

The Investigation and Prosecution of Regulatory Offences: Is There An Economic Case for Integration?

Garoupa, N; Ogus, A; Sanders, Andrew

DOI:

[10.1017/S0008197311000237](https://doi.org/10.1017/S0008197311000237)

License:

None: All rights reserved

Document Version

Publisher's PDF, also known as Version of record

Citation for published version (Harvard):

Garoupa, N, Ogus, A & Sanders, A 2011, 'The Investigation and Prosecution of Regulatory Offences: Is There An Economic Case for Integration?', *Cambridge Law Journal*, vol. 70, no. 01, pp. 229-259.
<https://doi.org/10.1017/S0008197311000237>

[Link to publication on Research at Birmingham portal](#)

Publisher Rights Statement:

© Cambridge University Press

Eligibility for repository checked July 2014

General rights

Unless a licence is specified above, all rights (including copyright and moral rights) in this document are retained by the authors and/or the copyright holders. The express permission of the copyright holder must be obtained for any use of this material other than for purposes permitted by law.

- Users may freely distribute the URL that is used to identify this publication.
- Users may download and/or print one copy of the publication from the University of Birmingham research portal for the purpose of private study or non-commercial research.
- User may use extracts from the document in line with the concept of 'fair dealing' under the Copyright, Designs and Patents Act 1988 (?)
- Users may not further distribute the material nor use it for the purposes of commercial gain.

Where a licence is displayed above, please note the terms and conditions of the licence govern your use of this document.

When citing, please reference the published version.

Take down policy

While the University of Birmingham exercises care and attention in making items available there are rare occasions when an item has been uploaded in error or has been deemed to be commercially or otherwise sensitive.

If you believe that this is the case for this document, please contact UBIRA@lists.bham.ac.uk providing details and we will remove access to the work immediately and investigate.

The Cambridge Law Journal

<http://journals.cambridge.org/CLJ>

Additional services for *The Cambridge Law Journal*:

Email alerts: [Click here](#)

Subscriptions: [Click here](#)

Commercial reprints: [Click here](#)

Terms of use : [Click here](#)



THE INVESTIGATION AND PROSECUTION OF REGULATORY OFFENCES: IS THERE AN ECONOMIC CASE FOR INTEGRATION?

Nuno Garoupa, Anthony Ogus and Andrew Sanders

The Cambridge Law Journal / Volume 70 / Issue 01 / March 2011, pp 229 - 259

DOI: 10.1017/S0008197311000237, Published online: 15 March 2011

Link to this article: http://journals.cambridge.org/abstract_S0008197311000237

How to cite this article:

Nuno Garoupa, Anthony Ogus and Andrew Sanders (2011). THE INVESTIGATION AND PROSECUTION OF REGULATORY OFFENCES: IS THERE AN ECONOMIC CASE FOR INTEGRATION?. The Cambridge Law Journal, 70, pp 229-259
doi:10.1017/S0008197311000237

Request Permissions : [Click here](#)

THE INVESTIGATION AND PROSECUTION OF REGULATORY OFFENCES: IS THERE AN ECONOMIC CASE FOR INTEGRATION?*

NUNO GAROUPA**, ANTHONY OGUS*** AND ANDREW SANDERS****

I. INTRODUCTION

In relation to criminal justice, there is a long-standing assumption, based primarily on the separation of powers doctrine, that law-making, investigation and determination of liability should be undertaken by different institutions. The separation of investigation and prosecution is a more recent phenomenon, although apparently inspired by similar ideas.¹ The approach to institutional design in regulatory law has been significantly different, since in general regulatory agencies exercise not only investigatory and prosecuting powers, but they are also involved, to some extent, in rule-making and adjudication. Indeed, current legislative developments in the UK are taking the process of integration even further, enabling regulatory agencies themselves in many sectors to impose financial sanctions for certain types of contravention (as in the Regulatory Enforcement and Sanctions Act 2008, discussed later).²

The reasons for different arrangements in the criminal justice and regulatory spheres are not obvious, given in particular that the criminal justice system plays an important part in regulatory enforcement.³

* We are grateful to two anonymous referees and the editor, John Bell. Roya H. Samarghandi has provided excellent research assistance. Garoupa acknowledges financial support by FCT, PPCDT/JUR/55752/2006. The usual disclaimers apply.

** University of Illinois.

*** University of Manchester and Erasmus University of Rotterdam.

**** University of Birmingham.

¹ For the sake of exposition, we assume the duality of the two functions of investigation and prosecution. We recognize that it is difficult to make a pure separation between investigation and prosecution because investigators are sometimes involved in prosecution decisions while prosecutors are sometimes involved in investigation. In fact, as we explain later, in most systems there is a continuum of arrangements from full separation to full integration. In addition, we refer to “criminal justice” as a short-hand for police-initiated processes. We recognise that in the UK many “regulatory” bodies (e.g. the HSE) sometimes prosecute in the criminal courts and that the full range of criminal sanctions then become available.

² This is already the case in many civil law jurisdictions. Note the police also have similar sanctions, on a more modest scale (see later).

³ Again for the purpose of exposition, we assume that there is a clear distinction between “regulatory” crimes and “mainstream” crimes, but the qualitative conclusions still hold in the more realistic context of a convoluted and debatable distinction. See, for example, G. Lamond, “What is a Crime?” (2007) 27 O.J.L.S. 609 and A. Ogus, “Regulation and its Relationship with the Criminal Justice Process” in H. Quirk, T. Seddon and G. Smith (eds.), *Regulation and Criminal Justice: Innovations in Policy and Research* (Cambridge 2010) for advocates, and S. Tombs and D. Whyte “A Crisis of Enforcement: the Decriminalisation of Death and Injury at Work” (2008)

Do the virtues of separation vary simply according to the moral content of the law? Are they outweighed by considerations of administrative convenience in the regulatory sphere? Is accountability more problematic for the police, the main enforcement agent for “mainstream” criminal law, than for regulatory agencies? Can we provide an economic explanation for these differences?

The use of economic analysis in understanding law is by now well established notwithstanding the standard critique of the rational model.⁴ There is also a vast literature on criminal law and economics.⁵ It is therefore timely to assess, from an economic perspective, the extent to which investigation and prosecution should be separated. Economics has developed appropriate techniques to assess if a particular (legal) policy is cost-effective. It seems only natural we apply such methodology to understand the advantages and the disadvantages, the benefits and the costs, of particular institutional arrangements.

In this paper, we focus on the relationship between investigation and prosecution. In this context, there is a sharp difference between the regulatory and criminal justice approaches. We offer economic explanations for the differences. We begin with an historical account of the approaches taken in the criminal justice and regulatory spheres, primarily in England and Wales. The focus is on the non-economic arguments underpinning these approaches as, historically, these arguments were mostly not economic in nature. We then develop an economic framework and apply it successively to the criminal justice and regulatory processes.

II. INVESTIGATION AND PROSECUTION OF “MAINSTREAM” CRIME IN ENGLAND AND WALES

A. Historical Background

1. Before the police

Before the introduction of professional police forces in the early-mid 19th century, there was no formal separation between investigation and prosecution, and little separation in practice: criminal investigation was the responsibility of the victim in most cases, as was the instigation and conduct of prosecutions.⁶ This is, at first sight, puzzling, as all parishes

(Centre for Crime and Justice Studies Briefing no 6, Kings College London) and A. Sanders “Reconciling the apparently different goals of criminal justice and regulation: the ‘freedom’ perspective” in H. Quirk et al., *op. cit.* for opponents, of these distinctions.

⁴ See, for example, A. Ogus, *Costs and Cautionary Tales: Economic Insights for the Law* (Oxford 2006).

⁵ See N. Garoupa (ed.), *Criminal Law and Economics* (Aldershot 2009).

⁶ In the second half of the 18th century in Staffordshire, for example, only 13% of the prosecutions were brought by anyone in public office: D. Hay and F. Snyder (eds.), *Policing and Prosecution in Britain, 1750–1850* (Oxford 1989).

had to appoint constables. But in most towns and rural areas the office of constable was an unpaid chore that rotated around eligible parishioners. Constables sometimes arrested suspected offenders when victims asked them to do so, but they rarely investigated themselves, nor was it usual for them to prosecute.⁷

Even the semi-professional police forces which operated in several urban areas in the early 19th century did little to effect prosecutions which were, in consequence, rare.⁸ This was a matter of concern for the state, which increasingly sought to encourage victims to prosecute through reimbursement of expenses and rewards, mainly in relation to property offences.⁹ But this was of limited effectiveness,¹⁰ partly because prosecution often remained a financial risk¹¹ and partly because local sympathies were often divided between criminals and their victims, and so if the latter prosecuted they sometimes risked reprisals.¹²

2. *The police*

The main functions of the professional police after their introduction in England and Wales in the early-mid 19th century were order maintenance and crime prevention. But it seems to have been accepted from very early on that their achievement required effective prosecution arrangements. Thus in 1838, less than 10 years after the Metropolitan Police were created, a Parliamentary Select Committee referred to the police as “qualified to act as prosecutors ... a duty which is now very generally and beneficially devolved upon them ...”.¹³

In reality large numbers of cases – even murders – went uninvestigated and/or unprosecuted for many more years.¹⁴ Though as the 19th century wore on the police increasingly seemed to regard themselves as suitable for filling the gap between victims and the courts, it is not clear by what mechanism or ideology this happened.¹⁵ All we know is that it was a gradual process. The argument that the police should take prosecutions over from victims in all cases had not been accepted

⁷ D. Phillips, *Crime and Authority in Victorian England* (London 1977), ch. 4; J. King, *Crime, Justice and Discretion in England, 1740–1820* (Oxford 2000), p. 75.

⁸ For example, many thieves in Essex sold stolen goods in London at that time. But the London-based Bow Street Runners charged Essex victims for their services if they became involved: the usual rule was “parties must pay” (L. Radzinowicz, *A History of English Criminal Law and its Administration*, vol. 2: *The Movement for Reform* (London 1956), p. 263).

⁹ King (2000), note 7 above, ch. 2 and 3.

¹⁰ C. Emsley, *Crime and Society in England, 1750–1900*, 2nd ed. (Longman 1996), ch. 8.

¹¹ As indicated by the fact that by the end of the 18th century the majority of prosecutors were still not legally represented: J. Beattie, *Crime and the Courts in England, 1660–1800* (Oxford 1986).

¹² D. Hay, “Controlling the English Prosecutor” (1983) 21 *Osgoode Hall Law Journal* 165–186.

¹³ Quoted by L. Radzinowicz, *A History of English Criminal Law and its Administration*, vol. 4: *Grappling for Control* (London 1968), p. 196.

¹⁴ Hay and Snyder, note 6 above; Phillips, note 7 above, p. 114.

¹⁵ Emsley, note 10 above, ch. 8.

even by the 1920s.¹⁶ They were not given any specific prosecution powers in the 19th century, and still have not been; nor are they under any legislative duty to investigate alleged crimes or to prosecute.¹⁷

3. Calls for public prosecutors

Shortly after the police were established, calls began to be heard for independent public prosecutors – by, for example, Royal Commissions and a Parliamentary Select Committee 1845, 1855 and 1875.¹⁸ There was a combination of reasons for this. First, the onus on victims to prosecute was an unfair burden. Secondly, while police powers (*e.g.* to arrest) were increasing, the courts were losing their supervisory jurisdiction over law enforcement officials, particularly over whether or not to prosecute. Thirdly, even if prosecution decisions were generally being made in good faith, policy could still vary greatly from locality to locality, from individual to individual and from offence to offence. Finally, there was criticism of the competence of police prosecutors: it was said that large numbers of criminals were going unpunished.

The principal argument against a public prosecution service was its cost; but there was also the same fear of creating excessive executive power that had obstructed the creation of a national police force (and, indeed, any publicly organised police forces) for many years. This reflected a tradition that celebrated the humbling of authoritarian monarchies. Thus anything that would increase the power of the executive, even at the indirect expense of the courts and of Parliament, was resisted.¹⁹

Linked to the argument against increasing executive power was the desire to preserve the right of direct access to the criminal courts for private citizens. Private prosecution, it was often said at the time, would be more difficult or impossible if there was a body charged with the public duty to prosecute. While it might be thought that private prosecutors could use the law oppressively, prosecutions could only proceed if permitted by the “grand jury”. In the 18th century this protection seems to have worked reasonably well, as grand juries at that time threw out around 25% of cases.²⁰

Eventually, a compromise solution was reached. The office of the Director of Public Prosecutions (DPP) was created in 1879. This was a small, and therefore inexpensive, organisation which remained the only public prosecution service, in relation to police-enforced crime, for over

¹⁶ A. Bodkin, “Prosecution of Offenders: English Practice” (1928) 1 *Police Journal* 353.

¹⁷ However, a refusal or failure to investigate can be held to be *Wednesbury*-unreasonable or a breach of a tortious duty under common law. See A. Sanders, R. Young and M. Burton, *Criminal Justice* (Oxford 2010), chs. 7, 13.

¹⁸ S. Uglow, “Independent Prosecutions” (1984) 11 *Journal of Law and Society* 233–245.

¹⁹ Hay, note 12 above.

²⁰ *Ibid.*

100 years. It dealt with very few cases. By 1981 it still consisted of only around 70 lawyers plus support staff, and most of its work concerned policy and cases where police officers were the suspects.²¹ The rights of private prosecution were preserved, though in 1908 the DPP was given the power to take over, and then (if the DPP wished) drop, private prosecutions. However, this power is very rarely exercised.

4. *The Crown Prosecution Service*

At various times in the 19th century there were experiments with different prosecution arrangements, usually involving local lawyers undertaking the more serious cases. By the mid-20th century there were prosecuting solicitors' departments (employed by police authorities, local authorities or, more rarely, by the police themselves) in all large cities and some rural areas. And, where there were no prosecuting solicitors' departments, local solicitors and barristers were nonetheless instructed on a case-by-case basis in serious and complex cases. But even in London in the 1970s, where prosecuting solicitors were employed directly by the Metropolitan Police, police officers still prosecuted large numbers of cases in magistrates' courts.²²

Prosecuting solicitors were used in different ways by their police forces. In many areas, there was a traditional solicitor-client relationship: the police used them almost exclusively as court advocates, to prosecute cases that the police wanted prosecuted, and with very little discretion to drop or reduce charges. In other areas, the police sought their advice in most cases where there was a prima facie case, and they generally followed that advice; this gave lawyers effective control over decisions in large numbers of non-prosecutions as well as prosecutions.²³ This latter approach was the one favoured by the Royal Commission on Criminal Procedure,²⁴ and formed the basis for the establishment of the Crown Prosecution Service (CPS) in 1985.

The CPS was established as a national service with the DPP as its Head. Prosecutors in each CPS area are under the direction of a Chief Crown Prosecutor. The head office deals with a small number of very serious and sensitive cases and various policy issues that are not the concern of this paper. The CPS was initially given the right to prosecute all cases, and the power to drop any of those cases it wished. But the

²¹ Royal Commission on Criminal Procedure, *The Investigation and Prosecution of Criminal Offences in England and Wales: The Law and Procedure* (HMSO 1981), para. 155–164.

²² For a survey see M. Weatheritt, *Survey of Prosecuting Solicitors' Departments (RCCP Research Study no. 11)* (HMSO 1980); for general discussion see J. Sigler, "Public Prosecution in England and Wales" (1974) *Criminal Law Review* 642.

²³ A. Sanders, "Prosecution Decisions and the Attorney-General's Guidelines" (1985) *Criminal Law Review* 4.

²⁴ Royal Commission on Criminal Procedure, *Report of the Royal Commission on Criminal Procedure* (HMSO 1981).

power to charge (*i.e.* to initiate proceedings) remained with the police, as did the power not to do so. The Criminal Justice Act (CJA) 2003 finally transferred the power to charge from the police to the CPS, though the police retain the right not to prosecute in most cases (see later).

B. The Legal and Constitutional Arguments for Separation of Investigation and Prosecution

By the mid 19th century the cost argument for separating investigation and prosecution no longer operated but the other arguments could still be invoked. Moreover, we should recall that in the 18th century there was a form of separation in that prosecution was the responsibility of the victim/police/prosecutor and of the grand jury. But grand juries, who we have seen were at least partially effective in gatekeeping prosecutions in the 18th century, became ineffective, so that by the late 19th century they were rejecting only 4% of cases.²⁵ Whatever the reasons for this, the arguments for separation in this period were effectively discussions about who or what should undertake the function of the grand jury.

The modern arguments for separation can be summarised under three heads. The first involves abuse of power. There is broad agreement in the literature on police discretion that prosecution policy is used, to a greater or lesser extent, as an instrument of broader enforcement policy.²⁶ This discretion might not in principle be problematic under adversarial “opportunity” systems such as those adopted in England and Wales. However the discretion can be abused if the policy itself is corrupt or if corrupt, unlawful or oppressive means are used to further an otherwise acceptable policy.

Widespread police corruption and malpractice were often alleged in the 19th century. There were seen to be three main sources. First, the increasing prevalence and scale of awards of costs to prosecutors in successful cases made it financially advantageous to prosecute; and created incentives to succeed at all, even unlawful, costs. Second, promotion was based, in part, on successful prosecutions. Third, officers could be bribed to overlook crime.²⁷

Scandals continued throughout the 20th century. For example, vice squad officers in London had been known to take bribes in exchange for overlooking offences of prostitution, obscene publications, gambling and the like. There have also been cases of prosecutions of innocent

²⁵ Hay, note 12 above.

²⁶ Sanders, note 23 above.

²⁷ P.B. Kurland and D.W.M. Waters, “Public Prosecutions in England 1854–9” (1959) 9 *Duke Law Journal* 493.

people aimed at covering up police misconduct through discrediting complainants.²⁸ And promotion and other benefits still accrue on the back of large numbers of successful promotions. For example, police forces have targets to achieve in “bringing offenders to justice”, and it is now well documented that when these targets are in danger of being missed, the police seek easy convictions, no matter how trivial the offence or deserving the suspect, instead of pursuing cases of greater priority.²⁹

The use of wrongful means to further acceptable policies has been dubbed “noble cause corruption” by a Met Police Commissioner. It includes giving repeated “warnings” to offenders who would almost certainly re-offend because those offenders would act as informers, enabling the police to apprehend other criminals; and prosecuting people against whom there are weak cases in order to deter them, or in the hope of a guilty plea; sometimes the weaknesses are covered up by fabrication or planting of evidence.³⁰

The second and third heads, consistency and competence, do not require much elaboration. The Royal Commission on Criminal Procedure noted hugely varying caution (warning) rates across different police forces, although this would seem to be an inevitable product of a system of local policing. There are at least two problems relating to competence. First, it connotes the ability to identify weaknesses that should lead either to cases being dropped or to seek evidence to strengthen them. Second, it arises in relation to the ability to present cases in court. The police are not legally trained for this purpose. There was evidence of large numbers of cases being dismissed at the end of the prosecution case because the police were insisting on prosecution without seeking the prior advice of prosecutors.³¹ For the Royal Commission, this was an indication that investigators would necessarily be less dispassionate than those, especially lawyers, independent of the investigating process.³²

C. Degrees of Separation

The recommendations of the Royal Commission in 1981 that an independent prosecution service be established for the reasons set out above went largely unchallenged. However, the extent of separation in fact adopted was somewhat limited and controversial. The Commission recommended that the line dividing the functions should be drawn not, as in many European jurisdictions, at the end of the investigation,

²⁸ Sanders, Young and Burton, note 17 above, chs. 7, 12.

²⁹ R. Flanagan, *Review of Policing: Final Report* (SO 2008).

³⁰ See generally Sanders, Young and Burton, note 17 above, *passim*.

³¹ Royal Commission (1981), note 14 above.

³² *Ibid.*, para. 6.29.

leaving all prosecution decisions to independent prosecutors, but rather *after* the police had decided whether or not to prosecute. Many commentators argued that this would give the CPS too little power over prosecutions, and empirical research has supported this criticism: weak cases were still being prosecuted, inconsistency remained, and the police still did as they wished in relation to warnings.³³ Only in 2003 was it eventually conceded that prosecutorial independence required the initial power to prosecute. Even so, the police retain the right not to prosecute (that is, to issue warnings or take no action at all). Although, as before the establishment of the CPS, the police can, and often do, consult prosecutors about difficult cases, they need not do so except in the most serious cases. And they usually do not do so. The ratio of prosecutions to formal warnings is around 5:1 in adult cases, and is approximately equal in juvenile cases. The CPS sees very few of these warned cases.³⁴

We should note two curious counter-developments. In one, the police have been given powers to dispose of relatively minor cases themselves by issuing “penalty notices for disorder” (PNDs) under the Criminal Justice and Police Act 2001. These are used increasingly, and for more serious matters than the motoring offences for which fixed penalties have traditionally been used. Thus in 2006, over 100,000 PNDs were issued for “notifiable” (non-motoring) offences. PNDs now constitute 7% of all detections, and the range of offences for which PNDs can be issued has been extended far beyond public order matters.³⁵ But if PNDs are an example of the police regaining lost power, the opposite has happened in relation to multi-million pound frauds. In the Serious Fraud Office (SFO) accountants, police officers and lawyers work together in teams, enabling investigation and prosecution to be handled as seamlessly as possible. But unlike in regulatory agencies (see later), lawyers – rather than investigators or administrators – take the lead.³⁶

The price for the increased formal power of the CPS is the deployment of some prosecutors in selected large police stations so that they can authorise charges (or not) on the spot in order to minimise delay and offer speedy advice. As a consequence, it may be difficult for prosecutors to resist cultural “capture” by the police. As with the SFO, lawyers formally lead in respect of police prosecutions. But in other respects CPS prosecutors are not in the position of SFO lawyers: they still do not lead in respect of police investigations (unlike in many

³³ Sanders, Young and Burton, note 17 above, ch. 7.

³⁴ *Ibid.*

³⁵ R. Young, “Street Policing after PACE” in E. Cape and R. Young (eds.), *Regulating Policing* (Oxford 2008).

³⁶ R.M. White, “Investigators and Prosecutors or, Desperately Seeking Scotland: Re-formulation of the ‘Philips Principle’” (2006) 69 M.L.R. 143.

jurisdictions, at least in serious cases, including Scotland); and many prosecution decisions are taken on police territory and in accord with timetables set by the police.

III. INVESTIGATION AND PROSECUTION OF REGULATORY OFFENCES IN ENGLAND AND WALES

A. Historical Background

Although the criminal justice system has always played a significant role in regulatory enforcement, the arrangements for investigation and prosecution have been markedly different from those for “mainstream” crime. Admittedly, until the “administrative revolution” which occurred in the middle of the 19th century in England and Wales, the prosecution of regulatory offences, just like all other kinds of offence, was a matter for individuals. The only officials with responsibility in this respect were local justices of the peace. While the latter could be stirred into action to deal with offences against property or the person,³⁷ their standing within the local community meant that they were generally reluctant to prosecute local traders for regulatory offences, for example those committed under the early Factories Acts.³⁸

A new approach was taken in the 1833 Factories Act, the primary concern of which was to overcome problems of enforcement.³⁹ That was perceived to require a central government agency with all-embracing powers to make rules and standards, to initiate prosecutions in local magistrates’ courts, and themselves to act as magistrates, thus with power to impose fines and even imprisonment. In fact this latter power proved to be over bold. Subsequent legislation in 1844 removed the conferring of judicial authority on officers of the agency, but they retained the power to prosecute, as well as to issue remedial orders (which nevertheless required a court order if they failed to achieve compliance).⁴⁰

This system, inaugurated in 1844 for the Factory Inspectorate, with powers to investigate contraventions, issue administrative orders and prosecute criminal justice offences, but not themselves to impose sanctions, became the standard pattern within England and Wales for a large variety of regulatory agencies, including those dealing with the

³⁷ D. Hay, *Albion's Fatal Tree: Crime and Society in Eighteenth Century England* (Harmondsworth 1977).

³⁸ H. Arthurs, “Without the Law”: *Administrative Justice and Legal Pluralism in Nineteenth-Century England* (Toronto 1985).

³⁹ Royal Commission on Practices and Proceedings of the Courts of Common Law, *5th Report* (1833–34).

⁴⁰ P. Bartrip and P. Fenn, “The Administration of Safety: the Enforcement Policy of the Early Factory Inspectorate, 1844-1864” (1980) 58 *Public Administration* 87.

mines, the railways, public health, the Poor Law, and impure food.⁴¹ With some exceptions, the tradition is still maintained today: prosecutions are undertaken by national agencies, such as the Health and Safety Executive, the Environment Agency, and the Revenue and Customs Prosecution Office⁴² but also by local authorities responsible for public health and hygiene, housing, trading standards, food and drugs, and child welfare.⁴³ The 1981 Royal Commission, as we have seen, played a key role in relation to mainstream crime, by recognising as a general principle the importance of independent legal expertise in the decision to prosecute and also the division of functions between the police and prosecutor. Although it considered the question of prosecution by non-police, and in particular regulatory, agencies and commissioned a study of practices within this area,⁴⁴ it did not recommend any significant changes to the institutional arrangements.⁴⁵

B. Administrative Organisation and Prosecution

The extent to which, and the circumstances in which, resort is had to prosecutions varies enormously across regulatory agencies and this is reflected in the administrative structures within agencies. In most agencies, there is a legal division which is responsible for assessing and preparing the case for prosecution, and these lawyers may indeed become involved in a case at an early stage in the enforcement process.⁴⁶ But, in some, the investigating officers themselves prepare and prosecute the case in the magistrates' courts, whether or not they have legal training.⁴⁷ If a prosecution is undertaken in the Crown Court, a qualified solicitor or barrister takes responsibility for appearing in court and such a person may be from the agency's legal division or else contractually engaged for the purpose. In all these respects these arrangements are very similar to the police arrangements prior to the 1981 Royal Commission that were described earlier.

There is evidence that legal officers within the organisation can adopt a focus on a case different from that of investigating officers, or at least investigating officers without legal training.⁴⁸ The focus of the legal officer tends to be on matters relating to what is required to secure

⁴¹ Arthurs, note 38 above.

⁴² This prosecutes for Her Majesty's Revenue and Customs (HMRC). See C. below.

⁴³ B.M. Dickens, "Control of Prosecutions in the United Kingdom" (1973) 22 I.C.L.Q. 1.

⁴⁴ K.W. Lidstone, R. Hogg and F. Sutcliffe, "Prosecutions by Private Individuals and Non-Police Agencies" (Royal Commission on Criminal Procedure, Research Study 10, 1981).

⁴⁵ See White, note 36 above.

⁴⁶ Lidstone *et al.*, note 44 above.

⁴⁷ K. Hawkins, *Law as Last Resort: Prosecution Decision-making in a Regulatory Agency* (Oxford 2002).

⁴⁸ J. Rowan-Robinson, P. Watchman and C. Barker, *Crime and Regulation: A Study of the Enforcement of Regulatory Codes* (Edinburgh 1990); R. Kagan, *Regulatory Justice* (London 1978).

a conviction, for example evidence, burden of proof and (where required) *mens rea*; the investigating officer tends to be more interested in the characteristics of the offender of the particular contravention, and whether a prosecution is deemed necessary to secure compliance.⁴⁹

In order to illustrate many of these points, and points of difference, it is instructive to consider an extended example. The Financial Services Authority (FSA) was established in 1997 to regulate financial markets and deal with financial crimes. It now works under the authority of the Financial Services and Markets Act 2000 (FSMA). It regulates banks, insurance companies, financial advisers, the mortgage business (since 2004) and general insurance intermediaries (since 2005). The Regulatory Decision Committee (RDC) of the FSA is responsible for prosecution and adjudication of liability, and has oversight of investigation by the enforcement teams.

Two cases illustrate some of the problems of this type of structure. In the first case, life insurer Legal & General (L&G) was determined by the FSA to have committed widespread mis-selling of with-profits endowment mortgage policies between 1997 and 1999. The RDC fined L&G £1.1m. On appeal, the Financial Services and Markets Tribunal found that that the RDC had insufficient evidence of widespread mis-selling, though it upheld part of the RDC's determination, and so it reduced the original penalty to £575,000. In the second case, the FSA had investigated the roles of various city traders in the placing of a spread bet. The FSA decided that they were guilty of market abuse under the FSMA 2000 but chose to recommend fines rather than prosecution. The RDC then imposed fines of £750,000 on one individual and £100,000 on another. On appeal the Tribunal found that, although there was evidence to support the claim that the spread bet did take place, they were not guilty of non-disclosure, nor of market abuse, because there had been no obligation to disclose in the applicable regulations. The Tribunal subsequently stated that the FSA's decisions in these cases were unreasonable, stating that there was no institution "independent of the enforcement division [of the FSA] to assist the [Regulatory Decisions] Committee in its decision-making".⁵⁰

As a result, the FSA established an independent review of its enforcement and adjudication process in 2005. The subsequent "Strachan Report"⁵¹ criticised the lack of transparency of the RDC and proposed a reformed structure for the FSA with more significant external accountability (for example, by giving accused people and firms a fairer

⁴⁹ C. Abbot, *Enforcing Pollution Control Regulation* (Oxford 2009).

⁵⁰ Details of both cases (the L&G case, and *Paul Davidson & Ashley Tatham v the FSA*, aka "the Plumber" case) can be found on the FSA website: <http://www.fsa.gov.uk/pages/Library/Communication/PR/2005/082.shtml>.

⁵¹ FSA, *Enforcement Process Review (The "Strachan Report")* (2005). The report is available at http://www.fsa.gov.uk/pubs/other/enf_process_review_report.pdf

hearing). The review led to the creation of a Litigation and Legal Review Unit within the enforcement division, hence promoting some separation between investigation and prosecution. The unit is charged with assessing cases before they are referred to the RDC, still the regulator's main sanctioning body. Other important recommendations (there were 44 in total), included (i) the correspondence between the RDC and the enforcement team should no longer be private, but should be disclosed to the party under investigation; (ii) the enforcement case team should no longer have direct access to the RDC after the conclusion of a discussion meeting, without the representatives of the organisation under investigation being present; (iii) any further submissions to be made, either by the case team or the organisation involved, would be disclosed to the other party; (iv) the RDC should be removed from settlement negotiations; (v) the introduction of an explicit discount structure in the level of financial penalty that would otherwise have been recommended by the RDC for organisations which would settle early in the proceedings.⁵² In summary, the report identified many of the disadvantages of an integrated structure of investigation and prosecution, particularly in an organisation, like the FSA, with powers to determine liability. The main consequence of the report was a formal separation of investigation, enforcement and determination of liability within the FSA itself.

C. The Policy Arguments for Integrating the Investigation and Prosecution of Regulatory Offences

The justifications for integrating the investigation and prosecution of regulatory offences do not seem to have been widely discussed, even though, as we have seen, the arrangements have become different in relation to mainstream crime. However, some arguments can be gleaned from policy documents and other sources.

The 1981 Royal Commission devoted only a short passage to the issue. The justifications for separation articulated in relation to mainstream crime, control of abuse, consistency and competence, while relevant to the regulatory context, were apparently outweighed by other considerations. These were not fully articulated by the Commission but seemed to have centred on resources, although there was also an indication that within the regulatory sphere the functions of investigation and prosecution were seen as not being easily separable.

The study of prosecution by non-police agencies undertaken for the Royal Commission was more expansive, though it should not be taken as having been endorsed by the latter. The authors of the study

⁵² *Ibid.*

suggested that the control of abuse and consistency arguments might apply with greater force to the police, as contrasted with non-police agencies.⁵³ In particular, they doubted whether the “excessive zeal” of the police in prosecuting had any real application to regulatory agencies which were notoriously reluctant to prosecute, as is evident now⁵⁴ as it was then. In addition, the technical aspects of regulatory requirements and their complexity meant that the experience of the investigating agency was of greater importance, particularly given the broad discretion on whether prosecution was to be undertaken. Prosecution by the police was considered to be more routine when the threshold of sufficient evidence had been attained.

However, the justifications for integration have not been universally accepted. The Robens Committee, whose Report had been instrumental in establishing the modern system of Health and Safety of Work regulation, had adopted a different stance. They perceived there to be a tension between an officer’s role in advising and educating regulatees regarding their regulatory obligations and a prosecutorial role: if the officer adopted both roles, this might inhibit the effectiveness of the advice and education.⁵⁵ It is noted too that the Australian Law Reform Commission was happy to endorse the separation of investigation and prosecution on the grounds that it helped to ensure consistency of treatment across different areas of regulation,⁵⁶ thus implying that consistency outweighed other considerations. Further, we have seen in relation to the FSA that the concerns within an integrated organisation about “seepage” between investigation and prosecution (and indeed adjudication) are as applicable to regulatory agencies as to criminal justice agencies. And the in-house prosecutors employed by the main tax collection and enforcement agency (Her Majesty’s Revenue and Customs: HMRC) were so ineffective that a new prosecution agency was established for the HMRC.⁵⁷ The particular scandal that precipitated this was the failure to disclose participating informants who were central to a high profile prosecution, the lawyers involved not being powerful enough (*vis-à-vis* the enforcement officers) to either know about this or to disclose it.⁵⁸

⁵³ Lidstone *et al.*, note 44 above.

⁵⁴ See *e.g.* Abbot, note 49 above.

⁵⁵ A. Robens, (1972) *Report of the Committee on Safety and Health at Work 1970–1972* (HMSO 1972).

⁵⁶ Australian Law Reform Commission (ALRC), *Principled Regulation: Civil and Administrative Penalties in Australian Federal Regulation* (1995), paras. 6.26, 6.35.

⁵⁷ The Revenue and Customs Prosecution Agency (RCPO). For details see White (2006), note 36 above. In January 2010 the RCPO became a Division of the CPS (see the CPS website for details).

⁵⁸ M. Levi, “Policing Fraud and Organised Crime” in T. Newburn (ed.), *Handbook of Policing* (Cullompton 2008).

D. Regulatory Enforcement and Sanctions Act 2008

For many years the industry and business lobby has argued that regulation imposes excessive burdens on enterprise. Government responded by establishing units such as the “Better Regulation Executive” (now part of the Government Department of Business, Innovation and Skills). In our context the most significant development was the events following the Hampton Report (2005).⁵⁹ Its proposals aimed to make regulation more “efficient” by reducing the “regulatory burden” on companies and by making sanctions effective. The Hampton inquiry was the initiative of HM Treasury (rather than the Home Office), and so its proposals prioritised efficiency over effectiveness.⁶⁰ The follow-up Macrory Report (2008) argued that regulation should be transparent, targeted, effective and proportionate, that a wider range of non-court sanctions should be created, and that agencies should be prepared to take strong deterrent action when less coercive sanctions do not work (thus leaving prosecution as a last resort).⁶¹ This is entirely consistent with the compliance approach we have seen is characteristic of regulatory agencies. It was, however, taken to extremes by the Regulators’ Compliance Code (RCC),⁶² which states that: “Regulators should recognise that a key element of their activity will be to allow, or even encourage, economic progress and only to intervene when there is a clear case for protection.”⁶³ The “benefits” must justify the “costs”. Prosecutions, unlike in the Macrory Report, are not mentioned once. This approach inevitably requires administrators and enforcement officials, rather than lawyers, to control the whole enforcement process, including whether or not to prosecute.

For our purposes, the most significant outcome of all this was the Regulatory Enforcement and Sanctions Act 2008 (RESA 2008). Part 3 of RESA creates a set of “administrative sanctions” that a range of enforcement bodies will be able to impose. These agencies include the HSE, FSA and so forth. Sanctions include monetary penalties (fixed and variable), enforcement undertakings, and cessation, compliance and restoration notices. Again, some agencies already use some of these sanctions (*e.g.* we have seen that the FSA can impose monetary penalties; and the HSE makes considerable use of enforcement notices). Perhaps providing a wide range of bodies with the ability to impose monetary penalties will be the most innovative and punishment-like

⁵⁹ P. Hampton, *Reducing Administrative Burdens: Effective Inspection and Enforcement* (HM Treasury 2005).

⁶⁰ D. Whyte, “Gordon Brown’s charter for corporate criminals” (2007/8) 70 *Criminal Justice Matters* 31; Tombs and Whyte, note 3 above.

⁶¹ R. Macrory, *Regulatory Justice: Sanctioning in a post-Hampton world*, Consultation Document (Cabinet Office 2008).

⁶² Department for Business Innovation and Skills, *Statutory Code of Practice for Regulators* (2007).

⁶³ *Ibid.*, para. 3.

sanctions introduced by RESA. However, it is too early to tell, as agencies will only be authorised to use these powers if they can demonstrate that they will conform to the principles set out in the Code.⁶⁴ The idea of administrative financial penalties is a return to the 19th century idea, mentioned above, that health and safety agencies should have punitive powers. However, even monetary penalties had been previously provided to some agencies, as we shall see in the next section.

IV. ECONOMIC FRAMEWORK

Against this background of contrasting approaches, and an incomplete and inconclusive debate on the respective merits of separation and integration respectively, we now come to explore economic considerations. We start by looking at the different costs and benefits of integration. We then discuss the extent to which the net balance between the costs and benefits of integration is different for “regulation” than for mainstream criminal justice.

We consider the costs and benefits of integration of investigation and prosecution which seems to be the most relevant dimension in the context of regulation vis-à-vis mainstream criminal justice. In some specific situations, we also consider the possibility of integrating adjudication or the determination of liability. It will be noted in the discussion when we are considering the three stages within one agency.

An economic theory that attempts to explain, or at least shed light on, integration and the contrasting approach taken to it in the mainstream criminal justice and regulatory enforcement contexts involves identifying the nature of the costs and benefits of integration versus separation. For sake of exposition, our discussion focuses on the duality of integration and separation, though we realize that in most systems there is a continuum of arrangements ranging from full separation to full integration that include intermediate arrangements such as coordination with different degrees of control.⁶⁵ The second step is to speculate, on the basis of this analysis, whether the benefits of integration outweigh the costs for regulatory crimes but not for “mainstream” crimes.

⁶⁴ C. Abbot, “The Regulatory Enforcement and Sanctions Act 2008” (2009) 11 *Environmental Law Review* 38.

⁶⁵ T. Krone, “The Limits of Prosecution Authority” (Australian Law Reform Commission) (2003) (http://www.aic.gov.au/conferences/other/krone_tony/2003-11-regnet.html) identifies six different possibilities. Our dual model is about integration (types 1, 5 and 6 in his notation) and separation (types 2 to 4 in his notation, where the different control exercised by the prosecutor is a palliative to the shortcomings we have identified). See also F. Guerrieri, “Law and Order: Redefining the Relationship between Prosecutor and Police” (2001) 25 *Southern Illinois University Law Journal* 353 (arguing for empowerment of prosecutors to supervise police behaviour and ensure proper standards of disclosure, that is, for further integration in our model).

A good starting point is the economic theory of vertical integration that provides useful insights from industrial organisation.⁶⁶ It explains why different stages of the production and distribution of a particular good and service should be integrated in the same firm. Traditionally, economists disliked vertical integration because it increased market power, therefore potentially monopolizing two markets at the same time (upstream and downstream). Later, with the development of transaction cost theory, economists recognized that a significant benefit from integration is the reduction of transaction costs.⁶⁷

A. Benefits of Integration

1. Reduction of transaction costs

Successful law enforcement, in economic terms, requires the benefits of enforcement to exceed the costs (an approach incorporated, as we saw earlier, in RESA 2008). Good decision-making regarding prosecution plays a key role in law enforcement and successful prosecution is determined, among other things, by the nature of investigation and fact discovery. Hence coordination between all these stages helps to achieve successful law enforcement. Integration improves coordination, therefore saving resources or reducing the likelihood of inappropriate condemnations and sanctions given the quality of investigation and prosecution. The costs borne by the system to achieve such coordination are called in economics transaction costs. These transaction costs are reduced by integration.

Obviously these transaction costs are zero when the same person investigates and prosecutes (full integration raises other problems that will be considered in the subsection on costs of integration). However, due to cognitive limitations, full integration might not be always feasible, therefore relying on other solutions may be necessary (teams with leaders, hierarchies with delegation, specialized agencies, etc).

One very relevant example concerns distortions from multiple tasks that cannot be objectively assessed. A single agency internalizes this problem by effectively centralizing decision-making, and hence it could reduce complexity and uncertainty. In other words, investigation and prosecution could be actually easier to understand and more predictable when they are not clouded by conflicting goals or when allocating responsibility is easier and less diffuse (in a single agency rather than

⁶⁶ The economic literature on vertical integration goes back to the 1930s, but the main foundations of the current theory were established in the late 1970s and early 1980s. For a comprehensive survey and critical discussion of the literature, see F. Lafontaine and M. Slade, "Vertical Integration and Firm Boundaries: The Evidence" (2007) 45 *Journal of Economic Literature* 629–685.

⁶⁷ On the empirical literature about transaction cost theory, see also Lafontaine and Slade, note 66 above.

apportioned across different agencies). Consistency in investigation and prosecution within the limited range of cases taken by an integrated agency is easier to achieve than if dispersed among several agencies.

A second aspect to be considered is the reduction of transaction costs caused by conflicts over residual property rights (for example, between the police or any other agency that conducts investigation and the prosecutors).⁶⁸ A single agency gets all the residual property rights over the payoff from achieving success in sentencing as well as over the losses from failing to accomplish the socially perceived goals, hence reducing transaction costs caused by disputes over who gets praised or criticized for the achievements.

2. *Specialisation*

Since transaction costs are reduced and coordination is improved within a single agency, specific investment in highly specialized assets is facilitated. In fact, the conventional problem with specific investment is that, without appropriate coordination, the return could be lost or expropriated by other agents (the so-called hold-up problem, that is, due to the possibility of expropriation, specific investment is reduced therefore hurting the underlying productive activity). In the context of a single enforcement agency, the hold-up problem is less likely and investment in specific assets is more secure.

Specific investment in the enforcement context could be seen at different levels: (i) specialisation in field of law (such as competition or tax law), (ii) specialisation in specific industries or sectors (such as telecommunications or banking), (iii) specialisation in process techniques or in tasks, (iv) specialisation in the nature of enforcement (for example, either in advice regarding compliance or in taking a “hard line” approach to prosecution).

Whatever the level or type of specialisation one can have in mind, the relevant problem is that the added value of pursuing such strategy could be lost when afterwards the output (either in terms of fact discovery or in prosecution) has to be delivered to an agency or a court that does not enjoy the same level of specialisation, does not share the same knowledge or does not have the same ability to appreciate the work done. Integration reduces this possibility. Clearly such advantage

⁶⁸ See, for example, D. Richman, “Prosecutors and their Agents, Agents and their Prosecutors” (2003) 103 *Columbia Law Review* 750–832 and R. Barkow, “Institutional Design and the Policing of Prosecutors: Lessons from Administrative Law” (2009) *Stanford Law Review* 61. Krone, note 65 above identifies the costs of having the prosecution relying on the information solely gathered and provided by the police as a major disadvantage of separation (the Prosecutor Fiscal’s Office in Scotland is the example provided and discussed in his article). However, the same was true of the pre-CPS model in England and Wales. The SFO is an example of an integrated approach that does not suffer from these problems.

is more relevant when the object of the investigation, prosecution and sentencing requires highly technical intellectual capacity or a very particular framework appropriately tailored for the case. In other words, the more investment in specificity is required by investigation and prosecution, the more persuasive is the argument for integration.

Another important consequence of the reduction of the costs imposed by potential hold-ups is the ability for a single agency to explore economies of scale that can allow expansion of core competences. In other words, the resources saved by appropriate and valuable specialisation can be used to develop new and innovative approaches as well as enhance the ability of the single agency for further action. On one hand, this could support a more effective agency. On the other hand, it could create jurisdiction problems by empowering the agency to act outside of the core competences.⁶⁹

3. Fast track intervention

A third advantage of integration is the ability to develop fast track approaches, that is to reduce the procedures required for processing individual cases. One possibility is the mitigation of procedural formalism between the different stages of the process (investigation, discovery, prosecution and sentencing) when this can take place within a common framework managed by a single individual or a single agency.⁷⁰ Another likely consequence is that intervention can be operated at earlier stages, or at least, presumably it is easier to intervene *ex ante facto*, precisely because formalism is mitigated and decisions can be taken quicker and faster.

Faster interventions have benefits and costs. The benefits include the ability to intervene *ex ante facto*, hence reducing the likelihood of production of harm⁷¹, and to provide a speedier solution, presumably minimizing the disruption of investigation and prosecution. The costs are essentially derived from the implementation of rushed decisions that effectively apply a lower standard of proof, and hence might result in errors. In an ideal world, formalism is designed to respond optimally to this trade-off. Fast track could be the optimal response for a subset of cases whereas more formalism would be the optimal response to most cases in mainstream criminal law. In the real world, excessive

⁶⁹ For example, the jurisdictional conflicts between industry-specific regulatory agencies and antitrust authorities can be understood from this perspective. See P. Barros and S. Hoernig, "Sectoral Regulators and the Competition Authority: Which Relationship is Best" (2004) CEPR Working-Paper 4541.

⁷⁰ We are considering the extreme case here. For intermediate arrangements, the benefits and problems detected with fast track interventions are mitigated.

⁷¹ In terms of harmful consequences, it is obvious that certain regulatory crimes can have much more harmful effects than certain kinds of pretty crime. Nevertheless, there is a general perception that crime, on average, is more harmful than regulatory offences.

formalism might be a consequence of bureaucratic interests and rents derived by the legal professions. In this scenario, fast track interventions are potentially more beneficial.

This does not mean that a separated agency cannot develop fast track approaches. For example, the police through warning and restorative justice processes can obviously develop different tracks, and assign many cases to be the object of quick and limited interventions. However, fast track interventions are necessarily bounded by some institutional arrangements that are imposed on enforcement agencies. If prosecution has to take place in court, and the court imposes the sanction, fast tracking is very limited by nature. If on top of that, the separated agency has a limited influence in the sequence of investigation, prosecution and sentencing, that further limits the ability to engage in fast tracking.

4. "Monopoly power"

Some form of bargaining between the investigator or prosecutor and the defendant plays an increasingly important role in relation both to crime (plea-bargaining) and regulation (negotiating compliance). A single agency supervising or operating in all stages of the process increases "monopoly power" when bargaining with a defendant. Clearly individual officials within a single agency are in a much better position to engage in this phenomenon since they do not have to negotiate or coordinate their policies with equivalents in other agencies. Therefore the decisions taken by the single agency are enhanced by the strong bargaining power they have. By comparison, in a multiple agency situation, diversion solutions such as plea-bargaining can be undermined by the lack of ability of credible commitment by the prosecutor or by potential squabbling between prosecutor, police and judge.⁷²

B. Costs of Integration

1. Error costs

Two individuals or institutions making or reviewing decisions are likely to make fewer errors than when the decision is made by a single individual or institutions. Errors in enforcement policy are therefore likely to increase with integration and these include both type I errors (that is, defendants are inappropriately prosecuted) and type II errors (that

⁷² Since plea-bargaining is prevalent without "monopoly power" the latter is clearly not a necessary condition, but one that simply enhances diversionary power. See, further, N. Garoupa and F. Stephen, "Why Plea-Bargaining Fails to Achieve Results in So Many Criminal Justice Systems: A New Framework for Assessment" (2008) 15 *Maastricht Journal of European and Comparative Law* 319.

is, defendants inappropriately escape prosecution).⁷³ The impact is different though depending on the costs of these errors. In the regulatory context, these costs are substantially different and predominantly lower than in the context of criminal law.⁷⁴ Since criminal prosecution and conviction generate a significant cost in terms of stigma and reputation, the cost of type I error is considerably higher in mainstream criminal law (to the point of being the focus of criminal procedure designed to minimize such cost by virtue of imposing a high burden of proof).⁷⁵ For serious crimes, the costs of type II errors are also very significant both in terms of reduced general deterrence (hence more crime in the future) and reduced incapacitation (the same individual is likely to undercomply more frequently). Hence error costs due to integration can vary across different areas of enforcement activity.

One source of enforcement error is capture.⁷⁶ Due to the repeated game nature of the actions taken by an integrated single agency and the weakness of external controls (when there are several agencies, each is subject to a degree of external accountability to the others), the possibility of capture by defendants is higher. The costs of capture can have repercussions in all stages, by contaminating the investigation and fact discovery and by distorting prosecution. The consequence of capture may be type I errors (individuals or firms are overly or unjustly punished due to pressure of competitors or other vested interests) or type II errors (individuals or firms are acquitted or punished too softly).⁷⁷ In both cases it seriously undermines successful enforcement policy. Obviously, separation of agencies does not fully resolve the problem, since monitoring across agencies is not perfect and, for example, non-prosecuted cases are not seen by “higher” agencies.

Another source of enforcement error is behaviour biases in the sense that decision making within given agencies tends to favour either a more strenuous approach, leading to type I errors, or else to a more lenient approach, leading to type II errors. Behavioural biases are more likely when a single agency integrates all stages since there is no system of balance and checks, that is, external accountability is weaker. One important example is the possibility of significant hindsight bias, that

⁷³ Although potentially related, we should make a clear distinction between errors in prosecution (as mentioned explicitly in the text) and errors in adjudication (defendants inappropriately convicted and defendants inappropriately acquitted).

⁷⁴ Notwithstanding that the costs of errors in terms of damage to human health and property resulting from regulating civil use of, for example, nuclear energy may be very high.

⁷⁵ Regulatory action consisting of inappropriately issuing notices to stop an industrial process under environmental legislation can generate significant type I error costs.

⁷⁶ Capture refers to situations where an agency created to pursue public interest goals acts in favor of private interests or adopts the agenda of special interests. Corruption is an extreme example of capture. For a broad discussion, see K.N. Hylton and V. Khanna, “A Public Choice Theory of Criminal Procedure” (2007) 15 *Supreme Court Economic Review* 61.

⁷⁷ These consequences are aggravated if adjudication is also controlled by the integrated agency.

is, an assessment bias due to the natural inclination to avoid discussion about the merits of the investigation.

There are two ways behavioural biases matter. First, by distorting decisions within the agency, since the evaluation of evidence, the assessment of facts or the consideration of appropriate prosecution is no longer as objective as it should be. For example, the single agency might make serious mistakes because it is unable to impartially or dispassionately assess the merits of each case. A costly investigation or a cumbersome fact discovery might lead to prosecution for the simple reason that the agency is unable to accept that such costs are sunk and the decision to follow should be based on the merits, not on the resources already spent. There is a second way behavioural biases matter. Once the single agency realises that some decisions were wrongly driven by internal biases rather than objective assessment of merits, there could be a tendency for covering-up rather than exposing flaws. This is quite likely when the reputation of the agency is at stake and responsibility cannot be shifted away.

2. *Weak accountability*

Weak accountability fosters capture.⁷⁸ Single agencies are less accountable because interaction between agencies is missing. Capture can be external and internal to the single agency.

With respect to external capture, one should consider the costs in terms of valuable resources socially wasted by attempts to capture the decision-maker (and the costs of capturing more than one agency are higher than those of capturing a single agency).⁷⁹ As we know from the economic literature on regulation, there are possible palliatives to undercut the possibility of capture. However, in that case, one should add the costs of deterring and punishing capture. These costs can be substantial in the contexts of very specific knowledge or when stakes are very large.

Another version of capture is the creation of rents by the single agency, that is, rather than external capture, we can observe internal capture. In this example, the problem is not undermining appropriate sentencing due to external pressure, but manipulating the investigation and prosecution to achieve internal goals of appropriating more resources (larger budget, higher salaries for the enforcers, more fringe benefits, even ideological goals). Since external accountability is

⁷⁸ On the economic literature on accountability within the principal-agent model, see D.E.M. Sappington, "Incentives in Principal-Agent Relationships" (1991) 5 *Journal of Economic Perspectives* 45–66.

⁷⁹ This known by economists as "rent-seeking activity": unproductive activities for which an economic rent is sought, that is, an income or profitability above (perfectly competitive) market equilibrium.

weaker when all stages of the process are integrated, the risk of such rent-seeking emerging and eventually prevailing is higher. Again, measures can be taken to reduce this possibility or overcome the problem, but they are costly.

Weak accountability also generates other costs. Due to the fact that external controls are weak and results cannot be easily assessed since specific knowledge is required to understand the actions and decisions of a single agency, excessive capacity financed by the taxpayer becomes a serious concern. There might be a temptation to develop oversized bureaucracies, consuming too many resources and imposing a heavy burden on the Government's budget. An important consequence, as we have seen, could be the manipulation of investigation and prosecution to justify excessive resources consumed by the bureaucracy.

One possibility of mitigating weak accountability within a single agency is to introduce external inspections. Inspections can investigate application of rules and policies, effectiveness, value for money, etc and therefore detect capture or mitigate problems generated by integration. Presumably external inspections are more valuable when the costs of weak accountability are higher and when investigation and prosecution are performed by a single agency. Therefore, given a certain standard for accountability one would think that where there are separate agencies there would be less need for inspectorates than where there is an integrated agency. To a certain extent, separate agencies reduce the need for an inspectorate, but we discuss later how far this approach is actually adopted in England and Wales.

3. Behavioural effects

Depending on how a single agency is structured and organized, having regard in particular to internal mechanisms of control, integration might exacerbate moral hazard (usually shirking, but any kind of hidden or unverifiable action that reduces productivity) and adverse selection problems (due to some hidden or unverifiable set of attributes, investigation and prosecution attract less productive individuals).⁸⁰

It is true that diffuse responsibility and lack of coordination between agencies, when it is difficult to apportion responsibility for failures or mistakes across all the agencies involved, is likely to induce shirking or free-riding behaviour. However, in an integrated agency, team production without an appropriate incentive structure could also create moral hazard. The balance depends very much on the relative merits of internal and external mechanisms of control. For a single agency, we

⁸⁰ See, among others, the canonical articles by B. Holmström, "Moral Hazard and Observability" (1979) 10 *Bell Journal of Economics* 74–91 and S. Shavell, "Risk Sharing and Incentives in the Principal Agent Relationship" (1979) 10 *Bell Journal of Economics* 55–73.

need an internal mechanism of control since external accountability is weaker. A single agency with an adequate hierarchy and an internally accountable team leadership should be able to reduce the problems. However, an oversized bureaucracy, with the risk that it will be unmotivated and dysfunctional, could well exacerbate them. The design of the agency will be crucial in addressing these problems, as is evident from our earlier discussion of the HMRC and FSA. There is no reason to think that motivated regulatory agencies and unmotivated police forces is an inevitable outcome. Within each structure (separated or integrated), moral hazard issues can be addressed by developing appropriate strategies.⁸¹

A similar trade-off applies to adverse selection. A single specialised agency can motivate and attract high human capital in a specific area. At the same time it can attract individuals who are not intrinsically motivated to investigate or prosecute. It could also attract more risk-averse individuals who prefer weaker external accountability and safer enforcement strategies that an integrated framework would foster. The balance between these two opposite effects will govern the set of incentives within a single agency. The rules and standards applied, for example, in the recruitment of human capital will inevitably influence the magnitude of an adverse selection effect.

These behavioural effects are relevant since they partially influence errors and miscarriages of justice, and thus can increase error costs. The costs of moral hazard and adverse selection are measured in terms of how much appropriate investigation and prosecution is undermined. As emphasised, they are not distinctive of an integrated structure, but rather can be augmented or reduced depending on the appropriate incentives within the agency.

The fact that external accountability is weaker when an agency is integrated certainly amplifies the consequences of moral hazard and adverse selection. Nonetheless, an appropriate institutional design of instruments to address motivation and a more significant emphasis on recruiting high human capital could minimize their impact.

C. The Balance of Costs and Benefits of Integration

In economic terms, the case for integration thus turns on the optimal responses to the several trade-offs we have identified. For example, the cost of coordination between investigation and prosecution efforts might be higher in relation to crimes that are difficult to investigate and

⁸¹ For example, by considering the extensive literature on intrinsic and extrinsic motivation. See R. Bénabou and J. Tirole, "Intrinsic and Extrinsic Motivation" (2003) 70 *Review of Economic Studies* 489–520.

prosecute, thus justifying the emergence of a single agency, as with the SFO discussed earlier.⁸²

Krone (2003) has suggested that integrated arrangements are more correlated with private intervention and have become more fashionable. Our approach can explain that observation. Integrated arrangements are easily tailored to specific industries or violations. It could also be a sign of rent-seeking by potential victims and potential violators of the law as we have explained before. Finally, the complexity of many areas of economic activity requires simplified procedure and faster decisions.

V. APPLICATION OF THE FRAMEWORK TO CRIMINAL JUSTICE AND REGULATORY PROCESSES

Here we apply the economic framework developed in the last section to the institutions and settings of mainstream criminal justice and regulatory enforcement respectively, to ascertain whether, and to what extent, there are significant differences between them, thus justifying the contrasting approach taken to the prosecution process. We also relate the analysis to the non-economic arguments outlined in sections II and III.

A. Benefits of Integration

1. Reduction of transaction costs

The principal idea, that coordination across multiple tasks reduces transactions costs and that this might imply integration of two organisations where the tasks are sufficiently related, on the face of it applies to both mainstream crime and regulation, since investigation and prosecution are related. It reflects also the non-economic goal of consistency highlighted in the report of the Royal Commission on Criminal Procedure.

The fact that regulatory agencies are generally smaller, and have a much narrower jurisdictional scope than the police might suggest that inconsistency of decision-making is in that context less of a problem, and that therefore the benefits of integration would not be that large. However, these considerations may be outweighed by two others. First, there is often a significant technical content to regulatory offences, implying that information costs and error costs are lower if a single

⁸² Following Krone, note 65 above, other examples include the Special Prosecutor's Office in the USA as a body of investigation and prosecution for independent pursuit of complaints against high office holders; the reform of the South African National DPP's Office for organized crime, or white collar crimes involving complex accounting transactions in Australia.

agency is involved. Secondly, as both the theoretical⁸³ and empirical literature⁸⁴ reveal, in most regulatory sectors prosecution is very much a last resort, with the enforcement agencies relying on other, less costly, means of enforcement in the great majority of cases. Given such an approach, which is not true for mainstream crime, it can be argued that the agency responsible for prosecuting the small percentage of cases selected for this process should be very familiar with the processes used for the large majority of those contravening, a condition less costly to achieve if the same agency is used for both processes.

Inconsistency of decision-making by the police might suggest the merits of a single agency, but existing problems in England and Wales are perhaps more the consequence of the fact that we do not have a national police force. Therefore reorganisation of the police might have a greater impact on inconsistency than integration of investigation and prosecution.

The second idea, that transaction costs arise from conflicts in the rights to payoffs and that integration would reduce such costs, might have a limited application in England and Wales in comparison with some other countries. To the extent that payoffs do exist, they might nevertheless assume greater importance for regulatory enforcement agencies, than for the police, because the reputation of the police covers a very much broader range of offences than a sectoral regulatory agency and therefore success with prosecution might have a higher profile for the latter.

2. *Specialisation*

The rules enforced by regulatory agencies tend to be more specific, more complex and, as we have seen, more technical, than those applicable to mainstream crime, thus suggesting greater benefits from integration in that context. Of course, the very generality of most principles of criminal law necessarily involves discretion by prosecutors, thus creating obstacles to consistency, as discussed above. On the other hand, this might be balanced by another observed feature that discretion as to whether or not to prosecute seems to play a much greater role in the regulatory sector. More generally, the scale economies argument would seem to have a greater degree of justification for the regulatory sphere, compared to mainstream crime, because the greater diversity and generality of criminal justice issues would require a very large organisation for the scale economies to be achievable.

⁸³ I. Ayres and J. Braithwaite, *Responsive Regulation: Transcending the Deregulation Debate* (Oxford 1992).

⁸⁴ Hawkins (2002), note 47 above; Abbot (2009), note 49 above.

The gains in specialisation are usually perceived as significant in the regulatory context. For example, the recent internal reforms developed by the FSA, as we have discussed before, were designed precisely to solve enforcement deficiencies without diminishing specialisation. The same is true of the SFO, though this should be understood not as a regulatory body but as a highly specialised policing body for one complex type of crime.

3. *Fast track intervention*

One generalisation on fast track intervention is easy to make: the benefits are large in relation to unlawful acts which generate large social costs and which can be prevented or deterred by an early intervention. It is difficult to make comparisons between regulation and crime on this basis alone because both cover a wide spectrum between, at one extreme, relatively trivial offences and, at the other, very serious offences. Nevertheless, activities which give rise to regulatory contraventions are often continuous and in that respect susceptible to preventative intervention, or repeated, in which case the intervention can have a deterrent effect. While many criminal offences are repeated, continuous crimes are much less frequent than continuous regulatory contraventions. In this respect, therefore, and subject to the question of the amount of harm caused, the benefits of early intervention are likely to be greater in the regulatory sphere.

We can also assume that because of the stigma attaching to mainstream criminal convictions, the error costs arising from a wrongful conviction are likely to be higher than those consequent on a wrongful regulatory condemnation.⁸⁵ If that is right, then the costs of fast track intervention are likely to be greater for criminal justice than for regulation, because the earlier the intervention, the greater the likelihood of an erroneous decision.

4. *“Monopoly power”*

Our economic analysis indicates the important benefits which can arise from a single agency bargaining with a defendant. Although plea bargaining plays an increasingly important role in the criminal justice system, it features even more prominently in regulatory enforcement⁸⁶ and with the growing phenomenon of co-regulation involving some form of collaboration in rule-making and enforcement between

⁸⁵ When a criminal conviction results from a regulatory offence, the stigma might apply equally. However, regulatory crimes are viewed by many (e.g. Lamond, note 3 above) as not “real” crimes. While we do not necessarily subscribe to this view, the fact that it is held by many affects the stigma attaching. Further, our argument also concerns the penalties imposed by regulatory enforcement that are “administrative”, and not criminal in nature.

⁸⁶ J. Black, “Talking About Regulation” (1998) Public Law 77–105.

regulators and regulatees,⁸⁷ that latter characteristic is becoming even more significant.

It should be noticed too that the agency costs of a negotiated solution are likely to be higher in relation to mainstream crime where defendants are typically poor, as compared with the regulatory contraventions where defendants are often firms or relatively wealthy individuals.⁸⁸

B. Costs of Integration

1. Error costs

Whether the level of competence exercised in the process would be higher or lower with an integration of the investigatory and prosecuting functions has featured in policy discussions of the issue. As we have seen, consistency and competence have featured in the non-economic arguments for separation. The consequences of (relative) incompetence, mistakes in the prosecution process, are dealt with in our economic framework. We there indicate why in general error costs of both type I (inappropriate prosecutions) and type II (inappropriate non-prosecutions) are likely to be higher in relation to mainstream crime, in particular serious crime, than for most regulatory enforcement. The non-integration of investigation and prosecution can serve as a constraint on such costs.

Integrated agencies, such as the police, may in principle be “captured” by lawbreakers, leading to type II errors, and indeed some police units sometimes are.⁸⁹ But since regulatory agencies deal principally with corporate defendants and representatives of industry, they are, as the literature indicates⁹⁰ for that reason typically more vulnerable to capture than the police. For example, the “revolving doors” phenomenon – regulators hired by the regulated industry – has been discussed and widely documented in the context of regulatory agencies, though less so for police and prosecutors.⁹¹ The empirical evidence also suggests that errors arising from behavioural biases in the enforcement personnel or agency policies are also more likely to arise in the regulatory context, most observed agencies perceiving it to be their

⁸⁷ N. Gunningham and P. Grabosky, *Smart Regulation* (Clarendon 1998); R. Baldwin and J. Black, “Really Responsive Regulation” (2008) 71 M.L.R. 59.

⁸⁸ For a detailed economic analysis, see Garoupa and Stephen, note 72 above.

⁸⁹ The main problem here is “undercover” policing where the police by definition integrate, to some extent, with the policed. See Sanders, Young and Burton, ch. 6, note 17 above.

⁹⁰ P.J. Quirk, *Industry Influence in Federal Regulatory Agencies* (Princeton 1981); G. Slapper and S. Tombs, *Corporate Crime* (London 1999).

⁹¹ M. Maggetti, “De facto Independence after Delegation: A Fuzzy-set Analysis” (2007) 1 Regulation and Governance 271–294. But see Garoupa and Stephen, note 72 above discussing the problems posed by prosecution of wealthy defendants.

responsibility to achieve compliance through persuasion and advice rather than the imposition of penalties.⁹²

These error costs were at the heart of the Strachan report on the FSA.⁹³ Its recommendations suggested mechanisms to reduce the possibility of enforcement mistakes while keeping the advantages of a specialised regulatory agency. They try to secure checks and balances across investigation, prosecution and the determination of liability inside the FSA. Thus these functions were kept in the one single agency, but were separated in order to avoid excessive error costs.

2. Weak accountability

Although the plurality of police forces in England and Wales necessarily constitutes an obstacle to accountability, the legal and the political mechanisms of police accountability are stronger than in relation to regulatory agencies. Many examples could be given. Each police force is accountable to a Police Authority, made up in part of local councillors; and individual police forces, as well as the service as a whole, are scrutinized by the Inspectorate of Constabulary. Arguably the CPS is another body to whom the police are accountable, insofar as a large proportion of police decisions are scrutinized by CPS and the material collected by police forms the basis of CPS prosecutions. And to the extent that a relatively high proportion of police enforcement decisions result in prosecutions that are scrutinised by the courts, there is some accountability to the courts.

In part, this all arises from the higher public profile of the police, relative to that of other agencies, but there are also more important reasons. It has never been easy to place regulatory agencies within the traditional accountability framework of English public law, based primarily on the Diceyan obsession with judicial review: since a very small proportion of agency enforcement decisions result in prosecutions, there is very little accountability to the courts. Nor are there Inspectorates or CPS- and Police Authority-equivalents looking over the shoulders of most such agencies.⁹⁴ Other accountability structures, for example the Ombudsman and the public audit institutions, have emerged in a fragmented and piecemeal fashion. It has been argued that the very diversity of the arrangements can in aggregate exercise powerful constraints on the agencies,⁹⁵ but it is more likely that the incoherence of the arrangements tends to dilute their impact.⁹⁶

⁹² Abbot, note 49 above.

⁹³ FSA, note 51 above.

⁹⁴ Sanders, Young and Burton, note 17 above ch. 7.

⁹⁵ C. Scott, "Accountability in the Regulatory State" (2000) 22 *Journal of Law and Society* 38–60.

⁹⁶ C. Graham, "Is There a Crisis in Regulatory Accountability?" (1995) Discussion paper 13, Centre for Regulated Industries.

An interesting aspect of accountability relates to the role of third parties. In the criminal justice context, increasing attention is being given to victims of crime. The focus of attention is both in relation to restorative justice and more orthodox prosecution decisions. Thus there is now a Code of Practice for Victims of Crime which creates obligations for the police (and other criminal justice agencies) in relation to victims, and court challenges from victims to the police are increasing. This undoubtedly creates pressure on the police in the way they exercise their law enforcement functions. In the regulatory sphere, there is not the same attention to victims. It is true that third parties may play a significant role in monitoring both to detect contraventions and to oversee the enforcement efforts of regulatory agencies, but in some sectors they are deprived of the right themselves to undertake the enforcement processes⁹⁷ and regulatory agencies have no obligations under the Victims Code.⁹⁸

3. Behavioural effects

General comparisons between police and regulatory agencies regarding behavioural effects are not easy to make. It may be that since regulatory agencies are typically more specialised they can attract more skilled and knowledgeable officials than the police. Although that might reduce the moral hazard and adverse selection problems, it exacerbates the “revolving doors” phenomenon. This has very little counterpart in the police force, and predictably can reduce the motivation to enforce conscientiously. Further, the range of values and motivations held by regulatory enforcement officials is almost certainly more diverse than that held by police officers, where the “law and order” goal predominates. Important behavioural effects have been detected in the FSA, as we have seen before. The recent changes in the internal organisation of the FSA aim to promote better mechanisms to reduce moral hazard, minimize adverse selection, and therefore enhance successful enforcement.

VI. CONCLUSIONS

We have developed an economic framework to explore the extent to which investigation and prosecution should be separated, and which can inform scholarly debate and legal policymaking in general. The economic argument is that different types of crime *and the behaviours that give rise to them* impact differentially on society; thus enforcement

⁹⁷ A. Ogus, *Regulation: Legal Form and Economic Theory* (Oxford 2004), pp. 86–87.

⁹⁸ Sanders, Young and Burton, note 17 above, ch. 13; D. Whyte, “Victims of Corporate Crime” in S. Walklate (ed.), *Handbook of Victims and Victimology* (Cullompton 2007).

policy and practice (broadly defined) should vary from one type of crime or offence to another.

We have identified the relevant benefits and costs that should be considered when designing investigation and prosecution processes and structures. As regards benefits, we have focused on the reduction of transaction costs, the gains from specialisation, the possibility of fast track interventions, and the advantages of “monopoly power” in the context of negotiated compliance. With respect to costs, we have discussed error costs, the consequences of weak accountability, and the problems posed by behavioural effects.

Just as the economic approach yields useful insights into enforcement policy,⁹⁹ it can do the same in relation to the design of investigation and prosecution processes and structures.¹⁰⁰ Our framework could easily explain why it is possible and consistent to promote separation of investigation and prosecution in the criminal justice system and yet favour integration in regulatory law. However, the example of the SFO indicates that this conclusion should be refined a little more. It is not so much that separation is most appropriate for traditional criminal justice, and integration for regulatory justice, as that separation is most appropriate for relatively straightforward crime, and integration for more complex matters.

Our discussion has identified several other matters for further consideration. The problems we document in relation to HMRC and the FSA arise in part from their integrated nature – or, at least, from the ways in which integration was operationalised. Different kinds of solution were put in place (including, for HMRC, the “separated” model), yet the integrated solution adopted for the SFO – with lawyers taking the lead – was not adopted. It may be that over time this will be seen to be a mistake. These examples also show that the accountability problems we identify in integrated models have not been addressed in public policy discussions. It is a paradox that there are several (admittedly flawed) mechanisms of accountability in relation to the police, but virtually none in relation to other agencies.

Our framework is important for the application of the recently enacted Regulatory Enforcement and Sanctions Act (RESA) 2008. This will significantly enlarge the enforcement powers of many regulatory agencies when it is implemented. In particular, it will enable agencies to impose administrative financial penalties for contraventions without resort to the criminal justice process. This is a further shift away from the separation model adopted for “mainstream” crime, and it reinforces the integrative approach taken in regulatory law. It is not clear

⁹⁹ Abbot, note 49 above.

¹⁰⁰ See Barkow, note 68 above, for a discussion about the appropriate design of policing and prosecution activities in the United States.

that the thinking underlying RESA 2008 addresses the costs and benefits of this expanded integrated approach. Our model identifies the costs imposed by integration. We discuss possible palliatives to reduce these costs, and therefore promote the additional benefits from integration. The implementation of RESA 2008 should pay close attention to these important palliatives, in particular the institutional design of regulatory agencies and the development of significant external accountability. As we have seen, RESA 2008 creates a potentially powerful set of administrative sanctions similar to many of those already possessed by the FSA. It will be important to ensure that in all agencies the process leading up to the determination of violations, and the imposition of sanctions following such a finding, does not have the same flaws that existed in the FSA.