9435-x>
- Andersson, F. N. (2016). "A Blessing in Disguise? Banking Crises and Institutional Change", *World Development*, 83:135-147.
- Australian Securities and Investment Commission, (2014) *Penalties for Corporate Wrongdoing*, Report 387. Sydney.
- Bakir, C. (2009). "The Governance of Financial Regulatory Reform: The Australian Experience", *Public Administration*, 87(4), pp.910-922.
- Becker, G. S. (1968). "Crime and Punishment: An Economic Approach", *Journal of Political Economy*, 76(2): 169-217.
- Becker, G. S. and Stigler G. (1974). "Law Enforcement, Malfeasance and Compensation of Enforcers", *The Journal of Legal Studies*, 3(1): 1-18.
- Blau, B. M., Brough, T. J. and Thomas, D. W. (2013). "Corporate lobbying, political connections, and the bailout of banks", *Journal of Banking and Finance*, 37(8): 3007-3017.

Brame, R., Bushway, S. D. and Paternostyer, R. (2003). "Examining the Prevalence of Criminal Desistance", *Criminology*, 41(2):423-448.

Bushway, S., Johnson, B. D. and Slocum, L. A. (2007). "Is the magic still there? The use of the Heckman Two-Step Correction for Selection Bias in Criminology", *Journal of Quantitative Criminology*, 23(2): 151-178.

Bouchard, M. (2007). "A Capture-Recapture model to estimate the size of criminal populations and the risks of Detection in a Marijuana Cultivation Industry", *Journal of Quantitative Criminology*, 23(3): 221-241.

Bryant, P. G. and Echard, J. W. (1991). "Price fixing the probability of getting caught", *The Review of Economics and Statistics*, 73(3): 531-536.

Cecchini, M., Aytug, H. Koehler, G. J. and Pathak, P. (2010). "Detecting Management Fraud in Public Companies", *Management Science*, 56(7), 1146-1160.

Chao, A., Tsay, P. K., Lin, S., Shau, W. Y. and Chao, D.Y. (2001). "The Applications of Capture-Recapture models to epidemiological data", *Statistics in Medicine*, 20(20): 3123-3157.

Collins, M. F. and Wilson, R. M. (1990). "Automobile Theft: Estimating the Size of the Criminal Population", *Journal of Quantitative Criminology*, 6(4): 395-409.

Dal Bo, E. (2006). "Regulatory Capture: A Review", *Oxford Review of Economic Policy*, 22(2): 203-225.

Danisewicz, P. McGowan, D. Onali, E. and Schaeck, K. (forthcoming) "The real effects of banking supervision: Evidence from enforcement actions", *Journal of Financial Intermediation*, <https://doi.org/10.1016/j.jfi.2016.10.003>

Dechow, P. M. Ge, W. Larson, C. R. and Sloan, R. G. (2011). Predicting Material Accounting Mis-statements”, *Contemporary Accounting Research*, 28(1): 17-82.

Delis, M. C. , Staikouras, P. and C. Tsoumas C. (2017). “Formal enforcement actions and bank behaviour”, *Management Science*, 63(4): 959–987

Dyck, A. Morse, A. and Zingales, L. (2013). “How Pervasive is Corporate Fraud?”, working paper. Chicago Booth.

Economist magazine, (2016). “Culture Wars”, Jan 16th. London.

Ehrlich, I. (1996). “Crime, Punishment, and the Market for Offenses”, *Journal of Economic Perspectives*, 10(1): 43-67.

Feinstein, J. S. (1990). “Detection Controlled Estimation ”, *Journal of Law and Economics*, 33(1): 233-276.

Ferran, E. (2011). “The Break-up of the Financial Services Authority”, *Oxford Journal of Legal Studies*, 31(3), pp.455-480.

Godwin, A. Howse, T. and Ramsey, I. (2017). “A Jurisdictional Comparison of the Twin Peaks Model of Financial Regulation”, *Journal of Banking Regulation*, 18(2): 103-131.

Gneezy, U. and Rustichini, A. (2000). “A Fine is the Price”, *Journal of Legal Studies*, 29(1): 1-17.

Georgosouli, A. (2011). “The FSA's 'Treating Customers Fairly' (TCF) Initiative: What is So Good About It and Why It May Not Work”, *Journal of Law and Society*, 38(3): 405-27.

Goodhart, C. (2017). “Has regulatory reform been misdirected?”, *Journal of Financial Regulation and Compliance*, 25(3): 236 – 240.

Gordon, F. and Squires, D. (2008). “The Deterrent Effect of UK Competition Enforcement”, *De Economist*, 156(4): 411-432.

Group of 30, (2015) *Banking Conduct and Culture. A Call for Sustained and Comprehensive Reform*, Washington, USA.

Gropper, D. M., Jahera Jr. J. and Park, J. C. (2013). “Does it help to have friends in high places? Bank stock performance and congressional committee chairmanships”, *Journal of Banking and Finance*, 37(6): 1986-1999.

Hahn, P. R. Murray, J. S. and Manolopoulou, I. (2016). A Bayesian partial identification approach to inferring the prevalence of accounting misconduct”, *Journal of the American Statistical Association*, 111(513): 14-26

Hanratty, P. (1997). *The Wallis Report on the Australian Financial System: Summary and Critique*, Research Paper No.16, Department of the Parliamentary Library, Canberra.

Harrington, J. E. and Chang? M.H. (2009). “Modelling the birth and death of cartels with an application to evaluating competition policy”, *Journal of the European Economic Association*, 7(6): 1400-1435.

Heckman, J.J. (1979). “Sample selection bias as a specification error”, *Econometrica*, 47(1): 153-161.

House of Commons, House of Lords (2013), *Changing banking for Good. Report of the Parliamentary Commission on Standards Vol.1: Summary and Conclusions and recommendations*, HL Paper 27-1, HC 175-1. London

House of Commons Committee of Public Accounts, (2016). *Financial services mis-selling: regulation and redress. Forty-first Report of the Session 2015-16*, HC 847, London.

Huggins, R. and Hwang, W.H. (2011). “A review of the Use of Conditional Likelihood in Capture -Recapture Experiments”, *International Statistical Review*, 79(3): 385-400.

Johnson, P. E. Grazioli, S., Jamal, K. and Berryman, R. G. (2001). “Detecting deception: adversarial problem solving in a low base rate world”, *Cognitive Science*, 25(3): 355-392.

Jolly, G. M. (1965). “Explicit estimates from Capture-Recapture data with both death and immigration-stochastic model”, *Biometrika*, 52(1): 225-247.

Kuang, Y. F. and Lee, G. (2017). “Corporate fraud and external social connectedness of independent directors”, *Journal of Corporate Finance*, 45: 401-427.

Lancaster, T. and Imbens, G. (1996). “Case-control Studies with contaminated Controls”, *Journal of Econometrics*, 71(1-2): 145-160.

La Porta, R., Lopez-De-Silanes, F. and Shleifer, A. (2006). “What works in Securities Laws?”, *The Journal of Finance*, 61(1): 1-32.

McCraw, T. H. (1975). “Regulation in America: A Review Article”, *Business History Review*, 44(2): 159-183.

Ormosi, P. L. (2014). “A Tip of the Iceberg? The probability of Catching Cartels”, *Journal of Applied Econometrics*, 9(4): 549-566.

Phillips S. J. and Elith, J. (2013). “On Estimating probability of presence from use – availability or presence – background data”, *Ecology*, 94(6): 1409-1419.

Poirer, D. J. (1980). “Partial Observability in Bivariate Probit Models”, *Journal of Econometrics*, 12(2): 209-217.

Polinsky, A. M. and Shavell, S. (1979). “The Optimal Trade-off between the Probability and Magnitude of Fines”, *The American Economic Review*, 69(5): 880-891.

Polinsky, A. M. and Shavell, S. (1984). “The Optimal Use of Fines and Imprisonment”, *The Journal of Public Economics*, 24(1): 89-99.

Polinsky, A. M. and Shavell, S. (1992). “Enforcement Cost and the Optimal Magnitude of Probability of Fines”, *The Journal of Law and Economics*, 35(1): 133-148.

Rossmo, D. K. and Routledge, R. (1990). “Estimating the Size of Criminal Populations”, *Journal of Quantitative Criminology*, 6(3): 293-314.

Skinner, C. P. (2015). “Misconduct Risk”, *Fordham Law Review*, 84, pp.1559-1610

Stigler, G., (1964). “Public regulation of the securities markets”, *Journal of Business* 37(2): 117–14.

Stigler, G. J. (1970). “The Optimal Enforcement of Laws”, *Journal of Political Economy*, 78(3): 526-536.

Tan, D. T., Chapple, L. and Walsh, K. D. (2017) “Corporate Fraud Culture: Re-examining the corporate governance and performance relation”, *Accounting and Finance*, 57(2): 597-620.

Taylor, M. (1995). *Twin Peaks: A Regulatory Structure for the New Century*, (pamphlet), Centre for the Study of Financial Innovation, London.

Thompson, D. and Abbot, M. (2000). Australian Financial Prudential Supervision: An Historical View”, *Australian Journal of Public Administration*, 59(2), 75-88.

Wang, T. Y. (2011). “Corporate Securities Fraud: Insights from a New Empirical Framework”, *Journal of Law, Economics, and Organisation*, 29(3): 535-568.

Wang, T. Y., Winton, A. and Yu. X. (2010). “Corporate Fraud and Business Conditions: Evidence from IPOS”, *The Journal of Finance*, 65(6): 2255-2292.

Wooldridge, J. M. (2002). *Econometric Analysis of Cross Section and Panel Data*, Cambridge, MA: MIT Press.

Veltrop, D. B. and de Haan, J. (2014). “I Just Cannot Get You Out of My Head: Regulatory Capture of Financial Sector Supervisors”, DNB Working Paper, No.410, January, De Nederlandsche Bank NV, Amsterdam.

Yu, F. and Yu, X. (2011). “Corporate Lobbying and Fraud Detection”, *Journal of Financial and Quantitative Analysis*, 46(6): 1865-1891.

Zakolyunina, A. A. (2018) “How common are intentional GAAP Violations? Estimates from a Dynamic Model”, *Journal of Accounting Research*, 56(1): 5-44

Zingales, L. (2015). “Does Finance Benefit Society?”, *Booth Working Papers*, No.15-09, Working paper 117, The University of Chicago.

Appendix

Table 4 - Variables Considered

Variable name	Description	Units	Data source
Capture	Whether a firm has received a Final Notice in the sample period.	A dummy variable recorded in each year of the sample	FSA/FCA Final Notices. ASIC press releases and enforcement updates. ASIC Enforcement Undertakings, enforcement summaries and press releases.
Market	The financial market in which the firm primarily operates. This definition focuses on the firms regulated activities	One of eight classifications including banking, consumer credit, insurance, investments, payments, stockbroker/ asset management/corporate finance /hedge funds and not know	
Type of offence	The classification of the offence. These are not mutually exclusive.	One of six classifications including market abuse, fraud and theft, mis-selling, reporting and compliance, money laundering and complaints handling.	
Offence duration	The 'relevant time period' of the offending as defined within the FSA/FCA Final Notice. Alternatively, the time between the first period of offending and the end of the offending.	Days	
Punishment	The outcome of the Final Notice. Multiple outcomes are commonly reported.	One of six punishments including public censure, fines individual prohibition, variation of regulatory permissions, disgorgement of profits and other measures.	

Regulated firm numbers	The number of regulated financial firms	Number of financial firms	ASIC and FSA/FCA. Annual Reports and Accounts
Regulator Net Assets	The assets net of liabilities to provide a perspective on regulators resources	(£m or equivalent)	
Regulator net operational costs	Net operational costs	(£m or equivalent)	
Regulator employees	The regulatory workforce size	Number of employees	
Regulator Employee costs	The costs of the regulatory workforce.	Total employment costs of the regulator.	
Main Stock Market Index Change %	Change in the appropriate stock market.	Change in the FTSE 100 or ASX 200	Market websites.

Table 5 - Descriptive statistics of conduct offences in Australia and the UK

Final Notice/ order year issued	Australia			UK		
	Cases	Average duration of offence (days)	Average fine £ equivalent†	Cases	Average duration of offence (days)	Average fine £
2002	57	753.34	3,621	15	747.80	913,000
2003	47	1,027.28	5,987	48	635.00	573,750
2004	28	938.92	4,013	89	639.33	804,516
2005	53	1,232.74	0	48	505.39	1,045,366
2006	18	635.53	0	203	335.00	483,044
2007	10	1,652.70	0	144	514.26	242,159
2008	18	886.22	0	218	612.65	431,731
2009	11	746.27	0	182	738.30	816,587
2010	7	973.86	130,665	251	546.55	1,050,650
2011	29	1,278.34	32,188	145	676.44	1,168,463
2012	18	2,799.06	99,455	161	692.83	5,431,582
2013	30	1,449.63	924,204	138	529.36	11,672,332
2014	33	1,519.39	474,254	113	566.83	33,489,179
2015	19	836.63	22,919	114	813.44	21,559,410
2016	26	1,692.00	610,237	182	562.41	12,77,635
Type of Offence %*	Australia	UK	Punishments %*	Australia	UK	
Reporting/compliance	68.4	55.7	Public censure	0	3.5	
Complaints handling	0.0	3.7	Prohibition	8.0	21.7	
Market abuse	6.8	8.2	Fine	7.0	27.4	
Fraud and theft	8.0	17.3	Variation/cancelled regulatory permissions	26.0	51.6	
Mis-selling	49.6	13.8	Disgorgement	41.9	1.6	

Money laundering	0.0	1.3	Other punishment	63.5	0.1
Other offence	2.7	1.4			
Client funds	7.0	0.0			

*percentages do not add up to 100% as more than one type of offence or punishments may be relevant to any one case. † AUS\$ values are converted in £ for yearend annual exchange rates for comparison. For Australia between 2005 and 2009 no fines where levied in Enforcement Undertakings.

Table 6 - Firm level offending and reoffending

Offending firms		2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016
UK	First offence	14	37	57	36	186	128	168	125	183	96	114	96	78	61	151
	Second Offence	0	1	2	4	5	2	4	8	19	14	14	18	22	21	13
	Total offences	14	38	59	40	191	130	172	133	202	110	128	114	100	82	164
Australia	First offence	40	25	25	24	16	8	9	4	6	16	13	20	24	11	19
	Second Offence	0	1	2	1	1	0	3	2	1	2	5	5	2	3	2
	Total offences	40	26	27	25	17	8	12	6	7	18	18	25	26	14	21
Total Regulated firms		2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016
UK	Number	41,791	42,901	53,830	53,172	53,375	54,346	55,182	56,000	57,058	58,918	60,991	66,870	84,596	n/a	n/a
	% offending	0.03	0.09	0.11	0.08	0.36	0.24	0.31	0.24	0.35	0.19	0.21	0.17	0.12	n/a	n/a
Australia	Number	15,245	15,451	15,181	15,340	15,674	16,067	15,480	18,105	21,063	20,960	20,639	20,093	14,783	20,105	21,063
	% offending	0.26	0.17	0.18	0.16	0.11	0.05	0.08	0.03	0.03	0.09	0.09	0.12	0.18	0.07	0.03

Table 7 - Covariates Descriptive statistics

UK	Main Stock Market Index Change %	Regulator Net Assets	Regulator net operational costs	Regulator employees	Regulator Employee costs
2003	11.66	3.30	208.26	2,288	198.30
2004	6.74	17.80	224.70	2,312	119.30
2005	15.92	22.60	246.30	2,356	158.30
2006	9.49	21.60	272.20	2,610	196.50
2007	2.31	13.70	263.70	2,659	199.90
2008	-30.90	6.80	304.70	2,535	197.80
2009	18.66	-19.50	346.50	2,730	208.60
2010	-1.59	17.10	384.30	3,150	269.10
2011	5.88	36.80	450.80	3,337	293.10
2012	3.47	58.00	474.70	3,502	314.00
2013	11.97	40.00	528.20	3,596	326.90
2014	-2.26	22.70	434.50	2,511	278.80
2015	-4.67	-17.10	452.70	3,155	337.00
2016	17.22	-23.10	479.00	3,285	307.80
Australia	Main Stock Market Index Change %	Regulator Net Assets†	Regulator net operational costs†	Regulator employees	Regulator Employee costs†
2003	8.99	19.80	172.60	1,396	96.21
2004	22.72	15.20	196.20	1,531	109.63
2005	17.37	12.00	208.00	1,570	123.86
2006	18.72	20.00	218.00	1,471	127.31
2007	11.81	42.00	256.00	1,610	147.73
2008	-41.41	99.00	274.00	1,669	164.68
2009	31.15	133.00	295.00	1,698	168.81
2010	-2.69	186.00	387.00	1,932	199.31
2011	-14.46	201.62	385.45	1,862	221.50
2012	13.36	220.36	384.49	1,716	217.85
2013	13.73	215.31	411.27	1,832	231.58
2014	0.80	189.55	405.45	1,773	237.05
2015	-1.93	179.95	354.29	1,609	237.05
2016	6.45	46.17	371.22	1,627	198.52

† AUS\$ values are converted in £ for yearend annual exchange rates for comparison.

Table 8 - Estimated capture recapture models and the corresponding DiD estimates

Model	AICc	AICc Weights	Model Likelihood	Num. Par	DiD	S.E.
$\phi(./.)(./.) \rho(\text{did})$	844.4221	0.46229	1	11	-0.5392	0.8191
$\phi(./.)(./.)\rho(\text{did})$ fines	845.8099	0.23097	0.4996	12	-0.2418	0.9205
$\phi(./.)(./.) \rho(\text{did})$ fine amount	846.6009	0.15552	0.3364	13	-0.1991	0.9179
$\phi(./.)(./.)\rho(\text{did})$ all covariates	846.7647	0.14329	0.31	18	1.233	1.0843
$\phi(./.)\rho(\text{did})$	852.5534	0.00793	0.0172	8	-1.1858	0.7099
$\phi(\text{g}^*\text{t}) \rho(\text{did})$	909.8406	0	0	18	-0.7136	0.914
$\phi(\text{did}) \rho(\text{did})$	917.5198	0	0	8	-0.4239	0.8119
$\phi(\text{t}) \rho(\text{t})$	926.2518	0	0	27		
$\phi(.)\rho(\text{t})$	927.0209	0	0	15		
$\phi(\text{j}) \rho(\text{j})$	928.441	0	0	4		
$\phi(\text{j}) \rho(\text{t})$	929.076	0	0	16		
$\phi(.)\rho(.)$	930.253	0	0	2		
$\phi(.)\rho(\text{j})$	930.6389	0	0	3		
$\phi(\text{j}) \rho(.)$	932.1454	0	0	3		
$\phi(\text{j}) \rho(\text{j}^*\text{t})$	933.8564	0	0	30		
$\phi(\text{j}^*\text{t}) \rho(\text{t})$	938.4066	0	0	41		
$\phi(\text{t}) \rho(.)$	938.4595	0	0	15		
$\phi(\text{t}) \rho(\text{j})$	938.6586	0	0	16		
$\phi(.)\rho(\text{j}^*\text{t})$	939.2417	0	0	29		
$\phi(\text{t}) \rho(\text{j}^*\text{t})$	940.8587	0	0	41		
$\phi(\text{j}^*\text{t}) \rho(\text{j})$	949.6203	0	0	30		
$\phi(\text{j}^*\text{t}) \rho(\text{j}^*\text{t})$	952.2445	0	0	53		
$\phi(\text{j}^*\text{t}) \rho(.)$	952.7271	0	0	29		

