

Northern Ireland Legal Quarterly

Volume 63 Number 4

EDITOR

PROFESSOR SALLY WHEELER



Queen's University
Belfast

School of Law

Contents

Options between legislative intervention and judicial collaboration: improving the effectiveness and coherence of EU law?	
<i>Mel Kenny</i>	435
Corporate political connection as a determinant of corporate governance in Hong Kong	
<i>Boyce Yung and Philip Lawton</i>	449
Taking the minimum content seriously: Hart's liberalism and moral values	
<i>Gavin Byrne</i>	477
The problem of, and with, financial crime	
<i>George Gilligan</i>	495
Modernising environmental regulation in Northern Ireland: a case study in devolved decision-making	
<i>Sharon Turner and Ciara Brennan</i>	509
The Sword of Damocles: who controls HSBC in the aftermath of its deferred prosecution agreement with the United States Department of Justice?	
<i>Justin O'Brien</i>	533
Leading the charge? Payments for single use carrier bags in Wales	
<i>Lori Frater and Robert G Lee</i>	551

Options between legislative intervention and judicial collaboration: improving the effectiveness and coherence of EU law?

MEL KENNY*

*Professor in Consumer and Commercial Law,
Leicester De Montfort Law School*

Abstract

This article reflects on ways to improve the effectiveness and coherence of EU law in the light of consolidation proposals emerging inter alia from the (Draft) Common Frame of Reference initiative. After supplying a map of effectiveness (section 2) and ways in which gaps in effectiveness have been addressed (section 3), the paper reflects on the reform options: ranging from legislative intervention to jurisdictional realignment (section 4). The paper illuminates the hard cases (Mangold and Abbey National) at the margins of effectiveness (section 5); cases which alert us to more collaborative ways in which to address the 'variable geometry' and fragmentation of EU and 'Europeanised' law. The paper argues for a reconceptualisation of effectiveness and challenges the EU fixation with top-down, legislative intervention.

Keywords: EU law; effectiveness; coherence; Europeanised private law; judicial collaboration

1 Introduction

This paper confronts two intertwined and essentially contested phenomena: the doctrine of effectiveness in EU law and the interface of EU law to national private law. The paper reflects on the chimerical quality of effectiveness and the compromised 'coherence' of EU law in the light of consolidation strategies and alternative suggestions for jurisdictional realignment. The acid question is whether there are more practical solutions than legislative intervention or judicial realignment; whether judicial collaboration would do more to improve the effectiveness and coherence of EU law?

Effectiveness in EU law is a fluid concept. At one level it refers to the applicability of EU norms, at another to the judicial protection of rights granted through those norms. At another level it may be seen as a legal doctrine, taking its place among a battery of interlocking doctrines, at a wider level as a governance concept, reinforcing and shaping legal doctrines.¹ Notwithstanding these multiple roles, there are important caveats to effectiveness, principally in the 'horizontal' application of EU law and the extent to which effectiveness can be invoked to obtain redress as between the 'public' and 'private' domains

* I am grateful to Professor James Devenney (Exeter) and to anonymous referees for comments on drafts of this text. The usual disclaimer applies.

1 A Arnall, 'The Principle of Effective Judicial Protection in EU law: An Unruly Horse?' (2011) 36 *European Law Review* 51; M Ross, 'Effectiveness in the European Legal Order(s): Beyond Supremacy and Constitutional Proportionality?' (2006) 31 *European Law Review* 476; W van Gerven, 'Of Rights, Remedies and Procedures' (2000) 37 *Common Market Law Review* 501.

of EU law.² Beyond these dimensions, the interplay between national and EU courts produces a ‘variable geometry’ of effectiveness.³ Similarly, the interplay of a ‘wildly unsystematic’ body of EU directives⁴ and national legal doctrine and statute⁵ has produced a new body of ‘Europeanised’ private law. Predictably, just as claims that EU law promotes a ‘complete system’⁶ of judicial protection rang hollow, so did the proposition that the traditional pattern of legislative intervention produced coherence attract scrutiny. A number of reform proposals followed:

- legislative intervention in the name of ‘greater coherence’:⁷ initiatives variously advocating a shift to ‘maximum’ harmonisation;⁸ a fully fledged EU Civil Code pursuant to the Draft Common Frame of Reference (DCFR) initiative;⁹ and/or an optional ‘Common European Sales Law’ (CESL) applicable to cross-border transactions;¹⁰
- judicial realignment: involving such steps as abandoning the Court of Justice of the European Union’s (CJEU) interpretative monopoly under the reference mechanism of Article 267 of the Treaty on the Functioning of the EU (TFEU);¹¹ and/or allowing private parties a right of appeal to the CJEU in questions of interpretation;¹²
- deeper judicial collaboration, in recognition of a ‘coordinate legal order’, between national and supranational courts.¹³

These proposals are evaluated in this piece. After mapping effectiveness (section 2) and ways in which gaps in effectiveness have been addressed/cemented (section 3) the paper reflects on the legislative and jurisdictional options (sections 4 and 5). The paper proceeds

- 2 D Leczykiewicz, “‘Where Angels Fear to Tread’: The EU Law of Remedies and Codification of European Private Law” (2012) 8 *European Review of Contract Law* 47. Same author: ‘Enforcement or Compensation? Damages Actions in EU Law after the Draft Common Frame of Reference’ in J Devenney and M Kenny (eds), *The Transformation of European Private Law* (CUP 2013 forthcoming).
- 3 J H H Weiler, ‘The European Court, National Courts and References for Preliminary Rulings – The Paradox of Success: A Revisionist view of Article 177 EEC’ in H G Schemers et al (eds), *Article 177 EEC: Experiences and Problems* (North Holland 1987) 366, at 371.
- 4 S Weatherill, ‘The Consumer Rights Directive: How and Why a Quest for “coherence” has (largely) failed’ (2012) 49 *Common Market Law Review* 1279, at 1280.
- 5 M Kenny, J Devenney and L Fox O’Mahony, ‘Conceptualising Unconscionability in Europe: In the Kaleidoscope of Private and Public Law’, in M Kenny, J Devenney and L Fox O’Mahony, *Unconscionability in European Private Financial Transactions: Protecting the Vulnerable* (CUP 2010) 377.
- 6 Case C-50/00 P *Unión de Pequeños Agricultores v Council* (UPA) [2002] ECR I-6677 para 40 ‘a complete system . . . where natural or legal persons cannot . . . directly challenge Community measures . . . they are able . . . either indirectly to plead the invalidity of such acts before the Community . . . or . . . national courts . . .’.
- 7 European Commission, ‘Action Plan on a More Coherent European Contract Law’ COM(2003) 68.
- 8 Consumer Policy Strategy 2002–2006, COM(2002) 208 final, (2002) OJ C137/2; Consumer Policy Strategy 2007–2013: Empowering Consumers, Enhancing their Welfare, Effectively Protecting Them COM(2007) 99 final.
- 9 C von Bar and E Clive, *Principles, Definitions and Model Rules of European Private Law: Draft Common Frame of Reference* (DCFR) (OUP 2010).
- 10 Introduction of the Optional Instrument, The Common European Sales Law, 11 October 2011 <http://ec.europa.eu/justice/newsroom/news/20111011_en.htm>.
- 11 AG Villalón, Case C-173/09 *Elchinov v Natsionalna zdravnoosiguritelna kasa*, 5 October 2010, nyr Casenote: A P van der Mei (2011) 48 *Common Market Law Review* 1297.
- 12 M Schillig, ‘Inequality of Bargaining Power Versus Market for Lemons: Legal Paradigm Change and the Court of Justice’s Jurisprudence on Directive 93/13 on Unfair Contract Terms’ (2008) 33 *European Law Review* 336.
- 13 C F Sabel and O Gerstenberg, ‘Constitutionalizing an Overlapping Consensus: The ECJ and the Emergence of a Coordinate Constitutional Legal Order’ (2010) 16 *European Law Journal* 511.

to illuminate hard cases at the interface of EU and national law: German reaction¹⁴ to *Mangold*¹⁵ and the English position in *Abbey National*;¹⁶ a juxtaposition disclosing more collaborative ways of managing the Europeanisation process.

2 Mapping effectiveness

Every year a new cohort of law students is inducted into the central articles of faith on the effectiveness of EU law. On the one hand, these concern the forms of action: direct actions (Articles 258 and 259 TFEU); judicial review (Article 263 TFEU); preliminary references (Article 267 TFEU); and actions for non-contractual liability (Article 340 TFEU). On the other hand, these also concern the effects of EU law: direct effect and supremacy; the effects of directives;¹⁷ and state liability. EU law seminarians are soon apprised of the CJEU's duty to ensure the uniform application of EU law and the national courts' duty of 'sincere cooperation'.¹⁸ But effectiveness may also be seen from a compliance perspective:¹⁹ a perennial governance issue given that the EU operates via a system of indirect administration.²⁰ Yet, for the EU law seminarian, the main focus bears on Article 267 TFEU and state liability.²¹ Necessarily, this fixation, given the ad hoc nature of litigation and the variation in national remedies, further compromises the overall coherence of EU law.²²

Yet the more critical EU law seminarian will also be sceptical of the consistency of EU case-law itself; arguing that the jurisprudence fails to disclose a clear methodology. While one may be tempted to agree with Ole Lando that the logic of EU law is one of the simple assertion of bourgeois values,²³ closer inspection produces greater nuance. Weiler famously elaborated three strains of case-law,²⁴ and, subsequently, eras of 'public/private'²⁵ and 'constitutionalisation'²⁶ could be added to those original strains. Does clear judicial policy emerge? Can *Defrenne* be convincingly juxtaposed with the trampling of collective rights in *Laval/Viking/Rijffert*?²⁷ Similarly, the treatment of public interest claims in *Altmark/Kobll/Decker* does not sit easily with the support for the relevant rights in *Schmidberger/Diego Cali*

14 GCC judgment of 6 July 2010, *Mangold-Urteil EuGH Honeywell -Az.2 BvR 2661/06*. Rs. EuZW 2010, 828. Translation <www.bverfg.de/entscheidungen/rs20100706_2bvr266106en.html>.

15 Case C-144/04 *Mangold v Helm* [2006] ECR I-9981.

16 *OFT v Abbey National plc* [2009] UKSC 6. [2010] 1 AC 696.

17 P Craig and G de Búrca, *EU Law* (4th edn, OUP 2008), respectively at 279–87, 287–96 and 296–300.

18 F Mancini, 'The Making of a Constitution for Europe' (1989) 26 *Common Market Law Review* 595, at 606: the national legal order reliant on the CJEU's 'unlimited patience'.

19 F Snyder, 'The Effectiveness of European Community Law: Institutions, Processes and Techniques' (1993) 56(1) *Modern Law Review* 19, at 19–20.

20 *Ibid* 22.

21 Only recently has work emerged on the correlation of compliance and Article 267 TFEU: M Broberg and N Fenger, *Preliminary References to the European Court of Justice* (OUP 2010) 39–46.

22 Snyder (n 19) 51: '(D)ifferences in national remedies affect the extent to which individuals can rely . . . on rights derived from Community law'.

23 O Lando, 'European Contract Law after the Year 2000' (1998) 35 *Common Market Law Review* 821, at 825: 'the guardians . . . of our law . . . grew up in well to do bourgeois homes . . . the legal values of the European brotherhood of lawyers are very similar'.

24 J H H Weiler, 'The Transformation of Europe' (1991) 100 *Yale Law Journal* 2403.

25 H Schepel and W Sauter, *State and Market in European Law* (CUP 2009).

26 Sabel and Gerstenberg (n 13).

27 Case 43/75 *Defrenne v Sabena (No II)* [1976] ECR 455; Case C-438/05 *International Transport Workers' Federation and Finnish Seamen's Union v Viking Line ABP and OÜ Viking Line Eesti (Viking)* [2007] ECR I-10779; Case C-341/05 *Laval un Partneri Ltd v Svenska Byggnadsarbetareförbundet and Others (Laval)* [2007] ECR I-11767; Case C-346/06 *Rijffert v Land Niedersachsen* [2008] ECR I-1989.

Figli.²⁸ Moreover, what is the relationship between the market-focused, precedential case-law and the revisionist, non-autolimitation jurisprudence?²⁹ EU law seminarians may question whether this law transports anything as grand as '*une certaine idée de l'Europe*'³⁰ or whether the CJEU is simply making it up as it goes along. Thus the 'beneficiaries' of the CJEU's case-law are frequently tragic figures, contributing to the Canon but, ultimately, unsuccessful.³¹ Unsurprisingly, national courts have, as a result, become wary of referring questions to what one President of the Bundesgerichtshof dismissed as a 'court of generalists'.³² Equally, perhaps more positively, weaker member states, though they could be found liable, could not be compelled into remedial action; the judicial enforceability of EU law could be decoupled from its judicial enforcement.³³

The effectiveness narrative thus emerges in a crucible of pressures.³⁴ For national systems, EU law represents a multidimensional challenge for their constitutions and (especially?) their constitutional courts.³⁵ Moreover, the 'coherence' of EU secondary law has always been compromised:³⁶ sectoral 'patchwork', 'minimum harmonisation' directives disrupting the residual provisions of national laws; and diversity in interpretation of relevant legal concepts' providing 'an unstable foundation for the internal market'.³⁷ Given these caveats, the heretical question is whether there can be an inappropriate fixation with effectiveness and coherence? Effectiveness is thus a multifaceted doctrine concerning the forms of action (principally Article 267 TFEU) and effects of EU law (direct effect and supremacy) necessary to render EU law effective. Yet a fixation with judicial (state liability) rather than administrative effectiveness and the lack of a guiding CJEU methodology has compromised effectiveness. In the next section, attention turns to the ways in which the gaps in this model have been addressed.

28 Case C-280/00 *Altmark Trans GmbH and Regierungspräsidium Magdeburg v Nahverkehrs-gesellschaft Altmark GmbH (Altmark)* [2003] ECR I-7747; ECJ, Case C-158/96 *Kohll v Union des Caisses de Maladie* [1998] ECR I-1931; Case C-120/95 *Decker v Caisse de Maladie des Employés Privés* [1998] ECR I-1831, Case C-112/00 *Schmidberger v Austria* [2003] ECR I-5659, Case C-343/95 *Diego Cali & Figli Srl v Servizi ecologici porto di Genova SpA (SEPG)* [1997] ECR I-1547.

29 Cases 56 & 58/64 *Consten SARL & Grundig GmbH v Commission* [1966] ECR 299; Case 8/74 *Procureur du Roi v Dassonville* [1974] ECR 837; Cases C-267 & 268/91 *Criminal Proceedings against Keck & Mithonard* [1993] ECR I-6097.

30 P Pescatore, 'The Doctrine of Direct Effect: An Infant Disease of Community law' (1983) 8 *European Law Review* 155, at 157. H Rasmussen, *On Law and Policy in the European Court of Justice: A Comparative Study in Judicial Policymaking* (Martinus Nijhoff 1986); same author, 'Between Self-Restraint and Judicial Activism: A Judicial Policy for the European Court' (1988) 13 *European Law Review* 28. Cf M Cappelletti, 'Is the European Court of Justice Running Wild?' (1987) 12 *European Law Review* 3.

31 D. Nassimian, '... And We Keep on Meeting: De-Fragmenting State Liability' (2007) 32 *European Law Review* 819, at 835: '[T]he mere existence of judicial redress routes does not necessarily lead to ... the required level of protection.'

32 G Nobbe, Leipzig University, 21 January 2000, 'Der Bundesgerichtshof– Innenansichten zur Struktur, Funktion und Bedeutung' <www.uni-leipzig.de/bankinstitut/files/festvortrag.pdf>.

33 Case 387/97 *Commission v Greece* [2000] ECR I-5047; D Chalmers et al, *EU Law* (2nd edn, CUP 2010) 345.

34 J Devenney and M Kenny, 'Unfair Terms, Surety Transactions and European Harmonisation: A Crucible of Europeanised Private Law?' (2009) *Conveyancer and Property Lawyer* 295.

35 G Maduro, *We the Court: The European Court of Justice and the European Economic Constitution* (Hart 1998) 175.

36 M Amstutz, 'Zwischenwelten. Zur Emergenz einer interlegalen Rechtsmethodik im europäischen Privatrecht' in C Joerges and G Teubner (eds), *Rechtsverfassungsrecht* (Nomos 2003) 237. Same author, 'In-Between Worlds: Marleasing and the Emergence of Interlegality in Legal Reasoning' (2005) 11 *European Law Journal* 766.

37 Weatherill (n 4) 1279.

3 Addressing (or cementing) the gaps in effectiveness?

A. DIRECT EFFECT AND SUPREMACY

The gaps in effectiveness were soon exposed.³⁸ In *Kingsgate*, directives ‘merely restating’ treaty principles were held to possess horizontal effect.³⁹ Equally, *Marshall* and *Foster* expanded interpretation of the ‘emanation of the state’.⁴⁰ Meanwhile, ‘indirect effect’ imbued directives with yet broader effects,⁴¹ converting judicial protection itself into a ‘general principle of law’.⁴² Directives thus possessed ‘indirect horizontal effect’.⁴³ Distinctions within secondary law⁴⁴ were further eroded with the advent of ‘incidental effect’.⁴⁵ However, national procedural autonomy and divergence in remedies compromised the effectiveness of these doctrines⁴⁶ such that: ‘the complete and uniform application of Community law may . . . be crowded out, despite its direct effect and supremacy’.⁴⁷

B. STATE . . . V PRIVATE LIABILITY

State liability was also crucial in elaborating effectiveness.⁴⁸ *Brasserie du Pêcheur* extended and revised the criteria,⁴⁹ which are fulfilled where member state or EU institution *manifestly and gravely* disregards the scope of their discretion.⁵⁰ Yet no sooner is effectiveness enhanced in one dimension than another fracture erupts: state liability exposing the limits to liability in the private dimension; *Courage* elaborating that damages were only available against the individual in competition proceedings.⁵¹

C. ARTICLE 267 TFEU

Article 267 TFEU has played a critical role in elaborating effectiveness.⁵² The mechanism provides that the CJEU *interprets* and the national courts *apply* the law;⁵³ promoting both flexibility⁵⁴ and efficiency.⁵⁵ Moreover, referral depoliticises issues in that cases are raised

38 A Albers-Llorens, ‘Keeping Up Appearances: The Court of Justice and the Effects of EU Directives’ (2010) Cambridge Law Journal 455, at 456.

39 Case 96/80 *Jenkins v Kingsgate Ltd* [1981] ECR 911, paras 19–22.

40 Case 152/84 *Marshall (No 1)* [1986] ECR 723; Case C-188/89 *Foster v British Gas* [1990] ECR I-3313.

41 Case 14/83 *Von Colson and Kamann v Land Nordrhein-Westfalen* [1984] ECR 1891.

42 Case 222/84 *Johnston v RUC* [1986] ECR 1651.

43 Ibid paras 9 and 13.

44 Case C-316/93 *Vaneetveld v SA Le Foyer* [1994] ECR I-763; AG Jacobs at 32: ‘The consequences [of horizontal direct effect] . . . could . . . be cushioned by limiting the temporal effect’.

45 Case C-443/98 *Unilever Italia SpA v Central Food SpA* [2000] ECR I-7535; Case C-194/94 *CIA Security International SA v Signalson SA* [1996] ECR I-2201. Case C-201/02 R (*ex parte*) *Wells v Secretary of State for Transport* [2004] ECR I-723, paras 57–8.

46 D Chalmers et al, *EU Law* (CUP 2006) 390–1; Case 39/73 *Reve Zentralfinanz* [1976] ECR 1989 para 5.

47 P J G Kapteyn and P VerLoren van Themaat, *Introduction to the law of the EU* (Kluwer International 1998) 559.

48 Joined Cases C-6 & 9/90 *Franco v Italian Republic* [1991] ECR I-5375 para 28: ‘inherent to the Treaty’.

49 Joined Cases C-46 & 48/93 *Brasserie du Pêcheur & Factortame III* [1996] ECR I-1029, para 51: ‘the rule of law . . . must be intended to confer rights . . . the breach must be sufficiently serious; and there must be a direct causal link between the breach . . . and the damage’. M Ross, ‘Beyond *Franco*’ (1993) 56 *Modern Law Review* 55; H Toner, ‘Thinking the Unthinkable? State Liability for Judicial Acts after *Factortame III*’ (1997) 17 *Yearbook of European Law* 165.

50 Ibid *Brasserie du Pêcheur*, paras 55–7.

51 Case C-453/99 *Courage v Crehan* [2001] ECR I-6297.

52 Case 107/76 *Hoffmann-La Roche v Centrafarm* [1977] ECR 957, para 5.

53 Case C-364/92 *SAT Fluggesellschaft v Eurocontrol* [1994] ECR I-43, para 9: ‘a non-contentious procedure’.

54 Craig and de Búrca (n 17) 461, 528–9.

55 Chalmers et al (n 33) 152.

by private litigants. Furthermore, as the more contentious *application* of EU law remains a national matter, referral upgrades the status of the national court, while shielding the CJEU from political backlash. Simultaneously, however, instability is injected into judicial policy.⁵⁶ Equally, the importance of referral is also due to the limits of judicial review.⁵⁷ Despite criticism,⁵⁸ the traditional, restrictive approach to Article 263 TFEU was confirmed in *UPA* and *Jégo-Quéré*,⁵⁹ in which the virtues of a *complete system of remedies* were famously asserted. Clearly, this 'completeness' has always been limited: the individual unable to show direct and individual concern cannot, alternatively, insist on a reference.

CILFIT defined the two cases in which national courts would not have to refer: first, where a matter has already been clearly decided (*acte éclairé*);⁶⁰ second, where the correct application of EU law is 'so obvious as to leave no scope for any reasonable doubt' (*acte clair*).⁶¹ National courts can only decline referral where they are 'convinced that the matter is equally obvious to the courts of the other Member States and to the Court of Justice', taking the 'state of evolution' of EU law into account.⁶² One can understand the resulting dilemma facing both national court, where delay may speak against referral,⁶³ and CJEU; where both overly abstract and overly specific rulings can be criticised either for a lack of guidance, or for assuming the referring court's jurisdiction.⁶⁴ Notwithstanding attempts to clarify the margins of discretion,⁶⁵ the parameters of referral remain uncertain: dependent neither on the importance of the issue, nor the extent to which the issue depends on interpretation, but on whether the relevant interpretation is open to doubt.

D. STATE LIABILITY AND ARTICLE 267 TFEU

Köbler further refined discourse by addressing liability for failure to refer. In the original reference the CJEU registry had intervened, challenging the referral,⁶⁶ whereupon the Austrian court withdrew the reference, deciding that the national rules were justified. Köbler sought review, damages and a new referral. While admitting a *potential* liability, the CJEU held that liability for failure to refer 'can be incurred only in the exceptional case where the court has manifestly infringed the applicable law'.⁶⁷ This test involved 'the degree of clarity and

56 J Bengoetxea, N MacCormick and L Soriano, 'Integration and Integrity in the Legal Reasoning of the European Court of Justice' in *The European Court of Justice*, G de Búrca and J H H Weiler (eds) (OUP 2001) 43, at 47; J Bengoetxea, *The Legal Reasoning of The European Court Of Justice* (Clarendon 1993). Cf Rasmussen (n 30).

57 Craig and de Búrca (n 17) 460–501, 551–16, 528–30; Case 25/62 *Plaumann & Co v Commission* [1963] ECR 95; Cases 789 and 790/79 *Calpak SpA et Società Emiliana Lavorazione Frutta SpA v Commission* [1980] ECR 1949, para 9.

58 Case C-50/00 *P UPA* (n 6); AG Jacobs, para. 59.

59 Ibid. AG Jacobs, para 102; judgment, paras 43–5; Case C-263/02 *P Commission v Jégo-Quéré & Cie SA* [2004] ECR I-3425, paras 29–39.

60 Case 283/81 *CILFIT* [1982] ECR 3415, para 14.

61 Ibid para 16.

62 Ibid para 20. Cf Case C-461/03 *Gaston Schul* [2005] ECR I-10523, AG Colomer, para 58.

63 T Tridimas, 'Knocking on Heaven's Door: Fragmentation, Efficiency and Defiance in the Preliminary rulings procedure' (2003) 40 Common Market Law Review 9. Sir T Bingham MR, 'The National Judge's View' in M Andenas (ed), *Article 177 References to the European Court: Policy and Practice* (Butterworths 1994) 43, at 46.

64 A Arnall, '*Arsenal Football Club plc v Matthew Reed*' (2003) 40 Common Market Law Review 753, at 769.

65 M Broberg, '*Acte Clair* Revisited: Adapting the *Acte Clair* Criteria to the Demands of the Times' (2008) 45 Common Market Law Review 1383, at 1388–9; contrasting AG Jacobs in Case C-338/95 *Wiener SI GmbH v Hauptzollamt Emmerich* [1997] ECR I-6495; AG Tizzano in Case C-99/00 *Criminal Proceedings against Lyykeskog* [2002] ECR I-4839; and AG Colomer in Case C-461/03 *Gaston Schul Douane-expéditeur BV v Minister van Landbouw, Natuur en Voedselkwaliteit* [2005] ECR I-10513.

66 Following Case C-15/96 *Schöningh-Kougebetopoulou v. Freie und Hansestadt Hamburg* [1998] ECR I-47.

67 Case C-224/01 *Köbler v Austria* [2003] ECR-I 10239, para 53.

precision of the rule infringed, whether the infringement was intentional, whether the error of law was excusable . . . the position taken . . . by a Community institution and non-compliance by the court in question with its obligation to make a reference'.⁶⁸ Thus the CJEU held the infringement insufficiently serious.⁶⁹ While important implications result,⁷⁰ *Köbler* begs the question of the type of judicial error that will prove sufficiently serious. Though the abuse of *acte clair* was identified as serious, no further clarity was supplied.⁷¹

Further clarification was anticipated in *Traghetti*. The case originally concerned damages for unfair competition; claims rejected at first instance and on appeal. The applicants appealed to the Italian Supreme Court, which confirmed the previous decisions, whereupon an action was brought for errors in interpretation and failure to refer. Two questions were referred: whether failure to refer could lead to liability; and whether the Italian liability rules were compatible with EU law. Once again the CJEU registry challenged the referral and the tribunal withdrew the first question. In its judgment the CJEU confirmed *Köbler*,⁷² spurning the opportunity to extend or revolutionise liability for failure to refer.⁷³

E. PROVISIONAL CONCLUSIONS

The steps taken to address gaps in effectiveness have contributed to a variable geometry, exemplified in the piecemeal elaboration of 'indirect', 'indirect horizontal' and 'incidental' effects of directives. Meanwhile, the effectiveness of EU law has remained compromised by national procedural autonomy and national remedies. Similarly, the extension of liability into the private dimension has been constrained. Meanwhile, the place of Article 267 TFEU and the CJEU monopoly on interpretation has injected instability into judicial policy. Finally, the elaboration of judicial liability for failure to refer remains compromised, further undermining the coherence of legal remedies.

4 Reaction: between legislative intervention and jurisdictional realignment

A. LEGISLATIVE INTERVENTION

Given the *incompleteness* of EU law and the emergence of the unsystematic body of EU secondary law, many find the case for legislative intervention, for such measures as an EU Civil Code, compelling.⁷⁴ To this end the European Commission charged the Study Group on a European Civil Code (SGECC) to investigate the available options.⁷⁵ Six volumes of a DCFR subsequently emerged in 2009.⁷⁶ Meanwhile, in 2010, the Directorate-General for Justice established an 'Expert Group' to revise and restructure elements of the DCFR

68 Case C-224/01 *Köbler* (n 67) para 55.

69 Ibid paras 122–4. B Beutler, 'State Liability for Breaches of Community Law by National Courts: Is the Requirement of a Manifest Infringement of the Applicable Law an Insurmountable Obstacle?' (2009) 46 Common Market Law Review 773; S Drake, 'State Liability under Community Law for Judicial Error: A False Dawn for the Effective Protection of the Individual's Community Rights?' (2004) Irish Journal of European Law 34, at 35.

70 P J Wattel, '*Köbler*, *CILFIT* and *Welthgrove*: We Can't Go on Meeting Like This' (2004) 41 Common Market Law Review 177, at 178–9: 'if a national highest Court wants to avoid the real risk of making its government liable, it had better ask for a preliminary ruling . . . in . . . every case involving a question of EC law'.

71 A Arnall, 'The Use and Abuse of Article 177 EEC' (1989) 52 Modern Law Review 622; Drake (n 69) 49–50.

72 Case C-173/03 *Traghetti del Mediterraneo SpA v Italy* [2006] ECR I-5177, paras 46 and 32.

73 Wattel (n 70) 182.

74 Green Paper on Policy Options for Progress Towards a European Contract Law for Consumers and Businesses, Brussels 1 July 2010 COM(2010) 348 final. (see option 7).

75 M Kenny, 'The 2004 Communication on European Contract Law: Those Magnificent Men in their Unifying Machines' (2005) 30 European Law Review 724–42.

76 DCFR (n 9).

relevant for contract law.⁷⁷ Finally, in October 2011, proposals were tabled for a regulation-based 'optional instrument' for a CESL, recycling the 2008 proposal on a consumer rights directive as an opt-in second regime of cross-border contract law.

It is beyond the remit of this paper to address all aspects of this *Professorenrecht* save to note, first, that a *legislative* response to the lack of systematicity in EU law at its interface with national private law is under review.⁷⁸ Second, that it has been German academics, notwithstanding (or because of?) the contributions of the German judiciary to the networked case-law, who have driven this process forward.⁷⁹ Thirdly, that the CESL initiative, to the extent that it would drive further fragmentation between purely domestic and Europeanised private law, would ultimately serve as a stalking horse for wider codification. Indeed, as the European Law Institute has already confirmed in its 2012 Statement on the CESL, the CESL already requires comprehensive upgrading.⁸⁰ Finally, codification would extend EU liability rules into the private law domain; begging further questions as to which norms of EU law, hitherto unenforceable against individuals, would henceforth become so enforceable;⁸¹ codification thus threatens to unpredictably realign the whole reach of EU law.

Meanwhile, a more pragmatic response, a shift to measures of 'maximum' harmonisation has been adopted to target the 'unsystematic' legislative *acquis*, exemplified inter alia in the 2011 Consumer Rights Directive.⁸² Yet review of the legislative history of this measure describes the emergence of a 'mouse of a directive'.⁸³ While the original 2008 proposal advocated replacing the eight consumer rights directives with a single instrument,⁸⁴ Directive 2011/83 simply replaced the doorstep (Directive 85/577) and distance-selling directives (Directive 97/7) and modestly amended the Unfair Terms (Directive 93/13) and Sale of Consumer Goods and Associated Guarantees Directive (Directive 99/44). Revealingly, no resort was made to the array of interpretative aids of the DCFR and the Expert Group's Feasibility Study. A *quid pro quo* emerges, in recognition of the fact that: '[c]oherent EU law comes at a cost: incoherent national law'.⁸⁵

77 Expert Group of a Common Frame of Reference in the Area of European Contract Law, Commission Decision 2010/233/EU (2010) OJ L105/109. Subsequently, the feasibility study: <http://ec.europa.eu/justice/contract/files/feasibility_study_final.pdf>.

78 B Markesinis, 'Why a Code is Not the Best Way to advance the Cause of European Legal Unity' (1997) 5 European Review of Private Law 519; P LeGrand, 'Antivonbar' (2006) Journal of Comparative Law 1.

79 C von Bar, 'Die Funktionen des Gemeinsamen Referenzrahmens aus der Sicht der Verfasser des wissenschaftlichen Entwurfs' in M Schmidt-Kessel (ed), *Der Gemeinsame Referenzrahmen. Entstehung, Inhalte, Anwendung* (Sellier 2009) 23–33; H Schulte-Noelke, 'The Way Forward in European Consumer Contract Law: Optional Instrument Instead of Further Deconstruction of National Private laws' in C Twigg-Flesner (ed), *The Cambridge Companion to European Union Private Law* (CUP 2010) 131–46.

80 Statement of the European Law Institute on the Proposal for a Regulation on a Common European Sales Law COM(2011) 635 final <www.europeanlawinstitute.eu/fileadmin/user_upload/p_eli/Publications/S-2-2012_Statement_on_the_Proposal_for_a_Common_European_Sales_Law.pdf>.

81 Leczykiewicz (n 2).

82 Directive 2011/83/EU on consumer rights (2011) OJ L304/64.

83 Weatherill (n 4) 1290.

84 Directive 85/577 to protect the consumer in respect of contracts negotiated away from business premises [1985] OJ L372/31; Directive 90/314 on package travel [1990] OJ L158/59; Directive 94/47 on the purchase of the right to use immovable properties on a timeshare basis [1994] OJ L280/83 (replaced by Directive 2008/122/EU, [2009] OJ L33/10); Directive 97/7 on the protection of consumers in respect of distance contracts [1997] OJ L144/19 (amended by Directive 2007/64/EU, [2007] OJ L319/1); Directive 98/6 on consumer protection in the indication of the prices of products offered to consumers [1998] OJ L80/27; Directive 98/27 on injunctions for the protection of consumers' interests [1998] OJ L166/51; Directive 99/44 on certain aspects of the sale of consumer goods and associated guarantees [1999] OJ L171/12.

85 Weatherill (n 4) 1315.

B. JURISDICTIONAL REALIGNMENT

While Article 267 TFEU has provided the catalyst for constitutionalisation, an important caveat has been the CJEU's monopoly over EU law interpretations. As the CJEU underscored in *Rheinmühlen I*, the constitutionalisation of EU law is limited: national provisions do not trump the lower courts' ability to refer; the procedural autonomy of national courts is bypassed to promote effectiveness.⁸⁶ In *Elchinov* Advocate-General Villalón called for a nationalisation of Article 267 TFEU jurisdiction, asserting that the growth of the EU had transformed the national judicial role such that the higher courts now constituted a 'keystone' in judicial cooperation.⁸⁷ Furthermore, EU law had reached a 'level of maturity which allows it to ensure its own practical effectiveness before the (national) courts'.⁸⁸ In drawing these conclusions the Advocate-General found his case supported by the availability of flanking doctrines: *Köhler*-based state liability;⁸⁹ *Commission v Italy*-based infringement actions;⁹⁰ and *Kühne & Heitz*-style review of administrative decisions.⁹¹ Yet the court declined to follow the Advocate-General, maintaining that: 'national courts are . . . free to exercise . . . discretion at whatever stage of the proceedings they consider appropriate'.⁹² In contrast, an alternative, more modest jurisdictional realignment might be the institution of an interpretative right of appeal from the national court to the CJEU.⁹³ Yet just as the nationalisation of Article 267 TFEU would undermine judicial collaboration, so might the institution of such a right appear disproportionate.

C. CAVEATS

Given the caveats identified to legislative intervention and jurisdictional realignment, we need to turn to the hard cases at the margins; to reassess the depth of judicial collaboration. In the next section, the variation in judicial approaches is elaborated, an analysis marking ways in which the effectiveness and coherence of EU law could be improved.

5 Margins of effectiveness

A. MANGOLD: COLLABORATIVE APPLICATION OF EU LAW

In *Mangold*, it was held that where directives implemented 'general principles' of EU law they possessed horizontal direct effect. German labour law thus infringed Directive 2000/78/EC, notwithstanding that the implementation date had not expired.⁹⁴ The directive established a non-discrimination right recognised as a 'general principle of Community law',⁹⁵ outranking any national policy discretion. The national legislator, as an

86 Case 166/73 *Rheinmühlen-Düsseldorf v Einfuhr- und Vorratsstelle für Getreide und Futtermittel* [1974] ECR 33 para. 4: 'a rule of national law . . . cannot deprive the inferior courts of their power to refer to the Court'.

87 Case C-173/09 *Elchinov* (n 11), AG Opinion, para 23.

88 Ibid AG para 31.

89 Ibid AG para 24.

90 Ibid AG para 24; Case C-129/00 *Commission v Italy* [2003] ECR I-14637.

91 Ibid AG para 27; Case C-453/00 *Kühne & Heitz v Produktschap voor Pluimvee en Eieren* [2004] ECR I-837.

92 Ibid Judgment paras. 26, 27 and 31.

93 Schillig (n 12).

94 §14(3) TzBfG (*Gesetz über Teilzeit und befristete Arbeitsverträge*). Directive 2000/78/EC establishing a general framework for equal treatment in employment and occupation [2000] OJ L303/16.

95 Case C-144/04 *Mangold* (n 15) para 75. 'Editorial: Horizontal Direct Effect – A Law of Diminishing Coherence' (2006) 43 Common Market Law Review 1; D Schiek, 'The ECJ Decision in *Mangold*: A Further Twist on Effects of Directives and Constitutional relevance of Community Equality Legislation' (2006) 35 Industrial Law Journal 329, at 333–4.

'agent' of EU law, could not pass measures incompatible with general principles.⁹⁶ Despite criticism,⁹⁷ this approach was confirmed in *Kücükdeveci* where the relevant national measures were also disapplied.⁹⁸ Despite criticism of incompatibility with the German Constitutional Court's (GCC) *Lisbon* ruling,⁹⁹ the GCC confirmed this approach,¹⁰⁰ underscoring the importance of non-discrimination¹⁰¹ and elaborating the balancing necessary between CJEU interpretation and GCC review:

ultra vires review must be exercised reservedly by the [GCC] . . . [I]n each case . . . the task and status of the independent supranational case-law must be safeguarded. This means, on the one hand, respect for the Union's own methods of justice . . . Secondly, the [CJEU] has a right to tolerance of error. It is hence not a matter for the [GCC] in questions of the interpretation of Union law . . . to supplant the interpretation of the [CJEU] with an interpretation of its own.¹⁰²

Intriguing is the 'tolerance of error' allowed the CJEU: the national court must desist from ruling that the CJEU selected the wrong interpretation, or replace its own interpretation for that of the CJEU. National procedural autonomy emerges from this jurisprudence not as an absolute but as subject to general principles; the principle of effective judicial protection being recognised as a general principle of EU law.¹⁰³ What is revealed in such cases is a profound collaboration in which national and supranational tribunals: 'agree to defer to each other's decisions'.¹⁰⁴ Sabel and Gerstenberg trace further collaboration in *Schmidberger*¹⁰⁵ and *Omega*,¹⁰⁶ noting the *reciprocal monitoring* by national courts and private parties of the EU legal system.

The case for improving judicial coordination is further enhanced given the implications of EU accession to the European Convention on Human Rights (ECHR). In this regard Protocol 8 TEU specified that accession should 'make provision for preserving the specific characteristics of the (EU) and (EU) law'¹⁰⁷ thereby focusing attention on the relationship between CJEU and the ECHR. Predictably, in its discussion document of 5 May 2010, the CJEU asserted its monopoly: '[t]o maintain uniformity in the application of EU law and to

96 *Mangold* (n 15) paras 77–8.

97 Chalmers et al. (n 33) 538–40. Case C-427/06 *Birgit Bartsch v Bosch und Siemens Hausgeräte (BSH) Altersfürsorge GmbH* [2008] ECR I-7245; Case C-227/04 *P Lindorfer v Council of the European Union* [2007] ECR I-6767. Critique: Case C-411/05 *Felix Palacios de la Villa v Cortefiel Servicios SA* [2007] ECR I-8531; AG Mazák, para 94, P Craig, 'The Legal Effect of Directives: Policy, Rules and Exceptions' (2009) 34 *European Law Review* 349.

98 Case C-555/07 *Kücükdeveci*, 19 January 2010 nyr paras 30 and 56. Casenote (2010) 47 *Common Market Law Review* 1161.

99 GCC judgment of 30 June 2009 <www.bverfg.de/entscheidungen/es20090630_2bve000208en.html>; A Fischer-Lescano, C Georges and A Wonka, 'The German Constitutional Court's Lisbon Ruling: Legal and Political Science Perspectives', ZERP Discussion Paper 1/2010 <www.zerp.uni-bremen.de>.

100 GCC (n 14).

101 C O Lenz, 'Erfreuliche Momentaufnahme– Zum Mangold-Urteil des GCC' (201) *EWS* 9, 1.

102 GCC (n 14) para 66.

103 Case C-432/05 *Unibet v Justiekansler* [2007] ECR I-2271, para 37; N Reich, 'Rights without Duties? Reflections on the State of Liability Law in the Multilevel Governance System of the Community: Is There a Need for a More Coherent Approach in European Private Law?' EUI Working Paper Law 2009/10; C Kakouris, 'Do the Member States Possess Judicial Procedural Autonomy?' (1997) 34 *Common Market Law Review* 1389.

104 Sabel and Gerstenberg (n 13) 511.

105 Case C-112/00 *Schmidberger* (n 28) para 93: '[T]he national authorities were . . . entitled . . . to consider that the legitimate aim . . . could not be achieved . . . by measures less restrictive of intra-Community trade.'

106 Case C-36/02 *Omega Spielballen- und Automatenaufstellungs-GmbH v Oberbürgermeisterin der Bundesstadt Bonn* [2004] ECR I-9609, paras 37 and 38.

107 Protocol 8 TEU, Article 1.

guarantee the necessary coherence of the Union's system of judicial protection'¹⁰⁸ adding: 'the possibility must be avoided of the ECHR . . . [deciding] on the conformity of an EU act with the Convention without the [CJEU] first having had an opportunity to give a definitive ruling'.¹⁰⁹ Yet such a 'reservation' of jurisdiction is questionable, amounting to a privileging of the CJEU.¹¹⁰ More persuasively, the case can be made that the ECHR should be bound to uphold the central rights to a fair trial and effective remedy under Articles 6 and 13 ECHR. However, the stability of judicial protection on offer can be questioned: are national courts up to the task? How sophisticated are national approaches to referrals under Article 267 TFEU?

B. *ABBEY NATIONAL*: THE 'NATIONAL APPLICATION' OF EU LAW?

Abbey National, in disclosing an almost mirror-image of the judicial technique in *Mangold*, flags up the limits of judicial collaboration. The case arose pursuant to complaints to the Office of Fair Trading (OFT)¹¹¹ on the extent of bank charges.¹¹² Many consumers challenged the fairness of these charges under the Unfair Terms in Consumer Contracts Regulations 1999 (UTCCR); regulations transposing the Unfair Terms' Directive (UTD) (93/13/EEC).¹¹³ The UTCCR adopted a full 'copy-out' transposition to 'reflect more closely the wording of the Directive'.¹¹⁴ Regulation 6(2), implementing Article 4(2) UTD, provided:

In so far as it is in plain intelligible language, the assessment of fairness of a term shall not relate—

- (a) to the definition of the main subject matter of the contract, or
- (b) to the adequacy of the price or remuneration, as against the goods or services supplied in exchange.¹¹⁵

As CJEU case-law confirms, the UTCCR must be interpreted in the light of the purpose of the UTD.¹¹⁶ The question facing the court pursuant to Regulation 6(2) UTCCR¹¹⁷ was thus whether the terms and charges represented core 'subject matter' or 'adequacy of price' terms exempt from review, or were peripheral terms and subject to the unfairness test. While the freedom of the court to pursue a pro-consumer application of the UTCCR had, previously, been heavily constrained in *Kleinwort Benson*,¹¹⁸ *First National* had appeared to

108 CJEU Discussion document on certain aspects of the accession of the EU to the ECHR of 5 May 2010, para 8, <http://curia.europa.eu/jcms/upload/docs/application/pdf/2010-05/convention_en_2010-05-21_12-10-16_272.pdf>. T Lock, 'Walking on a Tightrope: The Draft ECHR Accession Agreement and the Autonomy of the EU Legal Order' (2011) 48 Common Market Law Review 1025.

109 CJEU (n 108) para 9.

110 N Reich, 'Beitritt der EU zur ERMK—Gefahr für das Verwerfungsmonopol des EuGH' (2010) EuZW 641.

111 J P Devenney, 'Gordian Knots in Europeanised Private law: Unfair Terms, Bank Charges and Political Compromises' (2011) 62 Northern Ireland Legal Quarterly 33. OFT, *Personal Current Accounts in the UK* (OFT 918, April 2007) 2; OFT, *Personal Current Accounts in the UK: A Market Study* (OFT 1005, July 2008).

112 OFT's joint reply and defence (11 November 2007) <www.oft.gov.uk/shared_ofi/personal-current-accounts/OFTs-joint-reply-and-def.pdf>.

113 Directive 93/13/EEC on unfair terms in consumer contracts [1993] OJ L95/29; M Kenny, 'Orchestrating Sub-prime Consumer Protection in Retail Banking: *Abbey National* in the Context of Europeanised Private Law' (2011) 19 European Review of Private Law 43; C Twigg-Flesner, 'Time to Do the Job Properly: The Case for a New Approach to EU Consumer Legislation' (2010) 33 Journal of Consumer Policy 355.

114 Explanatory Notes, UTCCR.

115 Regulation 6(2), UTCCR.

116 Case 14/83 *Von Colson and Kamann* [1984] ECR 1891, para 26; Case C-106/89 *Marleasing* [1990] ECR I-4135, para 8; *Office of Fair Trading v Foxtons* [2009] EWCA Civ. 288, para. 42; *Director General of Fair Trading v First National Bank plc* [2001] UKHL 52, para 31.

117 UTCCR 1999, SI 1999/2083; Regulation 6(2).

118 *Kleinwort Benson Ltd v Lincoln City Council and Others* [1998] UKHL 38; [1999] 2 AC 349.

settle the matter in favour of a restrictive reading of Regulation 6(2) UTCCR.¹¹⁹ Yet, while the High Court and Court of Appeal followed *First National*,¹²⁰ the Supreme Court distinguished the case.¹²¹ Analysis of the legislative history of Article 4(2), it was held, uncovered intent to enhance contractual freedom rather than contractual fairness.¹²² Moreover, the extent of the charges, representing over 30 per cent of revenue, proved that the terms were core parts of the bargain and therefore exempt from an assessment for fairness.¹²³ Nevertheless, and somewhat opaquely, the court refrained from holding that bank charges could *never* be challenged:

I do not believe any challenge to the fairness of the Relevant Terms has been made on the basis that they cause the overall package of remuneration paid by those in debit to be excessive . . . It may be that, if and when the OFT challenges the fairness of the Relevant Terms, issues will be raised that ought to be referred to Luxembourg.¹²⁴

This confusion begs the question, especially stark in the wake of *Tragbetti*, of why reference was not made to the CJEU.¹²⁵ Here, the Supreme Court's assertion that, where a court of last resort was unanimous that appeal be allowed, a point should be treated as *acte clair*,¹²⁶ drastically reformulates *CILFIT*. More fundamentally, however, the Supreme Court made an important concession for, as Devenney observes:

the conclusion that the Court of Appeal wrongly applied its interpretation of Article 4(2) is dependent on a *particular view* of the core/ancillary terms dichotomy which, surely, is partly a question of law. There is also a wider issue: this argument . . . concedes that the Supreme Court may have been wrong on the question of interpretation!¹²⁷

The judgment thus concedes its own fragility. Moreover, the Law Commissions of England and Wales and of Scotland, alarmed by the approach taken, subsequently dealt with exempt terms in their 2012 Issues Paper, highlighting the conflicting elements in *Abbey National* which allow differing interpretations of the UTCCR, a complexity subsequently augmented with *Ashbourne Management Services*¹²⁸ to produce a kaleidoscope of unfairness approaches: seen in *First National* and *Abbey National* in bank charges; in *Singh* and *Kufner* in suretyships' law,¹²⁹ *Newham v Khatun* in social housing¹³⁰ and *Bairstow v Eves* in agents' fees¹³¹ and making

119 *Director General of Fair Trading v First National Bank plc* [2001] UKHL 52, [2002] 1 AC 481 (HL), cl 8, para 2. Cf *DGFT v First National Bank plc* [2000] 1 WLR 98 (HC), *DGFT v First National Bank plc* [2000] QB 672 (CA); E Macdonald, 'Scope and Fairness of the Unfair Terms in Consumer Contracts Regulations: *DGFT v First National Bank*' (2002) 65 *Modern Law Review* 763.

120 *OFT v Abbey National plc* [2008] EWHC 875 subsequently: [2009] EWCA Civ 116. E MacDonald, 'Bank Charges and the Core Exemption: *OFT v Abbey National plc*' (2008) 71(6) *Modern Law Review* 987.

121 *Abbey National* (n 16) per Lord Walker, para 43.

122 H Collins, 'Good Faith in European Contract Law' (1994) *Oxford Journal of Legal Studies* 229.

123 *Abbey National* (n 16), per Lord Walker, para 41.

124 *Ibid* per Lady Hale, para 91.

125 Law Commissions, Issues Paper, *Unfair Terms in Consumer Contracts: A New Approach*, 25 July 2012 <http://lawcommission.justice.gov.uk/docs/unfair_terms_in_consumer_contracts_issues.pdf> paras 5.64–7.

126 *Abbey National* (n 16) per Lord Walker para 49.

127 Devenney (n 111) 51.

128 *OFT v Ashbourne Management Services* [2011] EWHC 1237 (Ch), [2011] ECC 31. Law Commissions (n 125) paras 5.74–83 and 5.93.

129 *Bank of Scotland v Singh*, 2005 (unreported, 17 June 2005, HC) suretyships outside scope of UTCCR. *Bardays Bank v Kufner* [2008] EWHC 2319; suretyships within scope of UTCCR, Devenney and Kenny (n 34) 297.

130 *London Borough of Newham v Khatun* [2004] EWCA Civ 55.

131 *Bairstow Eves London Central Ltd v Smith* [2004] EWHC 263, Gross J, para 25: 'Reg. 6(2) must be given a restrictive interpretation; otherwise a coach and horses could be driven through the Regulations'.

it difficult to demarcate exempted from assessable terms.¹³² Moreover, as the Law Commissions note, given that the ultimate arbiter of the UTCCR is the CJEU, any future interpretation may change the way the UTD should be applied,¹³³ thus the 'current law' lulls traders into a false sense of security.¹³⁴ Moreover, any future CJEU interpretation is liable to be informed by the pro-consumer tenor of CJEU case-law exemplified in *Océano*¹³⁵ and *Pannon*.¹³⁶ The Law Commissions recite the shortcomings of *Abbey National* in graphic terms: the scope of the exemption is 'unacceptably uncertain' and there is an urgent need to render UTD implementation 'certain enough'.¹³⁷ The Law Commissions conclude that a new approach, affording a *higher* level of protection by *narrowing* the scope of the exemption, is needed.¹³⁸ However, whether the Law Commissions' solution to these shortcomings, reliance on improving the *transparency* and *prominence* of the relevant terms,¹³⁹ is equal to this task is open to considerable doubt, given the limits of the information model in protecting consumers, especially in financial services.¹⁴⁰ Yet the position in which the Law Commissions were placed by *Abbey National* was invidious: left to perform a poorly choreographed veiled dance between contradictory coordinates; ensuring the *clear and precise* implementation of the directive,¹⁴¹ *restricting* the exemption pursuant to *Abbey National*, appearing to faithfully invoke the cause of consumer protection as expressed in *Océano* and *Pannon*, and, while appearing to invite, seeking to evade *Traghetti*-liability by relying on transparency and prominence to effectively defuse the importance of Regulation 6(2).

C. SYNTHESIS

Rather than the 'tolerance of error' attached to the collaborative approach in *Mangold*, the 'national application' of EU law in *Abbey National* led to a multiplication of error. In private law terms, the judgment left an uncertain demarcation of the scope of exemption, a demarcation 'difficult to reconcile' with *Océano*. While the Law Commissions have explicitly underscored the fragility of the judgment and the countervailing tenor of CJEU case-law, *Abbey National* appears implicitly reversed even within the modest framework provided by the new Consumer Rights Directive.¹⁴² Meanwhile, in public law terms, the spectre of state liability for failure to refer arises, strengthened by the Law Commissions' interventions. Spectacular in this interplay are the 'hospital passes' passed on by both the Supreme Court

132 Law Commissions (n 125) para 1.17.

133 Ibid para 7.73.

134 Ibid para 8.12.

135 Joined Cases C-240–244/98 *Océano Grupo Editorial SA v Rocio Murciano Quintero* and *Salvat Editores SA v José M Sanchez Alcón Prades and Others* [2000] ECR I-4941, para 26; Casenote: T Corthaut, (2002) 8 Columbia Journal of European Law 293; Case C-473/00 *Cofidis v Fredout* [2002] ECR I-10875 para 33.

136 Case C-243/08 *Pannon GSM Zrt v Erzsébet Sustikné Györfi* [2009] ECR I-4713, para 35: 'the national court is required to examine, of its own motion, the unfairness of a contractual term where it has available to it the legal and factual elements necessary . . . Where it considers such a term to be unfair, it must not apply it, except if the consumer opposes that non-application'.

137 Law Commissions (n 125) para 8.13.

138 Ibid para 7.75.

139 Ibid paras 8.2–3, 8.22–4, 8.25–8, 8.40–2, 8.70–3.

140 P Rott, 'Effective Enforcement of Consumer Law: The Comeback of Public Law and Criminal Law', 64–81, and C Willett and M Morgan-Taylor, 'Recognising the Limits of Transparency in EU Consumer Law', 143–63, in J Devenney and M Kenny (eds), *European Consumer Protection: Theory and Practice* (CUP 2012).

141 Case C-144/99 *Commission v Netherlands* [2001] ECR I-3541, para 17; Case C-478/99 *Commission v Sweden* [2002] ECR I-4147, para 18: '[I]t is essential that the legal situation resulting from national implementing measures be sufficiently precise and clear'.

142 Articles 19 and 22, Consumer Rights Directive (n 82).

and the Law Commissions. The cause of legislative intervention is aided by such 'national applications' of EU law.

6 Conclusions

The 'complete system' of judicial protection averred to in *UP4* contains important gaps which lend the effectiveness of EU law a chimerical quality. Similarly, EU legislative intervention has provided an unstable foundation for the internal market. Unsurprisingly, calls for enhanced effectiveness and greater coherence have followed. This paper has argued that further legislative intervention cannot be expected to deliver greater coherence and that measures of judicial realignment are also unequal and/or disproportionate to the task. Indeed this survey questions the attraction of 'greater coherence' and a 'complete system of legal remedies'. By contrast the chimerical quality of effectiveness exists for a number of important reasons: the polycontextual nature of legal doctrine across the EU; subsidiarity and the need for the respect of national identities. EU law in this light requires something more than the responsorial psalm of orthodox functionalism, requiring instead discretion, competition, margins of appreciation *and* incompleteness; and this, not least, to accommodate the instability of CJEU judicial policy and the outer margins of the effectiveness of state liability in the context of the 'bailout states'. In this regard, a critical review of what sort of internal market is being created and how far it is appropriate to surrender national regulatory autonomy is needed.¹⁴³

Instead, this paper has advocated enhancing collaboration to improve the judicial interplay and interface between national and EU law. Yet the juxtaposition of *Mangold* and *Abbey National* reveals higher national courts travelling in opposite directions, begging the questions of whether greater sophistication in national approaches to EU law can be promoted and the extent to which, in the wake of *Tragbetti*, state liability for failure to refer might play an important (if in part symbolic?) role in this. Fundamentally, these cases flag up the need for taking the administrative dimension of effectiveness more seriously. Such a strategy would focus on enhancing the European Commission's compliance role, encouraging robust rather than 'copy-out' national implementation and strengthening judicial methodology at national and EU level.

In terms of policy development, the way events have conspired to support the introduction of a CESL and, by implication, the wider cause of codification, most recently exemplified in the European Law Institute's lobbying on the need to significantly upgrade the CESL, is sobering; especially so as a handful of practical steps might otherwise obviate the need for intervention altogether. Moreover, given that extension of the EU regime into broad areas of private law would entail the uncertain application of liability rules between private parties *inter se* and further reduce the margins of discretion and the field of application of national law, such intervention would produce further demarcational instability at the nexus of EU and national constitutional and private law. The danger of such policy development is clear with, once again, pragmatism and method being subverted by the imperative of the European grand design. Equally troubling, however, is the domestic policy context: caught between the allure of exiting the EU and the countervailing perception of EU obligations as essentially elective items, the branches of UK governance appear unable to resist a legally delusional and damaging dalliance with non-Europe.

¹⁴³ Weatherill (n 4) 1279.

Corporate political connection as a determinant of corporate governance in Hong Kong

BOYCE YUNG

Government Counsel, Department of Justice, The Government of Hong Kong Special Administrative Region

PHILIP LAWTON

Law School, Lancaster University

Abstract

This article reports a study which uses a unique dataset compiled from listed companies in Hong Kong to demonstrate the relationship between corporate political connection with the corporate structure, ownership background and industry type of companies. The study shows that companies with political connection tend to be larger companies while Chinese family-controlled companies and more regulated companies have a higher level of political connection. Identifying the inadequacies of the existing theories in explaining corporate governance in Hong Kong, the article suggests adopting corporate political connection as a determinant of corporate governance in Hong Kong and elsewhere.

1 Introduction

Companies are increasingly aware of the importance of politics. In the modern economy, law and policy have become important factors for the operation and even success of a business. They may either promote or restrict the development of the business. As suggested by Fisch, being politically ‘naïve’ can be commercially costly. Companies therefore need to develop their political strategy and accumulate their political capital in the same way they manage other business assets.¹

An important way to develop a company’s political capital and accumulate political influence is through political connection. Corporate political connection (CPC) can take many forms. In this article, CPC refers to political connection of companies *through directors*. The board of directors has been regarded as proxy of corporate activities in research.² This is because the board functions as the controlling mind of a company and owes the duty of care to serve in the interest of the company. As the court commented in the *dictum* of Neville J in *Bath v Standard Land Company*: ‘Directors are the brains and the only brains of the company which is the body, and the company can and does act only through them.’³

1 J E Fisch, ‘How Do Corporations Play Politics? The Fedex Story’ (2005) 58 Vanderbilt Law Review 1495, at 1569–71.

2 A Agrawal and C R Knoeber, ‘Do Some Outside Directors Play a Political Role?’ (2001) 44 Journal of Law and Economics, 179, at 181.

3 *Bath v Standard Land Company* [1910] 2 Ch 408.

While the board of directors acts as the brain in making and supervising decisions necessary for corporate management,⁴ other stakeholders have less dominant roles.

CPC may be in the form of taking formal positions in political bodies or establishing informal relationships with government officials. For the purpose of this article, CPC refers to the former. It is a formal form of CPC through corporate directors through formal appointment or election into government policy-making or advisory bodies. Directors may directly influence policy or law-making through these formal positions which confer the right to information, speech and/or vote in the decision-making process. From a research perspective, formal CPC can be more objectively assessed and measured. The information about such connection is commonly disclosed and easily accessible.

2 CPC in Hong Kong

When corporate directors take up political appointments, there are always questions regarding their motives and the consequence of their political involvement on their companies. Many studies revealed that politically connected firms can benefit from favourable government policies and awards of government tenders, resulting in more profits and higher share prices.⁵ Some researchers take a wider perspective of corporate governance and point out that there is a relationship between political systems and corporate ownership structures. For instance, Roe pointed out that in social democratic countries, business owners strive to hold more shares to offset the power they lose in complying with labour policies that favour workers. Thus concentrated structure has been prevalent in these countries.⁶

In democratic countries like the USA, CPC is also concerned with political contributions, campaign finance and lobbyists.⁷ These activities, however, are not as relevant to non-democratic Hong Kong where election campaigns and political contributions are not main entrances for CPC. CPC through political donation is far less popular in Hong Kong and does not cause as much concern. Possibly, companies find it not cost effective to influence public policy through electoral donation in a relatively undemocratic government structure.⁸ In addition, corporate political donation may irritate minority shareholders or politicians when the donation is revealed in the annual audit accounts. In Hong Kong, law does not specifically regulate corporate political donation. Election law only regulates general corrupt and illegal conduct and prohibits improper use of election donations.⁹

4 L E Mitchell and T A Gabaldon, 'If I Only Had a Heart: Or, How Can We Identify a Corporate Morality' (2002) 76 Tulane Law Review, 1645, at 1657.

5 See, for example, A O Krueger, 'The Political Economy of the Rent-Seeking Society' (1974) 64 American Economic Review, 291, at 303; A O Krueger, 'Virtuous and Vicious Circles in Economic Development' (1993) 83 American Economic Review, 351, at 355; M J Olson, *The Rise and Decline of Nations* (Yale University Press 1982); M J Olson, *Power and Prosperity: Outgrowing Communist and Capitalist Dictatorships* (Basic Books 2000); A Shleifer and R W Vishny, 'Corruption' (1993) 108 Quarterly Journal of Economics, 599, at 599–618; and A Shleifer and R W Vishny, 'Politicians and Firms' (1994) 109 Quarterly Journal of Economics, 995, at 995–1025.

6 M J Roe, *Political Determinants of Corporate Governance: Political Context, Corporate Impact* (Oxford University Press 2003).

7 T K Kuhner, 'The Separation of Business and State' (2007) 95 California Law Review, 2353, at 2354.

8 Here, 'undemocratic' refers to the lack of direct elections in government decision-making bodies. The Chief Executive is elected by an Election Committee with small membership. The top decision-making body, the Executive Council, is composed of the Chief Executive, top government officials and unofficial members appointed by the Chief Executive. Only half of the Legislative Council members are elected by popular votes. Most of the District Board members are elected but the board plays only an advisory role by giving opinions to government at district level.

9 Ss 14 and 15, Elections (Corrupt and Illegal Conduct) Ordinance, Cap 554, the Laws of Hong Kong.

CPC in Hong Kong takes another form: it is very common for corporations to participate in politics through controllers and directors directly joining public services or indirectly influencing policy-making with their socio-political connections. Needless to say, participation of corporate owners or managers in politics exists elsewhere around the globe. Many political leaders in Western democratic countries were elite business people before being elected to political positions. George W Bush and his father, George H W Bush, both oil company owners and then elected Presidents of the USA, are obvious examples. However, unlike in other parts of the world, in Hong Kong, political leaders – except chief executives and their appointed principal officials – enjoy unique flexibility in their political participation in that they are able to hold their political positions *as well as* their own full-time work. For instance, unofficial members in the Executive Council (its relationship with the Chief Executive is equivalent to that of a Cabinet with the Prime Minister in the UK) all serve part time. Over half of these unofficial members are renowned business owners, directors and managers. Besides them, Legislative Councillors, many of whom are from the business sector, can also retain their own full-time work.¹⁰

This article examines this particular mode of CPC in Hong Kong, representative of a typically non-democratic Chinese society, which is at the same time a major world trade centre with an interestingly complex political background.

3 Characteristics of CPC companies in Hong Kong

Three hypotheses are set to examine the characteristics of CPC companies in Hong Kong in relation to *structure*, *ownership* and *industry types*.

HYPOTHESIS 1 (H1): CPC AND NON-CPC COMPANIES ARE STRUCTURALLY DIFFERENT

It is expected that large-scale companies (i.e. companies with higher capitalisation or total asset value) have more resources for CPC and they are also likely to extract more rents through CPC due to bigger market share than their smaller counterparts. Previous studies have contradictory results over the topic: Faccio's worldwide CPC research finds supportive evidence that CPC is more widespread among larger corporations.¹¹ Faccio investigates CPC in 47 countries and finds that stock prices increase significantly when a board member of the firm enters politics. Additionally, firm value increases more when that board member is elected prime minister, rather than as a member of the parliament.¹² Faccio also finds that CPC has more significant impact on company value in countries with more corruption, lower quality of legal environment and less freedom of the press.¹³

Another Thai study shows that if the firm controllers take office in the government, their firms would have a higher market valuation than other firms in the country. The political power of the owners who are political leaders accounts for the extraordinary incremental gain in market valuation and share price of the firms.¹⁴

There are also studies showing that CPC affects firm value in less corrupt countries. For example, Goldman et al find CPC related to the stock price of the company. Investigating the S&P 500 companies in the USA, their study finds a positive abnormal stock return following the announcement of the nomination of a politically connected individual to the board. When the Republican won the 2000 presidential election, companies connected to

10 For instance, only 16 out of 60 legislative councillors were full-time legislators in 2008: *Annual Report, 2008–2009*, The Legislative Council, 94–124, <www.legco.gov.hk/general/english/sec/reports/a_0809.pdf>.

11 M Faccio, 'Politically Connected Firms' (2006) 96(1) *American Economic Review* 369, at 369–71.

12 Ibid 384–5.

13 Ibid 380.

14 P Bunkanwanicha and Y Wiwattanakantang, 'Big Business Owners in Politics' (2007) SSRN eLibrary, 26–7.

the Republican Party increased in value, and companies connected to the Democratic Party decreased in value.¹⁵

Nee and Opper's research of CPC in China, however, shows that while political connections provide the strongest competitive advantage on the market for government contracts, company size does not create a general advantage. This implies that firm size or value and CPC are not necessarily related in China.¹⁶ Whether the scale of a company matters in Hong Kong CPC is as yet unanswered by the literature.

Also related to corporate structure is its concentration pattern. Plenty of previous studies have discussed the impact of politics on ownership pattern. For example, Roe suggests that social democracies impede diffused ownership, which explains why many developed countries practising social democracy still exhibit concentrated shareholding.¹⁷ Coffee suggests that America reached dispersed structure mainly because the government allowed 'enlightened self-regulation' of US markets in the late eighteenth century. The 'paternalistic supervision' of markets by governments of other countries, on the other hand, has hindered the development of a dispersed structure.¹⁸ Gourevitch and Shinn also see the strong impact of interest-group politics on corporate governance. They predict that different coalitions among interest groups, namely owners, managers and employees, lead to different governance outcomes: diffusion or concentration.¹⁹ With reduced political pressure for minority and labour protection through political activities, it is beneficial for politically connected companies to release some of their shares while still maintaining sufficient controlling power. As this is theoretically probable but not empirically proven, the present study tries to test whether CPC companies do have lower concentration than non-CPC companies in Hong Kong.

Therefore, H1 tests whether CPC companies are characterised by a significant structural difference from non-CPC companies in terms of *scale* (as measured by a company's *total asset value* and *capitalisation*, the calculation methods of which will be discussed in later sections) and *concentration*.

HYPOTHESIS 2 (H2): CHINESE FAMILY COMPANIES HAVE MORE CPC THAN OTHER COMPANIES

Wong anticipates that Chinese family companies will split up over time due to the disintegrative effect of succession.²⁰ But the disintegration process of these companies seems to take a longer time than that of other companies as found by Lawton's analysis of winding-up petitions of companies in Hong Kong.²¹ In another study, Lawton also notes that the Chinese perception of business as personal and familial property rights will last for

15 E Goldman, et al, 'Do Politically Connected Boards Affect Firm Value?' (2009) 22(6) *Review of Financial Studies*, 2331, at 2331–4.

16 V Nee and S Opper, 'Political Connections in a Market Economy' (2007) 50 *CSES Working Paper Series* No 1, 30.

17 M J Roe, 'Political Preconditions to Separating Ownership from Corporate Control' (2000) 53 *Stanford Law Review*, 539, at 601–2.

18 J C J Coffee, 'The Rise of Dispersed Ownership: The Roles of Law and the State in the Separation of Ownership and Control' (2001) 111(1) *Yale Law Journal* 5–11.

19 See: P A Gourevitch and J J Shinn, *Political Power and Corporate Control: The New Global Politics of Corporate Governance* (Princeton University Press 2005), 23, 60, 281 and 295.

20 S L Wong, 'The Chinese Family Firm: A Model' (1985) 36(1) *British Journal of Sociology* 58, at 69.

21 P Lawton, 'Modelling the Chinese Family Firm and Minority Shareholder Protection: The Hong Kong Experience 1980–1995' (2007) 49 (5–6) *Managerial Law: International Journal of Law and Management*, 249, at 266.

generations if not perpetuity.²² Lawton also notes that many large listed Hong Kong firms retain the characteristics of small family businesses such as personalism and paternalism. Chinese Confucian thought endorses family as an economic unit. The Chinese perceive that property rights, including corporate ownership, can be maintained for generations. The Chinese cultural traits described by Lawton have explained why Hong Kong, as a Chinese society, has been dominated by family firms. The stress on personal and family relationships is translated into the unwillingness of family firms to relinquish control even after floatation.²³

Culturally, Chinese family controllers would be highly resistant in relinquishing their dominance over their companies because Chinese merchants had historically learnt the importance of politics on their businesses. For a very long period in the history of China, merchants lacked political power and were oppressed in society. In imperial China, the emperor had almost unrestrained power to control the economy. According to Rozman, there was virtually no limit for the government to intervene in the economic activities of merchants.²⁴ The government's intervention was given further legitimacy by Confucius' thinking which advocates that the government should be entrusted with power to redistribute wealth in society.²⁵ It endorses the emperors' dominance over both politics and economy so as to fulfil their obligations to ensure peace, prosperity and justice among people.²⁶ Knowing that politics had direct influence on their businesses, Chinese merchants sought to build up stable and close relationships with officials in an attempt to increase their political power and thereby benefit their businesses. They tried to merge their interests with those of state officials, for example, by inviting officials to become business partners.²⁷ In so doing, they managed to influence public policy and decision-making in favour of their businesses through unofficial channels.²⁸

CPC is a way by which controllers acquire power to maintain dominance. It is therefore hypothesised that Chinese family companies have a generally higher degree of CPC than other companies.

HYPOTHESIS 3 (H3): MORE REGULATED INDUSTRIES HAVE MORE CPC THAN LESS REGULATED ONES

Faccio finds that CPC is less common in the presence of more stringent regulations that set limits on the business activities of public officials to avoid political conflicts of interest,²⁹ which is quite expected. These regulations are imposed on officials irrespective of the type of industry they are connected with. But there are industry-specific regulations set by government on the functioning of individual industries such as stipulated conditions for entry into the industry, amount of capital investment, charges, maximum profit allowed etc. As found by Boubakri et al, companies operating in regulated industries are more likely to be politically connected.³⁰ It is anticipated that the more an industry is affected by

22 P Lawton, 'Berle and Means, Corporate Governance: Chinese Family Firm' (1996) 6 Australian Journal of Corporate Law 348, at 372.

23 Ibid 353–4, 378.

24 G Rozman, *The Modernization of China* (Free Press 1981), 107–8, 139.

25 X Yao, *An Introduction to Confucianism* (Cambridge University Press 2000), 138.

26 H G Frederickson, 'Confucius and the Moral Basis of Bureaucracy' (2002) 33(6) Administration & Society 610, at 613.

27 T K Wu, 'An Interpretation of Chinese Economic History' (1952) 1 Past and Present, 1, at 9.

28 The imperial period includes both 'Feng Chien' period (from Ch'in dynasty in 221 BC to the Opium War in 1840) and 'Semi Feng Chien' period (from 1840 to the end of Qing in 1911) as delineated by ibid. 7.

29 M Faccio, 'Politically Connected Firms' (2006) 96(1) American Economic Review 369, at 369.

30 N Boubakri et al, 'Political Connections of Newly Privatized Firms' (2008) 14(5) Journal of Corporate Finance 654, at 672.

Hypothesis	Pairs of company groups being compared		Characteristics being compared
1	CPC companies	Non-CPC companies	Corporate structure <ul style="list-style-type: none">• total asset value• capitalization• ownership concentration
2	Chinese family Companies	Non-Chinese family companies	Degree/level of CPC
3	More regulated companies	Less regulated companies	

Table 1: Characteristics of company groups under comparison

government policies and monitoring regulations, the more CPC the industry will seek to establish in order to gain the benefits of being close to, or even part of, the regulation-setting authority. It is also anticipated that through hypothesis testing Hong Kong industries that are more attracted to CPC will be revealed.

SUMMARY OF THE HYPOTHESES

The three hypotheses aim to compare the characteristics of three pairs of contrasting company groups in order to find out whether there are significant differences within each pair of groups in relation to CPC, as summarised in Table 1.

Hypothesis 1 examines whether there are any significant differences in respect of variables related to corporate structure (i.e. total assets, capitalisation and ownership concentration) between companies with and without CPC.

Hypotheses 2 and 3 examine whether there are any significant different levels of CPC between two other pairs of company groups with contrary corporate governance features (namely, Chinese-family verses non-Chinese-family companies, and companies under more-versus-less government regulation).

4 Methodology

SAMPLE SELECTION

Several criteria have been considered in selecting the research sample on which tests are to be conducted:

- (a) The sample should be composed of Hong Kong companies to reveal the characteristics of CPC in the local context.
- (b) These companies together should play a major role in the Hong Kong economy so that their special features related to CPC may significantly impact on the local society and its economy.
- (c) There should be a variety of companies in the sample as regards structure, ownership mode, monitoring mechanism and industry types so that rich research data can be obtained to test the hypotheses formulated.
- (d) The sample should provide reliable, transparent and retrievable data for systematic statistical analysis.

Public companies included in the Hang Seng Hong Kong Composite Index (HSKCI) were found to be the best sample that can satisfy all of the above criteria. They are subject to disclosure requirements under listing regulations and rules and thus can provide readily

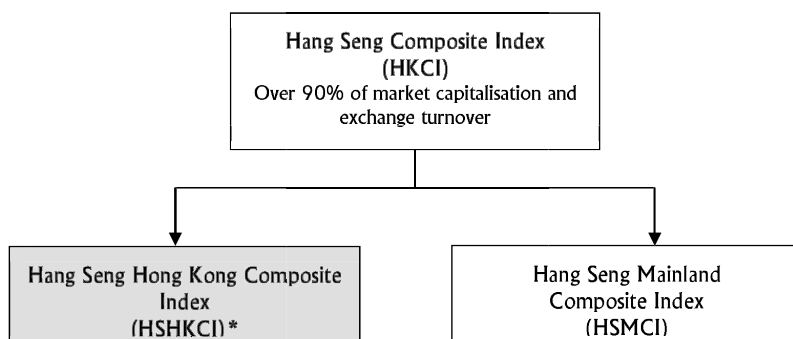


Figure 1: Components of HSCI as at 31 December 2005

*Sample of the study: HSHKCI

accessible and reliable data not easily obtainable from non-public companies. They are much more varied in different aspects than private companies, which are typically concentrated and controlled by the sole owner or a few business partners. More importantly, the HSHKCI companies are highly representative of the Hong Kong economy, which is characterised by a particularly active stock market where average citizens invest much of their life savings in stock activities.

HSHKCI is the Hong Kong component of the Hang Seng Composite Index (HSCI), which consists of the top 200 listed companies in terms of market capitalisation. The companies comprise over 90 per cent of the market capitalisation and exchange turnover of the stocks listed on the Main Board of the Stock Exchange of Hong Kong.³¹ HSCI is further divided into two sub-indexes: HSHKCI, the chosen sample, which includes HSCI constituent companies listed on the stock exchange of Hong Kong; and the Hang Seng Mainland Composite Index (HSMCI), which includes HSCI constituent companies listed in Hong Kong but with major earnings gained from Mainland China.³² HSMCI companies are therefore excluded from the research sample (see Figure 1).

It is important that the reference year of the sample should be set within the past decade to capture the most recent and relevant features of CPC in Hong Kong. It should also be a comparatively stable year financially so that the findings will not be skewed by temporary economic turmoil such as the internet bubble in 2000, the Asian financial crisis in 2003, the stock and property boom in 2007, the financial tsunami in 2008 and the credit and banking crisis afterwards. Amidst the other financially more stable years, 2005 – the middle year of the recent decade – is chosen as the reference year of which corporate data are to be drawn from HSHKCI for analysis.

As at 30 December 2005, HSHKCI represents 55.2 per cent of the market capitalisation and 48.9 per cent of the exchange turnover of the stock listed on the Main Board.³³ There were altogether 102 companies listed in HSHKCI in 2005.³⁴ Among them, Samson Holdings was listed in November 2005 and so did not have complete financial data for the whole year of 2005. In order not to distort the overall statistical analysis, it is excluded from the sample, which ultimately contains 101 companies of HSHKCI in 2005.

31 Press Release, 'HSI Services Announces Index Review Results', HSI Services Limited, 10 February 2006, Appendix 4, at: <www.hsi.com.hk/HSI-Net/static/revamp/contents/en/news/pressRelease/20060210e.pdf>.

32 Ibid 3.

33 Ibid. Appendix 5.

34 Ibid.

DATA COLLECTION

Full annual reports of the 101 HSHKCI companies published for the financial year of 2005 are used for data collection considering their reliability and accuracy as information sources. Under the Main Board Listing Agreement, a listed company is required to publish an interim report and a full annual report to its shareholders every year within the prescribed time.³⁵ The full annual report containing complete financial statements and reports of a company is published at the end of each financial year of the company, either at the end of June or December of each calendar year. Accordingly, the full annual reports of the 101 companies for analysis were published in December 2005 or June 2006. The financial statements contained in these reports were audited and endorsed by qualified accountants following professional accounting standards and stringent requirements prescribed by the Listing Rules. Since the professional accounting standards of Hong Kong are benchmarked against the International Accounting Standards, the figures in the annual reports are comparable domestically and internationally.

All listed companies have now made their annual reports available on their company websites.³⁶ Where necessary data are not provided by the annual reports, they are obtained from the 'Company/Securities Profile' webpage of the Hong Kong Exchanges and Clearing Ltd (HKEx).³⁷ Relevant data given in both the annual reports and the HKEx website are cross-checked between the two sources to ensure the accuracy of the data used for the study.

The following data are thus collected:

- (a) Company profile: the full annual report contains biographies or profiles of members of the board of directors (including executive directors, non-executive directors and independent non-executive directors) of a listed company. Information of each of the board members includes:

- i) ethnic background;
- ii) familial relationship, if any, with other board members;
- iii) percentage of shares held;
- iv) position(s) held in the company;
- v) public and community services;
- vi) industry type of each company.

The above information provides useful data for the analysis of Chinese family companies, concentration pattern and industry category.

- (b) Financial data: the following financial data of each company as at 30 December 2005³⁸ are collected from either its annual report or the HKEx website:

35 Rule 13.46(1) of the Main Board Listing Rules provides that an issuer shall send to its members and securities holders an annual report within a prescribed time. Note 1 to the rule further provides that the annual report must be in the English language and must be accompanied by a Chinese translation. Rule 13.48(1) provides that an interim report be issued for the first six months of each financial year. See *Rules Governing the Listing of Securities on the Stock Exchange of Hong Kong Limited*, ch 13, <www.hkex.com.hk/eng/rulesreg/listrules/mbrules/documents/chapter_13.pdf>.

36 A listed company must make its annual and interim report available on its website in order to obtain from HKEx a waiver from sending the reports in both English and Chinese to its members: 'The Stock Exchange's Announcement on 17.1.2001', Hong Kong Exchanges and Clearing Ltd, 2001, <www.hkgem.com/aboutgem/e_release010117.pdf>.

37 *Investment Service Centre*, Hong Kong Exchanges and Clearing Ltd, www.hkex.com.hk/invest/index.asp?id=company/profilemenu_page_e.asp.

38 The date 31 December 2005 was not counted because it was a Saturday, when the stock market was closed.

- i) share price;
- ii) number of shares issued;
- iii) total assets.

The financial data are used in the calculation of capitalisation and total asset value for the analysis of the scale: capitalisation and total asset value.

DEFINITIONS AND MEASURING METHODS

The research involves a list of variables that need to be delineated unambiguously with clear definitions and measuring methods in order for tests to be performed on the data. These are as follows.

(a) Chinese family company

Cox defines a *family* to mean 'people related by blood or marriage or lifetime commitment', and a *family firm* to mean 'any enterprise in which more than one family member has a significant investment (financial or emotional) or significant participation in the operation or management decisions of the enterprise'.³⁹ However, 'emotional' investment is not always discernible, and whether more than one member of a family has a significant 'financial' investment is not made known in firms held by family trusts. As a matter of fact, many family businesses are held by the company founder, whose family members hold only insignificant amounts of shares. To improve on Cox's definition, the study has adopted only the latter part of it, i.e. only the part on significant participation in corporate decisions is considered. Adding the notion of Chinese ethnicity and making the definition more concise, the term *Chinese family company* is defined in this study as *any enterprise which has a controlling shareholder who is an ethnic Chinese and which has at least two members from the family of the controller sitting on the board*.

(b) More regulated industries

Before distinguishing more regulated industries from less regulated ones, it is necessary to first identify what kinds of industries are involved in the companies under investigation. To this end, the sample of 101 listed companies is categorised into seven industry types by reference to the definitions given by the Heng Seng Industry Classification System (HSICS).⁴⁰ They are: public utilities and transportation; communications; financials; properties and construction; industrial and consumer products; hotels and entertainment; conglomerates and others. While all of them are subject to different kinds of government monitoring regulations, the first four industry types are under particularly close government watch and are highly sensitive to government policies and regulations for various reasons (see Table 2).

It can be seen from the above classification that government regulation is vital to the first four industries in that they either involve scarce resources such as energy and land (in industries (a), (b), (d)) or are costly to people if the systems fail (in industries (c) and (d)).

39 E S Cox, *The Family Firm as a Foundation of Our Free Society: Strengths and Opportunities* (Cornell University Family Business Research Institute 1996) <www.fambiz.com/Orgs/Cornell/articles/real/ed_cox.cfm>.

40 The HSICS is a comprehensive industry classification system designed for the Hong Kong stock market. Prompted by the listing of a wide variety of companies in different industries in Hong Kong, it meets the need for a detailed industry classification that reflects stock performance in different sectors. HSICS caters for the unique characteristics of the Hong Kong stock market while maintaining international compatibility with mapping to international industry classification systems. See 'Overview: Hang Seng Indexes' <www.hsi.com.hk/HSI-Net/HSI-Net>.

Table 2: Classification of industries and the respective regulations and policies

	Classification	Regulations and policies
1	Public utilities and transportation	Public utilities include distributors of electricity, gas and water and related public utilities. Public transportation involves providers and operators of all kinds of transportation services such as rail and roads facilities and services. Both kinds of companies provide basic urban living necessities that greatly influence the everyday life of the general public. Since these companies enjoy monopoly status in Hong Kong, they are closely monitored by the government on various business operations and fare level by franchises and on maximum return or profit by agreements. For example, Kowloon Motor Bus (1933) Company Ltd (KMB), a subsidiary of Transport International Holdings, operates the franchised public bus services provided in Kowloon and New Territories districts according to the terms and conditions laid down in the franchise granted by the Government of Hong Kong under the Public Service Bus Ordinance, Cap 230. Pursuant to cl 25 of the franchise granted to KMB, the Chief Executive in Council may review the scale of fares to be charged by the company and may determine and adjust the fares at any time as s/he deems fit. ¹ The China Light and Power (CLP) Holdings Ltd, the electricity monopoly in Kowloon and New Territories districts in Hong Kong, is regulated by the Hong Kong government under a Scheme of Control Agreement under which allowed shareholders are permitted annual return on average net fixed assets of 9.99 per cent for those investments financed by borrowings and for those financed by shareholders' funds. ²
2	Communications	This refers to operators of telecommunication networks such as broadband or mobile service providers and media engaged in broadcasting (e.g. of television and radio programmes) or publishing (e.g. of newspapers and magazines) activities. As radio frequency spectrum is a scarce resource, and the communications industry can exert enormous influence on the society, the relevant sector is subject to strict licensing requirements for entering and staying in the business. Publishing companies are relatively less controlled by the government. But the 'Communications' sector as a whole is one of the industries under more stringent government scrutiny. The telecommunications operators are regulated and licensed by the Telecommunications Authority under the Telecommunications Ordinance, Cap 106, and the broadcasting operators are regulated and licensed by the Broadcasting Authority under the Broadcasting Authority Ordinance, Cap 391, the Broadcasting Ordinance, Cap 562, and Part IIIA of the Telecommunications Ordinance, Cap 106. Under both licensing regimes, the operators are subject to various statutory requirements and licensing conditions which cover a wide range of operating matters including programme content and shareholding change of the controllers.
3	Financials	Financials refer to banks, insurance companies and other financial service providers. Like the communications industry, financial institutions are subject to a stringent regulatory framework so that the government can ensure that they hold sufficient capital and that savings and investments of society can be protected. Hong Kong maintains a three-tier system of deposit-taking institutions, namely, licensed banks, restricted licence banks and deposit-taking companies. They are collectively known as authorised institutions. Hong Kong has one of the highest concentrations of banking institutions in the world. 68 of the largest 100 banks in the world have an operation in Hong Kong. As at February 2010, there were 145 licensed banks, 26 restricted licence banks and 28 deposit-taking companies in business. ³ The Banking Ordinance provides the legal framework for banking supervision in Hong Kong. As provided in s 7(1) of the ordinance, the principal function of the Hong Kong Monetary Authority (HKMA) is to 'promote the general stability and effective working of the banking system'. The HKMA monitors closely the continuing development of banking practices, market environment as well as international regulatory standards and considers in consultation with the banking

		industry whether any change to the Banking Ordinance is necessary. The Banking (Amendment) Ordinance 2005 (BAO 2005) establishes a legislative framework for implementing the revised capital adequacy framework released by the Basel Committee on Banking Supervision (Basel Committee) in Hong Kong. Authorised institutions have to comply with the provisions of the Banking Ordinance which, among other things, require them to maintain adequate liquidity and capital adequacy, to submit periodic returns to the HKMA on the required financial information, to adhere to limitations on loans to any one customer or to directors and employees, and to seek approval for the appointment of directors and chief executives, and for controllers. ^{iv}
4	Properties and construction	This covers a variety of companies related to the property and building sectors: companies owning and developing properties, producers and wholesalers of building materials, constructors of commercial and residential buildings, and providers of services to construction companies. According to Article 7 of the Basic Law, land is a state property under the management of the government in Hong Kong. Properties, either for accommodation, investment, or both, play a central part in Hongkonger's lives. Much of people's life savings and monthly incomes are spent on paying the instalments of their properties. The property sector is greatly affected by government policies in the areas of land supply, town planning, property development etc. New land parcels are sold, usually by auction, for specific developments through long-term land leases. The government retains the title of the land, and the lessees have the property use, income and transfer rights. ^v Land sales have historically constituted a significant portion of government revenues – in some years, proceeds from land sales could be as much as 38.2 per cent of total revenues. ^{vi} On average, between 1991 and 2001, more than 30 per cent of the HKSAR government's revenue was related to real estate. ^{vii} The property development is therefore closely related to government policy. In addition, the property development is subject to various laws and regulations such as the Buildings Ordinance and Fire Safety Ordinances administered by the Buildings Department and Fire Services Department, which ensure the safety of the properties in Hong Kong.
5	Industrial and consumer products	This industry genre comprises the manufacturers and distributors of a wide range of products: machinery and equipment, electronic parts and products, vehicles, household goods, clothing and accessories, foods and drinks, health and personal care services and products, and farming and fishing goods.
6	Hotels and entertainment	This leisure-specific genre is related to hotel operators and management companies as well as providers of entertainment services, leisure facilities, photographic services and equipment, restaurants and bars.
7	Conglomerates and others	Conglomerates are diversified companies engaged in three or more businesses classified in different sectors, and others are sporadic companies engaged in industries not classified elsewhere.

- i Franchise, Transport Department, 2006 <www.td.gov.hk/filemanager/en/content_389/kmb%20franchise_1.8.2007%20-%2030.6.2017.pdf>.
- ii *Annual Report: Chairman's Statement* (China Light and Power Holdings Ltd, 2008) <www.clpgroup.com/Abt/Res/Pub/ARnSER/Documents/2008/eng/E102.pdf>.
- iii *Monthly Statistical Bulletin: Number of Authorized Institutions and Local Representative Offices*, Hong Kong Monetary Authority <www.info.gov.hk/hkma/eng/statistics/msb/attach/T0301.xls>.
- iv *Banking Policy and Supervision*, Hong Kong Monetary Authority <www.info.gov.hk/hkma/eng/bank/three_tier/three_tier_f.htm>.
- v H X Bao and S Z Zhou, 'The Land Market in Hong Kong: Price Index and the Relationship with the Property Market' (SSRN eLibrary 2008)
- vi This happened in 1980–1981. The average for 1947–1948 to 1984–1985 was 20 per cent. In the ten years from 1991–2000, land sale proceeds averaged 12 per cent of government revenues. C Chang et al, *Property Market Overvaluation, Toebolds, and the Winners' Curse: Evidence from Hong Kong Land Auctions* (SSRN 2007), 7.
- vii A Smart and J Lee, 'Financialization and the Role of Real Estate in Hong Kong's Regime of Accumulation' (2003) 79(2) *Economic Geography* 153, at 154.

They are overseen closely by regulators and are therefore defined as 'more regulated industries' in this study. As for the other industries, the government relies mainly on market regulation apart from basic laws and guidelines. They are defined as less regulated industries in comparison with the above more regulated group.

(c) Corporate scale

As explained earlier, the scale of a company is measured by its *total asset value* and *capitalisation*. Total asset value – the value of a company's possessions including all its fixed capital (land, building, equipment, raw materials etc.) and liquid capital (money) – is a good yardstick for measuring the scale of a company.⁴¹ Besides calculating how much capital has been put into the company, another measuring tool for corporate scale is its capitalisation⁴² (share price x no of shares issued as at 30 December 2005) – i.e. how valuable the company is in the market in the year under study.

(d) CPC

In Faccio's study, CPC is identified if 'at least one of [the company's] large shareholders (anyone controlling at least 10 per cent of voting shares) or one of its top directors (CEO, president, vice-president, or secretary) is a member of parliament, a minister, or is closely related to a top politician or party'.⁴³ In other words, a company is considered to be politically connected if *one of its large shareholders or top directors either holds a political position or is closely related to a top politician or party*. This definition of CPC can be further improved.

First, a 'large shareholder' holding 10 per cent or more of voting shares does not necessarily participate in corporate decisions if s/he is not a board member. Second, directors are collectively responsible for the company. Board decisions are generally arrived at by voting of the whole board rather than by 'top directors' only. Last but not least, being 'closely related' to a top politician or party is a vague concept. Faccio herself admits that 'the necessity of relying on publicly available sources for information on close relationships such as friendship or well-known cases of relationships with political parties produces an incomplete picture'.⁴⁴ While this might be an interesting aspect of CPC for qualitative observation, it is certainly not objectively reliable for her quantitative study or for the statistical analysis of the present study.

For a more reasonable and reliable definition, this study identifies a company as having CPC (i.e. formal CPC) if *one or more of its directors occupy formal positions in public service institutions*. The term 'director' is broadly defined in statute as 'any person occupying the position of director by whatever name called'.⁴⁵ In this study, all directors including executive, non-executive, independent and alternate directors are included for the purpose of calculation of CPC. Shadow directors are not included as they are not disclosed in

41 R S Demsetz and P E Strahan, 'Historical Patterns and Recent Changes in the Relationship between Bank Holding Company Size and Risk' (1995) 1 Economic Policy Review 2, at 15.

42 The Securities and Exchange Commission of the United States uses the public float of the companies to measure the size of companies to determine whether a company is a smaller reporting company (i.e. a company of less than \$75m of capitalisation under Regulation S-K as amended in 2008) which qualifies for less stringent reporting requirements. See also M L Ettredge et al, 'The Effects of Firm Size, Corporate Governance Quality, and Bad News on Disclosure Compliance' (forthcoming) Review of Accounting Studies 30.

43 M Faccio, 'Politically Connected Firms' (2006) 96(1) The American Economic Review 369, at 370–2.

44 M Faccio, *Politically Connected Firms* (SSRN 2004), 4.

45 S 2, Companies Ordinance, Cap 32, the Laws of Hong Kong.

company annual reports and it is impracticable to obtain sufficient information to confirm whether someone is a shadow director.⁴⁶

Directors' political connection is considered representative of corporate political connection because they are the people appointed to exercise powers over all business and affairs of the company.⁴⁷ They are also representative of the controlling shareholders' interests because the latter can select their preferred candidates and share the power of the directors through the company's appointment mechanism. The law imposes no restrictions on what a company's memorandum or articles may stipulate in the appointment of directors.⁴⁸ The articles of association normally provide that the first directors are named by the founder of a company.⁴⁹ Subsequent directors are appointed by ordinary meetings.⁵⁰ Controllers, who have controlling shareholding and thus greater decision power at ordinary meetings, are able to appoint their nominated candidates as directors to serve their interests.

As for 'public service institution', McGregor defines the term as a body in society which 'concerns itself with the achievement of public objective and the implementation of public policy'.⁵¹ A company in which directors serve in the following public service institutions is counted as having CPC.

Hong Kong Public Service Institutions:

- the Executive Council (the highest policy decision-making body of Hong Kong);⁵²
- the Legislative Council (the law-making body of Hong Kong);⁵³
- the Election Committee (a body established by the Hong Kong Basic Law for the election of the Chief Executive);⁵⁴
- statutory bodies (bodies established by statutes, e.g. Hospital Authority, Broadcasting Authority, Equal Opportunities Commission etc);
- public bodies (bodies established by the government for performing specific public functions, e.g. Education Commission, Trade Development Council etc);
- district councils (district level consultative bodies);
- government consultative committees and advisory boards (consultative bodies established by the government for consultation on specific policy areas);
- interest groups (e.g. political parties, trade unions, professional bodies);

Chinese Public Service Institutions:

- National People's Congress (the highest organ of state power);
- Local People's Congresses (regional organs of state power);

46 The term 'shadow director', in relation to a company, means a person in accordance with whose directions or instructions the directors or a majority of the directors of the company are accustomed to act': *ibid.*

47 Article 82, Schedule 1, Companies Ordinance, Cap 32, the Laws of Hong Kong.

48 R R Pennington, *Company Law* (7th edn, Butterworths 1995), 713. The Companies Ordinance, Cap 32 provides that any articles may adopt all or any of the regulations contained in Table A, which is a standard article set out in Schedule 1, Companies Ordinance, Cap 32, the Laws of Hong Kong.

49 Article 80, Schedule 1, Companies Ordinance, Cap 32, the Laws of Hong Kong.

50 Article 96, *ibid.*

51 E B McGregor, Jr, 'The Institution of Public Service' (1982) 42(4) *Public Administration Review* 304, at 305.

52 Article 43, the Basic Law of Hong Kong.

53 Article 66, the Basic Law of Hong Kong.

54 Article 45 and Annex I (Method for the Selection of the Chief Executive of the Hong Kong Special Administrative Region), the Basic Law of Hong Kong.

CPC in Hong Kong	
Level of connection	Score
Executive Council Member, Legislative Council Member	4
Election Committee Member, Statutory Body Member, Public Body Member	3
District Council Member, Government Consultative Committee Member	2
Interest group chair/executive member	1
CPC in China	
Level of connection	Score
National	4
Provincial (including autonomous government)	3
Municipal/prefecture	2
District/county township	1
(a) Score of each director = aggregate score of each director's connection with each level of public service institutions	
(b) Score of each company = aggregate score of all directors (each found in (a) above) in each company	

Table 3: Scorecard – weighing the level of CPC of companies in Hong Kong and China

- National Committee of Chinese People's Political Consultative Conference (consultative body on major national policies);
- regional committees of Chinese People's Political Consultative Conference (CPPCC) (consultative bodies on major regional policies).

(e) Degree/level of CPC

Few previous studies have systematically measured the degree of CPC. Besides comparing between CPC and non-CPC companies, i.e. distinguishing the sample companies into politically connected and unconnected ones, this study also tries to gauge the degree of CPC by assigning different weightings to different levels of CPC, as illustrated by the scorecard in Table 3.

The scorecard divides different public service institutions of Hong Kong and Mainland China each into four levels according to the extent of their political influence on public policy and law-making. For example, the Executive and Legislative Councils, being the supreme policy-making and law-making bodies in Hong Kong respectively, represent the uppermost local political authorities. Membership in them implies the highest level of political connection. They are assigned the highest score of 4. Memberships in institutions of less political influence ranging from statutory bodies to regional consultative committees and interest groups⁵⁵ receive the scores of 3, 2 and 1 respectively. Thus, the different CPC weightings are used to reflect how deeply corporations have penetrated into the core of political power and serve as the measuring tool for the degree of CPC in the study.

In Mainland China, the People's Congress and CPPCC exercise their power at all levels: there are national, provincial, municipal and county congresses and CPPCC committees. Like CPC in Hong Kong, membership at different levels of public service institutions in

⁵⁵ For interest groups, chairship or executive membership is demanded by the study considering their large number of members and mere membership, who do not have much political influence as do those who hold key positions in the interest groups.

Total asset value	Count	Mean (HK\$m)	Mean rank	<i>p</i>
CPC companies	80	99,389	54.35	0.025*
Non-CPC companies	21	20,780	38.24	

Table 4: Difference in total assets between CPC and non-CPC companies

Capitalisation	Count	Mean (HK\$m)	Mean rank	<i>p</i>
CPC companies	80	52,320	55.49	0.003**
Non-CPC companies	21	9,305	33.90	

Table 5: Difference in capitalisation between CPC and non-CPC companies

China are assigned different scores reflecting their extent of influence, ranging from 4 at the national level to 1 at the county level.

5 Findings

Results of the tests are presented in the following sections.

CPC AND CORPORATE STRUCTURE

As mentioned earlier, corporate structure is measured by its *scale* in terms of *total assets* and *capitalisation* and *ownership concentration* in the study. A statistical comparison of each of these variables between the 80 CPC companies (with CPC score > 0) and 21 non-CPC companies (with CPC score = 0) in the sample of 101 listed companies was conducted to discover any significant differences of corporate structure in the two company groups.

(a) Total asset value

Data of the total asset value of the companies under comparison are non-parametric. Therefore, U-test was performed and the results shown in Table 4 were obtained.

As seen in Table 4, CPC companies have a much higher average total asset value of HK\$99,389m, which is 4.78 times more than the average asset total value of \$20,780m of non-CPC companies. U-test results confirm that the mean rank difference is statistically significant at 2.5 per cent. The finding shows that, as expected, CPC companies tend to have larger total assets than non-CPC companies.

(b) Capitalisation

Total assets measure the scale of a company by the concrete value of its possessions. Another tool to measure company value and thus its scale is capitalisation, the total market value of the issued shares of a company. It represents how valuable the company is in the market and how much market share it occupies and is considered one of the most influential variables in measuring market size in organisational studies.⁵⁶

The data of capitalisation are also non-parametrically distributed. Comparison of the mean ranks of the CPC and non-CPC company groups by U-test demonstrates that their difference in capitalisation is highly significant, as shown in Table 5.

56 M J Chen and D C Hambrick, 'Speed, Stealth and Selective Attack: How Small Firms Differ from Large Firms in Competitive Behaviour' (1995) 38(2) *Academy of Management Journal* 453, at 455.

Table 5 shows that CPC companies have a mean capitalisation of HK\$52,320m, which is 5.60 times higher than that of the non-CPC companies (HK\$9,350m). The mean rank difference of the two groups is highly significant at 0.3 per cent.

It is true that all listed companies in Hong Kong are relatively large. The above tests, however, are able to detect significant differences in scale between the CPC and non-CPC listed companies in the sample. The previous studies have contradictory results over whether CPC is more widespread among larger-scale companies.⁵⁷ As little research in Hong Kong corporate literature has focused on the subject before, the research finding of this study, which proves that company scale matters in Hong Kong CPC, can help to fill the literature gap.

It is logical to reason that larger-scale companies have more human and financial resources for CPC. With a bigger market share than smaller companies, they are also likely to extract more rents through CPC. Moreover, holding formal posts in the government and other public service institutions is a symbol of prestigious status in Chinese society,⁵⁸ which is especially attractive to larger companies for maintaining and promoting their brand names. Above all, naturally the more assets a company possesses, the more it is concerned about safeguarding them against unfavourable government policies. A direct way to do so is establish CPC through formal participation in policy-making or consultative institutions so that it can influence government policies affecting its business.

The finding of this study is different from Nee and Oppen's research on CPC in China, which finds no relationship between company size and CPC.⁵⁹ It is believed that Hong Kong, though part of China, has different political traditions and practices from China, where company size is a negligible issue in establishing CPC when compared with other critical factors such as 'guanxi' (i.e. personal relationships) and the political background of companies.⁶⁰

(c) Ownership concentration

Besides company scale, ownership concentration is another aspect of corporate structure for comparison between CPC and non-CPC companies in this study. Berle and Means have provided a classical theory on ownership concentration by categorising companies into five major types according to the level of control.⁶¹ Applying the classification system to the sample companies investigated in this study, we can see a clear picture of the concentration pattern typical of companies in Hong Kong, as depicted in Table 6.

According to Berle and Means, only Type 5 companies which have no controlling shareholders can be considered 'widely held'. In other words, an overwhelming majority of 98.02 per cent of the listed companies in Hong Kong have concentrated ownership. Even when adopting the broader definition suggested by La Porta et al that counts a company as widely held if no ultimate owner controls 20 per cent or more of its shares (which includes both Type 4 and 5),⁶² still over 91 per cent of the sample companies are classified as concentrated ownership in structure.

57 M Faccio (n 43) 370; Nee and Oppen (n 16) 30.

58 Y Li, *The Structure and Evolution of Chinese Social Stratification* (University Press of America 2005), 23–48.

59 Nee and Oppen (n 16) 30.

60 U C Braendle et al, 'Corporate Governance in China: Is Economic Growth Potential Hindered by Guanxi?' (2005) April 25 SSRN eLibrary, 12–13.

61 A A Berle Jr and G C Means, *The Modern Corporation and Private Property* (Transaction Publishers 1932/1991), 67.

62 R La Porta et al, 'Corporate Ownership around the World' (1998) 54(2) *Journal of Finance* 471, at 478.

Control type		Concentration*	No of companies	%	Accumulated %
Type 1	Complete control	> 80%	2	1.98	1.98
Type 2	Majority control	> 50–80%	48	47.52	49.50
Type 3	Minority control	> 20–50%	42	41.58	91.08
	Control through legal device				
Type 4	Joint minority management	> 5–20%	7	6.93	98.02
Type 5	Management control	< 5%	2	1.98	100.00
		Total	101		

Table 6: Control types of sample companies under Berle and Means' classification

*Largest shareholding held by a single shareholder

Ownership Concentration	Count	Mean (%)	Standard deviation	<i>p</i>
CPC companies	80	44.22	20.40	0.000***
Non-CPC companies	21	56.87	10.84	

Table 7: Difference in ownership concentration between CPC and non-CPC companies

According to Berle and Means, a company's structure gradually evolves from Type 1 to Type 5, i.e. from concentrated to dispersed ownership, when shareholding is broken up as a result of inheritance or death.⁶³ Leech attributes such a change more to intense market competition, when shareholding is broken up as a result of the issuing of more securities for raising capital.⁶⁴ However, both theories seem not applicable to the Hong Kong case. Many large Hong Kong public companies have already experienced death of corporate founders and the rapid economic growth of the 1970s to 1990s when competition was keen and the need for capital investment was great. Yet, corporate ownership remains highly concentrated. In fact, Hong Kong has the highest ownership concentration in Asia.⁶⁵

Despite the overall concentrated ownership structure of Hong Kong companies, it is useful to find out whether CPC further pushes companies towards the concentrated end or significantly lowers the degree of concentration. After all, the relationship between CPC and ownership concentration is a scarcely explored subject in past literature. The data of ownership concentration of the CPC and non-CPC company groups under study being parametric, U-test was conducted with the results shown in Table 7.

As expected, both CPC and non-CPC companies display high degrees of ownership concentration (with on average around half of the company shares owned by the largest controller), which is typical of Hong Kong companies as discussed above. However, the degrees of concentration of the two groups of companies are found to be dissimilar. CPC companies have statistically lower average concentration (44.22 per cent (SD: 20.40)) than non-CPC companies (56.87 per cent (SD: 10.8)). The difference is extremely significant with *p*-value lower than the 0.01 per cent level.

63 Berle and Means (n 61) 66.

64 D Leech, 'Corporate Ownership and Control: A New Look at the Evidence of Berle and Means' (1987) 39(3) Oxford Economic Papers 534, at 537.

65 Gourevitch and Shinn (n 19) 18.

The finding that CPC companies have significantly lower ownership concentration than non-CPC companies confirms the prediction of the study. Through participating in the development of public policies, controlling owners of CPC companies can greatly amplify their control over other stakeholders (minority shareholders, employees etc.) and even capital markets. As aptly argued by Morck et al, 'political influence is plausibly related to what one controls, rather than what one owns'.⁶⁶ With increased control of the company and the market, controlling shareholders do not need to actually own as big a proportion of shares as their counterparts who have no political connection.

Another probable explanation is related to the scale of CPC companies. As found from the previous statistical analysis, CPC companies are generally larger in scale than non-CPC companies in Hong Kong. In discussing minority expropriation, Lang vividly illustrates how the controlling shareholder of a large corporate group can 'steal' from minority shareholders through pyramiding, which allows him/her to hold only a small portion of shares of different companies but gain majority control of each of the companies.⁶⁷ In another work he co-authored with Claessens and Djankov on the ownership and control of all listed corporations in nine East Asian economies, it was discovered that pyramiding is a common phenomenon in Hong Kong.⁶⁸ Large controlling companies at the top of these pyramids have magnified control despite low ownership. This helps one to understand why the generally large-scale-CPC companies, many of which are involved in pyramiding, have lower ownership concentration than the smaller non-CPC companies.

Some analyses attribute concentrated corporate structure to family ownership. For example, Goo and Weber observe that public companies have typically emerged from companies owned by families.⁶⁹ The owner families regard floatation of their companies' shares as merely a means of raising capital for the companies. Giving up the companies' control has never been the intention of floatation. The owner families will continue to keep a shareholding sufficient for controlling the companies.

The phenomenon of public companies found and controlled by family companies is not unique in Hong Kong but is common around the world.⁷⁰ There is, however, great variation in the degree of ownership concentration in different economies.⁷¹ It seems that the clinging on to controlling power in the company is less obvious in non-Chinese family companies than in Chinese family companies, as will be discussed in the next section.

66 R Morck et al, 'Corporate Governance, Economic Entrenchment, and Growth' (2005) 43 *Journal of Economic Literature* 655, at 655.

67 L H P Lang, *Governance and Expropriation* (Edward Elgar 2005).

68 S Claessens et al, 'The Separation of Ownership and Control in East Asian Corporations' (2000) 58 *Journal of Financial Economics*, 81, at 92–3.

69 S H Goo and R H Weber, 'The Expropriation Game: Minority Shareholders' Protection' (2003) 33 *Hong Kong Law Journal*, 71, at 71.

70 A majority of listed companies in Western and Eastern Europe, South and East Asia, the Middle East, Latin America, and Africa are publicly traded companies. See, e.g: M Burkart et al, 'Family Firms' (2003) 58(5) *Journal of Finance* 2167, at 2167–2201; M Faccio and L H P Lang, 'The Ultimate Ownership of Western European Corporations' (2002) 65(3) *Journal of Financial Economics* 365, at 365–395; S Claessens et al, 'The Separation of Ownership and Control in East Asian Corporations' (1999) SSRN; La Porta et al (n 62) 471–517; M Pagano and A Roell, 'The Choice of Stock Ownership Structure: Agency Costs, Monitoring, and the Decision to Go Public' (1998) 113(1) *Quarterly Journal of Economics* 187, at 187–255; A Shleifer and R W Vishny, 'A Survey of Corporate Governance' (1997) 52 *Journal of Finance* 737, at 738–83. Even in the US and the UK, some of the largest listed companies are controlled by families, for example, Wal-Mart, Ford in the US and J Sainsbury in the UK. See Anonymous, 'The World's 250 Largest Family Businesses' (2004) <www.familybusinessmagazine.com/topglobal.html>.

71 Gourevitch and Shinn (n 19) 18.

Controllers' background	Count	%	Capitalisation (in HK\$m)	%	Total asset (in HK\$m)	%
Chinese family	67	67.68	1,795,321	40.98	2,848,729	33.96
HK Chinese	6	5.94	100,877	2.30	116,891	1.39
HKSAR government	1	1.01	84,616	1.93	113,666	1.36
Mainland Chinese	7	7.07	267,600	6.11	1,144,574	13.65
Taiwan Chinese	1	1.01	14,646	0.33	23,821	0.28
British	7	6.93	1,832,678	41.83	3,824,754	45.60
American	2	2.02	12,093	0.28	8,227	0.10
European	2	2.02	43,967	1.00	24,551	0.29
Australian	1	1.01	8,083	0.18	2,108	0.03
Malaysian	4	4.04	91,020	2.08	138,009	1.65
Indonesian	1	1.01	9,614	0.22	18,307	0.22
Jewish	2	2.02	120,513	2.75	123,877	1.48
	101	100.00	4,381,028	100.0	8,387,513	100.00

Table 8: Company controllers' backgrounds, total assets and capitalisation

(d) Hypothesis 1 confirmed

H1: CPC and non-CPC companies are structurally different.

Statistical findings of total asset value and capitalisation show that CPC companies are of significantly larger scale than non-CPC companies. Findings of ownership concentration also show that CPC companies are significantly less concentrated than their counterparts. These results prove that CPC and non-CPC companies are structurally different in terms of company scale and concentration pattern. H1 is therefore *confirmed*.

CPC AND CHINESE FAMILY BUSINESSES**(a) Chinese family businesses**

Hong Kong-based large companies are commonly Chinese family businesses, as shown in the summary of controllers' backgrounds of the 101 listed companies under study in Table 8.

Due to Hong Kong's background of Chinese ethnicity and British colonialism, local Chinese and British companies together occupy 82.18 per cent (Chinese: 40.98 per cent, British: 41.83 per cent) of the total capitalisation and control 79.56 per cent (Chinese: 33.96 per cent, British: 45.6 per cent) of the total assets of all listed companies under study. The large market share of British companies is attributable mainly to one of its member corporations, the giant HSBC. In terms of number, a great majority of the listed companies in fact have a Chinese background (81.20 per cent, N=82) as shown from the first five rows of controllers' background in Table 8. Among them, most are Chinese family companies, which represent 67.68 per cent of all companies under study, 33.96 per cent of the total assets, and 40.98 per cent of the total capitalisation.

The above figures show the prevalence of Chinese family companies and the considerable market share these companies occupy in Hong Kong. Previous studies have revealed the cultural characteristic of Chinese family business controllers, who see the

	Ownership	Count	Mean CPC score	Mean rank	<i>p</i>
HK CPC	Chinese-family	67	8.69	55.23	0.040*
	Non-Chinese family	34	6.53	42.66	
China CPC	Chinese family	67	3.33	55.48	0.019*
	Non-Chinese family	34	1.56	42.18	
Total CPC	Chinese family	67	12.01	55.88	0.018*
	Non-Chinese family	34	8.09	41.38	

Table 9: Company controllers' backgrounds, total assets and capitalisation

business as personal and familial property, to hold fast to their controlling power.⁷² It would be interesting to find out if these influential corporations in Hong Kong seek more CPC to secure their power than other companies.

U-test was performed on the non-parametric data of the CPC degrees of the 67 Chinese-family companies and 34 non-Chinese family companies. Table 9 indicates a significant difference between the CPC levels of the Chinese family and non-Chinese family company groups as measured by their aggregate CPC scores in Hong Kong and China ($p = 0.018^*$). When regional connections are considered separately, the test similarly yields significant CPC differences between the two groups in both Hong Kong ($p = 0.040^*$) and China ($p = 0.019^*$). In all these three findings, the Chinese family company group shows higher mean CPC scores than its counterpart.

Why are Chinese family businesses more connected with politics? The above-mentioned assumption of their eagerness to maintain control over familial property can be a good explanation. Such eagerness could originate from the Chinese cultural emphasis on familial loyalty, a philosophical concept raised by Confucius that has moulded the thinking and behaviour of the Chinese people for thousands of years. In brief, the Confucian ideal of family is a paternal hierarchy, where the forefather is the source of authority over a large extended family system, where filial piety towards parents and ancestors and loyalty to family and even clan members are central values, and where the needs and glory of the family have priority over those of the individuals.⁷³ The Confucian ethic seems to have extended to other social contexts as well. Chinese businesspeople would strive to protect their family businesses and make sure that they do not fall into the hands of non-family members. Ruskola further points out that the kinship logic of Chinese family firms is against the individual personality of corporations stressed by the Western 'nexus of contract' theory. In Chinese societies, clan corporations justify their profit-seeking at the expense of others by emphasising their fiduciary duties to maximise the collective interests of their families and extended families.⁷⁴

Besides the ideological-philosophical background of the Chinese, the traditional low social status of business people in China also plays a role. First, they were not as influential as landlords or farmers in the old days, when the economy relied heavily on farmland and agriculture. Apart from farmers, scholars were respected for their knowledge, and craftsmen for their skills. Merchants, however, were perceived by the general public as selfish and

⁷² Lawton, 'Berle and Means' (n 22) 372; Lawton, 'Modelling' (n 21) 266.

⁷³ A N Licht, 'Legal Plug-Ins: Cultural Distance, Cross-Listing, and Corporate Governance Reform' (2004) 22 Berkeley Journal of International Law, 195, at 214–5.

⁷⁴ T Ruskola, 'Conceptualizing Corporations and Kinship: Comparative Law and Development Theory in a Chinese Perspective' (2000) 52 Stanford Law Review, 1599, at 1607.

More regulated industries (no of companies = 52)	HK CPC			China CPC			Total CPC
	D	ID	Sub-total	D	ID	Sub-total	
Public utilities and transportation (5)	5.20	8.40	#13.60	2.80	0.00	2.80	#16.40
Properties and construction (22)	3.86	5.23	9.09	2.82	1.59	4.64	13.73
Financials (17)	3.82	5.18	9.24	0.24	1.88	2.12	11.35
Communications (8)	2.25	7.13	9.38	0.38	1.38	1.75	11.13
Average	3.73	5.81	*9.62	1.60	1.50	*3.19	*12.81
Less regulated industries (no of companies = 49)	HK CPC			China CPC			Total CPC
	D	ID	Sub-total	D	ID	Sub-total	
Conglomerates and others (18)	5.20	8.40	#13.60	2.80	0.00	2.80	#16.40
Hotels and entertainment (9)	3.86	5.23	9.09	2.82	1.59	4.64	13.73
Industrial and consumer products (22)	3.82	5.18	9.24	0.24	1.88	2.12	11.35
Average	2.25	7.13	9.38	0.38	1.38	1.75	11.13

Table 10: More regulated industries v less regulated industries and their CPC scores in Hong Kong and China (D = executive and non-executive directors; ID = independent directors)

corrupt, making profit out of exploitation. There were in general four social classes in ancient China. From top to bottom these were: scholars, farmers, craftsmen and merchants, with merchants ranking lowest on the four-tier social ladder.⁷⁵ Because of Chinese merchants' inferior social status in past history, their traditional fear of and great respect for government officials in the highly hierarchical society, and their adherence to political power for raised social status and, more importantly, protection of business, especially family business, all these point to a greater need for political connection in Chinese family companies than other companies.

(b) Hypothesis 2 confirmed

H2: Chinese family companies have more CPC than other companies.

The test comparing the CPC level of Chinese family companies verses non-Chinese-family companies shows that Chinese family businesses do have more CPC – in Hong Kong, in China and in Hong Kong and China as a whole – than the other companies. H2 is also *confirmed*.

CPC AND GOVERNMENT REGULATION

Referring to the definitions used by the HSICS, the 101 companies were categorised into seven industry types under the more regulated and less regulated group as shown in Table 10.

75 D Twitchett, 'A Confucian's View of the Taxation of Commerce: Ts'ui Jung's Memorial of 703' (1973) 36(2) Bulletin of the School of Oriental and African Studies, University of London 429, at 438.

	Industries	Count	Mean CPC score	Mean rank	<i>p</i>
HK CPC	More regulated industries	52	9.62	56.63	0.044*
	Less regulated industries	49	6.20	45.02	
China CPC	More regulated industries	52	3.19	54.59	0.169
	Less regulated industries	49	2.24	47.19	
Total CPC	More regulated industries	52	12.81	56.67	0.044*
	Less regulated industries	49	8.45	44.98	

Table 11: Difference in degree of CPC between more regulated and less regulated industries

(a) More and less regulated companies

Table 10 shows that the industries under more government regulations (regarding entry into industry, licence conditions, licence renewal etc.) have higher average Hong Kong, China and total CPC scores than industries under fewer regulations (scores with *).

Among the more regulated industries, ‘public utilities and transportation’ has the highest Hong Kong CPC score as well as total CPC score (score with #). This is a reasonable outcome as the industry is composed of monopolised companies of natural resources and public transportation such as Hong Kong and China Gas, China Light and Power Holdings, and MTR Corporation. These companies are subject to the strictest regulations that govern their price-setting and limit their profit boundaries. Since they are so stringently controlled by the local legislature, Hong Kong CPC becomes paramount for them to increase their say in shaping relevant policies in the Executive Council, Legislative Council and policy consultative bodies.

Table 10 also shows that ‘properties and construction’ has the highest average China CPC score among all the highly regulated industries. An explanation for this can be found in the annual reports of these companies. For example, the largest property developer in Hong Kong, Cheung Kong Holdings had 20 new acquisition and joint development projects in 2005, seven (i.e. over one-third) of them were in Mainland China.⁷⁶ In the same year, the company had 21 properties under development in Hong Kong but 44 (i.e. more than twice those in Hong Kong) in Mainland China. The second largest local property developer, Sun Hung Kai Properties Ltd reported that its Hong Kong land bank in the financial year of 2005/2006 was 42.4m ft², which is a slight growth of 1.2 per cent compared with the previous financial year. Within the same period, its China land bank was 19.8m ft², a massive 132.9 per cent increase from the previous year.⁷⁷ These figures show a clear trend of rapid business expansion of the property development sector from Hong Kong to Mainland China. Given the limited land resources in Hong Kong, such a development direction is inevitable. This makes CPC in China strategically important for the Hong Kong properties and construction industry.

⁷⁶ *Annual Report 2005* (Cheung Kong (Holding) Ltd 2006) <www.ckh.com.hk/eng/investor/annual/2005eng.pdf>.

⁷⁷ *Annual Report 2005/2006* (Sun Hung Kai Properties Ltd 2006) <http://shkp.com/data/investors/reports_detail/7/7/7_152_en.pdf>.

Finally, Table 11 indicates that for each of the seven industries, more and less regulated ones alike, Hong Kong CPC scores are consistently higher than their corresponding China CPC scores. Notwithstanding the growing impact of Chinese politics on Hong Kong businesses after the change of sovereignty in 1997, Hong Kong public and corporate policies, after all, have the most direct and determining influence on companies based in Hong Kong. Hong Kong CPC is therefore the primary means of political connection for more regulated corporations in Hong Kong.

U-test results comparing the more regulated and less regulated company groups shown in Table 11 above confirms that the higher level of Hong Kong and total CPC scores of more regulated industries are statistically significant at the same level of 4.4 per cent. China CPC scores between the two groups are, however, not significantly different. As explained in the previous paragraph, more regulated industries are under strict local regulations. While they have the tendency of building somewhat more China CPC than less regulated industries as an indirect means of influencing local regulations, at the same time they naturally seek significantly more Hong Kong CPC than the less regulated group to directly protect their business interests.

(b) Hypothesis 3 partially confirmed

H3: more regulated industries have more CPC than less regulated ones.

Findings show that more regulated industries have significantly higher Hong Kong CPC and total CPC levels than less regulated ones. But the difference in the China CPC level between the two groups is insignificant, showing that industries subject to strict local regulations are more obviously reliant on local CPC than CPC with the mainland government. H3 is partially confirmed.

6 Implications

CPC AS A NEW APPROACH TO UNDERSTANDING CORPORATE GOVERNANCE

In recent years, politics has played an increasingly important role in corporate governance. Especially after the financial crises in 1998 and 2008, the market economy has become more in need of government intervention for stimulation and assistance.⁷⁸ Companies are also playing a more public role in society. Not only do they have clearly defined public responsibilities and purposes recognised by law, but society also expects them to be accountable to the public for their actions. In this sense, they are both private associations and public bodies.⁷⁹ This public role is intensified under globalisation, with international commercial arrangements often decided by multinational enterprises rather than governments alone. The new world trade order displays the dual role of corporations as

78 For example, as at February 2009, the US government has put US\$789bn into rescuing the market, which exceeds the entire cost of the Iraq War fought since 2003. G Hitt and J Weisman, 'Congress Strikes 789 Billion Stimulus Deal', Wall Street Journal, 12 February 2009, A1. In September 2008, a rumour spread by text message triggered a brief run on the Bank of East Asia as depositors feared that the bank in Hong Kong had large exposures to Lehman Brothers and AIG. The government in response guaranteed from October 2008 through year-end 2010 the repayment of all customer deposits held in all banks and authorised financial institutions in Hong Kong. The guarantee, which was backed by the government's Exchange Fund, applied to both Hong Kong-dollar and foreign-currency deposits, including those held in Hong Kong branches of overseas institutions. The original protected sum was HK\$100,000 (US\$12,898) per depositor per banking institution. 'Recent Developments in Asian Deposit Guarantee Programs' (2008) Asia Focus, Country Analysis Unit, Federal Reserve Bank of San Francisco <www.frbsf.org/publications/banking/asiafocus/2008/Asia_Focus_Deposit_Insurance_Oct_08.pdf>.

79 A Gamble and G Kelly, 'The Politics of the Company' in J Parkinson et al (ed), *The Political Economy of the Company* (Portland 2001), 27.

both private and public commercial actors in the political, social and economic arenas.⁸⁰ The complex interactions between government and corporations and the connections between the political and business sectors are topics of growing relevance to the contemporary world, especially when they affect corporate governance at both micro and macro levels.

This is particularly true in Hong Kong. Since the 1997 handover to China, the fight for universal suffrage among democrats, the growing dissatisfaction of the public towards the Hong Kong and mainland governments and the frequent protests of various sectors of the society, all show that Hong Kong society is increasingly politicised. Even the government acknowledged the fact and introduced the political appointments system of senior officials in 2002 to help tackle political issues.⁸¹ Both political authorities and corporations need to handle CPC shrewdly and diplomatically and make necessary adjustments in order to meet public demands.

As shown from the above quantitative analysis, *large-scale companies*, *Chinese family companies*, and *more regulated companies* are proved to have more CPC than the respective opposite company groups. These three types of companies share a common feature. They are all the most influential companies in that they have more resources, dominate the local business sector, and/or belong to industries that have the greatest impact on the daily lives of the general public. In other words, the most influential companies in Hong Kong are also companies that are most connected politically by taking up official positions and by informal contacts with political authorities. The combination of political and business power gives rise to privileged, excessively dominant business controllers who upset the power balance in corporate governance. The great contrast in power balance is reflected in the wide gap between the rich and the poor as shown by the high Gini score of Hong Kong. The United Nations Development Programme investigated income inequality worldwide in 2007 and reported that Hong Kong topped all other advanced economies as the region with the biggest gap between the rich and the poor. Hong Kong had a Gini score of 43.4,⁸² with the richest 10 per cent of the population receiving 34.9 per cent of the city's total income but the poorest 10 per cent sharing only 2 per cent of it.⁸³ The government cannot afford to lose control of the development of CPC, or serious problems in companies and in society as a whole can be expected.

RE-EXAMINATION OF CORPORATE GOVERNANCE THEORIES

Besides pointing to a new approach to understanding corporate governance, the results of this study have also helped to re-examine some of the existing corporate governance theories.

(a) Market thesis

The market thesis predicts that as a company develops or the market matures, shareholding will break up through inheritance or issues of more securities, resulting in the market gradually changing towards diffuse ownership.⁸⁴ Berle and Means consider dispersed ownership a natural product of market force; for as corporations expand, they require an

80 B Ahunwan, *Globalisation and Corporate Governance in Developing Countries* (Transnational Publishers Inc 2003), 35.

81 D Tsang, Press Release, 'Chief Secretary's Speech', 18 March 2002 <www.info.gov.hk/gia/general/200203/18/0318152.htm>.

82 The Gini coefficient is named after the Italian statistician, Corrado Gini, with zero signifying absolute income/wealth equality and 1 absolute inequality. In the United Nations Development Programme report, Gini scores range from 0 (absolute income equality) to 100 (absolute income inequality).

83 B Einhorn, 'Countries with the Biggest Gaps between Rich and Poor', *Business Week*, 16 October 2009.

84 See Berle and Means (n 61) 66; and Leech (n 64) 537.

increasing amount of capital that exceeds the resources of any single individual or family⁸⁵ and that can only be raised through selling shares to numerous small investors.⁸⁶ A diffused ownership structure is necessitated for large public corporations due to the demand for more capital to meet market competition.

From a market perspective, Hong Kong is equipped with most requisites for a dispersed system. First of all, it is a highly developed economy. Its gross national income (GNI) per capita ranked 28th in the world in 2005.⁸⁷ Hong Kong stock market's capitalisation ranked seventh and third in the world and Asia respectively in 2009.⁸⁸ As for corporate governance, in spite of all the aforesaid problems related to its concentrated family business model, Hong Kong has maintained a relatively high standard in its legal and corporate governance systems among countries practising ownership concentration.⁸⁹ However, as seen from the statistical evidence of this study, Hong Kong is a sophisticated market with a vast majority of highly concentrated companies. The empirical finding shows that existing theories, especially those developed in the context of the US/UK markets, may not be applicable to Hong Kong and other countries.

As observed by La Porta et al, the diffuse corporate ownership of the US/UK is highly exceptional. Outside of the two countries, large firms usually have ultimate controlling owners.⁹⁰ Morck et al's study further confirms that most large corporations in the world are controlled by very wealthy families. It is common for them to have super voting rights, pyramidal control and cross shareholding, through which they can control a considerable proportion of a country's economy.⁹¹ For instance, in Hong Kong, the Chinese family businesses have their cultural and psychological persistence in corporate ownership. No matter how developed the market is, it is not easy for them to transform into diffuse ownership. It follows that the convergence theory (i.e. the prediction that all economies will converge into the US/UK diffuse ownership model) is unlikely to be realised through market adjustments. An ideological metamorphosis of most countries is a prerequisite if the theory is possible at all.

(b) Law thesis

La Porta et al argue that the quality of law is a robust determinant of ownership patterns.⁹² Their study reveals that there is positive correlation between good legal minority shareholder protection and dispersed ownership and vice versa.⁹³

However, this correlation does not apply to Hong Kong. Hong Kong has relatively good legal minority shareholder protection. In the Corporate Governance Index of FTSE Institutional Shareholder Services (ISS), Hong Kong ranks fourth among 24 markets in the world and the first among the Asian markets. It ranks just slightly lower than Singapore in

85 Berle and Means (n 61) 59.

86 Ibid 2–7.

87 World Bank, 2006 <<http://siteresources.worldbank.org/DATASTATISTICS/Resources/GNIPC.pdf>>.

88 *Market Capitalisation of the World's Top Stock Exchanges* (Securities and Futures Commission 2007) <www.sfc.hk/sfc/doc/EN/research/stat/a01.doc>.

89 C K Low, 'A Road Map for Corporate Governance in East Asia' (2004) 25 *Northwestern Journal of International Law and Business* 165, at 168.

90 La Porta et al (n 62) 471–517.

91 Morck et al (n 66) 657.

92 La Porta et al (n 62) 511.

93 Ibid 505–11.

terms of corporate governance culture and international audit and accounting standards.⁹⁴ It also ranks in the top quartile on the World Bank's rule of law index in 2004 and 2009.⁹⁵ Yet, Hong Kong has remained an economy of highly concentrated ownership, as suggested by the findings reported in this article.

La Porta et al also suggest that the common law system has more minority shareholder protection and that good minority protection is correlated with dispersed ownership.⁹⁶ The presumed link between the common law system and dispersed ownership concentration is contradictory to the real situation of Hong Kong, which has inherited the British common law system since its colonisation more than a hundred years ago.

This CPC study suggests a new perspective for ownership concentration and corporate governance. Although the overall concentration pattern of Hong Kong companies under the common law system remains high, the ownership of CPC companies is proved to be statistically less concentrated than that of non-CPC companies. In the context of CPC companies, less concentrated corporate structure does not necessarily guarantee more minority shareholder protection because controlling shareholders exist in almost all of these less concentrated companies.

The fact that corporate governance cannot benefit from a reduced yet still high concentration level in CPC companies points to two interesting questions. First, is there a threshold or optimum point on the concentration continuum to be crossed for any economies to function as effectively as the US/UK model? And even if there is, is the US/UK model really the best model to be pursued by any countries? While the first question is worth more research efforts in the area, the answer to the second question seems to be negative, judging from the results of this study.

The transplantation of a legal system is not equivalent to transplanting the legal spirit. Borrowing laws and regulations from the Anglo-American and common law systems does not guarantee dispersed ownership or good law. If the implanted laws contradict local values, customs and existing institutional make-up, the laws are likely to be changed or twisted to suit local needs.⁹⁷ Therefore, when the same legal system is introduced to other countries, different outcomes can be expected.

(c) Politics thesis

The political theory developed by Roe suggests that the political reaction of majority shareholders to the demand of government for social equality and stable employment in social democratic European countries has resulted in the persistent concentrated ownership, which helps majority shareholders to gain more control to resist the political pressure of raising costs and foregoing profit-maximising opportunities for social causes.⁹⁸ The theory is obviously inapplicable to the Hong Kong case, which is quite opposite to the example raised by Roe.

Hong Kong has a non-democratic government and controlling shareholders' interests are not in conflict with those of the government. The controlling shareholders do not resist

94 The rankings of ACGA and ISS are quoted from the speech of M Wheatley, 'Corporate Governance' (Securities and Futures Commission 2006) <www.sfc.hk/sfc/doc/EN/general/general/press_release/06/mw_060116_hkioid.pdf>.

95 'Worldwide Governance Indicators Country Snapshot' (World Bank 2011) <http://info.worldbank.org/governance/wgi/sc_country.asp>.

96 R La Porta et al, 'Law and Finance' (1998) 52 *Journal of Finance* 1131, at 1138.

97 K Pistor et al, 'The Evolution of Corporate Law: A Cross-Country Comparison' (2002) 23 *University of Pennsylvania Journal of International Economic Law* 791, at 797.

98 Roe (n 17) 539, 543 and 594.

but actively participate in politics and cooperate with the government as far as possible in exchange for political benefits. Such reaction does not result in further increasing their shareholding but in the political phenomenon of CPC. As found by the statistical analysis of the study, CPC companies have significantly less concentrated ownership than non-CPC companies. It is probably because CPC companies have available political channels to maintain their controlling power through CPC. In sharp contrast to Roe's theory, the political reaction of controlling shareholders in Hong Kong is associated with lower rather than higher ownership concentration. However, a commonality of the European and Hong Kong cases is that politics does make a difference in corporate ownership structure.

Previous approaches to the study of corporate governance are largely confined by *the market thesis*, *the law thesis* and *the politics thesis* which do not recognise the potential significance of CPC. CPC could be a new approach to understanding the political aspects of corporate governance and exploring solutions to related problems.

7 Conclusion

This article has explored the characteristics of CPC companies in Hong Kong in relation to their corporate structure, Chinese ownership background and industry type. In general, CPC companies tend to be larger in scale, and companies with Chinese family ownership and in more regulated industries tend to have a higher degree of political connection. The study attempts to anatomise the complex issue of political connection established by companies and contends that past theories and approaches are inadequate to deal with corporate governance problems. Hong Kong, as do other countries in the world, has its own unique and complex social backgrounds and conditions. The simple convergence of law is not penicillin for every economy. The world needs to adopt a more political perspective that takes into account the interaction of cultural, historical, legal and political factors. The study of CPC as a determinant of corporate governance provides a possible direction for future research on corporate governance in Hong Kong and elsewhere.

Taking the minimum content seriously: Hart's liberalism and moral values

GAVIN BYRNE

*Birmingham Law School**

In spite of its immense reputation, H L A Hart's legal philosophy is rarely examined as a single, coherent body of work. *The Concept of Law* tends to be discussed on its own, separate from Hart's liberal critiques of substantive law. Hart himself regarded these projects as distinct. This essay will unite Hart's descriptive project with his critical commentaries. As such it runs contrary to the manner in which Hart's work is normally presented and perhaps the way in which he intended it to be read. Yet, as Nicola Lacey's biography explains, Hart saw himself as a liberal perhaps above all else.¹ It is important, therefore, that his descriptive project fits with the various critical positions that he espouses elsewhere to form a coherent liberal whole. A central tension emerges in Hart's writings when they are read in this way. It relates to Hart's position on the status of moral values, an aspect of Hart's legacy that is largely overlooked.²

Much has been made of Hart's supposed scepticism in relation to the objective truth of moral values. Hart claimed to be non-committal on this issue.³ Yet Hart provides fact-based reasons for certain core moral values, a position that renders such values objectively true. This essay begins with an expository analysis; Hart's scepticism in this regard has been misunderstood. Hart was sceptical as to the usefulness of debate on the objective status of moral values, but his own position requires a commitment to such values. As such, many values that he espouses in his critical comments about law are best understood as objectively true.

Once this preliminary point is dealt with we move on to a central tension in Hart's writings and the core argument in this essay. Hart embraced a form of political liberalism that requires the existence of lasting moral values. On the other hand, Hart goes to great lengths to emphasise the contingent nature of the values discussed in chapter IX. This tension goes to the heart of liberal positions. Liberalism wishes to promote tolerance of multiple viewpoints and avoid moral absolutism, yet the only way in which this worldview

* I am immensely indebted to Sean Coyle for very helpful feedback on multiple versions of this paper. I would also like to thank Gordon Woodman, Steve Smith, Martin Borowski, James Lee and attendees at the jurisprudence section of the 2010 Society of Legal Scholars conference for comments on earlier drafts.

1 N Lacey, *A Life of H L A Hart: The Nightmare and the Noble Dream* (Oxford University Press 2004) 68, see also 36, 171 and 195.

2 Recent exceptions include A Perreault-Saussine, 'An Outsider on the Inside: Hart's Limits on Jurisprudence' (2006) 56 *University of Toronto Law Journal* 371; and S Delacroix 'Meta-Ethical Agnosticism in Legal Theory: Mapping a Way Out' (2010) 1(2) *Jurisprudence* 225

3 H L A Hart, *The Concept of Law* (2nd edn, Oxford University Press 1994) 254, see also 168.

can be advanced is by committing to lasting moral values such as the protection of fundamental freedoms. I argue that Hart's legal philosophy, viewed as a whole, provides us with a mechanism for the support of lasting liberal values about the specific content of law.

The manner in which Hart's position ultimately supports lasting critical moral values about the content of law casts doubt on his general approach to legal philosophy. In my concluding remarks I note how any contemporary Hartian faces a choice between following Hart's liberalism or following Hart's sharp distinction between analyses of what law is and what law ought to be.

1 Objective truth of moral values

A number of commentators discuss Hart's supposed scepticism about the objective truth of moral values. Leiter points to a tension between such scepticism and Hart's 'soft' positivism.⁴ Raz argues that certain weaknesses in Hart's concept of law result from his 'rejection of evaluative objectivity'.⁵ Perreau-Saussine suggests that Hart was uncomfortable with his 'doubts that . . . objective morality exists'.⁶ In this section I show that Hart's work supports the existence of objective moral values. A far bigger problem for Hart, as for all liberals, is the issue of whether such values are lasting.⁷

Hart never attached a specific label to his meta-ethical position. A number of commentators have done so. There is general consensus that Hart endorsed a naturalist meta-ethics.⁸ This labelling is useful as it helps to draw out some of the implications in Hart's claims. Many of the same commentators that categorise Hart as a naturalist, have also claimed that Hart is a sceptic about the existence of objectively true moral values. This suggestion makes little sense; such a position requires the existence of at least some objective moral values. In this section I demonstrate that objectivity about moral values is also the most natural reading of various claims that Hart makes. Locating Hart's meta-ethics within naturalism is useful for a second reason. The specific type of naturalism that Hart tacitly endorses has difficulty with the idea of moral values that are necessary or even lasting. It is the idea that certain core moral values might be hostage to change that really troubled Hart as a liberal.

The Concept of Law chapter IX is Hart's most sustained discussion of the relationship between law and morality. Hart explains why we can expect a degree of overlap between the substantive content of law and certain basic moral principles. In doing so he warns against the conclusion that legal systems must embrace certain values in order for their rules to count as law at all. This leaves the question of whether there are moral values that always *ought* to be embraced by a legal system, even if we accept that they do not need to be embraced in order for such a system to count as law. I return to this specific issue in the next section. For now let us consider how the argument in chapter IX and elsewhere displays the basic traits of reductionist naturalism about moral values.

4 Brian Leiter 'Legal Realism and Hard Positivism' in J Coleman (ed), *Hart's Postscript: Essays on the Postscript to The Concept of Law* (Oxford University Press 2001) 355–70, in particular 361–3.

5 J Raz, 'Two Views on the Nature of the Theory of Law: A Partial Comparison' in Coleman (ed) (n 4) 1–38, at 5–6.

6 Perreau-Saussine (n 2) 386–8.

7 The terms 'positive' and 'critical' morality are Hart's; moral philosophers speak of descriptive morality, the values a particular society holds, and normative morality, values we ought to hold. Discussion of 'moral values' herein should be understood in the normative sense.

8 See D Priel, 'Were the Legal Realists Legal Positivists?' (2008) 27(4) *Law and Philosophy* 309, at 330–2; Raz (n 5) 6; and Delacroix (n 2) 225–30. Priel notes that Hart looks to social practices in order to explain obligation. This is also a naturalist methodology, but the focus of this paper is on the nature of moral values themselves. Delacroix attributes 'bald naturalism' to Hart.

(a) *Source of Moral Values*

Reductionist naturalists hold that moral values come from non-moral facts about the world. Hart acknowledges that viable legal systems will have certain specific content in common with accepted moral values. 'Truisms' relating to the facts of our existence impact upon the substantive content of law and morality as systems of social ordering. Hart's explanation of reasons as to why we hold certain moral values is typical of reductionist naturalism. Survival is a 'good' that we seek to achieve. Facts about the physical world and human nature impact upon how that goal can be achieved. The content of core moral values reflects the pursuit of this good as impacted by the practicalities involved.⁹

(b) *Empiricism*

Reductionist naturalists seek empirical proof for matters of fact; the truth of a moral value is only accepted if there is empirical evidence for the underlying fact that it reduces to. Hart holds that 'truisms' of law and morality are not necessary, universal truths or teleological ends of man. His claims are expressly non-metaphysical.¹⁰ When we look throughout human history man actually *has* sought survival as a matter of observable fact; survival as a good to be pursued has played and continues to play a central part in all of our ordered affairs, how we interact and how we understand the world around us.¹¹ Similarly it is an empirical fact that human beings are vulnerable to physical harm. As a result, we have the generally accepted moral value that one should not physically harm another human being. Throughout his writings Hart embraced empiricism with regard to underlying non-moral claims used in support of moral arguments. In *Social Solidarity and the Enforcement of Morality*¹² Hart argues against the claim that unless the positive morality of a particular society is enforced through law that society is likely to disintegrate. Hart points out that this thesis lacks empirical evidence.

(c) *Scientific evidence*

Reductionist naturalists use the sciences to inform their moral positions.¹³ In the manner of reductionist naturalists, such as Peter Railton, Hart is willing to have his philosophy led by developments in physical and social sciences. Hart sees the issue of how we come to accept and adopt moral values as a question

9 Hart, *Concept* (n 3) 174–200. Hart follows Hobbes in 'lower[ing his] sights' to explain moral values. The fact that human beings seek survival is a common starting point; Hart emphasises the special place this holds in human endeavour: *ibid* 191–3. In this respect Hobbes is a forerunner of reductionist naturalism in meta-ethics.

10 *Ibid* 192.

11 *Ibid*. For Hart, that 'we are not a suicide club' must be presumed in any discussion of human ordered affairs; Peter Railton makes a similar claim in his explanation of reductionist naturalism, see 'Moral Realism' (1986) 95(2) *Philosophical Review* 163, at 181.

12 H L A Hart, *Essays in Jurisprudence and Philosophy* (Oxford University Press, 1983) 248–62. We can see Hart's distrust of moral claims that lack empirical bases elsewhere, see H L A Hart, *Law, Liberty and Morality* (Oxford University Press 1963) 58–60.

13 Reductionist naturalists are not 'scientists'; they do not hold that the sciences provide all answers to questions posed in moral philosophy. They simply use sciences to inform claims about value and rationality. See Peter Railton, 'Naturalism and Prescriptivity' (1989) 7(1) *Social Philosophy and Policy* 151, at 159–60.

for 'psychology and sociology'.¹⁴ This willingness to incorporate scientific fact into critical moral argument is a consistent feature of Hart's work.¹⁵

(d) *Revisability*

For the reductionist naturalist facts are contingent. The best that we can say about a matter of fact is that it is empirically verifiable as true for the time being. Should the truth value of an underlying fact change, the truth value of the corresponding moral value would also change.¹⁶ Moral values are thus revisable. For Hart, human beings might have evolved so as to be impervious to physical harm, resources might have been or might one day be limitless and human beings might have evolved or yet evolve so as to be radically different physically. Certain moral values are deeply embedded in our thought and language because of matters of fact, but Hart takes these facts to be contingently true.¹⁷

This brings us to the objectivity issue. Reductionist naturalism holds that moral values are objective; certain things are human aims, certain non-moral facts about the world enable us to pursue those aims. Moral values are reducible to a combination of objective non-moral aims and objective non-moral facts.¹⁸ If Hart really is a reductionist naturalist when it comes to moral values, he would need to hold that at least some such values are objective. Hart expressly avoided this issue.¹⁹ Yet there are a number of reasons why the values he discusses in chapter IX *should* be considered objective.

Hart treats the aim of human survival as an observable fact; human beings simply 'do wish to live, even at the cost of great misery'.²⁰ We are committed to this aim in our very

14 Hart, *Concept* (n 3) 193–94. Hart's review of J L Mackie's *Ethics: Inventing Right and Wrong* (Penguin Books 1977) is also consistent with this aspect of reductionist naturalism, see H L A Hart 'Morality and Reality', New York Review of Books (9 March 1977) 35. Hart endorses Mackie's explanation of moral values; mankind becomes disposed to act in particular ways because 'human beings are competitive with and vulnerable to one another, living in a world of scarce resources', morality becomes a 'device which by counteracting human egoism provides some of the essentials of human welfare', *ibid* 38. Hart is unconvinced that the alleged 'queerness' of ethical objects is tantamount to an argument against their existence. He also criticises Mackie for understanding the nature of morality as an intelligent creation with little explanation of the part played by human emotion. Railton's two major points of disagreement with Mackie are similar. Railton disagrees with the idea that there is anything 'queer' about moral values when taken to supervene upon non-moral goods, 'Moral Realism' (n 11) 171–7 and 183–4. Railton also disagrees with Mackie's conceptual point about moral facts requiring the existence of objective and *categorically* prescriptive facts. For Railton, emotion plays a part in how we get from needs and wants to moral values: *ibid* 183–9 and 200–4.

15 Hart, *Law, Liberty* (n 12) 68–9, and the postscript to *Punishment and Responsibility* (2nd edn, Oxford University Press 2008) 210.

16 Railton, 'Moral Realism' (n 11) 198–201, see also Railton, 'Naturalism' (n 13) 158–9. This does not require a theory of moral progress. Moral values transform to reflect changes in our environment and needs; there is no reason to assume that we are getting morally better, see Railton, 'Moral Realism' (n 11) 195.

17 Hart, *Concept* (n 3) 193–200.

18 Railton, 'Moral Realism' (n 11) 173–84. Railton claims that there is a 'reduction basis' for our beliefs as to what will produce a good or desirable sensation. This is a combination of the qualities of the individual, the object or phenomenon experienced and other factors such as context. This explains subjective wants. Railton posits the idea of an 'objectified subjective interest' to explain moral values. Objectified subjective interests are those that an idealised version of an actual individual *would* wish their non-idealised self to hold. Such interests are intrinsically good for the non-idealised individual. Objective *moral* goods involve interests of two or more people.

19 Hart, *Concept* (n 3) 168 and 253–4.

20 *Ibid* 192.

'structures of thought and language'.²¹ Yet Hart's claim about survival as a human aim goes far deeper. No discussion of morality makes sense unless we accept survival as a basic human goal. It simply has to be good for us to survive, because any discussion of what we ought to do as a group is premised on the idea that 'we are not a suicide club'. For there to be morality at all, our continued survival must be accepted as an objective 'good' rather than a matter of subjective preference.

Hart then discusses 'truisms' that impact upon the achievement of survival. Hart does not explicitly describe these as objective, but he presents them as such. Human vulnerability, approximate equality and limited altruism are said to be 'salient characteristics' of human nature; we do not merely *believe* these things to be true. Throughout the analysis Hart refers to these truisms as facts, 'men are . . . vulnerable to bodily attack', 'no individual is so much more powerful than others', 'men are not devils', 'human beings need food, clothes and shelter'.²² The most natural reading of these claims is that they do not depend on the conscience of an observer for their truth.

This, then, is the first reason why the moral values discussed in chapter IX should be thought of as objective. The whole thrust of Hart's argument is that core moral values are rooted in fairly banal facts about physical nature and our world, facts that both law and morality reflect. Objective truth simply means true in virtue of some set of facts about the world. In explaining the commonalities in substantive content throughout moral and legal codes Hart rooted these values in the world rather than in us.

The second reason is the inherent oddness of any alternative. In his analysis Hart makes it clear that he is providing *reasons* for moral values rather than *causes* of those values. Hart explains why we *do* hold certain 'universally recognized' core principles. This method of explaining the 'core of good sense' in natural law, also (perhaps inadvertently) *justifies* the moral values under discussion. Given that Hart provides a set of reasons for holding certain moral values, it seems impossible for him to accept any moral position according to which we ought not to hold these values provided the underlying facts continue to hold. Hart cannot argue against the view that we ought not to kill other human beings or that we *ought to* protect personal property without arguing against his own good reasons as to why we generally do hold these values.

To be a sceptic about the objectivity of these particular values, Hart would have to accept the counter-intuitive position that the reasons for such values are objective, but the values themselves might not be. If Hart wished to commit strongly to scepticism about all moral values, a very complex argument was required at this juncture. Such an argument is not provided.

Of course, Hart is free to accept that not all moral questions will have a single right answer; objective moral values may clash, which would involve a weighing-up process on the part of an individual.²³ This is not to deny that the moral values themselves are objective, it merely means that they conflict from time to time. Furthermore, Hart could accept that not *all* moral propositions have objective bivalence. There may be moral values that cannot be explained in the same way as those discussed in chapter IX. In this weaker sense, then, Hart may have been a sceptic when it comes to objective moral values, but scepticism about the objectivity of *all* moral values fits poorly with his analysis.

21 Hart, *Concept* (n 3) 192.

22 Ibid 194–9.

23 Railton notes that there may be different goods to weigh up each of which is objective and appealing. Railton distinguishes 'a good for A' (objective moral values) from 'the good for A' which an individual might decide upon by balancing objective values, 'Moral Realism' (n 11) 176.

Any tension in Hart's writings or discomfort that he may have had about the non-existence of objective moral values is easily addressed from within his own work. Hart's own comments reveal less scepticism about moral objectivity than they do scepticism about the very question of such objectivity. Quite simply, Hart felt that there was little to be gained in endless debate as to whether moral values ultimately come from within us or external facts about the world. Hart did not need to defend a position on objective moral values to respond to his critics.²⁴ In his review of Bernard Williams' *Ethics and the Limits of Philosophy*, Hart briefly addressed the issue of why meta-ethical arguments should matter; he was unconvinced that debates as to the objectivity of moral values were of practical significance.²⁵ A far bigger problem for the consistency of Hart's overall position relates to the contingency of moral values. This is a separate issue. It is an entirely tenable position that moral values are objective, yet contingent. This is a core commitment in the reductionist naturalism that we have been using to unpack Hart's meta-ethics. That moral values are 'revisable' accepts that they are objective because when true they relate to a set of facts about the world, but it also requires that they are contingent because those underlying facts are hostage to change. We have seen here that Hart accepts that moral values are revisable. This raises a particular challenge to liberalism. As I shall show in the next section, liberalism of the sort that Hart wished to endorse needs to be confident that some core moral values will hold from one generation to the next.

2 Revisable moral values and Hart's critical position

Hart comments critically on law. If there is to be consistency in Hart's position these moral claims come with the commitments attached to them in chapter IX. There are objective reasons for these moral values and as such they can be described as objectively true. Nevertheless, this truth is contingent. This contingency exposes an underlying tension between Hart's argument about the minimum content of natural law and his general liberal worldview.

A. THE HARM PRINCIPLE

Mill's harm principle is a key element in Hart's work. For Mill, 'the only purpose for which power can be rightfully exercised over any member of a civilized community, against his will, is to prevent harm to others'.²⁶ Hart endorses this approach in debate with Lord Devlin on the Wolfenden report.

Hart supports the recommendation that homosexual acts between consenting adults and conducted in private should not lead to criminal sanctions. Hart does so on the grounds that such acts do not cause demonstrable harm to others. Devlin claims that social harms would be caused by what was at the time an affront to popular positive morality. Hart demands empirical evidence.²⁷ He also argues that requiring an individual to repress their sexual orientation is likely to cause that individual pronounced psychological harm.²⁸

The harm principle is as objectively true as the claim that any good flourishing legal system will include laws against harming others if read in light of Hart's claims in chapter IX. The reasons for one provide reasons for the other. The same argument holds that these

24 Hart, *Concept* (n 3) 254. 'For practical purposes', as Hart puts it, a judge ought 'to make the best moral judgment he can on any moral issues he may have to decide' regardless of whether objective moral truths exist.

25 H L A Hart 'Who Can Tell Right From Wrong?', *New York Review of Books*, 17 July 1983, 49, at 52.

26 J S Mill, *On Liberty* (Basil Blackwell 1948) 8.

27 Hart, *Law, Liberty* (n 12) 68–9.

28 *Ibid* 21–2.

reasons, and hence this moral value, are contingently true. As Hart notes 'if men were to lose their vulnerability to each other there would vanish one obvious reason for the most characteristic of provision of law and morals: *Thou shalt not kill*'.²⁹ We can see an objective reason for the harm principle in chapter IX. This too would vanish in such an eventuality.

Hart speculates that we could, or could have, evolved so as to develop a hard exocrine skeleton that renders us impervious to any sort of physical harm. In this situation, the harm principle would be in jeopardy as a moral value, at least in relation to physical harm. In a less extreme form we can speculate as to more likely changes in the state of human vulnerability that might have a knock-on effect for the minimum content of natural law and for the harm principle as a critical moral argument about law. Respected scientific discussion is now taking place as to attainable forms of human immortality or a severe diminution in human vulnerability.³⁰ These advancements are far from imminent, yet they are sufficiently on our horizon to merit theoretical consideration.³¹ If facts surrounding human vulnerability were to change, so must the scope and shape of any moral argument about law that is based on the idea of human vulnerability under the account of reasons for moral values that Hart espouses.

Some types of harm discussed in *Law, Liberty and Morality* are distinct from the physical harm mentioned in *The Concept of Law*.³² Nevertheless, if we must hold open the possibility that our vulnerability to physical harm might change, we should also hold open the possibility that we could immunise ourselves from the sorts of psychological harm that Hart mentions in his debate with Devlin. There are corresponding medical advancements in this area too.³³

The same issue arises in relation to arguments based on scarcity of resources, approximate physical equality, limited altruism and limited understanding and strength of will. If these observable features of human nature are hostage to change, so is any moral argument about law's specific content that depends upon them. We can see how the nature and scope of these facts might change in light of recent research. Let us take the fact that resources are limited. In ongoing debates as to the benefits of genetic engineering in crops,

29 Hart, *Concept* (n 3) 195. What such an eventuality would do to man's pursuit of survival is moot. If we were physically invulnerable, it seems that our survival would be all but guaranteed and so the pursuit of this goal would lose significance. Hart does not discuss the interplay between vulnerability as a contingent truth and survival as a centrally important goal; it is beyond the remit of this essay to do so.

30 This is due to advances in nanotechnology (see R Freitas Jr, R Merkle, *Kinematic Self-Replicating Machines* (Landes Bioscience 2004)), cryonics (see B P Best 'Scientific Justification of Cryonics Practice' (2008) 11(2) *Rejuvenation Research* 493), computational neuroscience (see R Kurzweil, *The Singularity is Near: When Humans Transcend Biology* (Viking Press 2005)), and cybernetics (see K Warwick, M Gasson, B Hutt, I Goodhew, P Kyberd, H Schulzrinne and X Wu 'Thought Communication and Control: A First Step using Radiotelegraphy' (2004) 151(3) *Institution of Electrical Engineers Proceedings on Communications* 185). Hard exocrine skeletons for human use continue to be improved, see Cybernetic Inc's HAL® 5 type B.

31 See J P Dupuy 'Some Pitfalls in the Philosophical Foundations of Nanoethics' (2007) 32(3) *Journal of Medicine and Philosophy* 237; J Davis, 'Life-extension and the Malthusian Objection' (2005) 30(1) *Journal of Medicine and Philosophy* 27; C Lafontaine 'The Postmortal Condition: From the Biomedical Deconstruction of Death to the Extension of Longevity' (2009) 18(3) *Science as Culture* 297.

32 Space does not permit analysis of 'harm'. Joel Feinberg discusses its meaning in relation to the harm principle; see *The Moral Limits of the Criminal Law, Vol 1: Harm to Others* (Oxford University Press 1984) 31–7. For Feinberg, harm here means the wrongful setback of another's interest, at 36. Interest is understood as anything in which an individual has a stake, at 33–4. This will suffice as a working definition and includes psychological harm.

33 Consider the growth in antidepressant medications available and increased prescription of such medications, see 'Explaining the rise in antidepressant prescribing: a descriptive study using the general practice research database' (2009) *British Medical Journal* 339: bmj.b3999. The rise in use of antidepressants raises many issues, including the desirability of a society immune from emotional pain, see C Barber, *Comfortably Numb: How Psychiatry is Medicating a Nation* (First Vintage Books 2009).

those in favour have claimed that this has the capacity to end world hunger thus making the resource less limited.³⁴ In this situation, moral arguments about law's specific content that are premised on the idea that our resources are limited would no longer carry weight in relation to this resource. On the other hand, many resources are far scarcer than they were when Hart wrote *The Concept of Law*. In relation to these, scarcity-based arguments are likely to become more significant as critical commentary on law.³⁵

B. LIBERALISM

Hart's argument against Devlin borrows from Mill; this is far from the only example of Hart's liberalism. McCormick³⁶ and Ryan³⁷ each document the liberal underpinnings in Hart's moral critique of law, Lacey repeatedly asserts that Hart's politics were 'steadfastly liberal'³⁸ and, on numerous occasions, Hart used Mill and Bentham as starting points for critique of positive law.

Liberalism encompasses a wide variety of views. At its core it involves a tension between the vision of an open society and liberalism as a practically achievable political goal. The ideal of liberalism is based on individual freedom and value neutrality. It is for the individual to determine their own moral values and live their lives accordingly; no one person's moral outlook should trump that of another. We can see this vision in Hart's argument against Devlin; it is not for the moral majority to determine what is right in terms of sexual practices. Nevertheless, one cannot meaningfully critique law from a liberal perspective and avoid committing to ideas that ought to be protected as lasting values. Typically, liberal critique involves commitment to the protection of fundamental freedoms such as freedom of religion, freedom of expression, freedom of association and freedom of conscience.

Hart's liberalism cannot avoid this tension. Hart cites the French Declaration on the Rights of Man and the United Nations Declaration on Human Rights as the great liberal achievements in positive law.³⁹ To endorse these documents one must hold that the values contained within them are in *some* sense lasting. If the truth of all moral values were to be revisable, in the manner of the harm principle, then it seems difficult to support the fact that some are enshrined in lasting bills of rights and constitutions. Such enshrinement would be a bad idea if the general moral values articulated therein were hostage to radical change. While legislation and precedent are sources of law that we might expect to change over time with changes in underlying facts about the world, declarations on human rights, bills of rights and constitutions are different. Their very purpose is to present lasting, fundamental values that will act as guidelines for legislators far beyond our own generation, come what may in the future.

Hart's philosophy may be able to support objective moral values. This presents no problem to a reductionist naturalist. Yet Hart needs some liberal values to be lasting too if he wishes to endorse 'the great liberal achievements' or any instrument that permanently

34 See N Borlaug, 'Genetically Engineered Food Could Help End World Hunger' in J Torr (ed), *Genetic Engineering: Opposing Viewpoints* (Greenhaven Press 2000) 129–36. This claim is debatable; see B Halweil, 'Genetically Engineered Food Will Not Help End World Hunger' in *ibid* 137–46.

35 For one example of such arguments see B Saul, 'Climate Change; Resource Scarcity: Towards an International Law of Distributive Justice' in R Lyster (ed), *In the Wilds of Climate Law* (Australian Academic Press 2010) 71–94.

36 N MacCormick, *H L A Hart* (2nd edn, Stanford University Press 2008) 17–22.

37 'Hart and the Liberalism of Fear' in M Kramer, C Grant, B Colburn and A Hatzistavrou (eds), *The Legacy of H L A Hart* (Cambridge University Press 2008) 315–30.

38 Lacey (n 1) 68.

39 H L A Hart, 'Are there Any Natural Rights?' (1955) 64(2) *Philosophical Review* 175, at 89.

protects fundamental freedoms. Necessary moral truths do present an inherent problem for reductionist naturalists.⁴⁰

We can see that Hart was concerned about the possibility that moral values might not last. He accepts that there is a serious issue to address when it comes to the concerns of ordinary people who attempt to live up to moral standards 'and to transmit them to their children'.⁴¹ The expression of this concern is a hint that what really troubled Hart about moral values is that they *remain* true from one generation to the next. In *The Concept of Law*, Hart '[sought] to evade [the] philosophical difficulties' relating to the status of moral values, a desire that he re-emphasised in the postscript. In the main text, the specific instance of a philosophical difficulty mentioned is whether moral values are 'immutable principles which form part of the fabric of the Universe' or 'expressions of changing human attitudes'.⁴² This too suggests that a concern at the forefront of Hart's mind, even then, was the question of whether we can say of some moral values that they will last.

In what follows, I argue that a broadly liberal, critical morality about law that cannot be revised to the point of abandonment is possible within Hart's work. Some such moral truths seem inevitable if we robustly commit to Hart's concept of law and his reductionist naturalist account of reasons for moral values.

3 Lasting liberal values about (Hart's concept of) law

A. A TRUTH ABOUT HUMAN NATURE THAT HART CANNOT ABANDON

Railton makes the following point:

Revisionism may reach a point where it becomes more perspicacious to say that a concept has been abandoned, rather than revised. No sharp line separates tolerable revisionism and outright abandonment . . .⁴³

We are concerned with the possibility of lasting values that might be used to critique law in Hart's philosophy. Hart never suggested that his concept of law was eternal; his aim was simply to clarify the meaning of the concept 'law' as we understand it.⁴⁴ Nevertheless, for the exercise at hand we must take Hart's account of the meaning of the concept as fixed; Hart's critical moral values about law must relate to *this* account of what law is if his work is to form a coherent whole. 'Fixing' Hart's account of law in this way may puzzle some readers as this is not what Hart himself intended to do. In the context of the exercise at hand, however, it makes perfect sense. We are concerned with uniting Hart's concept of law with his critical comments about law. For these purposes, the critical comments that he makes must be about *his* concept of law rather than any other. If there are facts about human nature or the world we live in that must persist for Hart's concept of law to work, such facts may be revisable for Hart but not to the point of abandonment. Otherwise, Hart's concept of law itself would need to be abandoned as would all of his critical comments about that concept.

This point has an important consequence if we take Hart's reductionist naturalism about moral values seriously. If a certain fact about human nature cannot be abandoned in Hart's legal philosophy, this fact may provide a reductive basis for objective moral values

40 Railton, 'Moral Realism' (n 11) 198–200 and 'Naturalism' (n 13) 158–66.

41 Hart 'Who Can Tell' (n 25) 52.

42 Hart, *Concept* (n 3) 168.

43 Railton 'Naturalism' (n 13) 159.

44 Hart does not prescribe rules for the valid use of the term 'law'. As such, he does not define it. Instead he tries to capture the meaning of this concept as we understand it: Hart, *Concept* (n 3) 213; Hart's private notes as reproduced in Lacey (n 1) 222–27; and D Priel 'Trouble for Legal Positivism' (2006) 12(3) *Legal Theory* 225.

about law that cannot be abandoned either. Hart never provided an account of human nature itself to do this work for us, but he did, of course, provide a detailed account of law. Let us consider a simple underlying fact about human nature that Hart's explanation of the concept 'law' requires.

Hart clarifies how law is not simply a set of commands backed by sanctions, but is instead a set of guides for behaviour.⁴⁵ These may be guides for private citizens in terms of obeying criminal law, guides for conduct in private affairs such as making a valid will or contract, or guides directed to officials as to how to conduct themselves in an official capacity. The common point is that legal rules, like any rules, are guides for human beings as to how to behave in various circumstances.

Under this understanding, most human beings, most of the time, must be capable of guiding their own behaviour. There may be *examples* of people who cannot do so or whose ability is limited in this regard. We frequently talk of those who do not have legal capacity, largely on the basis that their ability to control their behaviour is impaired in some way. There may also be instances in which this capacity to guide behaviour has been compromised, for example, if an individual is suffering from an epileptic convulsion. Hart's empirical approach to reasons for moral values can easily support the idea that an exception should be made where a non-moral fact such as evidence of diminished capacity is empirically observable. Similarly it may be the case that we will have to revise our notions of capacity and our claims as to what the substantive content of law ought to be as we discover more about things like brain function. Yet a revision to the point of abandonment of the general idea that human beings have a capacity to guide their own behaviour would require Hart to also abandon his concept of law. So long as we are interested in critical moral values in relation to what Hart calls law, a capacity on the part of humans to guide their own behaviour must remain in some way. Most elect to conform to rules but there are exceptions – a minority choose to commit criminal acts while most citizens avoid doing so. If the observable facts of social behaviour were different, if we were a society of angels (as suggested by Raz⁴⁶) who conscientiously and deliberately all adhere to the law, members of such a society are capable of guiding their own behaviour albeit that they seem likely to always choose to do so in accordance with criminal law. In a possible set of social facts that Hart imagines – a society of metaphorical 'sheep' that follows rules blindly without very much in the way of reflection – we would still be capable of choosing to behave in a way that conforms to legal rules even if we did so in an unthinking fashion.⁴⁷ Hart's account of the concept of law *could not* work if the observable facts of the matter were such that all human beings were automata. If most human beings most of the time could not be said to self-govern in terms of their behaviour, then 'guides for human behaviour' would be meaningless. There could be no law for a society of robots which have been preprogrammed to act in a set number of ways, or a society of creatures which are enslaved to instincts that they cannot control.

Hart comes close to stating that, wherever there is law, we will observe this fact:

[A legal system is] dependent for its efficacy on the possession by a sufficient number of those whose conduct it seeks to control of the capacities of understanding and control of conduct which constitute capacity-responsibility.

45 Hart, *Concept* (n 3) 28–9, 89–91, 98, 115–17, 124–5.

46 J Raz, *Practical Reason and Norms* (Hutchinson & Co 1975) 159.

47 Hart, *Concept* (n 3) 117. See also 'Negligence, Mens Rea and Criminal Responsibility' in *Punishment and Responsibility* (n 15) 136–57. It counts as rule-following even if one has not exercised thought. Similarly one might fail to follow a rule because one has failed to think. This failure may be morally blameworthy.

For if a large proportion of those concerned could not . . . form and keep a decision to obey, no legal system could come into existence or continue to exist.⁴⁸

So there is *at least* one feature of human nature that we can be guaranteed to observe if Hart's account of the concept of law is to work.⁴⁹ I shall now demonstrate how some of the values that Hart embraces in his critical commentaries on law depend upon this feature of human nature. For each, I hypothesise as to how these claims and the underlying non-moral truth could be turned into a critical moral value about law along reductionist naturalist lines. Each of these values fits with a general liberal outlook.

This exercise comes with two important caveats.

The list is non-exhaustive. My argument does not deny the possibility that there may be other non-moral facts that simply must be true if Hart's concept of law is to hold.⁵⁰

What follows are mere outlines of moral values. Space does not permit a full defence of these values and that is not the aim of this essay. I merely illustrate how Hart's meta-ethics can be used to generate moral values that cannot be abandoned so long as one commits to his concept of law.

B. THE DOCTRINE OF FAIR OPPORTUNITY

Hart repeatedly embraced what he terms the 'doctrine of fair opportunity'; law ought not to impose punishment on an individual unless that individual has had a relatively unimpeded opportunity to make their behaviour conform to law. A good example is Hart's argument against Lady Wootton. Wootton argued that the issue of *mens rea* in criminal law should only be considered at the point of conviction, in order to determine the most socially useful outcome. The basis for this argument was a presumed link between the doctrine of *mens rea* and traditional retributivist theories of punishment. Wootton assumed that looking into *mens rea* at the point at which we determine guilt, and therefore looking into the defendant's past mental state rather than his or her present one, implies that sentencing is a means of achieving retribution for past wickedness. Hart makes the following point:

. . . a primary vindication of the principle of responsibility could rest on the simple idea that unless a man has the capacity and a fair opportunity or chance to adjust his behaviour to the law its penalties ought not to be applied . . . Such a doctrine of fair opportunity would not only provide a rationale for most of the existing excuses which the law admits in its doctrine of *mens rea* but *it could also function as a critical principle to demand more from the law than it gives.* [emphasis added]⁵¹

48 Hart, *Punishment and Responsibility* (n 15) 229.

49 Many would contend that law amounts to more than rules. My discussion does not exclude this possibility. The point is that at bare minimum there must be some human capacity for behaviour guidance for the understanding of law presented in *The Concept of Law*.

50 I have selected a non-moral fact that seems uncontroversial and that Hart mentions. Other facts that Hart might not be able to abandon include abilities to communicate, understand and reason. A Freudian might suggest that rule-following requires an ego; unfettered id will not provide a mechanism to follow rules other than the pleasure principle. See S Freud, *The Ego and the Id*, J Riviere and J Strachey (trans) (Hogarth 1962) and *New Introductory Lectures on Psychoanalysis*, J Riviere, J Strachey and A Richards (trans) (Penguin 1991) 102–12. Hart alludes to the possibility that psychological truths may be necessary for rule-following, *Concept* (n 3) 193. Any of these facts could be coupled with a non-abandonable good and used as a reduction basis for moral values. This is to say nothing of the range of values that might emerge if we were to follow Epstein in pressing Hart beyond his original parameters based on survival to include the maximisation of social welfare as a general human aim, see R A Epstein 'The Not So Minimum Content of Natural Law' (2005) 25 Oxford Journal of Legal Studies 219.

51 See Hart, *Punishment and Responsibility* (n 15) 158–85, at 181.

So positive law only ought to penalise individuals in situations in which they have been able to exercise free choice.⁵² This is a critical moral value to guide existing legal systems.

We can see a similar commitment to fair opportunity in Hart's debate with Devlin. Hart distinguishes the repression of sexual impulses from the repression of other desires.⁵³ One may be tempted to steal, but unless one happens to be a kleptomaniac, one can freely choose to repress this desire. One might repress one's sexual impulses by not acting upon them, but the impulse itself is 'a recurrent and insistent part of daily life'; it is not something that one can exercise control over.⁵⁴ Criminalising sexuality causes a specific type of harm *because* one cannot choose one's sexual orientation in the same way that one can choose to steal or refrain from doing so. The human capacity to control one's own behaviour plays an important role in this argument. This is, again, a critical moral argument about the specific content of law that depends upon the doctrine of fair opportunity.

It is possible to provide an objective reason for the doctrine of fair opportunity by combining the fact that individuals are by and large capable of guiding their behaviour with a non-moral good. For Hart, the 'general justifying aim' of punishment is to reduce crime.⁵⁵ It is consistent with this good that we should only punish individuals in instances where they have been able to govern their behaviour. If the good to be achieved is deterrent effect, those subject to law can only be deterred if they had opportunity to guide their behaviour one way or another. Nevertheless, the good in this reduction basis is one that we could revise, even to the point of abandonment, without abandoning Hart's concept of law. If the acts that we deem criminal were to cease occurring, there would be no need to deter people from them and thus no good to be achieved. In the hypothetical society of angels posited earlier, there may be no need for punishment at all. Hart's concept of law could still operate in such a society; there would still be a need for laws to govern contracts and wills. If we wish to establish a moral value that we cannot abandon so long as Hart's concept of law holds, we must look elsewhere.

Arguments that deny the idea that human beings are by and large capable of guiding their behaviour are another matter. If an argument were to proceed from the basis that human beings are generally *incapable* of behaviour guidance, it would be false, provided that we commit to Hart's concept of law and the idea that the truth of moral values depends upon non-moral facts. This falsehood is one that we could not abandon. To be by and large subject to law, is to be by and large capable of guiding one's behaviour.

Similar arguments have been made in instances where a specific group has been oppressed. In the movement towards emancipation some have occasionally, and egregiously, suggested that these groups should not be afforded certain rights on the grounds that the

52 For further examples of Hart's commitment to this doctrine see his discussion of strict liability in 'Acts of Will and Responsibility', *Punishment and Responsibility* (n 15) 90–112; the arguments in 'Intention and Punishment' on the capacity for punishment to 'goad' behaviour, *ibid.* 113–35, especially 134–35; the general argument in 'Punishment and the Elimination of Responsibility', *ibid.* 158–85 of which the Wootten debate forms a part; and various comments in 'Changing Conceptions of Responsibility', *ibid.* 186–209.

53 Hart, *Law, Liberty* (n 12) 21–2.

54 Lord Hope makes a similar point in *HJ (Iran) and HT (Cameroon) v Secretary of State for the Home Department* [2010] UKSC 31, paras 11, 14 and 21. The involuntary nature of an individual's sexuality formed part of a discussion as to the viability of asylum-seekers relocating within their home state and concealing their sexuality.

55 Hart's position combines Utilitarian and Retributivist elements. Hart's justification for punishment is the societal benefit in crime reduction; in this respect he is Utilitarian. Hart holds that a Retributivist element, punishing a specific individual for wicked acts, is important in achieving that aim. See 'Prolegomenon to the Principles of Punishment', and 'Postscript: Responsibility and Retribution' in Hart, *Punishment and Responsibility* (n 15) 8–13 and 230–7.

oppressed group is *generally* incapable of exercising rational choice with regard to its actions. Such arguments have been made even though the group in question was subject to other legal rules. This was one of a number of arguments put forth by those that opposed women's suffrage. Consider the following claim by James McGrigor Allan:

Woman never escapes from male control, direct or indirect, personal or impersonal, traditional or present. She is always ruled by some man, either living or governing from the grave . . . she embodies her ideal of masculine superiority in some man whose teachings . . . she accepts with implicit reverence, making him to all intents and purposes an infallible judge, from whose decision there is no appeal.⁵⁶

As far as Allan is concerned, women are always under male 'control'; affording women the right to vote is thus futile. This argument is not only wrongheaded but deeply offensive to modern minds. If we draw out the implications in Hart's work, arguments of this sort are objectively false. It is nonsensical for Allan to make this claim yet accept that women have capacity when it comes to entering a contract or committing criminal acts.⁵⁷ If it were the case that a woman's nature is to 'never escape male control', then no woman should be liable for failing to adhere to criminal law or bound by any contract. There could be no meeting of minds for a valid contract as the woman will have been controlled by some third party. There could be no *mens rea* for criminal liability as the accused's intentions would not be her own. Law can defensibly acknowledge that an individual might have capacity to perform certain acts and not others.⁵⁸ Law also can, and frequently does, recognise instances of undue influence by one person over another. Allan's claim about the nature of woman goes much further by stating that a woman is 'always ruled by some man'. If this *were* the case, then there could have been no law for women at the time as women could not individually guide their own behaviour in accordance with legal rules.

If we commit to Hart's account of law, this aspect of Allan's moral argument is false; furthermore we cannot abandon the belief that this argument is false. As a critical moral value that cannot be abandoned, we can say that no group that is generally subject to law should be denied a particular right, privilege or protection in law on the basis that the individuals making up that group cannot by and large guide their own behaviour.⁵⁹

C. THE PROMOTION OF HUMAN INDIVIDUALITY

The wrongness of Allan's assertion is based on inconsistency. Under Hart's concept of law, Allan cannot assert that a particular group is incapable of guiding its own behaviour at all

56 *Woman Suffrage Wrong in Principle and Practice* (Remington & Co 1890) 125–6. The judiciary was not immune to such attitudes; see *Chorlton v Lings* (1868) LR 4 CP 374, at 388, Bovill CJ accepted 'that fickleness of judgment and liability to influence' might have been rational grounds for the historic denial of women's suffrage.

57 Others made much of the fact that women were, at the time, not considered fully legally liable in support of arguments against women's suffrage. See the myopic writings of E B Bax, *The Legal Subjection of Men* (The New Age Press 1908) and *The Fraud of Feminism* (Grant Richards 1913). Bax argued that since women were afforded certain privileges they ought not to be afforded equality. Bax's arguments are as weak as Allan's, but he does not make the specific claim that I highlight.

58 Hart deals with this; see the doctrine of fair opportunity, discussed above and Hart's postscript to *Punishment and Responsibility* (n 15) 227–30. Many nineteenth-century arguments against women's suffrage considered women to have capacity in *some* respects but not others. This attitude was clothed in effusive, but horribly essentialist, language about feminine virtues as distinct from male ones; see J Bridgman and S Millns, *Feminist Perspectives on Law: Law's Engagement with the Female Body* (Sweet & Maxwell 1998) 11–27.

59 Certain groups are not fully subject to law's rules. We do not afford full rights or legal capacity to children and other individuals lack capacity. The concept of law which Hart gives us would need to be abandoned if these exceptional cases were to make up the populace generally as nobody subject to law would be capable of being bound by rules or capable of exercising rights.

and yet consider that group to have legal capacity. This is not yet an account of a moral value that cannot be abandoned along the lines of Hart's chapter IX analysis. Rather than a moral truth, it is a moral falsehood. Furthermore, for a moral truth of this sort to develop, we need a non-moral human aim that cannot be revised to the point of abandonment.

In *Law, Liberty and Morality* Hart also makes the following claim:

The unimpeded exercise by individuals of free choice may be held a value in itself with which it is *prima facie* wrong to interfere; or it may be thought valuable because it enables individuals to experiment – even with living – and to discover things valuable both to themselves and others.⁶⁰

Hart's language is tentative. He does not argue strongly for this value. Mill does. Mill argues that experimentation in the exercise of individual choice is to be encouraged; it is only through such experimentation that mankind can evolve and learn. The development of the individual character is, for Mill, 'the chief ingredient of individual *and* social progress' [emphasis added].⁶¹ Allowing 'originality in thought and action' is good for the state as a whole. It is through such freedom that 'new truths' can be discovered and criticism of existing attitudes develops. Like Hart, Mill holds that custom is at the heart of our moral and legal rules. Mill, however, warns that:

[t]he despotism of custom is everywhere the standing hindrance to human advancement, being in unceasing antagonism to that disposition to aim at something better than customary . . .⁶²

Hart cannot abandon the fact that moral values are revisable, without giving up his account of moral values. Human beings need some way by which they can suggest improvements to substantive law that would meet the requirements of changes to these facts. Access is needed to a way of thinking that departs from custom if the laws and moral values that custom presents no longer help to achieve a specific good or goods. As a result, critical perspectives on law are non-moral 'goods' for law that cannot be abandoned.

Under reductivist naturalism, the primary means through which we revise our ideas about what is good for us or identify that existing ideas might require revision is through experimentation.⁶³ This suggestion fits well with Hart's claims about moral values generally. Such values do not submit to discoverability as metaphysical entities above and beyond our day-to-day living, in the manner that a theist might try to discover the will of God. Unavailable too is a teleological approach or any form of intuitionism.⁶⁴ Yet experimentation in thought and action of the sort described by Mill is suited to the revisability of moral values based on empirical input that we see in reductivist naturalism. Freedom of choice by individuals is a value that law ought to encourage because it is a 'good' for law that individuals have a capacity and an opportunity to experiment with alternatives to traditional approaches. If custom is to be critiqued rather than enjoy the despotic rule that Mill feared, individuality in thought and action is needed. Hart might

⁶⁰ Hart, *Law, Liberty and Morality* (n 12) 21.

⁶¹ Mill (n 26) 50.

⁶² Ibid 62.

⁶³ For Railton we adapt moral values to meet changes in the reductive basis through feedback gleaned from 'trial and error', 'Moral Realism' (n 11) 179–82; and 'Naturalism' (n 13) 152–8.

⁶⁴ As noted, Hart rejects 'metaphysical' and teleological approaches, *Concept* (n 3) 185–94. While Hart felt that intuitionist explanations of moral value are worthy of discussion, Hart 'Morality and Reality' (n 14) 37, intuitionism would fit poorly with the account in chapter IX.

accept that law could exist among a society of metaphorical sheep, but it seems impossible for any sort of critical morality to do so if we take his claims about moral values seriously.⁶⁵ To accept that critical morality is a non-moral good for law is to also accept that independent thought and the implementation of that thought into lifestyle choices is a non-moral good for law given Hart's other commitments.⁶⁶

The *degree* to which law should encourage experimentation might vary. This good is revisable on the basis of other contingent matters of fact. For example, there are powerful reasons why we discourage experimentation with hard drugs, many of which have to do with the harm that such experimentation can cause. Yet Hart cannot utterly abandon the value at stake. Human beings have a capacity to guide their own behaviour; they can slavishly follow custom, or they can be thoughtful and questioning towards it. There are many ways in which law's substantive content might be used to foster the latter attitude. One might argue that this good is achieved in Western liberal democracies through the protection of fundamental freedoms such as freedom of expression, freedoms of religion, belief and opinion and freedom to experiment sexually as alluded to by Hart. Such justification of fundamental freedoms is consistent with the liberal position that Hart endorsed as espoused by Mill and others.⁶⁷ The moral value behind this argument is one that can be revised but not to the point of abandonment.⁶⁸

D. HART'S THEORY OF MORAL RIGHTS

Liberal writers such as John Locke have taken the rights of man to be foundational properties that persist regardless of era. For Locke, rights are inalienable and a fundamental part of what it means to be human. This approach is not open to Hart, although rights are certainly important to him. In addition to his endorsement of rights-based constitutional democracy, he dedicates a significant amount of space to the meaning of a legally respected right and the related issue of moral rights as a tool for the evaluation of substantive law.⁶⁹

As Bayles notes, 'no [single] paper represents [Hart's] definitive views [on rights], which have changed over time'.⁷⁰ In Hart's early position, we can see a reductionist tendency. 'Are There Any Natural Rights?' concerns the nature of rights generally rather than specifically legal rights. Hart argues that 'if there are any moral rights at all, it follows that there is *at least* one natural right, the equal right of all men to be free'.⁷¹ This is not reductionist naturalism as the 'one natural right' is not an empirically observable, non-moral fact. Furthermore, the claim that there is at least one natural right suggests that there may be more. Yet, the argument as it developed reduced other rights to this general and equal right to freedom.⁷² The nearest thing to the approach that Hart would later adopt in *The Concept of Law* is a non-moral condition that Hart adds to the existence of this natural right:

65 Mill uses the same analogy to argue that such a society is undesirable (n 26) 60.

66 Railton alludes to this in relation to law and critique of law, 'Moral Realism' (n 11) 207, fn 42.

67 In addition to Mill, J Milton's *Areopagitica* (Deighton, Bell & Co 1973) contains a defence of press freedom similar to the one suggested here. See also the dissenting opinion of O W Holmes in *Abrams v United States* 250 US 616 (1919).

68 This does not render the principle more important than others; in order to work the contingently true harm principle should take priority in many instances. No hierarchy is implied. The point is that the value persists regardless of changes in contingent matters of fact.

69 See H L A Hart, 'Legal Rights' in *Essays on Bentham* (Clarendon 1982) 162–93, especially 190–93.

70 M Bayles, *Hart's Legal Philosophy: An Examination* (Kluwer Academic Publishers 1992) 141.

71 Hart 'Are there Any Natural Rights?' (n 39) 175.

72 This position is closer to non-reductive naturalism; the existence of a moral value must be accepted in order to explain other things. See N Sturgeon, 'Moral Explanations' in G Sayre-McCord (ed), *Essays on Moral Realism* (Cornell University Press 1988) 229–55.

This right is one which all men have if they are capable of choice; they have it qua men and not only if they are members of some society or stand in some special relation to each other.⁷³

Although Hart's concept of law requires that human beings generally are capable of this sort of choice,⁷⁴ it does *not* follow from the above statement that all human beings have an equal moral right to freedom wherever we commit to Hart's concept of law. Hart's claim about the 'one natural right' comes with the caveat 'if there are any moral rights at all'. The concept of a 'right', moral or otherwise, does not *necessarily* exist.

Hart dissociated himself from this position.⁷⁵ Although Hart did not provide specific reasons for this, it would have been difficult to reconcile the belief in a *natural* right with the anti-metaphysical stance that he later adopted in *The Concept of Law*. Hart's ultimate position is that a general theory of legal or moral rights is impossible. A large part of what we call rights can be explained through choice theory; the view that legal and moral rights are forms of protected individual choice. Hart maintained that 'ordinary legal rights' fit best with this account. According to Hart, fundamental rights that guarantee protections for the individual against their own state cannot be explained in this way. Hart referred to these as 'immunity rights'.⁷⁶

Immunity rights limit '[the legislature's] powers to make (or unmake) the ordinary law, where to do so would deny to individuals certain freedoms and benefits'.⁷⁷ Immunity rights include security of life and person, education, equality of treatment and freedom from arbitrary arrest.⁷⁸ Crucially for the purposes of this investigation, Hart notes that his analysis of rights has relevance beyond those that form part of existing law. The notion of an immunity right helps to describe existing legal rights. It also provides a means of critiquing law. As Hart notes 'law . . . is too important a thing to leave to lawyers';⁷⁹ some fundamental rights used to critique existing law are also based in the notion of 'immunity' rather than 'choice'.

The moral right of individuals to have certain immunities from the state is based on 'essentials of human well being'.⁸⁰ In chapter IX Hart linked the reasons for these sorts of moral values to contingent facts about human nature. By way of illustration, rights to security of life and person are immunity rights according to Hart.⁸¹ Their focus is to guarantee that individuals will not be subjected to certain types of harm or degradation. This type of right

73 Hart 'Are there Any Natural Rights?' (n 39) 175.

74 'Choice' and 'behaviour guidance' are not synonymous. Hart discusses the right to behave as one wishes provided that one is capable of exercising choice in the matter. In what follows on 'choice-based rights', Hart specifically refers to a right to act in accordance with choice, rather than a right to simply will an outcome. Reference to 'choice' should be taken to read 'choice where this manifests itself in behaviour'.

75 Hart, *Essays* (n 12) 17.

76 Hart, 'Legal Rights' (n 69) 190.

77 Ibid.

78 Ibid. The right to a fair trial, the right to privacy and the right to compensation for miscarriages of justice among others are also likely immunity rights. These examples are from the European Convention on Human Rights, see D J Harris, M O'Boyle and C Warbrick, *Law of the European Convention on Human Rights* (2nd edn, Oxford University Press 2009) 201–330, 361–424, 557–576, 750.

79 Hart, 'Legal Rights' (n 69) 192.

80 Ibid 190.

81 See United Nations Declaration on the Protection of All Persons from Being Subjected to Torture and Other Cruel, Inhuman or Degrading Treatment or Punishment, and Article 3 of the European Convention on Human Rights.

informs moral debate on the death penalty⁸² and the use of interrogation techniques, including torture.⁸³ For Hart, that human beings are physically vulnerable is a contingent truth. If this underlying fact were to cease being true, this would have serious, potentially fatal, consequences for moral arguments that security of person ought to be protected.⁸⁴

We could reduce many of these immunity rights to the sorts of non-moral, contingent facts about human nature that Hart identifies in chapter IX and non-moral goods such as survival or flourishing. That these immunity rights ought to be protected may be an objective moral truth, but it is revisable to the point of abandonment.

The non-moral fact that underlies choice-based rights is one that Hart cannot revise to the point of abandonment, on pain of giving up his concept of law.⁸⁵ Certain rights are reducible to 'choices'. If it were to be empirically observable that human beings cannot guide their own behaviour, then we would not only have to give up choice-based rights and any argument based on them for critiquing law, we would also have to give up Hart's concept of law itself. Nevertheless, two problems arise if we wish to turn the protection of choice-based rights into permanent moral values.

The first problem is that the mere existence of a human capacity to exercise choice that underlies choice-based rights will not turn into a proposition that we ought to respect such rights unless there is some non-moral good to be achieved by doing so. It may be possible to address this by coupling choice-based rights with an argument made earlier. It is a good, wherever there is law, for individuality in human thought and action to be encouraged, the argument might go. Choice-based rights give effect to human choices in thought and action and so, the argument would continue, choice-based rights ought to be protected as a moral value that we cannot revise to the point of abandonment. One might extend this argument to certain immunity rights on the basis that these also help to foster and encourage individuality through experimentation in thought and action.

The second problem concerns the very notion of 'rights'. Rights are useful tools that we employ in moral argument and in law. Yet, there would not appear to be any non-moral fact that Hart is committed to that would *require* the existence of a right, rather than some other tool, as a means of promoting this good.

What we can say is that the starting point for a system of choice-based moral rights exists wherever there is law. Whether such a system grows and flourishes is a matter of contingent fact, but the basis for this system simply must be there if Hart's concept of law holds. As a result, whenever we talk of moral rights that citizens have against other citizens we are talking about something that has its basis in a truth about human nature that we cannot abandon if we accept Hart's account of the concept of law. So long as we accept the idea of rights at all, there is the possibility for a lasting moral truth in the notion that choice-based rights ought to be protected as should those immunity rights that help to foster a spirit of critical reflection.

82 See W Schabas, *The Death Penalty as Cruel Treatment and Torture: Capital Punishment Challenged in the World's Courts* (Northeastern University Press 1996).

83 See, K J Greenberg (ed), *The Torture Debate in America* (Cambridge University Press 2006) and S Levinson (ed), *Torture: A Collection* (Oxford University Press 2004); in particular H Shue, 'Torture', *ibid* 47–60; A Dershowitz 'Tortured Reasoning', *ibid* 257–80; and E Scarry, 'Five Errors in the Reasoning of Alan Dershowitz', *ibid* 281–90.

84 Some argue that choice theory *can* account for 'immunity' rights. See P Graham, 'The Will Theory of Rights: A Defence' (1996) 15(3) *Law and Philosophy* 257; and H Steiner 'Are There *Still* Any Natural Rights?' in Kramer et al (n 37) 239–250.

85 As with moral values, no hierarchy is created between different types of rights as a result.

4 Conclusion: Hart's liberalism and non-evaluative description

The reading of Hart's work presented here places his claims about reasons for moral values to the fore. Read this way, it is possible for Hart's philosophy to support lasting, moral values about law of the sort that his liberal politics demands.

The methodology employed has built upon Hart's own claims about reasons for moral values. Nevertheless, the result rests uneasily alongside Hart's desire to distinguish descriptive accounts of law from claims as to what law's substantive content ought to be.

By linking the reasons behind moral values to matters of fact and human aims, as Hart does in chapter IX, questions about what law is and reasons for moral values become inextricably linked. Any conceptual analysis of law will make tacit assumptions about human nature and the world we live in; some of these facts must remain true if the analysis is to remain accurate. If one accepts Hart's account of reasons for moral values, these facts are likely to bring with them a number of objective reasons as to why the specific content of law ought to adhere to certain values. Some tools for what law ought to be are furnished by the account of the concept.

In reading Hart's works together, we are left with a choice. Our first option is to take his claims about reasons for moral values seriously. Under this approach we can link those values to a defensible liberal position, but this comes at a price; we must downplay or ignore the idea that descriptive accounts of law belong in a different realm to evaluative claims about law's content. Our second option is to strictly adhere to Hart's distinction between conceptual analysis of law and evaluative perspectives on law. This option requires us to disregard the reasons that he gives for moral values and leaves the core liberal tension identified earlier unresolved.

Each reading involves prioritising some aspects of Hart's work over others. For those that regard it as important to identify Hart as a positivist, it should be noted that the reading presented here does not prevent us from doing so. As John Gardner has pointed out, a commitment to positivism does not require the denial of a connection between law and morality, even a necessary one.⁸⁶ There is nothing in this reading that suggests the validity of a legal norm is anything other than its source. Faced with the choice of disregarding Hart's account of rights or his distinction between description and evaluation, there is more to be gained by taking the account of moral values seriously. The reading that I have presented allows us to contextualise Hart's work as an important part of a liberal tradition Hart wished to join. The alternative not only requires us to read his analysis of the concept of law as an enterprise distinct from his critical perspective, it requires us to accept a deep conflict between the claims in chapter IX about contingency of moral principles and his critical need for lasting moral values.

⁸⁶ John Gardner '5½ Myths about Legal Positivism' (2001) 46 *American Journal of Jurisprudence* 199, at 222–25.

The problem of, and with, financial crime

GEORGE GILLIGAN¹

University of New South Wales, Australia

Abstract

Financial crime is a term that is widely used, but it is a label or category that is bedevilled by definitional uncertainty and this uncertainty impacts upon how it is perceived and acted upon by law enforcement and other regulatory actors. This is perhaps not surprising and echoes many of the difficulties that have plagued efforts to counter white-collar crime. This article considers the definitional and other ambiguities that have permeated debates about both white-collar and financial crime. The analysis draws on a short survey which asked law enforcement and other regulatory actors in Australia and the UK whose responsibilities included countering behaviours that could be viewed as financial crime, what operational definitions of financial crime they employed in the course of their work. Results indicate that definitional uncertainty ensures that there are numerous understandings of what constitutes financial crime and no immediate prospect of a universal legal definition. However, there are some interesting classification developments for financial crime emerging from the business sciences literature and interdisciplinary approaches would seem to offer the most promise for categorising the suite of evolving behaviours that comprise financial crime.

White-collar crime and problems of definition

In 2012 as the world grapples with the ongoing fallout from the global financial crisis (GFC) – four years and counting so far – the grim prospects of sovereign default and a double-dip recession remain and there is an increased focus on the massive harms that can be wrought by white-collar crime, especially in the financial sector. That increased focus has been paralleled by heavy sentencing imposed upon some of those who have been convicted of what is increasingly referred to as *financial crime*. Perhaps the most well-known example of this is the 150-year term of imprisonment handed to Bernie Madoff by a New York court in June 2009 after he had pleaded guilty to 11 federal felonies he had committed via his wealth management business, which the court-appointed trustee estimated had cost his clients more than US\$18bn. This huge sentence is seen by some commentators as signifying a new post-GFC punitiveness towards financial crime, especially if the offenders are seen as high-ranking financial actors.² However, as with white-collar crime, there is a murkiness

1 George Gilligan, Senior Research Fellow, Centre for Law, Markets and Regulation, Faculty of Law, University of New South Wales; and Senior Research Fellow, Centre for Corporate Law and Securities Regulation, Melbourne Law School, University of Melbourne, Australia. His research interests centre on governance and regulatory theory and practice, especially in relation to financial services and markets. Email address: george.gilligan@unsw.edu.au.

2 C Hurt, 'Evil Has a New Name (and a New Narrative): Bernard Madoff' [2009] (4) Michigan State Law Review 947.

permeating what financial crime is, how it should be classified, how it should be measured, how it should be countered and how it should be punished. As the GFC repercussions continue it is informative to consider how the structural dilemmas associated with financial crime parallel (perhaps unsurprisingly), those that have clouded the white-collar crime discourse.

In the first decade of the twentieth century the American sociologist and eugenicist Edward Ross wrote of the *criminaloids*, whom he saw as wealthy and influential business people who cultivated public persona of respectability in order to conceal their manipulative and criminal personalities.³ However it was Edwin Sutherland who introduced the term *white-collar crime* in the 1940s. Sutherland's research was in many ways a research report on the criminal behaviour of 70 top US corporations, with many violations stretching back to the mid-nineteenth century.⁴ Nevertheless what has come to be generally understood as white-collar crime has been recognised as a social problem for much longer than a century. It seems to be 'an inevitable concomitant of business, trading and commerce', for example, forgery and counterfeiting posed regulatory problems for Roman and Byzantine administrations almost two thousand years ago.⁵ There are long-established traditions to counter market manipulation that can be traced back to the doctrine of *just price*, which was first adopted during the later Roman Empire by the Emperor Diocletian.⁶ The doctrine of just price held that markets had to be organised in certain ways in order to be just, and so it legitimised interference in the institutional structures of markets to achieve that aim and was backed by the ideology of church law. The common law offences of engrossing, regrating and forestalling were a product of this ideology, and were intended to prevent manipulation of markets in essential goods. It is important to remember this tradition of legitimated interference in markets when trying to understand how white-collar crime, and more latterly, financial crime, have come to be recognised and countered under more modern legal processes.

The label white-collar crime was introduced by Sutherland in the 1940s because he aimed to foster an integrated analysis of 'crime in the upper, or white-collar class'.⁷ He hoped that building a bridge between the economic and sociological disciplines would provide an improved theoretical framework within which to propagate his strong moral convictions about commercial, political and professional wrongdoing. Sutherland's long-held interest in the 'occupational crimes of persons of respectability and high social status' had been evident from research he undertook in the 1920s.⁸ Sutherland believed that higher social status not only facilitated differential implementation of the law, but also offered greater opportunities to commit crime. Sutherland sought to raise, at a very practical level, a greater awareness about the harms caused by those whom he saw as white-collar criminals. However, he also wanted to challenge established criminological theory because he saw white-collar criminals escaping the great criminological sampling net, as well as escaping prison and other sanctions of the legal process. Sutherland perceived the bases of criminological theory and sampling to be hopelessly biased along class lines, and its personal and social pathology as insupportable and unworkable.⁹

3 E A Ross, *Sin and Society: An analysis of Latter-Day Iniquity* (Mifflin 1907).

4 E Gross, 'Organisations as Criminal Actors' in P Wilson and J Braithwaite (eds), *Two Faces of Deviance: Crimes of the Powerless and Powerful* (University of Queensland Press 1978) 198–213, at 204.

5 H Croall, *White-Collar Crime* (Open University Press 1992) 8.

6 B M Mitnick, *The Political Economy of Regulation* (Columbia University Press 1980) 243.

7 E H Sutherland, 'White-Collar Criminality' (1940) 5 *American Sociological Review* 1–12, at 1.

8 K Schuessler, *Edwin H Sutherland On Analyzing Crime* (University of Chicago Press 1973) xix.

9 W G O Carson, *The Challenge of White-Collar Crime* (La Trobe University Press 1983) 2–3.

Many of the theoretical problems associated with the label white-collar crime stem from Sutherland's faith in differential association as an adequate general explanatory framework for all crime, not merely white-collar crime. Under differential association theory, Sutherland asserts that criminal behaviour is learned like any other behaviour and that many types of business career virtually necessitate some form of harmful or criminal activity.¹⁰ Sutherland was unable to adequately substantiate these assertions, so that even though his ideas on the differential implementation of legal processes upon high-status offenders are now widely accepted, differential association theory is not. Sutherland's faith in differential association theory meant that he was not rigorous in his attempts at definition of white-collar crime. In his initial 1940 paper his attempt at definition merited only a footnote.¹¹ In later works Sutherland returned intermittently to issues of definition, but he was unable to produce a universal standard which could gain wide acceptance.¹² As a consequence, problems associated with definition have weakened the legitimacy of white-collar criminology within the discipline by contributing to:

[A] continuing absence of conceptual clarity and consensus in regard to white-collar crime which has tended to stifle creative and replicative research. Paradoxically, this confusion was built into the concept of white-collar crime by Sutherland. By failing to specify precisely and fully what it was that he was concerned with, Sutherland left the door wide open for a barrage of speculative attempts to refine white-collar crime. If the goal were to stimulate definitional debate, Sutherland's rather nebulous definitions of white-collar crime were marvellously attuned to the task. If the aim were to create an aesthetically satisfying typology of mutually exclusive and homogenous categories of white-collar crime, then Sutherland clearly fell short of his goal.¹³

So, it can be seen that white-collar crime is something of 'a taxonomic zoo'.¹⁴ It is a jumbled concept which relies on a rather 'spurious correlation between role-specific norms and the characteristics of the occupants of these roles'.¹⁵ Sutherland himself acknowledged in 1946, in correspondence to one of his sternest critics, Paul Tappan, that 'the concept of white-collar crime is questionable in certain respects and I hope to elaborate on these in a later publication'.¹⁶ Sadly Sutherland never did elaborate on this matter before his death in 1950, but although his theorising bears inherent flaws, it has not prevented his work having influence:

The Sutherland legacy is not easily cast aside. The concept of white-collar crime is polemically powerful and, notwithstanding considerable imprecision, palpably self-evident.¹⁷

Nevertheless, since the early days of the white-collar crime debate there has been criticism of Sutherland's inconsistent theorising. Paul Tappan attacked the 'swampy dogma' of the white-collar crime concept, which depended on *loose, doctrinaire, invective descriptors*, that

10 J W Coleman, 'The Theory of White-Collar Crime, from Sutherland to the 1990s' in K Schlegel and D Weisburd (eds), *White-Collar Crime Reconsidered* (Northeastern University Press 1992) 53–77, at 54.

11 G Geis and C Goff, 'Edwin H Sutherland: A Biographical and Analytical Commentary' in G Geis (ed), *On White-Collar Crime* (Heath & Company 1982) 171–88, at 181.

12 G Geis, 'White-collar crime: What Is It?' in Schlegel and Weisburd (n 10) 31–52.

13 G Geis, and R F Meier, 'Controversy Regarding White-Collar Crime' in G Geis, and R F Meier (eds), *White-Collar Crime: Offenses in Business, Politics and the Professions* (The Free Press 1977) 253–59, at 253–4.

14 H C Finney and H R Lesieur, 'A Contingency Theory of Organizational Crime' in S B Bacharach (ed), *Research in the Sociology of Organizations*, vol 1 (JAI Press 1982) 263.

15 S P Shapiro, 'Collaring the Crime, Not the Criminal: Reconsidering the Concept of White-Collar Crime' (1990) 55 *American Sociological Review* 346–65, at 346.

16 Schuessler (n 8) xxi.

17 Shapiro (n 15) 357.

could lead to a fantastic array of ordinary business practices being classified as white-collar crime.¹⁸ By way of contrast, Aubert felt that criminology should shake off its subservience to legal agencies and that white-collar crime issues provided a focal point for such a development.¹⁹ This did not happen during the late 1950s and early 1960s when there was limited interest in white-collar crime and it remained on the periphery of academic research. However, since the mid-1960s interest in white-collar crime issues has increased and academics have offered differing theoretical approaches to the problems of definition. Quinney wanted to shift the focus from the social status of the offender to notions of occupational crime, which he hoped would stimulate empirical deviancy studies. He opposed Aubert's position, arguing that it was necessary to delineate more homogenous units for study by basing white-collar crime explanation upon the criminal law itself.²⁰ Edelhertz sought to broaden the definitional net of the white-collar crime label even further by removing any occupational criterion.²¹ This approach could be seen to confound the original purpose of the debate, which was to focus upon the harmful behaviour of high-status offenders. Schrager and Short stressed concepts of 'organisational crime', which emphasise the objective features of illegal behaviour.²² Their definitions are still dependent upon the criminal law's interpretation of harmful behaviour, but seek to extend legal sanctions upon the basis of impact without having to prove harm.

Pepinsky wanted to use the white-collar crime debate to move away from the criminal law and indeed the civil law as the medium for classifying undesirable behaviour, and instead adopt exploitation as the preferred measurement variable for the proscription of behaviours.²³ Such an approach of course raises even greater difficulties of definition and classification. Nevertheless, it is political power that ultimately determines the precision of definitions and the measurement of phenomena. Failing to acknowledge the influence of prevalent power relations hinders analysis of socially injurious behaviour by the powerful.²⁴ However, the 'human rights' criminology, as espoused by the Schwendingers, gained limited purchase upon academe in general or the white-collar crime debate in particular. Clarke is dismissive of the Schwendingers and other left-radicals, accusing them of trying to develop a discipline which would be 'divorced from reality' and stuck in an 'idealist limbo'.²⁵

Clinard and Quinney considered white-collar crime too vacuous a term, arguing that it consisted of two distinct types: occupational and corporate.²⁶ Cressey applauds their aims of greater classification, but considers them a failure.²⁷ He criticises Clinard and Quinney and Clinard's later collaborative work with Yeager,²⁸ because they perpetuate the ambiguities established by Sutherland regarding issues of corporate and individual responsibility and liability; in particular, how these are affected by the status of legal and

18 P W Tappan, 'Who Is The Criminal?' (1947) 12 *American Sociological Review* 96–102, 99–101.

19 V Aubert, 'White-Collar Crime and Social Structure' (1952) 58 *American Journal of Sociology* 263–71, at 265.

20 R Quinney, 'The Study of White-Collar Crime: Toward a Reorientation in Theory and Research' (1964) 55(2) *Journal of Criminal Law, Criminology and Police Science* 208–14.

21 H Edelhertz, *The Nature, Impact and Prosecution of White-Collar Crime* (US Government Printing Office 1970).

22 L S Schrager and J F Short, 'Towards a Sociology of Organizational Crime' (1978) 25 *Social Problems* 412.

23 H E Pepinsky, 'From White-Collar Crime to Exploitation: Redefinition of a Field' (1974) 65 *Journal of Criminal Law and Criminology* 225–33, at 227.

24 H Schwendinger and J Schwendinger, 'Defenders of Order or Guardians of Human Rights' in I Taylor, P Walton and J Young (eds), *Critical Criminology* (Routledge & Kegan Paul 1975) 113–46, at 122.

25 D H Clarke, 'Marxism, Justice and the Justice Model' (1978) 2(1) *Contemporary Crises* 27–62, at 44.

26 M B Clinard and R Quinney, *Criminal Behaviour Systems* (Holt, Rinehart & Winston 1973).

27 D R Cressey, 'The Poverty of Theory in Corporate Crime Research' in W S Laufer and F Adler (eds), *Advances in Criminological Theory*, vol 1 (Transaction Publishers 1989) 31–55.

28 M B Clinard and P C Yeager, *Corporate Crime* (Free Press 1980).

natural persons. Cressey situates the base cause of these problems of definition with Sutherland. He believes that Sutherland's work was not much of a theoretical contribution and finds it ironic that there was a great difference between what Sutherland said and what he did. This irony is compounded by the fact that Sutherland was something of an academic mentor for Cressey and that after Sutherland's death Cressey edited some of Sutherland's work into one of the standard textbooks on criminology. The disagreement between Cressey and Sutherland typifies the seeming inability of academics to develop an accepted definition of white-collar crime, or to agree on a theoretical basis for the study of white-collar crime. Carson believes that the criminological schools of the 1960s and 1970s for all their novel attempts at definition added little to the white-collar crime debate.²⁹ More than 20 years later, Shover and Hochstetler see theoretical and empirical research on white-collar crime as 'marginal to and largely unaffected by core developments in academic criminology'.³⁰ Agnew, Piquero and Cullen do not favour one definition of white-collar crime over another but argue that general strain theory can provide a useful theoretical lens through which to view white-collar crime phenomena.³¹

The debates about defining white-collar crime have often been polarised, with legal realists such as Tappan at one end of the definitional continuum and left idealists such as Pepinsky and the Schwendingers at the other. The legal realist position wants to focus only on 'crimes' and is plagued by problems associated with weak enforcement, ambiguity and differential censure. The left idealist perspective wants all 'undesirable behaviours' to be sanctioned and this seems an impossible task. Of course these problems of theoretical definition are accentuated by methodological and political difficulties. A three-day academic workshop in 1996 on issues of definition and white-collar crime organised by the National White-collar Crime Center canvassed many of these thorny definitional issues and emerged with the following imperfect but consensual definition:

Illegal or unethical acts that violate fiduciary responsibility or public trust, committed by an individual or organisation, usually during the course of legitimate occupational activity, by persons of high or respectable social status for personal or organisational gain.³²

Law enforcement agencies do not have the luxury of relative academic detachment, of course, and having acknowledged that the definition of white-collar crime is contested, in some cases have made their own specific attempts. For example the US Federal Bureau of Investigation (FBI) has defined white-collar crime as 'illegal acts which are characterised by deceit, concealment, or violation of trust and which are not dependent upon the application or threat of physical force or violence'.³³ This is obviously a relatively vague definition and is testimony to the practical operational realities that a law enforcement agency such as the FBI faces when it seeks to counter the myriad and pervasive behaviours that might fall under a white-collar crime canopy. Such pragmatic realities have motivated commentators such as Benson and Simpson to adopt an *opportunity perspective* to white-collar crime which stresses that there are different opportunity structures for different types of crime and that the opportunities to commit white-collar crime are not randomly or equally distributed throughout society, and so some people will have more opportunity to commit white-collar

29 Carson (n 9) 4.

30 N Shover and A Hochstetler, *Choosing White-Collar Crime* (Cambridge University Press 2006) 157.

31 R Agnew, N L Piquero and F T Cullen, 'General Strain Theory and White-Collar Crime' in S S Simpson and D Weisburd (eds), *The Criminology of White-Collar Crime* (Springer 2009) 35–60.

32 J Helmkamp, R Ball and K Townsend (eds), *White-Collar Crime?* (National White-Collar Crime Center 1996) 351.

33 US Department of Justice, Federal Bureau of Investigation, *White-Collar Crime: A Report to the Public* (Government Printing Office 1989).

crime than others.³⁴ A logical consequence of this approach is to emphasise situational crime prevention strategies to reduce the opportunities for white-collar offending via mechanisms such as improvements in surveillance, accounting protocols, compliance mechanisms and risk-management strategies in general. Such risk-oriented initiatives are widely implemented by organisations, professional associations, national regulatory bodies and multilateral regulatory actors, so that in many instances there can be multiple layers of prophylactic prevention strategies in place at local, regional and international levels. Nevertheless white-collar crime still remains prevalent and notoriously difficult to classify and measure, and, as discussed below, debates about financial crime face similar structural difficulties.

Financial crime – a case of *déjà vu*?

Justice Potter Stewart of the US Supreme Court once famously said of 'obscene material' that he might not know precisely what it was, but 'I know it when I see it'.³⁵ There are similar issues regarding financial crime when individuals and agencies are faced with the problem of defining such activity. Financial crime is a slippery concept, notably resistant to precise definition due to its blurring of activities and structures. So, over the years, there have been many definitions put forward by government agencies and other commentators, but finding a universally acceptable definition has been difficult.

The ubiquitous Wikipedia defines financial crime as: 'A crime against property, involving the unlawful conversion of the ownership of property (belonging to one person) to one's own personal use and benefit. Financial crime often involves fraud'.³⁶ This collapse into the category of fraud, which is in many senses more familiar, is more readily acknowledged under legal codes and by courts, and which therefore can be more easily delineated, is, as we shall see, a feature of debates about financial crime.

For the International Monetary Fund (IMF): 'Financial crime can refer to any non-violent crime that generally results in a financial loss, including financial fraud. It also includes a range of illegal activities such as money laundering and tax evasion'.³⁷ In the view of the government of Liechtenstein: 'There is no generally valid definition of financial crime. At a minimum, it includes money laundering, organized crime, and the financing of terrorism'.³⁸

The UK's Financial Services Authority (FSA) has stated that financial crime includes any offence involving money laundering, fraud or dishonesty, or market abuse and the reduction of financial crime interacts with the FSA's three other core statutory objectives – protecting consumers; market confidence; and public awareness.³⁹ In a further attempt at definition the FSA states that:

Financial crime is any crime involving money. More formally, the Financial Services and Markets Act 2000 defines financial crime 'to include any offence involving (a) fraud or dishonesty; (b) misconduct in, or misuse of information relating to, a financial market; or (c) handling the proceeds of crime'. The use of

34 M L Benson and S S Simpson, *White-Collar Crime: An Opportunity Perspective* (Routledge 2006).

35 *Jacobellis v Ohio*, 84 S Ct 1676 [1964]

36 Wikipedia, 'Financial Crimes' (CIFAS) <http://en.wikipedia.org/wiki/Financial_crime>.

37 International Monetary Fund, *Financial System Abuse, Financial Crime and Money Laundering – Background Paper* (IMF 2001) 3 <www.internationalmonetaryfund.com/external/np/ml/2001/eng/021201.pdf>.

38 Portal of the Principality of Liechtenstein, What is Financial Crime (2007) <www.liechtenstein.li/en/eliechtenstein_main_sites/portal_fuerstentum_liechtenstein/fl-wuf-wirtschaft_finanzen/fl-wuf-finanzdienstleistungen/fl-wuf-finanzdienstleistungen-finanzkriminalitaet/fl-wuf-finanzdienstleistungen-finanzkriminalitaet-was.htm>.

39 Financial Services Authority, *Fighting Financial Crime* (FSA 2012) www.fsa.gov.uk/about/what/financial_crime.

the term 'to include' means financial crime can be interpreted widely to include, for example, corruption or funding terrorism.⁴⁰

This is obviously a pretty broad church and attests to the difficulties of pinning down financial crime into neat classification, even for a large regulatory agency with the resources of the FSA, which has financial crime as one of its four key statutory objectives. Therefore, presumably countering financial crime is one of the FSA's key performance indicators, but if it is not specifically defined, classified and counted, how is it possible to measure and subsequently demonstrate improved regulatory performance on the issue? This is not to say that the FSA is not seeking to counter financial crime. In fact when pursuing its financial crime objective, the FSA's main focus is on firms' risk management, systems and controls. Rather it highlights the lack of empirics in debates about financial crime and one should have some sympathy for the FSA position. For example, in Australia, a search of the Australian Prudential Regulatory Authority (APRA) and Australian Securities and Investments Commission (ASIC) websites did not reveal any specific guides which focused on defining and classifying financial crime, although there are numerous publications by both the ASIC and the APRA which aim to improve how financial actors and markets function in Australia. The Australian Crime Commission (ACC) is not able to offer much definition or measurement specificity either:

Financial crimes can include a wide range of activities from fraud through to active manipulation of the stock market or laundering of the proceeds of crime . . . [and]

It is difficult to fully judge the extent of these activities.⁴¹

This definitional uncertainty at a national level is unsurprisingly reproduced in regional and international contexts: for example, the experience of the European Union (EU) after it formally acknowledged the increasing trend towards international financial crimes on a larger scale. It called for common definitions, common incriminations and common sanctions to be established throughout the EU for these crimes. A Framework Decision on combating fraud and counterfeiting of non-cash means of payment was adopted on 28 May 2001. Under this proposal, fraud and counterfeiting involving any form of non-cash payment is recognised as a criminal offence and punishable by effective, proportionate and dissuasive penalties in all EU member states. A Framework Decision on money laundering, the identification, tracing, freezing, seizing and confiscation of instrumentalities and the proceeds of crime was adopted on 26 June 2001. However, 'The Framework Decision deliberately avoids references to specific offences under the existing criminal law because they do not cover the same elements in the member states. Instead, the proposal merely lists the various types of behaviour, which should constitute criminal offences throughout the EU (i.e. theft or unlawful appropriation of payment instrument, receiving, obtaining, transporting, sale or transfer to another person of a stolen payment instrument in order for it to be used fraudulently etc).⁴²

The EU has repealed the 2001 Directive on Money Laundering and Terrorist Financing and agreed a new one in 2005, and the financial crime section of the European Commission has produced a string of reports that have largely focused on countering money laundering and terrorist financing.⁴³ The path towards harmonisation unsurprisingly has been difficult

40 Financial Services Authority, *PS 11/15 Financial Crime: A Guide for Firms* (FSA 2011) 61.

41 Australian Crime Commission, *Financial Crimes* (ACC 2011) 1.

42 EUROPA Justice and Home Affairs, *Financial Crime a Priority Area for Approximation of Criminal law in the EU* (EUROPA 2007) <http://ec.europa.eu/justice_home/fsj/criminal/financial/fsj_criminal_financial_en.htm>.

43 European Commission, *The Single Market, Financial Crime* <http://ec.europa.eu/internal_market/company/financial-crime/index_en.htm#3rdcountry>.

and there has been little progress on comprehensive definition and measurement of financial crime as a category. For example, Article 11 §4 of Directive 2005/60/EC contains an obligation for member states to inform each other, the ESAs and the European Commission of cases where they consider that a third country meets EU AML/CFT (anti-money laundering/combating the financing of terrorism) equivalence standards. The commission publishes this list on its website, but stresses that the assessment of third-country equivalence remains a member-state competence. Indeed, Directive 2005/60/EC does not grant the European Commission any mandate to establish a binding 'positive' list of equivalent third countries.⁴⁴

So there is underlying uncertainty globally regarding what financial crime is, and even in a more integrated region of the world such as Europe, a lack of specificity about financial crime offences. This lack of specificity is a problem with respect to establishing baselines about scale of activity and evaluation of law enforcement responses to such activity. These problems appear to be true for approaches towards financial crime by various regulatory actors and across a range of jurisdictions. Many of the behaviours of financial crime may be enacted in multilateral contexts, so this legal and law enforcement uncertainty can make national strategic and operational policing responses to multilateral financial crime more difficult.

The limited number of examples in the academic literature that try to define financial crime reflect the uncertainty shown by regulators. For example, Pickett and Pickett state that financial crime usually entails: 'The use of deception for illegal gain, normally involving breach of trust and some concealment of the true nature of the activities.'⁴⁵ Fleming notes that the array of offences that could constitute financial crime is 'potentially indistinct'.⁴⁶ Gilligan stresses how difficult it is to define and measure financial crime in comparison to *street crime* despite widespread recognition in the media and government policy and decision-making that it, i.e. financial crime, causes substantial harm.⁴⁷ Many academic writers (perhaps understandably given the measurement difficulties) simply accept financial crime as a label, and/or quote an attempt at definition by an organisation such as those listed earlier by the FSA or IMF, then move into a discussion of regulatory models, policy initiatives or other interventions such as AML, insider trading or terrorist financing that have been employed in the general area of corporate/financial crime.

Much of the academic writing on financial crime has come from the disciplines of law, sociology and criminology but, as discussed above, these contributions have not seemed able to move the debate on from relatively vague generalisations about financial crime being difficult to classify and measure. Some more recent contributions from the management field demonstrate potential for fleshing out what actually might constitute financial crime and begin the long process of mapping its content empirically. For example, Gottschalk suggests four major categories of financial crime – corruption, fraud, theft and manipulation – with delineated subsets of behaviours under these four categories – such as kickbacks, bribery, extortion and embezzlement under corruption.⁴⁸ This categorisation process has been built on through a study of Norway's largest business companies which

44 European Commission, *Common Understanding between Member States on the Procedure and Criteria for the Recognition of Third Countries Equivalence under Directive 2005/60/EC on the Prevention and use of the Financial System for the Purpose of Money Laundering and Terrorist Financing* <http://ec.europa.eu/internal_market/company/docs/financial-crime/3rd-country-common-understanding_en.pdf>.

45 K H S Pickett and J M Pickett, *Financial Crime, Investigation and Control* (Wiley 2002).

46 M Fleming, *FSA's Scale and Impact of Financial Crime Project (Phase One)*, Critical Occasional Paper Series 37 (FSA 2009) 2.

47 G Gilligan, 'Financial Crime in Australia' (2007) 27(1) *Journal of the Institute of Economic Affairs* 10–13, at 10.

48 P Gottschalk, 'Categories of Financial Crime' (2010) 17(4) *Journal of Financial Crime* 441–58, at 443.

invited company representatives to report on their organisation's experience with financial crime and asked them which categories of financial crime they were most vulnerable to. Unsurprisingly, numerous types of fraud (loan, credit, fee insurance) and theft (cash, identity) were considered the more prevalent vulnerabilities with other behaviours such as manipulation and extortion also prominent.⁴⁹

This type of bottom-up research probably presents the most fruitful pathway for identifying what does fall under the financial crime rubric and scoping out how much financial crime there may be under different categories and in different sectors. A mosaic model that gradually integrates such information will help to provide a more substantive empirical underpinning of how much of a problem financial crime is in contemporary society. This is a mammoth longer-term task and will require commitment not just from governments, law enforcement and regulatory agencies, but also from business and the professional service actors such as large accounting and consulting firms that service business needs. In order to make a contribution to this discourse I undertook a small-scale study of law enforcement and other regulatory actors in Australia and the UK about whether they utilised working definitions of financial crime and if so what those working definitions were.

How is financial crime defined and classified by law enforcement and regulators?

The study utilised email and telephone to conduct semi-structured interviews with relevant key actors in Australian and UK law enforcement agencies and other regulatory agencies whose operational remit would include countering behaviours that might reasonably be viewed as financial crime. The respondents were invited to comment on how they sought to define and classify financial crime and what they saw as the scale of financial crime. There were four core questions:

1. Does your agency use an operational definition of financial crime, and if so what is it?
2. What is the extent of financial crime in your jurisdiction?
3. If your agency does not use an operational definition of financial crime, what behaviours/offences that your agency seeks to counter could be categorised as financial crime?
4. If your agency does not use an operational definition of financial crime, what is the extent in your jurisdiction of the behaviours/offences that could be categorised as financial crime?

Table 1 details how many agencies were approached in each jurisdiction and the overall response rate of 59 per cent was reasonable given the empirical and other methodological difficulties surrounding the issue of financial crime. A selection of responses to the four specific questions is provided below. It is clear that in both jurisdictions police and other regulatory agencies canvassed had varying degrees of certainty about what their responsibilities for countering financial crime were.

The responses from the Australian sample are summarised first. One state attorney general's office offered this view:

We are not aware of any use of the term *financial crime* in our department. We use other bodies' definitions, such as the Australian Bureau of Statistics for Court statistics or the state police definitions for crime statistics. This provides around 70 groups, none are *financial crime*.

49 P Gottschalk and H Solli-Saether, 'Financial Crime in Business Organizations: An Empirical Study' (2011) 18(1) *Journal of Financial Crime* 76–92.

	Australia	UK	Australia + UK
Agencies approached	20	39	59
Agencies responded	14	21	35
Police forces approached	8	36	44
Police forces responded	5	20	25
Non-police forces approached	12	3	15
Non-police forces responded	9	1	10

Table 1: Responses from agencies approached in small-scale study of law enforcement and other regulatory actors in Australia and the UK

Another state-based regulatory actor gave these insights to the four questions:

- 1. No, however a working definition could be: ‘Any activity of a financial nature which might constitute an offence against the Criminal Code 1899 or an offence of money laundering against the Criminal Proceeds Confiscation Act 2002.
- 2. Unable to quantify the extent of ‘financial crime’.
- 3. Public sector corruption and fraud. Fraud conducted in the context of organised criminal activity i.e. major organised fraud or fraud committed in the context of other organised criminal activity e.g. fraud on financial institutions using false identities, fraudulent misrepresentations etcetera, money laundering.
- 4. See Q2 above.

Australian police forces that participated also made it clear that financial crime did not exist for them as a distinct operational category or key performance indicator, although they were extremely active in seeking to counter harmful behaviour in areas that reasonably could be considered as financial crime. For example:

- 1. No
- 2. We are not in a position to respond to this question. Perhaps a statistical gathering agency such as the Bureau of Crime Statistics and research can assist.
- 3. We investigate both Commonwealth and State ‘fraud related’ offences. It could be reasonably be asserted that the vast majority of fraud related offences contained in these subdivisions involve some financial aspect/loss to the victim.
- 4. We are not in a position to answer this question because we do not collect or have access to this type of information.

Another police force had this comment:

We have no operational definition. The Commercial Crime Division investigates serious crimes in business transactions (fraud, stock stealing, gold stealing, computer-aided crime). It has four sections (Computer Crime, Gold, Rural, Major Fraud). The Major Fraud squad investigates crimes related to fraud, stealing as a servant, identity fraud, special commodities.

This absence of financial crime as a separate category was echoed by another Australian police force:

We have no formal definition, it is not used in crime statistics and I am not aware of the use of the term in operations. We classify crimes according to the Standard Offence Table and so keep statistics of fraud, credit card fraud, computer fraud, identity theft and stealing offences.

So in Australia there seems to be no sense of financial crime as a distinct category amongst those law enforcement actors responsible for countering criminal behaviour. That is not to

say that police and other regulatory actors are not trying to prevent fraud and other financial crimes, rather that because the specific classification and sub-categories of financial crime do not exist yet, statistically there is no clear record of their overall efforts in this area. Similarly, there is no clear picture of the scale of the problem that they face as the empirical information is not being collated in ways that could inform performance measurement and policy development on the issue of financial crime.

Responses from agencies in the UK reveal a similar situation. For example, a large regional police force replied in this way:

1. Answer: No we do not.
2. Answer: As mentioned above, we do not specifically categorise financial crime in this police force area. In fact, throughout England and Wales there are many specific crime types or categories that might involve what might be termed 'financial crime'. For example, offences such as theft, deception and offences under the Fraud Act 2006 could involve a financial element, as indeed could offences involving drug trafficking, people trafficking, money laundering and so on. I could not honestly answer as to what the extent of such crime is in our area, although it is fair to say that we have examples of all of these offence types.
3. Answer: Please see the answer to 2 above. The main thrust is in relation to asset recovery (that is to say the recovery of assets from criminals who have been convicted of offences where we seek to remove the benefit of their crimes from them).
4. Answer: As above. We are unable to speculate on the extent of this because what might be termed 'financial crime' may feature in a whole host of different crime-types. Because we do not record this data under a specific category we could not accurately calculate this.

Similar views are evident in the response of a rural police force:

Question 1: We do not have an operational definition of the term 'financial crime'. Under the Proceeds of Crime Act 2002 any crime committed whereby the perpetrator has benefited financially could be construed as 'financial crime'.

Question 2: We are unable to state the extent of financial crime within our jurisdiction as the Force does not have an operational definition for the term 'financial crime' (as per response to question 1).

Question 3: Under the Proceeds of Crime Act 2002, we would look at all crimes where the perpetrator has benefited financially from his/her criminal acts e.g. through money laundering per se or through confiscation of assets post charge.

Question 4: See response to question 3.

Unsurprisingly a police force with both large cities and rural areas in its zone had a similar position:

Question 1. No there is no definition of financial crime per se as the financial investigation unit work primarily to the provisions of the Proceeds of Crime Act 2002 which have a bearing on the terms of reference for the department.

Question 2. The extent of financial crime is not measured as arguably there is a dimension of financial crime in all acquisitive crime which accounts for about 70% recorded crime.

Question 3. Distinct from a general heading of all acquisitive crime, financial crime is investigated predominantly to counter Money Laundering and Fraud.

Question 4. There is no known extent of financial crime let alone the behaviours that could be so categorised. This is not peculiar to us but a national issue.

This reply makes the important point that the lack of clarity around financial crime is a national issue (indeed it is an international one as well). An individual police force or regulatory agency has to operate within its state and national legal structures, as well as deal with behaviours that cross geographical jurisdictional boundaries and this means that its capacity for innovation on offence classification and measurement is limited. A potential national lead actor on this area is the FSA. However, its response to the specific four questions of this study was to reiterate that countering financial crime was one of its four core statutory objectives and state that: 'the extent of financial crime is difficult to determine'. So, in the UK it is clear that although there seems to be an increasing political and regulatory commitment to countering financial crime, there are still major difficulties of definition and evaluation of the scale of the problems. This is despite since January 2007 the FSA establishing the Financial Crime and Intelligence Division (FCID) and launching a dedicated fighting financial crime website to further raise the FSA's capability to counter financial crime.⁵⁰

This small-scale study indicates that the uncertainty and empirical lacuna surrounding financial crime is prevalent in Australia and the UK. It is likely that this is the case in most other jurisdictions as well. First principles would expect that better definition and classification are required in order to develop the necessary adaptive policing strategies to counter financial crime in Australia, the UK and elsewhere.

Conclusions

Many types of crime are relatively easy to define and measure. Much conventional crime falls into this category, for example, burglary, homicide, vehicle theft and arson. However, as we have seen, financial crime, like white-collar crime, is much harder to define and subsequently measure, so in comparison to more conventional or *street* crime there can be a lack of hard data. Compounding this empirical uncertainty is the ambiguity that can sometimes surround financial crime, because its effects can be more diffused. An example of this is the savings and loans scandal in the US in the 1980s and 1990s when more than 700 savings and loans associations failed. Initially most failures were assessed as careless management practices, but later US government reports 'strongly suggest that criminal activity in the form of fraud was a central factor in 70 to 80 per cent of these failures'.⁵¹ Add to this mix the operational reality that law enforcement resources to counter financial crime may be limited in many jurisdictions and it easy to see why much financial crime may not be counted in official crime statistics. This loop effect contributes to what may seem at times to be disjointed law enforcement responses to financial crime.

The emerging regulatory and law enforcement activity around the label *financial crime* has parallels to the white-collar crime discourse and seems likely to suffer from similar definitional ambiguity and ambivalence which can be expected to result in similar difficulties not only in measuring the scale of the activity, but also in evaluating policing and other regulatory responses to such activity. The irony is that all this ambiguity, ambivalence and uncertainty does not seem to be preventing the emergence of something of a growth industry around the label *financial crime*, as regulatory agencies, police bodies and industry actors establish specialist financial crime units, websites and advisory teams. For example, Deloitte has been expanding its advisory services in the area of financial crime.⁵² Similarly

50 <www.fsa.gov.uk/Pages/About/What/financial_crime/index.shtml>.

51 H N Pontell and K Calavita (1992), 'Bilking Bankers and Bad Debts, White-Collar Crime and the Savings and Loans Crisis' in Schlegel and Weisburd (n 10) 195–213, at 196.

52 Deloitte, *Financial Crime, Advisory Services* <www.deloitte.com/view/en_AU/au/services/forensic/financial-crime/advisory/index.htm>.

Dun & Bradstreet have sought to boost its advisory role to business by producing financial crime-related publications.⁵³

The evidence suggests that progressing the definition and measurement of financial crime is unlikely to be achieved by academics working in isolation. The earlier discussion indicates that this seems especially true for academics in the law, sociology and criminology disciplines, despite their long involvement in this area harking back to the 1940s and the legacy of Edwin Sutherland. Rather, an integrated multi-professional, multi-agency as well as multidisciplinary effort is required with an emphasis on *bottom-up research*. Business actors such as advisory firms are likely to have more empirical knowledge of the realities of financial crime in the business world. However, their obligations are largely to their client base and profit-maximisation strategies, and so they will need to be persuaded of the utility of broader-based research strategies. Similarly, police forces will have some awareness of the scale of the financial crime problems that businesses face and records of at least some of the financial offences suffered by individual citizens that have been reported in their jurisdiction. However, they are limited by their jurisdictional remit, whether state, urban or regionally based.

As discussed earlier there is less definitional uncertainty about fraud in comparison to financial crime and so, regarding fraud, there have been some initiatives which combine private and public sector efforts in scoping the problem. For example, in the UK, CIFAS is a not-for-profit association originally created in 1988 by a group of credit retail companies. It now has more than 250 member associations (both public and private agencies) covering an array of business sectors and maintains two substantial databases aimed at fraud prevention – the CIFAS National Fraud Database and the CIFAS Staff Fraud Database.⁵⁴ However, mapping out financial crime and organising data collection is an even greater challenge and so will need to be coordinated by national actors such as the FSA in the UK, or the ASIC and the ACCC in Australia. Realistically these types of agencies are the actors that possess the necessary powers and resources capable of the *heavy lifting* required to persuade advisory firms, police forces, professional associations and other groups to engage in extensive cross-sector and cross-disciplinary research on defining, classifying and measuring financial crime. Methodologies could build on work emerging from the management sciences field by researchers such as Gottschalk discussed earlier.⁵⁵ Also, the small-scale field study discussed above demonstrates that police and other regulatory actors in Australia and the UK do seem to be aware of the weaknesses of existing definitions and categorisations related to financial crime. However, given their existing range of responsibilities most are unable to see how they might contribute to improving the situation.

The political context is crucial and without requisite political commitment nothing is likely to change, except that likely effects of financial crime will continue to grow. Those interested in progressing debates about, and regulatory efforts against, financial crime should lobby to generate that political momentum. This is a big but not impossible challenge. For example, in November 2011, I appeared before the Commonwealth Parliamentary Joint Committee on Corporations and Financial Services as part of its remit of statutory oversight of the ASIC. Part of my evidence drew linkages between regulatory failure and inconsistent definition and measurement of fraud and financial crime. In its subsequent report the Joint Committee has recommended that:

53 Dun & Bradstreet, *Protection Against Financial Crime: Lessons from the Global Experience* (2008) <http://dnb.com.au/Header/News/Reports_and_White_Papers/Protection_against_financial_crimes/index.aspx>.

54 See CIFAS, UK Fraud Prevention Service <www.cifas.org.uk/>.

55 See n 48 and n 49 above.

ASIC take steps to use available information to collate and analyse definitions of, and approaches to, financial crime, with a view to developing standard definitions and classifications that can be used across the Commonwealth.⁵⁶

A recommendation from a Parliamentary Joint Committee is an influential pressure on a statutory agency such as the ASIC, but obviously it remains to be seen how much progress ASIC can make in the future on these thorny issues and whether it grasps the nettle of the research-coordinating role discussed above. Nevertheless, initiatives such as this, in Australia, the UK and elsewhere, allied with continuing profile-raising and field research by academics, business actors, law enforcement and other commentators on the utility of developing meaningful baseline data on financial crime can only be helpful in seeking to counter its harmful effects both to the economy and civic society.

⁵⁶ Parliamentary Joint Committee on Corporations and Financial Services, *Statutory Oversight of the Australian Securities and Investments Commission* (March 2012) 11, <www.aph.gov.au/Parliamentary_Business/Committees/Senate/Committees?url=corporations_ctte/asic/index.htm>.

Modernising environmental regulation in Northern Ireland: a case study in devolved decision-making

SHARON TURNER AND CIARA BRENNAN

*School of Law, Queen's University Belfast**

Introduction

For over a decade, controversy about the quality of environmental regulation has cast a shadow over the effectiveness of environmental governance in Northern Ireland. Most fundamentally this debate has centred on a crisis of confidence about the quality of regulation and a consensus that effective reform depends on the externalisation of this responsibility from central government. Not surprisingly, the causes of weak regulation were rooted in the eclipsing impact of the Troubles and the fossilisation of government that occurred during the decades of Direct Rule.¹ However, although the first steps towards meaningful reform were eventually taken under Direct Rule, the restoration of devolution and the stabilising power-sharing process has meant that the trajectory of regulatory reform has been largely shaped by a devolved administration. The purpose of this paper is to examine the nature and implications of that process. Pressure for regulatory reform is an issue that has confronted both configurations of Northern Ireland's power-sharing Executive. Despite its brief and tumultuous lifespan, the first Ulster Unionist Party (UUP)/Social Democratic and Labour Party (SDLP)-led administration was immediately faced not only with the evidence of serious regulatory dysfunction but also the first stage of what became a concerted civil society campaign for independent regulation. The collapse of power-sharing did nothing to quell this pressure. Instead, when devolution was restored five years later the new Democratic Unionist/Sinn Féin-led administration was faced once again with pressure for regulatory reform. However, this time the case for independent regulation was supported not only by civil society, but also the overwhelming majority of stakeholders to this governance process, including all but one of the parties sharing power and all but one of the industries subject to environmental regulation. This consensus was furthermore

* School of Law, Queen's University Belfast, 27–30 University Square, Belfast BT7 1NN. Email: s.turner@qub.ac.uk and cbrennan12@qub.ac.uk

1 The impact of the conflict on environmental governance has been discussed in detail elsewhere. See, for example, S Turner and K Morrow, *Northern Ireland Environmental Law* (Gill and Macmillan 1997); K Morrow and S Turner, 'The More Things Change, the More They Stay the Same? Environmental Law, Policy and Funding in Northern Ireland' (1998) 10(1) *Journal of Environmental Law* 41–59.

supported by an independent review of the arrangements for environmental governance that had been commissioned by the Direct Rule administration during the hiatus in devolution.²

Despite the cacophony of voices calling for structural reform, the newly restored devolved Executive proved unable to facilitate that process. The political dynamics of devolution combined with the arrangements for decision-making within the power-sharing Executive effectively enabled the largest party sharing power to unilaterally block reform. Although the new Democratic Unionist Minister for the Environment (then Arlene Foster) acknowledged the need for improvement in the quality of regulation and committed to investment in this context, her party opposed the externalisation of this function from the Department of the Environment (DOE) and thus the change viewed by their partners in the power-sharing Executive as the necessary foundation stone for credible regulatory reform. While this debate arguably demonstrated the coming of age of the environment as a mainstream political issue in Northern Ireland and, indeed, fostered a maturing of civil society in this sector, the devolved administration's handling of this issue provided a powerful and unsettling insight into the nature and culture of devolved government and governance. While the legacy of weak regulation inherited by the devolved administration can be blamed on years of conflict and political absenteeism during Direct Rule, the debate surrounding its modernisation makes clear that the very different but real democratic limitations inherent in the region's devolution settlement will also function to stifle political stewardship of the environment. Regrettably, but most inevitably, this will force judicial resolution of the core structural arrangements for delivering environmental regulation, and thus diminish democratic control of a key aspect of economic regulation in post-conflict Northern Ireland.

Facing the legacy of neglect

It is probably an understatement to say that when the Good Friday Agreement was signed in 1998 the state of the regional environment and the arrangements for its protection were at best marginal to the concerns of the negotiating parties, not least Northern Ireland's major political parties. The political challenges of peace-building and power-sharing not surprisingly dominated the agenda for Northern Ireland's first power-sharing Executive. However, despite these pressures and its brief lifespan, the first power-sharing Executive was nevertheless forced to immediately grapple with the consequences of decades of neglected environmental governance. In essence, the UUP/SDLP-led administration inherited a system of environmental governance defined by serious legislative antiquation and very weak regulation. However, the proverbial ink was scarcely dry on the Good Friday Agreement when it was also forced to confront the legal and financial consequences of devolved responsibility for that inheritance. During the 1990s, endemic failure by government in Northern Ireland to ensure the timely and complete transposition of EU Directives on the environment and failure to invest in the water and sewerage infrastructure necessary to ensure operational compliance with the Urban Waste Water Treatment Directive 91/271/EEC³ had led the EU Commission to commence numerous 'infraction' proceedings against the UK.⁴ By the time devolution was restored in 1998, UK central government and the new devolved administration faced a phalanx of serious and advanced

2 REGNI, *Foundations for the Future: The Review of Environmental Governance* (2007) <www.doeni.gov.uk/foundations_for_the_future_-_final_report.pdf>, discussed in more detail below. See also S Turner, 'Laying the Foundations for a Sustainable Northern Ireland' (2007) 58 Northern Ireland Legal Quarterly 422–58.

3 OJ L135/40.

4 S Turner, 'Transforming Environmental Governance in Northern Ireland: Part One: The Process of Policy Renewal' (2006) 18 Journal of Environmental Law 55–87.

EU enforcement proceedings. The profound impact of EU pressure in forcing the first power-sharing Executive to invest in an immediate and extensive process of environmental law reform has already been the subject of detailed analysis by the present authors.⁵ However, while the mutual exposure of UK central government and the devolved administration to the legal and potentially serious financial consequences of EU infraction action undoubtedly induced the political will necessary to deliver a major programme of legislative and infrastructural modernisation, a similar dynamic did not apply to the equally dysfunctional process of environmental regulation.

On one level the restoration of devolution also coincided with an unveiling of serious regulatory weaknesses. Within weeks of the signing of the Good Friday Agreement, the Northern Ireland Audit Office published the first detailed analysis of the state of water pollution regulation in the region.⁶ This report was essentially a searing analysis of failure by the DOE to discharge its responsibilities as Northern Ireland's environmental regulator; however, it was not an isolated publication. Instead it became the first in a series of highly critical reports published by the Audit Office, the Public Accounts Committee (PAC), the Northern Ireland Affairs Committee (NIAC) and the Criminal Justice Inspectorate (CJI) during the period 1998 to 2007 concerning the quality of environmental regulation in the region. Their cumulative critique laid bare a landscape of enduring and serious failure by the DOE to discharge almost every aspect of its duty as environmental regulator.⁷ The nature and scale of this dysfunction has already been the subject of detailed analysis by the present authors.⁸ Suffice it for present purposes to state that these reports revealed a catalogue of lax, fragmented, inconsistent and non-transparent regulation and a particularly problematic approach to enforcing the rule of environmental law. However, while the official scrutiny community operating in the region gradually documented the systemic scale of the regulatory dysfunction that pertained in Northern Ireland, it did not fall within the ambit of the EU enforcement action then underway. Thus, while the first devolved administration began to face significant internal pressure for regulatory reform, this did not extend to legal, financial or political pressure from the EU or Whitehall. In sharp contrast to the extensive modernisation programmes launched to overhaul the legislative framework governing the environment and the region's water and sewerage treatment infrastructure, the devolved administration's response to the evidence of seriously weakened environmental regulation revealed if anything, a willingness to exploit rather than resolve the problem.

The implications of weak regulation and in particular the centralised nature of this responsibility was cast into graphic relief by the manner in which the first power-sharing Executive handled the pressure to meet decades of unmet but not well aligned economic and environmental needs. On the one hand, the consolidating peace process and burgeoning property market on the island had fuelled escalating pressure for economic regeneration and development within the region. However, this pressure was arising in the context of a region with almost Victorian standards of water and sewerage infrastructure, which was itself the subject of concerted EU infraction action.⁹ The devolved Minister for

5 Turner (n 4).

6 Northern Ireland Audit Office (NIAO), *Control of River Pollution in Northern Ireland*, HC (1997–1998) 693.

7 Northern Ireland Assembly PAC, Third Report, *Control of River Pollution in Northern Ireland* (2001) <www.niassembly.gov.uk/public/reports/report3-00r.htm>; NIAO, *Areas of Special Scientific Interest*, HC (2003–2004) 499; House of Commons NIAC, *Waste Management Strategy in Northern Ireland*, HC (2004–2005) 349-I; House of Commons Committee of Public Accounts, *Northern Ireland's Waste Management Strategy*, HC (2005–2006) 74; NIAO, *Northern Ireland's Waste Management Strategy*, HC (2005–2006) 88.

8 S Turner, 'Transforming Environmental Governance in Northern Ireland. Part Two: The Case of Environmental Regulation' (2006) 18 *Journal of Environmental Law* 245–75, at 266–9.

9 Turner (n 4).

the Environment (then Dermott Nesbitt, UUP) was consequently faced with expressions of serious concern by the Executive Agency within his department responsible for environmental regulation (then called the Environment and Heritage Service) as to the legal and environmental implications of permitting further development in areas lacking appropriate sewage treatment, while at the same time the Executive Agency within his department responsible for development control in Northern Ireland (the Planning Service) faced intense pressure to facilitate economic regeneration. Nesbitt made an initial decision to impose a moratorium on all development in what became known as sewage 'hotspots' to enable the department to consider the legal, environmental, economic and operational implications of this situation. However, this review resulted not in a decision to restrain development or even to require developers to share the costs of ensuring temporary sewage treatment for proposed development, but a decision to constrain the environmental regulator. Nesbitt acknowledged that serious and potentially irreversible environmental damage would be caused by further development in areas without appropriate sewage treatment. Although the devolved administration had already signalled its commitment to a major programme of infrastructure investment to respond to EU infraction pressure, it was also clear this process would take years to deliver. Consequently, the minister announced that the most 'pragmatic' solution was to enable the regulator to 'alert' the planning authorities to the environmental implications of proposed development but to prevent it from lodging a formal objection to the granting of development consent.

Quite apart from the disquiet caused by evidence of weak environmental regulation being delivered by the DOE, this decision cast into sharp relief the even more fundamental problem arising from the centralised nature of responsibility for this function. Despite the creation of independent regulators across the rest of the UK and Ireland during the 1990s, successive Direct Rule administrations had resisted pressure for equivalent structural reform in Northern Ireland.¹⁰ This function thus remained the responsibility of the DOE, and although delivered by an Executive Agency possessed of its own resources and staff, it lacked any separate legal identity. The 'hotspots' debacle highlighted not only the governance implications of centralised regulation, but also the willingness by Northern Ireland's new political class to exploit the weaknesses inherent in these arrangements in order to facilitate even crude economic development and build political capital. However, it also marked a turning point in the relationship between government and the region's environmental non-governmental organisation (NGO) sector.

A maturing of civil society

In sharp contrast to their counterparts in the fields of human rights and equality in Northern Ireland and their sectoral counterparts in Britain, the NGO sector on the island of Ireland is relatively underdeveloped.¹¹ Environmental campaigning on the island has historically been characterised by an emphasis on highly localised campaigning;¹² however, in Northern Ireland a range of distinctive factors flowing from the dynamics of Direct Rule had additionally forged an NGO culture that valued access to government over robust public debate. In essence, the absence of accountability levers and the highly centralised nature of environmental governance under Direct Rule had created a strong disincentive to the high-

¹⁰ The reasons for which are discussed by Turner (n 4) 249–54.

¹¹ See L Leonard, *The Environmental Movement in Ireland* (Springer 2007) and J Barry and P Doran, 'Environmental Movements in Ireland: North and South' in T McDonagh, F Varley and S Shorthall (eds), *A Living Countryside* (Ashgate 2009) 321–41.

¹² Barry and Doran (n 11).

profile public campaigning on the environment common in other UK jurisdictions.¹³ Direct Rule ministers were also frequently absent leaving civil servants to assume a greater responsibility for decision-making.¹⁴ However, the small scale of government and the heightened role of officials led many within the civil service to personalise public criticism of government activity.¹⁵ Direct Rule also made it unattractive for many of the UK's leading ENGOs to fund anything other than a relatively small presence in Northern Ireland. The organisational risks and operational burdens posed by public campaigning were further exaggerated for this small sector due to the absence of a key participant in the public debate; namely, the regulator itself. As civil servants, officials within the regulator could not publicly contradict either the Minister for the Environment or the wider government and thus could not represent the public interest in the environment.

When devolution was restored, the spirit of political optimism that surrounded the new power-sharing Executive but also the widely understood fragility of this process initially compounded this sector's innate aversion to criticising government.¹⁶ However, the spectacle of Nesbitt's blatant political interference in the process of environmental regulation prompted Friends of the Earth (NI) to accuse him of 'gagging' the regulator, and to mount the first highly critical campaign of public opposition ever launched by an ENGO in Northern Ireland.¹⁷ Although this more confrontational style of political campaigning certainly alienated senior officials within the department and initially their ENGO counterparts,¹⁸ ultimately it catalysed an important process of maturing within the sector as a whole. In the face of the DOE's failure to formulate a meaningful reform agenda to respond to the by then mounting evidence of regulatory dysfunction, Friends of the Earth led the formation of a coalition comprising the region's nine largest ENGOs for the purposes of launching a united public campaign for the externalisation of this responsibility to an independent environmental agency. The formation of this coalition was certainly facilitated by greater investment in regional capacity by the major national ENGOs, which flowed from their expectation that devolution would create a more fruitful political context for public campaigning. However, the decision to collaborate not only protected individual groups from government pressure but enabled the members of the coalition to find their separate and collective public voice and with it came a process of maturing within the sector. But perhaps even more importantly, the high-profile, four-year campaign they waged provided not only a policy leadership completely lacking within central government, but was successful in bringing the environment from the remote margins to the mainstream of the post-conflict political agenda emerging in Northern Ireland.

Pressure for reform on the cusp of political transition

Despite the collapse of devolution in October 2002, the campaign coalition was formally launched in 2003.¹⁹ Most fundamentally it argued that credible environmental regulation

13 These observations are based on interviews with senior figures in the ENGO sector in the region and the direct professional observations of Professor Turner during the REGNI process and during an extended secondment to the DOE as a senior legal adviser.

14 Turner and Morrow (n 1).

15 See n 13.

16 Ibid.

17 Turner (n 4).

18 Interview with former Director of Friends of the Earth; see also J Barry, "'It Ain't Easy Being Green': Sustainable Development between Environment and Economy in Northern Ireland' (2009) 24 Irish Political Studies 45–66.

19 The coalition consisted of the Conservation Volunteers, Friends of the Earth, National Trust, Northern Ireland Environment Link, RSPB, Ulster Wildlife Trust, Wildfowl and Wetlands Trust, Woodland Trust and WWF.

required the externalisation of responsibility for this function to an independent environmental protection agency akin to those already established in the UK and Ireland. However, the coalition commissioned an independent expert analysis of the various options for structural reform and used the resultant report as the basis for a formal public consultation exercise to assess stakeholder support for externalisation of this function.²⁰ The responses received were also the subject of independent analysis, which reported overwhelming support for the transfer of this responsibility to an independent arms-length entity.²¹ Although stakeholders varied in their views as to the level of independence necessary for effective regulation, only one expressed strong opposition to externalisation per se; namely, the Ulster Farmers' Union, which argued that the existing system could be made to function effectively through a process of internal reform.²² However, as the campaign matured so too did the ENGO agenda. In 2005, the coalition hosted a Chatham House conference to discuss the options for and implications of structural reform, which was attended by senior officials from UK and Irish environmental regulators and core Northern Ireland departments engaged in environmental governance as well as senior representatives from the major regulated industries, the UK Sustainable Development Commission, UK Environmental Law Association, the UK judiciary, the office of the Northern Ireland Lord Chief Justice and leading academic commentators on this issue.²³ The meeting debated and supported the coalition's proposal that significant structural change should be preceded by an independent expert review of the region's wider arrangements for environmental governance, on the grounds that decades of Direct Rule and conflict had significantly distorted this landscape. Direct Rule ministers indicated their support in principle, but were initially unwilling to instigate such a major process of reform because of the expectation that devolution would be imminently restored. However, by 2005, against a hinterland of apparently intractable deadlock in constitutional negotiations, the Direct Rule Environment Minister (then Jeff Rooker) announced his support for the creation of an independent Environment Agency and his decision to commission the independent review of environmental governance demanded by ENGOs.²⁴

The Review of Environmental Governance (REGNI) formally commenced in January 2006.²⁵ Consistent with its terms of reference, it was conducted in a participative manner and involved recorded public meetings with an extensive range of the key stakeholders to this debate, spanning not only civil society but also business, industry, almost all of Northern Ireland's political parties, government advisory bodies, local and central government and the environmental regulator itself.²⁶ The final report of the review, *Foundations for the Future: A Review of Environmental Governance*, was published in June 2007.²⁷ Although its recommendations spanned the entire governance regime, the report's core findings concerned the process of environmental regulation. The review concluded that the

20 R Macrory, *Transparency and Trust: Reshaping Environmental Governance in Northern Ireland* (2004) <www.ucl.ac.uk/laws/environment/docs/NL_report.pdf>.

21 L Fawcett, *Environmental Protection Consultation: Analysis of Responses* (2004) <www.foe.co.uk/resource/reports/liz_fawcett_report.pdf> at 2. Responses were received from 42 NGOs, 39 individuals, 14 private sector companies and organisations, four political parties and three public sector organisations.

22 Ibid 8.

23 Under Chatham House rules individual delegates cannot be named.

24 Friends of the Earth Northern Ireland Newsletter, 'EPA Inquiry Begins' (2005) (11 Autumn) <www.foe.co.uk/resource/newsletters/northern_ireland_11.pdf>.

25 The panel was comprised of Professors Tom Burke (chair), Sharon Turner and Mr Gordon Bell (then the recently retired CEO of Liberty IT).

26 For a full discussion of the process see, Turner (n 2).

27 REGNI (n 2).

present institutional arrangements for environmental regulation in Northern Ireland did not reflect the standards expected of modern environmental governance. First and foremost, it pointed out that the regulator's ability to command public confidence depended on its ability to act, and to be seen to act, in a consistent, fair and predictable manner.²⁸ However, the review emphasised that the 'constant message' to the panel was of a loss of public confidence that this was in fact the case.²⁹ Secondly, the review pointed out that officials working within the regulator are exposed to both a real and perceived conflict of interest in that, as departmental civil servants, they are accountable, through the Departmental Permanent Secretary to Ministers; thus their first priority is to serve the minister not the environment.³⁰ It concluded that this situation inevitably raised the suspicion and the risk that regulatory judgment might be tailored to suit immediate political circumstances.³¹ Thirdly, the report concluded that extensive audit of the regulator's function carried out in recent years by the regional scrutiny community had created a persuasive body of evidence confirming the misplaced nature of the argument that the regulator's position within central government enabled it to exert greater influence on policy-making.³²

In view of the prime importance of restoring public confidence in this critical governance process, REGNI recommended that responsibility for regulation should be externalised to an independent environmental protection agency and set out proposals for how the new entity should be structured. However, *Foundations for the Future* also recommended that significant steps should be taken to strengthen the regulator's ability to enforce the rule of environmental law. In this regard, it stated that one of the key operational characteristics the new agency should endeavour to rapidly acquire and be recognised as having acquired is the capacity to prosecute non-compliance when appropriate.³³ To this end the review recommended the creation of an integrated enforcement office within the agency, with control of prosecutorial decision-making and its own dedicated legal staff.³⁴ However, the review also recommended that reform of the regulator be complemented by improved judicial handling of environmental prosecutions. Although it acknowledged that equivalent research on environmental sentencing to that undertaken in Great Britain did not exist in Northern Ireland,³⁵ the report nevertheless concluded that sufficient evidence of unacceptably low and inconsistent sentencing could be gleaned from official scrutiny analysis to merit the launch of a comprehensive programme of training by the Judicial Studies Board and consideration of the case for a specialised Environmental Tribunal for Northern Ireland.³⁶

Four months later a second independent analysis of the regulator was published that chimed strongly with the messages conveyed by REGNI, this time by the CJI following a year-long investigation of the DOE's use of its criminal justice powers.³⁷ Like REGNI, the CJI recommended a strengthening of the regulator's approach to enforcement. It

28 REGNI (n 2) para 5.2.

29 Ibid para 5.3.

30 Ibid para 5.11.

31 Ibid.

32 Ibid para 5.13.

33 Ibid para 5.37.

34 Ibid.

35 Ibid para 9.10.

36 Ibid paras 9.10–20.

37 The report contained an analysis of the performance of the department's three Executive Agencies, which included the environmental regulator but also the Planning Service and the Driver and Vehicle Agency: CJI, *Enforcement in the Department of the Environment* (2007) <www.cjini.org/TheInspections/Inspection-Reports.aspx>.

highlighted not only the regulator's overwhelming reliance on a compliance-orientated approach to regulation³⁸ but also that it shared the wider culture of the other executive agencies within the department in that it viewed enforcement as a peripheral activity.³⁹ The inspectorate strongly urged a significant rebalancing of regulatory effort to ensure a more explicit and robust approach to enforcement and particularly to criminal prosecution. To this end it made extensive recommendations designed to ensure a far more integrated and transparent approach to enforcement and build capacity through investment in specialised skills and the forging of more effective relationships with key agencies within the wider enforcement community. However, it was also clear that the inspectorate's report was influenced by both the political shift that had led to REGNI and also its recommendations. While the CJI stopped short of joining the review and indeed other scrutiny and advisory bodies who by this stage had articulated their support for externalisation of the regulatory responsibility,⁴⁰ its report nevertheless explicitly emphasised the need for clear procedures to ensure the independence of the enforcement function from political and other internal or external pressures.⁴¹ The inspectorate furthermore reflected not only an awareness of rising public and ENGO demand for more rigorous enforcement of environmental law⁴² but also an explicitly stated expectation that REGNI would lead to significant reform and the view that it was thus timely to recommend changes to the delivery of the regulator's enforcement function.⁴³

The impact of restored devolution

Although both REGNI and the investigation by the CJI had been commenced during Direct Rule, their final reports were launched into a totally changed political context. A breakthrough in constitutional negotiations in 2006 had led to the signing of the St Andrews Agreement and the restoration of devolution in early 2007. However, while the first power-sharing Executive was led by the relatively moderate UUP/SDLP, this time it was led by the political polar opposites of Sinn Féin and the Democratic Unionist Party (DUP). Successive Assembly elections held in the run-up to restoration had resulted in the decimation of the political middle ground and concentration of power in the extremes of Northern Ireland's political spectrum. The running of the d'Hondt process⁴⁴ resulted in Northern Ireland's largest party taking control of the DOE, thus leaving the DUP with responsibility for responding to the recommendations made by these reports. In June 2007, *Foundations for the Future* was formally presented to the new devolved Minister for the Environment (then Arlene Foster). Although the two largest parties in the power-sharing Executive were the only two who did not submit evidence to the review process, Sinn Féin had made public its support for the creation of an independent regulator; albeit on an all-

38 CJI (n 37) para 2.16.

39 Ibid xi.

40 House of Commons Select Committee on the Environment, *Environmental Issues in Northern Ireland*, HC 39 (Session 1990–1991); Northern Ireland PAC (n 7) para 42; NIAC (n 7) para 101; NIAO (n 7) paras 4.7–9; House of Commons Committee of Public Accounts (n 7) para 21; Waste Management Advisory Board, *Northern Ireland Waste Management Strategy Review Report* (2004) <www.doeni.gov.uk/nica/wastemanagementstrategyreview.pdf>.

41 CJI (n 37) ix.

42 Ibid paras 2.19, 5.3.

43 Ibid vii.

44 Through which executive responsibility is allocated under the power-sharing arrangement. See, generally, D Horowitz, 'Explaining the Northern Ireland Agreement: The Sources of Unlikely Constitutional Consensus' (2002) 32(2) *British Journal of Behavioural Science* 193–220.

island basis.⁴⁵ The DUP, however, had remained ominously non-committal as to its position on the status of the regulator.⁴⁶

In September 2007, the Northern Ireland Assembly made clear the strength of its support for independent regulation through the debate of a motion submitted by the Leader of the Alliance Party (David Ford, now Minister for Justice), 'call[ing] on the Executive to establish an independent Environment Agency',⁴⁷ which was adopted without division, and included the defeat of a DUP attempt to amend the motion so as to deflect and postpone an unequivocal statement of political support for independent regulation.⁴⁸ However, Foster waited a year before making her formal response to the review report, which came in the form of a statement to the Assembly made in May 2008.⁴⁹ She also used this opportunity to make her initial response to the recommendations published by the CJI. The minister informed the Assembly that she and her party took the role of environmental governance 'too seriously' to externalise responsibility for this function to an outside agency and thus that it would not be transferred to an independent entity.⁵⁰ For the reasons discussed below, the DUP's rejection of independent regulation was not unexpected; however, it was nevertheless intensely controversial. Quite apart from the fact that it signalled a halting of its government engagement with the ENGOs concerning the need for far-reaching regulatory reform, the reasons motivating this decision, and even more fundamentally the manner in which it was handled, provided an unsettling illustration of the style of governance that power-sharing had unleashed and also how its distinctive dynamics would shape the regional environmental agenda.

THE POWER DYNAMICS OF DEVOLUTION

First and foremost, the DUP's ability to impose what was effectively a unilateral decision, opposed by all of the other political parties sharing power, made real the well-documented anticipation that power sharing would lead not only to the carving up of power but also the replacement of one form of compromised accountability with another.⁵¹ The Executive Committee formed by the Northern Ireland Act 1998 to exercise executive authority under devolution comprises each of the departmental ministers and the First and Deputy First Ministers drawn from the five political parties sharing power.⁵² Although the Ministerial

45 Gerry Adams MLA, President of Sinn Féin provided his endorsement of the externalisation of regulation to an independent agency <www.foe.co.uk/resource/evidence/ni_epa_stakeholder_endorsement.pdf>.

46 DUP's 2007 election manifesto did, however, suggest the party was 'open' to considering the best mechanism through which to implement environmental law, at DUP, *Getting it Right* (2007) <www.dup.org.uk/pdf/DUPManifesto2007LR.pdf> at 57.

47 Minutes of the Northern Ireland Assembly, Private Member's Business, 25 September 2007 <<http://archive.niassembly.gov.uk/record/reports2007/070925.htm#4>>.

48 25 September 2007. Peter Weir MLA (DUP) sought an amendment of the motion which would have called on the Assembly to simply 'note' REGNI's recommendation for independent regulation but would have called for further work to be undertaken to identify the costs and benefits of structural reform before a decision could be taken.

49 Full Ministerial Statement on Environmental Governance <www.northernireland.gov.uk/news/news-doe/news-doe-may-2008/news-doe-270508-environment-minister-cuts.htm>.

50 Ibid 3.

51 In other words, that devolution involved the replacing of the democratic deficit inherent in Direct Rule with the democratic deficit posed by power-sharing. For further discussion of this issue, see, for example, R Wilford and R Wilson, *A Democratic Design? The Political Style of the Northern Ireland Assembly* (Constitution Unit, University College London 2001) <www.ucl.ac.uk/spp/publications/unit-publications/74.pdf>; R Wilson, *The Northern Ireland Experience of Conflict and Agreement: A Model for Export?* (Manchester University Press 2010); R Wilford, 'Northern Ireland: The Politics of Constraint' (2010) 63(1) *Parliamentary Affairs* 134.

52 C Knox, *Devolution and the Governance of Northern Ireland* (Manchester University Press 2010) 24.

Code of Practice⁵³ governing the committee's functioning is designed to inhibit the potential for ministerial 'solo runs', it nevertheless ensures only a very pale imitation of the Westminster concept of cabinet responsibility.⁵⁴ On the one hand, the code provides that ministers must bring issues to the attention of the Executive for collective consideration in certain situations, including any matter that 'cuts across the responsibilities of two or more Ministers' or, those deemed 'significant or controversial' by the First and Deputy First Minister acting jointly.⁵⁵ While the status of the regulator had certainly become a 'significant' and 'controversial' issue by 2008, the DUP's trenchant opposition to independent regulation made it unlikely that the DUP First Minister (then Ian Paisley) would have agreed to act jointly with his Sinn Féin Deputy First Minister (Martin McGuinness) to require Arlene Foster to bring this matter to the Executive Committee. However, decision-making concerning the status of the environmental regulator was also an inherently cross-cutting issue, particularly in the Northern Ireland context where regulatory responsibilities are exercised by at least three other government departments in addition to the Department of the Environment.

While there was little doubt that Arlene Foster was bound by the Ministerial Code to bring this matter to the committee for collective consideration, it was also clear that the DUP retained the power to impose a unilateral rejection of independent regulation despite universal political support for externalising this responsibility. Although the code requires the First and Deputy First Ministers to seek to ensure that decisions of the committee are reached by consensus, it also provides that where consensus proves impossible, a vote can be taken. However, where it is requested by any three ministers, the code also provides that the vote must be taken on a 'cross community basis', the rules for which require a weighted majority of both unionist and nationalist ministers.⁵⁶ It was already clear that consensus on this issue was unlikely. Because the DUP then held five of the committee's 12 ministerial positions, had the issue been put to a vote, the party could insist that it be taken on a cross-community basis and, because they comprised over 40 per cent of its unionist membership, could defeat both a majority and even unanimous support for independent regulation within the Executive Committee.

However, despite concerted efforts by members of the Alliance Party to force a formal clarification of how the decision to reject independent regulation had been made, the Executive Committee refused to confirm that it was an issue that should have been brought to its attention for collective consideration, or even to clarify whether a vote had been taken. During the Assembly debate of Foster's decision to reject independent regulation, Alliance Party MLAs⁵⁷ argued that this decision was inherently cross-cutting and thus sought clarification as to whether it had been brought to the Executive Committee for collective consideration.⁵⁸ In response, Foster stated that while the committee had been informed of her decision 'out of courtesy', oddly she did not consider the matter to be a cross-cutting

53 <www.northernireland.gov.uk/index/ministerial-code.htm>.

54 Knox (n 52) 25 and Wilford (n 51) 146.

55 Knox (n 52) para 2.4(i) and (vi).

56 Para 2.12 of the Code of Practice provides that a vote taking on a cross-community basis shall be governed by s 4(5) of the Northern Ireland Act 1998, which in turn provides that 'cross-community support', in relation to a vote on any matter, means either (a) the support of a majority of the members voting, a majority of the designated Nationalists voting and a majority of the designated Unionists voting; or (b) the support of 60 per cent of the members voting, 40 per cent of the designated Nationalists voting and 40 per cent of the designated Unionists voting.

57 Member of the Legislative Assembly (NI).

58 Assembly debate <<http://archive.niassembly.gov.uk/record/reports2007/080527.htm#4>> 7 (David Ford, Alliance Party) and 12 (Dr Farry, Alliance).

one⁵⁹ – despite emphasising in a later part of her statement that externalising this responsibility would affect the powers of several other Northern Ireland government departments.⁶⁰ A few days later, 30 MLAs supported another initiative mounted by David Ford MLA, this time a petition submitted under s 28B of the Northern Ireland Act 1998, which empowers Assembly members to seek a referral of a minister's decision to the Executive Committee where they are considered to have acted in breach of the code, or where a decision relates to a matter of public importance. Following consultation with Assembly parties, the Speaker formally confirmed that the matter had been deemed 'a matter of public importance' and thus that a valid petition had been lodged.⁶¹

Although MLAs were successful in forcing the Executive to formally consider the issue, the rules governing the matters that must be addressed by the committee when responding to an Assembly referral rendered this a pyrrhic victory. Consistent with s 28B, the committee simply confirmed its view that the minister's decision was not taken in contravention of the Code of Practice. It furthermore confirmed the committee's view that decisions relating to environmental governance were 'significant and controversial' and also that it 'had noted' Foster's decision to reject independent regulation. Crucially, the committee was not required to and did not take the opportunity to clarify whether the minister's decision had been taken with the Executive's support, or indeed even if a vote had been taken. Rather bizarrely, its response went on to state that future decisions by Foster's successor, concerning the implementation of the alternative reform agenda she proposed (outlined below), would be brought to the committee's attention on the grounds that they would require its specific approval under the Ministerial Code and s 28A of the Act. The extraordinary obfuscation surrounding how the decision to reject independent regulation was made effectively concealed the unedifying reality that had Foster been forced to comply with the terms of the Ministerial Code, the DUP could simply have used its tribal veto to impose a minority, unilateral position in the face of universal opposition by its partners in the power-sharing Executive. To add insult to political injury, Arlene Foster was furthermore shielded by her party from the consequences of this intensely unpopular decision by an immediate transfer to a new and more senior ministerial position leading the Department of Enterprise, Trade and Investment.⁶²

However, quite apart from what this process revealed as to the nature of power-sharing and in particular the very constrained form of democratic accountability it entailed, it also underlined the DUP's pronounced antipathy to participative governance.⁶³ Although Foster could not avoid making a formal written response to the recommendations of the CJI given that it was a report published by a statutory scrutiny body, the DOE did not provide the detailed response to the review of environmental governance that would normally have been expected upon receipt of a report commissioned by government. Indeed, it even failed to notify the chair of the review panel of the minister's intention to make a statement to the Assembly. However, Foster underlined her rejection of deliberative decision-making, and in particular the ENGO coalition, by also announcing that the Environment and

59 Ibid 7 (Minister for the Environment).

60 Ibid discussed below.

61 Under s 28B(3), confirmed on 9 June 2008, Speaker's Business <<http://archive.niassembly.gov.uk/record/reports2007/080609.htm#2>>.

62 13 days later, on 9 June 2008.

63 Although the DUP is not alone in this position amongst the region's political parties, its particular hostility to the participative governance has been addressed by: C McCall and A Williamson, 'Governance and Democracy in Northern Ireland: The Role of the Voluntary and Community Sector after the Agreement' (2001) 14(3) *Governance: An International Journal of Policy and Administration* 363; and V Bell, 'Spectres of Peace: Civic Participation in Northern Ireland' (2004) 13(3) *Social and Legal Studies* 408.

Heritage Service would be relaunched as the 'Northern Ireland Environment Agency' (NIEA).⁶⁴ The cynical suggestion that the ENGO coalition's demands for independent regulation could be satisfied through a simple rebranding exercise was widely criticised, but, combined with the DUP's decision to replace Arlene Foster as Environment Minister with Sammy Wilson, an avowed opponent of the environment lobby and vocal climate sceptic,⁶⁵ represented not only a powerful snub to the ENGO coalition but also an ultimately successful attempt to suppress this nascent lobby.

THE IMPACT OF 'HIGH' CONSTITUTIONAL POLITICS

Although Foster's rejection of the recommendation for independent regulation was largely an exercise in the use of crude political power, it was nevertheless clear that the decision was also strongly influenced by the 'high' politics of the constitutional transition, which concerned the commandeering of political power by the new administration but also the inherent fragility of power-sharing itself. From the outset of her statement to the Assembly, Foster characterised both REGNI and the CJI investigation as processes commenced during the Direct Rule period thus obviating a sense of political ownership as to their findings.⁶⁶ However, the pronounced hostility directed towards the recommendations made by REGNI also stemmed from the distinctive political dynamics that had characterised the constitutional transition. Whereas policy stagnation had for decades been the traditional angle of repose under Direct Rule, the administration that governed Northern Ireland following the collapse of power-sharing in 2002, and particularly during Peter Hain's tenure as Secretary of State from 2005–2007, was defined by the proactive development of a series of major policy initiatives, notably the abolition of the 11+ educational selection process, the introduction of domestic water charges based on the capital value of (then) escalating house prices, and proposals to significantly curtail Northern Ireland's historically lax rural development policies.⁶⁷ All touched core cultural and economic nerves within the region.⁶⁸

However, while the environment and the environment lobby were undoubted beneficiaries of this 'policy spring', the agricultural industry and landowners were notable losers, and at a time when advanced EU infraction pressure had also forced Northern Ireland to implement the EU Nitrates Directive on a 'total territory' basis, thus radically

64 Ministerial Statement (n 49) 3.

65 Sammy Wilson's hostility to the environment and climate lobby is well documented, for example, referring to the campaign to halt climate change as an 'hysterical pseudo-religion' at *BBC News Online*, 'Wilson Row over Green Alarmists', 5 September 2008 <<http://news.bbc.co.uk/1/hi/7599810.stm>>. His position in this regard has been furthermore covered in the *ENDS Reports* 401 and 402 and by *The Guardian*, 10 February 2009 <www.guardian.co.uk/politics/2009/feb/10/climate-change-sceptic-environment-minister>. Wilson's prioritisation of development over environmental protection in the field of planning is also well documented, for example: 'Government policy and decision making has been heavily influenced by the green lobby, which I believe has been detrimental to the British economy.' This position statement on the environment and a number of other areas is outlined on his DUP website <www.sammywilson.org/issues/>.

66 Ministerial Statement (n 49) 1.

67 The scale of opposition to more constrained rural development is outlined at <www.planningni.gov.uk/index/policy/policy_publications/planning_statements/pps14/pps14_background.htm>.

68 Some sense of the scale of the political furore surrounding the planned introduction of domestic water-charging is provided in the Direct Rule administration's summary of the initial consultation responses at: Department of Regional Development, *Integrated Impact Assessment of the Government's Proposals for Reform of Water and Sewage Services* (2005) <www.drdni.gov.uk/iaa_4-web_7_.pdf>, although this was also subject to intense media coverage and stakeholder litigation. Hain's decision to abolish the 11+ is set out at <www.nio.gov.uk/media-detail.htm?newsID=13172>. Intense unionist opposition to this policy even threatened to destabilise power-sharing. A flavour of this debate is provided by H McDonald, 'Ulster Unionists Demand Deal on 11plus before Backing Power Sharing', 12 February 2010, *The Guardian* <www.guardian.co.uk/politics/2010/feb/12/ulster-unionists-demand-11plus-powersharing>.

intensifying this industry's exposure to the costs and process of environmental regulation.⁶⁹ In part, policy modernisation was motivated by UK Treasury demands that Northern Ireland pay its way and by Hain's particular interest in the environment. However, it was also perceived as a mechanism for mobilising local public pressure to render detachment from devolved power increasingly uncomfortable for local political parties, thus assisting in leveraging an end to the deadlock in constitutional negotiations that had kept devolution in suspension for five years. Not surprisingly, when devolution was finally restored in March 2007, environmental policy modernisation and the environment lobby became the subject of a pronounced political backlash from local politicians. The rejection of the high-profile campaign for independent regulation, like the immediate stalling of the other major policy initiatives progressed during Hain's term in office,⁷⁰ became part of the political narrative that defined the restoration of devolution; namely, that power had been wrested from the resented Direct Rule administration. However, this decision was also used by the DUP to consolidate its particular relationship with one of its key political allies; namely, the agricultural industry, whose representative body (the Ulster Farmers' Union) had been alone in opposing the creation of an independent environment agency.

A final but critical factor militating against an immediate commitment to create an independent regulator concerned its potentially destabilising implications for the fragile arrangements created to support power-sharing. Although a reality only hinted at in the minister's statement, there was little doubt that externalising the Environmental and Heritage Service would have involved the removal of a significant part of the critical mass of the DOE thus undermining its viability as political portfolio. In addition, as Foster herself pointed out, three other Northern Ireland departments also exercised significant duties in this context.⁷¹ The creation of a coherent independent regulator would have confronted an inherently unstable administration with the immediate prospect of renegotiating how executive responsibilities should be reallocated across central government, whose original distribution had been informed not by concern to maximise policy synergies or regulatory outcomes, but to ensure that no one political party dominated the control of major policy sectors.⁷² Indeed, though many were sceptical of Foster's stated commitment to instigate an independent review of her decision in 2011, it was arguably an implicit recognition of the reality that a reconfiguration of central government would be more appropriate in the event that power-sharing proved sufficiently stable to survive at least the period of the first mandate.

A new reform agenda emerges

However, though Foster rejected the case for structural reform, she nevertheless appeared to acknowledge the need for improved regulation and used her statement to the Assembly to set out her own agenda in this regard. First, the minister acknowledged the need to improve the enforcement response to environmental crime⁷³ and, secondly, announced her commitment to ensuring that environmental regulation would henceforth be premised

69 91/676/EEC. OJ L375/1. The significance of this approach to implementation is discussed by Turner (n 4).

70 Amongst the very first announcements made by the new devolved administration was its decision to halt the imposition of domestic water charges and to launch an independent review of the policy, 10 May 2007 <www.iwrp-ni.org.uk/index/background.htm>. This was followed shortly afterwards by the launch of a review of Draft PPS 14 on rural development <www.planningni.gov.uk/. . . /ministerial_statement_pps14_251007.pdf>. The battle concerning the abolition of the 11+ continues to be a source of significant disagreement between Sinn Féin and unionists and remains unresolved to the present day.

71 Ministerial Statement (n 49) 2.

72 Knox (n 52), ch 2.

73 Ministerial Statement (n 49) 6.

firmly on the principles of 'better regulation'.⁷⁴ However, Foster also explicitly reassured the Assembly that she wished to 'see clear blue water between the role of the core Department as policy maker and legislator and the role of the NIEA as protector, regulator and enforcer'.⁷⁵ Thus, she implicitly recognised that restoring the regulator's credibility would depend fundamentally on how well the DOE could assure the public and regulated community of the impartiality of regulatory decision-making. Although her immediate successor (Sammy Wilson) dismissed further calls by the Assembly later in 2008 for a commitment to externalise regulatory responsibility,⁷⁶ arguing that a further review of the 'new' Agency was premature,⁷⁷ sufficient time has now elapsed to consider whether the Ulster Farmers' Union were justified in their conviction that effective regulation could be ensured through a process of internal reform.

SIGNS OF PROGRESS

For almost all of the period since Foster's statement, the DUP remained in control of the DOE. Despite the disintegration of the ENGO coalition following the powerful rejection of its reform agenda, Foster and her successors (Sammy Wilson and Edwin Poots) worked to implement a reform agenda, which produced important advances in delivering a more robust regulatory response to the region's distinctive problems with waste crime. Whereas the sophisticated standards of waste regulation required under EU law had been implemented in Britain by 1990, over a decade elapsed before the equivalent controls were operationalised in Northern Ireland and then only in response to EU enforcement action.⁷⁸ In addition to pressure to come to terms quickly with its new regulatory responsibilities, the regulator also faced an escalating problem of illegal transfrontier shipments of waste into the jurisdiction. By 2002, the combination of unprecedented volume of waste arising from Ireland's then 'tiger' economy, differences in the landfill taxes applying on each side of the Irish border, weak environmental regulation in both jurisdictions and an extensive land border had created the conditions for a highly profitable black market in illegal cross-border waste transport and dumping, which was being exploited on a significant scale by organised criminal gangs with possible paramilitary links.⁷⁹

In order to respond to criticism of the regulator's failure to deal effectively with this problem,⁸⁰ the Environment and Heritage Service established a dedicated Waste Crime Unit in 2004, which launched not only an intensive programme of criminal prosecutions⁸¹

⁷⁴ Ministerial Statement (n 49) 3–4.

⁷⁵ Ibid 6.

⁷⁶ At the end of 2008 another Alliance Party motion was adopted by the Assembly (although amended by the SDLP), reiterating its concern at the minister's failure to create an independent regulator, expressing its view that the present entity was unable to meet the environmental challenges facing the region and calling for externalisation by the end of the first mandate: 11 November 2008 <<http://archive.niassembly.gov.uk/record/reports2008/081111.htm#4>>.

⁷⁷ Ibid.

⁷⁸ The legislative framework for applying modern waste management licensing in the jurisdiction was introduced with the enactment of the Waste and Contaminated Land (NI) Order 1997, but the transfer of responsibility for waste regulation from local government to the Environment and Heritage Service was not operationalised until 2003, Turner (n 4).

⁷⁹ Discussed in particular by NIAC (n 7) para 88, Ev 203, Ev 6. Recent estimates suggest that between 2001–2004, approximately 8 per cent (approximately 250,000 tonnes) of the waste generated in Ireland was illegally dumped across the border in Northern Ireland. See T Hogan, 'Files will be sent to DPP on Illegal Dumping', *Irish Independent*, 7 May 2012 <www.independent.ie/national-news/files-will-be-sent-to-dpp-on-illegal-dumping-3101771.html>.

⁸⁰ Ibid NIAC. See also the discussion provided by the scrutiny reports on waste management (n 7).

but also began to make innovative use of new financial investigation powers contained in the Proceeds of Crime Act 2002 in order to strip the profits from serious waste crime. Following an expansion of the range of statutory bodies permitted to use these powers in 2005,⁸² the DOE had applied for the accreditation of officials within the Waste Crime Unit.⁸³ Working initially with the regional Assets Recovery Agency (ARA) from 2005–2008, the Environment and Heritage Service became the first environmental regulator in the UK to obtain a criminal confiscation order against a defendant convicted of serious waste crime,⁸⁴ and by 2008 was the first such body to employ fully qualified financial investigators. Foster's statement to the Assembly highlighted these successes and committed to translating them across the agency as a whole. To this end she committed to establishing a dedicated Environmental Crime Unit (ECU),⁸⁵ designed to provide the integrated enforcement capacity called for by REGNI and the CJI. In addition, she committed to a substantial programme of investment to support recruitment and training to develop its capacity.⁸⁶ Although the dawning age of national austerity ultimately inhibited the full investment promised by Foster,⁸⁷ her successor Sammy Wilson oversaw the creation of the new ECU later in 2008. This unit now leads the UK in the use of financial legislation to strip the profits from serious waste crime.⁸⁸ It has furthermore made concerted efforts to embrace the more sophisticated and intelligence-led approach to enforcement called for by the CJI, forged close partnerships with its Irish and UK counterparts and the specialised agencies engaged in tackling serious organised crime within the region,⁸⁹ and has invested significantly in information management systems designed to enable it to collate data concerning those engaged in waste crime.

81 There were 24 prosecutions for waste offences in 2004, 61 in 2005 and 116 in 2006. This fell to 68 in 2007, but increased slightly in 2008 to 74. These figures are based on information provided by the NIEA in 2011.

82 Pursuant to the Proceeds of Crime Act 2002 (References to Financial Investigators) (Amendment) Order 2005, No 386.

83 Accreditation was received from the National Policing Improvement Agency and further discussed by House of Commons NIAC, Third Report, *Fuel Laundering and Smuggling in Northern Ireland*, HC (2010–2012) 1504, Q410, Ev 72.

84 The case involved the illegal dumping of around 4000 tonnes of municipal waste from the Irish Republic in the early 2000s, which resulted in the making of a confiscation order of £80,868 in September 2007 following the defendant's conviction in May 2006 for two counts of keeping and disposing of waste [2007] NICC 53.

85 Ministerial Statement (n 49) 6.

86 Ibid 3. It should be noted Foster did not disaggregate between the investment directed towards better regulation and that directed towards the ECU, but committed to a total investment of £0.77m in 2008–2009 and £1.98m in 2010–2011.

87 While the ECU was designed to operate with a staff of 41, as of December 2011 only 25 staff were employed by the unit and the majority of vacancies existed at senior levels, although in early 2012 the new Minister for the Environment, Alex Attwood (SDLP), committed to hiring a further 11 staff. See *BBC News Online*, 'Northern Ireland environmental crime unit in recruitment drive', 23 December 2011 <www.bbc.co.uk/news/uk-northern-ireland-16321099>.

88 In partnership with the former Assets Recovery Agency (now the Serious Organised Crime Agency (SOCA)), NIEA secured a total of £833,120 in five confiscation cases. At the time of writing, NIEA's ECU has independently secured 10 confiscation orders subsequent to Crown Court convictions which total £862,300. In total, NIEA has through 15 confiscation cases secured £1,695,420 from convicted waste criminals. The confiscation regime has not yet been applied to environmental crime in Scotland thus far. Though the Environment Agency (EA) has recently started to make use of these powers, they have been used to a relatively lesser extent than in Northern Ireland. Information provided by the EA in March 2012 indicates that from 2006–2011, 38 confiscation orders have been made subsequent to EA waste prosecutions, but the vast majority (22) of these were in 2011 and 10 of those cases were for significantly smaller sums of under £5000. Source: information provided to the authors subsequent to requests made to the agencies under the Freedom of Information Act 2000.

89 In particular, the SOCA, HM Revenues and Customs, the Police Service of Northern Ireland and the Ports Authority.

The second significant advance achieved under the leadership of DUP Environment Ministers concerns the DOE's success in forcing Ireland to accept its primary responsibility under EU Regulation 1013/2006 on the Shipment of Waste⁹⁰ to repatriate waste illegally transported into the jurisdiction during the early part of this decade. Despite the DUP's constitutional hostility to the North South Ministerial Council, which had been created by the Good Friday Agreement to facilitate all-island co-operation in specified areas, Foster used this mechanism to negotiate a joint 'Road-Map' with her Irish counterpart to govern their respective responsibilities for the repatriation of this waste. The leveraging of agreement was undoubtedly aided by the application of legal pressure exerted by the EU Commission, which had commenced enforcement proceedings against the two governments due to their mutual failure to adequately control illegal transfrontier shipments of waste,⁹¹ and summonsed both to attend a trilateral meeting to discuss their plans for ensuring compliance. That said, the Road Map provided the basis of a successful defence to this litigation and the adoption of the formal bilateral framework agreement signed two years later by the two Environment Ministers under which Ireland formally accepted total responsibility for the cost of disposing of illegally dumped waste and 80 per cent of the cost of excavation, remediation and removal.⁹² Although the process of repatriation remains ongoing and is likely to take some years to complete, to date over 50,000 tonnes (out of an estimated 250,000) of waste have been transferred to Ireland at a cost of upwards of €30m to the Irish taxpayer.⁹³

A PARTIAL AND POLITICALLY MOTIVATED COMMITMENT TO ENFORCEMENT?

However, despite the undoubted strides made in the context of enhanced waste enforcement, Foster's reform agenda has not been an unmitigated success in so far as the agency as a whole is concerned. Despite its official title, the new ECU was effectively a relaunching of the Waste Crime Unit. Instead of delivering the promised integration of enforcement efforts, its creation has intensified the fragmentation of enforcement efforts, a problem raised as a serious weakness by the CJI only a year earlier⁹⁴ and acknowledged by the DOE.⁹⁵ Responsibility for enforcing water and nature conservation controls remained with the Water Management Unit and a unit within the Natural Heritage Directorate respectively,⁹⁶ while responsibility for waste enforcement was actually sub-divided between the new ECU and the Land and Resource Management Unit (LRMU), with the latter retaining responsibility for the enforcement of waste management licences.⁹⁷ However, the

90 OJ L190/1.

91 IMPEL UPDATE (European Network for Implementation and Enforcement), June 2007 <www.rijksoverheid.nl/bestanden/documenten-en-publicaties/brochures/2007/08/01/newsletter-enforcement-actions-third-edition/w986.pdf>.

92 <www.northernireland.gov.uk/news/news-doe/news-doe-june-2009/news-doe-120609-minister-announces-agreement.htm>.

93 Hogan (n 79).

94 CJI (n 37) 5, para 2.3.

95 DOE, *Action Plan in Response to Criminal Justice Inspectorate Report on Enforcement within DOE* (2008), at <www.cjini.org/CJINI/files/b1/b169ceda-4e37-4448-b4a4-40e23b3f2419.pdf>.

96 As this article went to press, the authors were made aware that the ECU is in the process of expanding its portfolio of enforcement activity to include referrals from other units within DOE dealing with areas such as built and natural heritage crime. The nature and significance of these very recent changes will be analysed in a forthcoming paper by the present authors.

97 Although officials within ECU have indicated that its work more latterly includes cases against the holders of waste licences, they have also confirmed that the defendants are largely associated with organised criminal networks or operating on the periphery of the legitimate waste industry. It is worth noting that ECU and LRMU are currently negotiating a protocol to separate their respective responsibilities for enforcement.

concentration of investment in enhancing the ECU's enforcement capacity has also created gaps and distortions in the agency's wider regulatory response, even in the context of waste enforcement. In 2007, a decade after the introduction of modern waste management standards, and a decade after the transfer of primary responsibility for waste regulation to the agency, the scale of its failure to implement this new regime was revealed by the CJI which reported that almost all of the landfill sites in Northern Ireland continued to operate without either planning permission or a waste management licence.⁹⁸ Several other scrutiny reports were furthermore highly critical of the DOE's failure to finalise a protocol delineating the respective responsibilities of local government and the regulator for dealing with fly-tipping.⁹⁹

However, despite the manifest need for balanced investment in supporting the agency in tackling the full spectrum of waste crime occurring in the region, and in coming to terms with its complex new responsibilities in this context, investment in enhanced waste enforcement has instead concentrated almost exclusively on enabling the ECU to respond to the enforcement failure for which the DOE was most severely criticised; namely, to take action against those responsible for illegal dumping on a commercial scale.¹⁰⁰ This very narrow focus, combined with the ECU's concentration on using resource-intensive financial investigation procedures as the primary means of sanctioning serious waste crime, has resulted in a dramatic decline in the rate of prosecution and in the number of cases the unit can investigate at any one time. Although its earlier incarnation as the Waste Crime Unit had pioneered the use of financial investigation, it had also demonstrated a capacity to take a relatively high number of prosecutions, which rose from a standing start in 2003 to 116 prosecutions in 2006. In effect, the outworking of Foster's investment in the ECU has resulted, on the one hand, in increased investment in a specific form of enforcement expertise but, on the other, in a steep decline in the number of prosecutions being taken in relation to waste crime. By 2010, only 35 waste prosecutions were heard by the Northern Ireland courts. A year later the CJI published a follow-up to its 2007 report,¹⁰¹ which expressed concern about this pronounced narrowing of waste enforcement effort. While it acknowledged the unit's improved investigative approach to serious waste crime, the CJI also expressed concern that the concentration on using resource-intensive confiscation of assets powers as the primary approach to sanctioning meant it could only handle 16 live cases at any one time.¹⁰² However, the CJI also raised important concerns about the negative collateral impacts of this strategy on the rest of the agency's enforcement capacity, particularly on enforcement concerning the regulated waste industry. The inspectorate

98 CJI (n 37) para 2.23.

99 This issue has been raised in a number of the waste scrutiny reports (n 7) and more recently by the CJI (n 37) para 2.30. Evidence to the NIAC inquiry into fuel laundering and smuggling in January 2012 heard one MLA refer to 'passing the parcel' in terms of the blurred line between council and NIEA waste enforcement responsibilities, minutes available at <www.publications.parliament.uk/pa/cm201012/cmselect/cmniaf/uc1504-vii/uc150401.htm> at Q432>. In May 2012, the Assembly's Environment Committee again highlighted the need for a fly-tipping protocol, although it recognised progress had been made with a pilot being rolled out between NIEA and six local councils, see <www.niassembly.gov.uk/Assembly-Business/Official-Report/Reports-11-12/15-May-2012/#a5>.

100 This focus is explicitly acknowledged by NIEA in a number of publications, for example, NIEA, *Compliance and Enforcement Report* (2011) <www.doeni.gov.uk/niea/compliance_and_enforcement_report_2008_to_2010.pdf> and in an NIEA briefing on establishing an integrated environmental crime unit from August 2008, see <www.doeni.gov.uk/niea/establishing_an_ect_-_december_2008_-_no_8.pdf>. It is furthermore emphasised by the CJI follow-up report in 2011, CJI, *Enforcement in the Department of the Environment Northern Ireland: A Follow-up Review of Inspection Recommendations* (2011) <www.cjini.org/CJNI/files/d7/d71473bc-2dc9-4ff5-b957-d410ff851852.pdf>.

101 CJI (n 100).

102 Ibid para 2.15–16.

highlighted not only that most of the LRMU's enforcement staff had been transferred to the new ECU,¹⁰³ but also that this diminution of enforcement capacity has forced what remained of the LRMU to rely excessively on a compliance-based approach to regulation, characterised by a failure to prosecute even where this was acknowledged to be the most appropriate regulatory response.¹⁰⁴ Similarly, although not raised by the inspectorate, the agency's continued failure to finalise a fly-tipping protocol with local government means that a significant category of illegal waste activity remains in a well-documented and strongly criticised regulatory limbo.¹⁰⁵

While action to tackle serious waste crime was undoubtedly justified, Foster's highly partial approach to investment in enhanced enforcement raises a number of important questions not only about its regulatory outcomes but also its motivation. That reform of the agency's enforcement function was not designed to achieve better environmental outcomes is reflected in the fact that the net impact of the investment made has been to weaken the regulator's capacity to take enforcement action against the regulated waste industry, and to bring little if any positive impact on enforcement beyond the realms of waste crime. Instead, the distinctive focus of reform conveys the impression of a politically motivated enforcement agenda. Despite the minister's assurances that she wished to see 'clear blue water' between the departmental policy core and the agency's role as regulator, the exclusive investment in enhanced capacity to tackle serious waste crime, and exaggerated focus on supporting asset recovery procedures to sanction this activity, strongly reflects Foster's articulation of the reform required in that her statement to the Assembly specifically prioritised tackling serious waste crime and made clear her view that 'assets recovery is far more effective than court fines as both a punishment and a deterrent'.¹⁰⁶ The impression of a politically driven enforcement agenda is further underlined by the fact that no additional investment was made in much-needed strengthening of the enforcement of environmental law in relation to legitimate economic activity, an approach that resonates strongly with the DUP's well-documented neoliberal economic agenda.¹⁰⁷ Moreover, the NIEA's strong focus on high-profile criminal confiscation proceedings arguably serves to bolster the party's profile in stemming the revenue stream to organised crime and paramilitaries, and thus in countering destabilising forces within the region.

A LIMITED CONCEPTION OF 'BETTER REGULATION'

The second major stream of reform launched by Foster concerned her commitment to ensure that environmental regulation would henceforth be firmly premised on the UK and EU principles of 'better regulation'.¹⁰⁸ However, in this context it was clear from the outset that the NIEA would be embracing a selective conception of this paradigm, and furthermore that reform would bear the imprint of a distinctive political vision. That the shift to better environmental regulation would focus strongly on lightening the burden of regulation for industry, and in particular for the sector most trenchantly opposed to more effective regulation, was made clear by the immediate commissioning (in collaboration with the Department of Agriculture and Rural Development) of an independent review of the

¹⁰³ These included the staff previously working on financial investigation and asset recovery litigation.

¹⁰⁴ CJI (n 100) para 2.15.

¹⁰⁵ Environment Committee (n 99).

¹⁰⁶ Ministerial Statement (n 49) 5.

¹⁰⁷ B Murtagh and P Shirlow, 'Devolution and the Politics of Development in Northern Ireland' (2012) 30 *Environment and Planning C: Government and Policy* 46–61; Barry (n 18).

¹⁰⁸ Ministerial Statement (n 49) 3–4.

administrative burdens falling on the region's agri-food sector.¹⁰⁹ No equivalent review has been instigated for any other regulated industry; but this central emphasis is reinforced by the White Paper on *Environmental Better Regulation* published by the DOE in 2011, which provides a more detailed articulation of its plans concerning the embedding of this regulatory paradigm.¹¹⁰ But what this relatively brief White Paper also reflects is that the DOE simply plans to import the approach to reform used by the UK government to embed this regulatory paradigm in England and Wales without any consideration of the practical implications of doing so in a regulatory context characterised by a very weak deterrent to non-compliance.

The White Paper formally embraces the principles of better regulation developed by the Hampton¹¹¹ and Macrory analyses,¹¹² upon which the UK paradigm rests. In addition, it reflects their emphasis on the need to ensure that risk regulation is supported by an appropriately calibrated system of regulatory sanctions in order to ensure the existence of an effective deterrent. The White Paper goes on to set out a sanctioning reform agenda that is almost entirely derivative of Macrory's recommendations to the UK government in that it commits to the introduction of civil penalties and to the development and clarification of the courts' criminal justice powers. However, in doing so it strongly reflects Arlene Foster's marked unwillingness to embrace an evidence-based reform agenda in this context. Although scrutiny of the quality of environmental regulation over the past decade has focused almost exclusively on the role of the regulator, successive reports have expressed concerns about judicial handling of environmental prosecutions.¹¹³ REGNI explicitly concluded that, while the extensive empirical research conducted in the rest of the UK on environmental sentencing had not been paralleled in Northern Ireland, sufficient evidence existed to raise serious concerns about the quality of environmental justice being delivered in the jurisdiction.¹¹⁴ However, despite evidence indicating a compromised deterrent to environmental crime, and thus the likelihood that embracing risk regulation would pose distinctive challenges for the region, the DUP's determination to ignore the fact and regulatory implications of this glaring problem were clear from the outset.

During the Assembly debate of her decision to reject independent regulation in which she set out her plans for reform, the minister explained that she had sought the views of Northern Ireland's Lord Chief Justice as to the review's conclusions concerning the problematic nature of judicial sentencing.¹¹⁵ The minister referred to the letter she had received in response from the Lord Chief Justice, in which he is quoted as stating that 'The report does not provide evidence to support this statement',¹¹⁶ and furthermore that 'The very least one might have expected would have been for the group to have sought my views before making this bold statement.'¹¹⁷ By quoting this statement the minister effectively enabled the region's most senior judge to challenge the reality of inappropriate judicial

109 M Dowling, D Graham and B Jack, *The Northern Ireland Agri-Food Better Regulation and Simplification Review* (2009) <www.doeni.gov.uk/index/protect_the_environment/local_environmental_issues/better_regulation.htm>.

110 This was developed under the leadership of the most recent DUP Environment Minister, Edwin Poots. <www.doeni.gov.uk/environmental_better_regulation_white_paper.pdf>.

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

112 R Macrory, *Regulatory Justice: Making Sanctions Effective Final Report* (Cabinet Office 2006).

113 House of Commons Select Committee on the Environment (n 40); NIAO (n 6) paras 8.15–17; PAC (n 7) para 36 and Minutes of Evidence paras 211–229; NIAO (n 7) para 3.25; CJI (n 37) paras 2.66–7.

114 REGNI (n 2) paras 9.10–20.

115 Minutes of the Northern Ireland Assembly, 27 May 2008 <<http://archive.niassembly.gov.uk/record/reports2007/080527.htm#4>>.

116 Ibid.

117 Ibid.

sanctioning in environmental prosecutions. While there is little doubt that Foster was seeking to harness an influential voice in undermining the REGNI recommendations, when she added her own view that 'assets recovery is more effective than are court fines as a punishment and a deterrent',¹¹⁸ the minister arguably exonerated the judiciary for refusing to use its powers to impose meaningful sanctions for environmental crime. In doing so she effectively punctured any incentive for officials to properly consider the implications of a weak deterrent for the principles of better regulation.

In the years since Foster's statement to the Assembly, the present authors have completed the first comprehensive empirical study of judicial sentencing for environmental crime in Northern Ireland. The confines of space prevent a detailed exposition of their findings, however, suffice it for present purposes to say that the research confirms, first and foremost, that REGNI's concerns about the quality of environmental justice were well founded. An analysis of 10 years of judicial sentencing from the late 1990s to the present day makes clear not only that sentencing for environmental crime in the jurisdiction is far below the statutory maximums in all key sectors, but also that it is remarkably out of line with the equivalent process in other UK jurisdictions.¹¹⁹ The research furthermore confirms that there is little if any meaningful deterrent to environmental crime in the region, and, in some contexts, arguably an incentive not to comply. While an eroded deterrent cannot be linked entirely to the absence of meaningful sanction, the experience of pronounced and entrenched resistance on the part of the judiciary to appropriately penalise environmental crime has undoubtedly played a pivotal role in achieving this outcome. This evidence consequently raises serious questions about the environmental and economic implications of the DOE's plans to import a regulatory paradigm designed for England and Wales where significant action has been taken over years to ensure that environmental crime is taken seriously by the judiciary and where the credibility of the regulator itself is not in question.

Conclusion

It is perhaps not surprising, given the foregoing analysis that the pressure for independent regulation has not gone away. While the first full mandate of devolved government has undoubtedly delivered some improvements in the quality of environmental regulation, there is significant evidence that reform of the agency's enforcement function has actually weakened its systemic capacity in this critical context. However, even more fundamentally, little has been done to rectify the perception of a politically captured regulator; if anything, quite the opposite. That independent regulation remains a live political issue was demonstrated vividly by the immediate resurfacing of this issue when control of the DOE passed from the largest party sharing power to one of the smallest following Assembly elections in 2011. Within weeks of his appointment, the new SDLP Environment Minister (Alex Atwood) announced his intention to publish a discussion paper to ascertain stakeholder views on the need to revisit this question. In August 2011, the DOE published *Environmental Governance in Northern Ireland: A Discussion Document*,¹²⁰ which set out various options for structural reform of the regulator including the creation of an independent entity structured along the lines recommended by REGNI.¹²¹ The department's analysis of the responses received revealed not only that support for independent regulation in the

118 Ministerial Statement (n 49) 5.

119 The authors presented these findings in a seminar delivered to the Northern Ireland Judicial Studies Board (29 September 2011) and to an Enforcement Summit hosted by the Minister for the Environment (25 June 2012).

They will also be the subject of a separate paper shortly to be submitted for publication by the current authors.

120 <www.doeni.gov.uk/env_gov_discussion_document.pdf>.

121 Ibid.

terms recommended by REGNI remained very strong (83 per cent),¹²² but also that the dynamics of the debate remained unchanged. The Ulster Farmers' Union remained the only stakeholder to object to structural reform. Meanwhile the Assembly reflected its support for Atwood's process by adopting another motion submitted by the Alliance Party, this time noting publication of the discussion document, reiterating its view that environmental regulation should be independent and calling on the new Environment Minister to externalise the function.¹²³

However, while the new minister has signalled his commitment to the creation of an independent environment agency,¹²⁴ it is far from clear that he will be able to deliver this outcome despite the more favourable composition of the new Executive Committee. Whereas it was prepared to fudge how the decision was made in 2008 to reject externalisation, the Executive Committee itself has forestalled a repeat of this scenario. Quite apart from any arguments concerning the cross-cutting nature of this decision, the committee's response to the Assembly's referral of the matter explicitly confirmed that this is a 'significant and controversial issue', thus, Atwood will be required to submit proposals to deviate from this decision for collective consideration. The DUP remains in a position to insist that a vote concerning the creation of an independent regulator must be conducted on a cross-community basis.¹²⁵ Crucially, though the balance of power within the Executive has more recently shifted in favour of those supportive of structural reform following electoral success by the Alliance Party, which resulted in its assumption of ministerial responsibility for the Department of Employment and Learning, but also the allocation of the new Department of Justice, to its leader (David Ford) – one of the most outspoken political proponents of independent regulation – even a decision by the two Alliance ministers to designate themselves as 'unionist' for the purposes of voting with the one UUP minister would still produce only 37.5 per cent of the 40 per cent support required amongst the unionists voting, thus enabling the DUP to block any move in this direction.¹²⁶

However, where concerted political campaigning and the restoration of local democratic accountability have failed to induce meaningful leadership by the devolved Executive, it is almost inevitable that Northern Ireland's judiciary will be left to force what should have been a politically led process of reform. Within months of Foster's decision to reject independent regulation, the Northern Ireland High Court ruled on a judicial review in which the NIEA's position within central government was used as the basis for challenging the legality of draft Area Plans published by the Northern Ireland Planning Service.¹²⁷ It was argued that the lack of functional separation between the plan maker and the authority designated under the Strategic Environmental Assessment Directive 2001/42/EC¹²⁸ (SEA) as the environmental consultee (namely, the Planning Service and the NIEA, both executive agencies of the DOE) subverted the core mechanism created by

122 DOE, *Synopsis of Responses to Discussion Paper entitled 'Environmental Governance in Northern Ireland'* (DOE 2011).

123 5 September 2011.

124 For example, at a Northern Ireland Green Party Conference in October 2011, see his keynote address at <<http://greenparty-ni.co.uk/green-party-ni-conference-2011/>>.

125 The DUP continues to hold five ministerial positions in the new Executive. See also n 53 and associated discussion.

126 See n 56 and associated discussion. It is also worth noting that the DUP and Sinn Féin have recently announced proposals to disband the Department of Employment and Learning as part of a planned rationalisation of the region's elaborate central government arrangement, thus restoring the balance of power reflected in the first mandate.

127 *Seaport (NI) Ltd, Magherafelt District Council and Others v Department of the Environment for Northern Ireland* [2007] NIQB 62.

128 [2001] OJ L197/30.

the directive to ensure environmental protection; namely, the environmental assessment of proposed plans and programmes. More specifically it was argued that these arrangements breached the implicit requirement that the assessment process should be based not only on consultation with the domestic authorities tasked with environmental responsibilities but also that they should be independent from the plan-making entity in order to ensure that the environmental information supplied is comprehensive and reliable.¹²⁹ The High Court agreed with Seaport's interpretation of the directive¹³⁰ and ruled that, if an independent environmental consultee did not exist, it may become necessary to create such an authority.¹³¹ On appeal, the Court of Appeal sought a preliminary ruling from the European Court of Justice as to whether the SEA Directive should be interpreted to require consultation with an independent environmental authority. The European Court did not embrace the High Court or Advocate General Bot's even more robust emphasis on the pivotal importance of consultation with an independent environmental authority as a prerequisite to ensuring the credibility and legality of the environmental assessment process.¹³² However, though it did not require that the entity be formally independent of the plan-making body, the court ruled that member states must ensure a sufficient degree of functional separation to enable the giving of an 'objective' opinion by the environmental authority.¹³³ More specifically, it stated that the environmental authority must have 'real autonomy' and thus administrative and human resources of its own,¹³⁴ but left it to the domestic court to make the final assessment as to whether these criteria could be satisfied in the circumstances of the case.

On one level, the European Court's ruling was a disappointment. The robust emphasis on the need for transparent, credible and thus independent consultation reflected in the High Court ruling and Advocate General's Opinion had raised hopes within the ENGO coalition that, despite its origins in a challenge to long overdue planning policy modernisation, this judicial review would ultimately provide an irresistible lever to force the externalisation of responsibility for environmental regulation. Without doubt it could always be argued that this case was an unlikely lever for achieving the creation of an independent environment agency called for by ENGOs. It was clear from the outset that even a European Court ruling requiring an independent environmental consultation body for SEA purposes could potentially have been satisfied by simply moving the officials responsible for area planning to another government department. However, there was also little doubt that explicit European Court confirmation of the need for an independent environmental consultation authority would have subjected the tribal veto of structural reform to important new pressure and potentially brought the political debate to a critical tipping point. That said the DOE's claim to have won this seminal legal battle is almost certainly premature.¹³⁵ While the European Court's ruling was regrettably brief, it nevertheless drew a proverbial line in the legal sand concerning the governance arrangements required to facilitate the discharge of a key function performed by modern environmental regulators. It furthermore transferred responsibility to the national courts to

129 The arguments and decision in this case are discussed in more detail by S Turner, 'The Strategic Environmental Assessment Directive: A Potential Lever for Independent Environmental Regulation in Northern Ireland?' (forthcoming) 24(2) *Journal of Environmental Law*.

130 [2007] NIQB 62.

131 *Ibid* para 17.

132 Case C-474/10, delivered 14 July 2011 (unreported).

133 *Ibid* para 42.

134 *Ibid*.

135 DOE issued a press release entitled 'DOE wins important European ruling', 20 October 2011 <www.northerireland.gov.uk/index/media-centre/news-departments/news-doe>.

make the final verification of whether the specific arrangements for functional separation at national level guarantee 'real autonomy' sufficient to enable the expression of an objective opinion. While the European Court was not prepared to comment directly on whether NIEA's status as an executive agency of the DOE enabled it to express an objective opinion on the environmental implications of plans proposed by another executive agency of the department, the Northern Ireland High Court squarely addressed this issue and in the negative. Weatherup explicitly ruled that as long as they remained part of the same department and legal entity, even a 'formal separation of roles' between these two executive agencies would not have satisfied him that sufficient separation was ensured to provide the nature and quality of consultation required by Article 6 of the SEA Directive.¹³⁶

Regrettably Seaport Investments withdrew the judicial review and so the Court of Appeal was not ultimately required to rule on whether NIEA's status within the DOE enabled it to express an objective opinion on the environmental impacts of plans proposed by another executive agency of the department. However, it remains open to others to rely on the European Court's ruling and to use the NIEA's position within central government as a means to challenge decision-making by the DOE. In the meantime, the evidence of the NIEA's position as a captured regulator continues to mount. Despite the fact that the NIEA's lack of independence has cast a significant shadow over the credibility of this critical player in the process of regional environmental governance, it enjoys less functional separation within the DOE today than it did when the ENGO campaign began. During Sammy Wilson's tenure as Environment Minister, the roles of Chief Executive of the NIEA and the Deputy Secretary¹³⁷ with responsibility for the department's core environmental and planning policy function were merged so that they are now held by the same official. In effect, the DUP has proved unwilling to maintain even the appearance of 'clear blue water' between the regulator and the departmental core. More recently the latest report from the CJI, published in 2011,¹³⁸ also reflects entrenched political resistance to creating a transparent mechanism to protect the NIEA's independence even to take criminal enforcement decisions. While the department put in place a protocol governing external inputs into these decisions following the CJI report in 2007, its 2011 report notes that this protocol relates only to third parties but not to internal or ministerial interventions.¹³⁹ The inspectorate accepted the minister's constitutional position in relation to control of decision-making within the department, but emphasised that staff must have the protection of 'a transparent decision-making process that is free from undue and inappropriate interference'¹⁴⁰ and emphasised that procedures must be put in place to appropriately record ministerial involvement in decision-making by the regulator.¹⁴¹ Not surprisingly, this report also reflects the CJI's more explicit support for the structural independence recommended by REGNI in that this time the inspectorate points to the existence of independent regulators in all neighbouring jurisdictions and to Atwood's recent decision to re-opening the debate in Northern Ireland.¹⁴² Thus, while it seems highly unlikely that the DUP will permit the externalisation of responsibility for environmental regulation in the interests of effective environmental protection or as a concession to democracy, it is only a matter of time before the High Court and Court of Appeal are called on once again to

136 Para 15.

137 In effect, the grade below Permanent Secretary.

138 CJI (n 100).

139 Ibid para 2.40.

140 Ibid para 2.43.

141 CJI (n 100) paras 2.43–6.

142 Ibid.

consider whether the NIEA's status as an executive agency within DOE complies with the rule of EU environmental law. Given the serious reservations already expressed by *Weatherup J* and the CJI, and the NIEA's reemerging into the departmental policy core, it seems inevitable that the judiciary will ultimately force the devolved administration to take the first meaningful step towards structural modernisation of Northern Ireland's environmental regulator.

The Sword of Damocles: who controls HSBC in the aftermath of its deferred prosecution agreement with the United States Department of Justice?

JUSTIN O'BRIEN*

*Centre for Law, Markets and Regulation, Faculty of Law,
The University of New South Wales*

Abstract

HSBC has entered into a \$1.92bn deferred prosecution with the Department of Justice in the United States to settle charges that the bank's compliance systems and corporate governance controls had failed to prevent money laundering and sanctions violations on an industrial scale. The violations spanned the globe and demonstrated fundamental flaws with the bank's business model. The article evaluates the terms of the settlement and explores the national and extra-territorial implications. It argues that the settlement, the largest ever imposed on a financial institution, marks a significant turning point in the use of criminal prosecution precisely because it occurred just as the still burgeoning London Interbank Offered Rate (Libor) manipulation scandal reaches a denouement.

1 Introduction

The US comprises the most significant export market for Mexican and Colombian drug cartels. Mexican democracy itself has been destabilised by cartel-sponsored corruption. Ongoing political violence, fuelled, in part, by the drug trade, has further weakened social and political capital. When the London-based global bank HSBC used advertising that claimed the importance of knowing when emerging markets have emerged it most certainly did not have the facilitation of the narcotics industry in mind. Yet this was precisely what occurred as a consequence of systemic compliance failures across the group. From the parent operation in London to affiliated entities in both the United States and Mexico, there was, according to an agreed Statement of Facts tabled in a New York federal court, a wanton disregard for the societal implications. When combined with identified inability to control money transfers to North Korea, Burma, Cuba and Sudan in violation of a United States-imposed sanctions regime, the global failure of compliance at the bank suggests deep structural problems with HSBC's core business model. Providing local businesses with a global imprimatur without strenuous checks to safeguard reputational capital has been shown to be an exceptionally dangerous strategy.

The \$1.92bn deferred prosecution agreement entered into by HSBC with US regulators contains the largest financial penalty ever imposed on a global bank by prosecutorial authorities in either a civil or criminal matter. The bank is required to disgorge \$1.256bn of profits. It will also pay a total of \$665m in civil penalties to regulatory agencies, including the Office of the Comptroller of the Currency (\$500m) and the Federal Reserve (\$165m). The

* I acknowledge the support of an Australian Research Council Future Fellowship.

payment to the Office of the Comptroller of the Currency is a partial, if not complete, vindication of the agency. Given the criticism of its oversight in an eviscerating report tabled to the United States Congress in July this year it is, at best, an equivocal endorsement of federal priorities.¹ Of even more significance, however, is the requirement that a corporate compliance monitor be appointed to a five-year term, in compliance with a template in operation since at least 2008.² Although ostensibly independent, the terms of engagement and accountability structures governing the design and implementation of the monitor's work plan make it abundantly clear that the holder is an agent of the Department of Justice, for which it makes no apology. 'HSBC is being held accountable for stunning failures of oversight – and worse – that led the bank to permit narcotics traffickers and others to launder hundreds of millions of dollars through HSBC subsidiaries, and to facilitate hundreds of millions more in transactions with sanctioned countries', noted the head of the Criminal Division of the Department, Lanny Breuer, in a broadly circulated circular.³

In a deferred prosecution, a corporation enters into an effective contract with the prosecutorial authority in which it accepts not to subsequently challenge an agreed narrative and engages in remedial action in exchange for a decision not to proceed with the charges. If there is no repetition of the complained of conduct within an agreed timeframe the charges are voided. Conversely, a violation allows for a filing of an indictment in which the statement of facts cannot be challenged. It is, therefore, an admission of guilt. It is closely allied to a non-prosecution agreement, which can also contain contractually agreed remedial action. In policy terms, the HSBC agreement is one of the most significant uses of the deferred prosecution mechanism since its application to deal with KPMG's development of abusive tax shelters in 2005.⁴ The KPMG prosecution had ended with the Department of Justice castigated in the Manhattan Federal Court. Judge Louis Kaplan condemned what he termed its unconstitutional conduct. He voiced grave concern that the prosecutors had 'put a gun to KPMG's head' by forcing it to end legal support for partners whose defence centred on the fact that they were following corporate-sanctioned objectives.⁵ The Department of Justice, stung by the criticism, retreated largely from forcing change on the financial sector, with the exception of active prosecution of sanctions violations and breaches of the Foreign Corrupt Practices Act of 1977. With the exception of the settlement with Lloyds Bank TSB, however, it had not imposed an external monitor on a major financial institution for

- 1 See Senate Permanent Sub-Committee on Investigations, *US Vulnerabilities to Money Laundering, Drugs, and Terrorist Financing: HSBC Case History* (US Congress, Washington DC, 17 July 2012), 306 (noting a 2008 examination in which the OCC records 'As the U.S. dollar clearing bank for the Global HSBC network, HBUS maintains numerous relationships with institutions worldwide . . . The bank does business with numerous customers in both High Intensity Drug Trafficking Area and High Intensity Money Laundering and Related Financial Crime Area locations. HBUS provides pouch services through several business units. Historically, pouch services are vulnerable to money laundering risk').
- 2 Craig Morford, 'Memorandum on Selection and Use of Monitors in Deferred Prosecutions and Non Prosecution Agreements with Corporations' (Department of Justice, Washington DC, 8 March 2008).
- 3 Department of Justice, 'HSBC Holdings Plc and HSBC Bank USA NA Admit to Anti-Money Laundering and Sanctions Violations, Forfeit \$1.256 Billion in Deferred Prosecution Agreement', Press Release, Washington DC, 11 December 2012.
- 4 Justin O'Brien, *Redesigning Financial Regulation* (John Wiley & Sons 2007) 123–169; for assessment of the Department of Justice's usage, see Government Accountability Office, *Corporate Crime: Preliminary Observations on DOJ's Use and Oversight of Deferred Prosecution Agreements and Non-Prosecution Agreements* (Washington DC, 25 June 2009).
- 5 *United States of America v Jeffrey Stein et al* S1 05 Crim 0888 (LAK, 26 June 2006). Kaplan further noted: 'Those who commit crimes – regardless of whether they wear white or blue collars – must be brought to justice. The government, however, has let its zeal get in the way of its judgment. It has violated the Constitution it is sworn to defend': at 3. The decision was upheld on appeal in 2008, see *United States v Stein* No 07–3042 (2d Cir 28 August 2008).

sanctions violations. Instead it had relied on the stated intention of institutions to reform the compliance function.⁶ This time, as they say, is different.

The success of the HSBC negotiations, significantly brought not in the Southern District of New York but in neighbouring Brooklyn, where HSBC Bank USA holds neither its head office nor conducts major business, served three core purposes. First, it expunged the debilitating error of judgment that informed the prosecution of the KPMG partners. Second, it signalled a determination by the Department to ensure that nascent state action, particularly in the sanctions violations space, did not usurp federal leadership in the setting of prosecutorial and regulatory priorities. The settlement came as Standard Chartered, another UK domiciled bank, agreed an overarching settlement of \$327m to draw to a conclusion litigation brought by a range of regulatory agencies, including the Criminal Division.⁷ The timing is far from incidental. It follows the success by New York Department of Financial Services in securing a \$340m settlement with Standard Chartered in August 2012 on broadly similar charges,⁸ which were dismissed at the time as the actions of a 'rogue regulator'.⁹ The strategic approach adopted by the New York Department of Financial Services followed a playbook made famous by Eliot Spitzer, the former State Attorney General.¹⁰ The decision not to require an independent monitor in the Standard Chartered case is, in part, linked to the fact that the violation amounted to a fraction of what was initially alleged by the New York Department of Financial Services.¹¹ Third – and most significantly – it repositioned the Department of Justice as a core moderator of regulatory priorities to use threatened prosecution as a catalyst for cultural change, not only

6 Similar settlements have been reached with a number of banks for sanctions violations involving Cuba and Iran, including – in descending monetary order – ING (\$619m), Credit Suisse (\$536m) Lloyds (\$350m) and Barclays (\$298m), see Carrick Mollencamp and Brett Wolf, 'HSBC to Pay Record \$1.9 US Billion Fine in Money Laundering Case', Reuters, 11 December 2012 <www.complinet.com/global/news/news/article.html?ref=160723>; see also Sharlene Goff, 'Barclays Fined \$298m Over Sanctions Breach', *Financial Times*, 17 August 2010 <www.ft.com/intl/cms/s/0/6918f646-a96b-11df-a6f2-00144feabdc0.html#axzz2F4BqZ9tM>. The settlements have prompted judicial scepticism, see, for example, Jean Eaglesham and Justin Baer, 'Barclays "Sweetheart Deal" Under Fire', *Financial Times*, 18 August 2010 <www.ft.com/intl/cms/s/0/dece7c62-aa51-11df-9367-00144feabdc0.html#axzz2F4BqZ9tM>.

7 Department of Justice, 'Standard Chartered Bank Agrees to Forfeit \$227 Million for Illegal Transactions with Iran, Sudan, Libya, and Burma', Press Release, Washington DC, 10 December 2012.

8 Department of Financial Services, 'Statement from Benjamin M Lawskey, Superintendent of Financial Services, Regarding Standard Chartered Bank', Press Release, New York, 14 August 2012.

9 Kishore Mahbubani, 'A Lawskey Unto Himself, or Why New York Erred on StanChart', *Financial Times*, 12 August 2012 <www.ft.com/intl/cms/s/0/f4c6b142-e2d5-11e1-bf02-00144feab49a.html#axzz24PyoHpnq>.

10 For discussion of Spitzer's strategy (and the accountability deficit at its heart), see Justin O'Brien, 'The Politics of Enforcement, Eliot Spitzer, State-Federal Relations and the Redesign of Financial Regulation' (2005) 35 *Publius The Journal of Federalism* 439; see also Jonathan Macey, 'Wall Street in Turmoil: State-Federal Relations Post Eliot Spitzer' (2004) 70 *Brooklyn Law Review* 117, at 120–2; Kulbir Walha and Edward Filusch, 'Eliot Spitzer: A Crusader Against Corporate Malfeasance or a Politically Ambitions Spotlight Hound? A Case Study of Eliot Spitzer and Marsh & McLennan' (2005) 18 *Georgetown Journal of Legal Ethics* 1111, at 1114–15.

11 Standard Chartered, 'Standard Chartered Reaches Final Settlement With US Authorities', Press Release, 12 December 2012. The release notes that the investigation by the Office of Foreign Assets Control found 'that while SCB's omission of information affected approximately 60,000 payments related to Iran totaling \$250 billion, the vast majority of those transactions do not appear to have been violations of the Iranian Transactions Regulations'. Over the entire period from 2001 to the end of 2007, it found that approximately \$24m of transactions processed on behalf of Iranian parties and a total of \$109m on behalf other sanctioned entities from other countries (Burma, Sudan and Libya) appeared to be in violation of sanctions laws. Over the same period, SCB New York processed \$139trn in US dollar payments.

within individual entities but also across sectors, at both national and global level.¹² As such, it signals its return as a pivotal, if unpredictable, force in financial regulation.

The Standard Chartered and HSBC settlements reflect the growing centrality of deferred prosecutions as the preferred prosecutorial tool of choice.¹³ The expansion of the measure reflects both its strengths and limitations.¹⁴ On the one hand, it avoids the possibility of broader collateral damage. In the United States, a criminal conviction of a financial services firm would automatically trigger licence revocation. This could have devastating consequences for the individual institution indicted and the livelihoods of those who work for them.¹⁵ Moreover, the licence revocation of a major bank or financial services firm deemed to be of regional or global significance could have an immediate effect on the stability of the global financial system. Indeed, these factors were explicitly noted by the Department of Justice in justifying the decision to delay prosecution. British regulators have gone further. Andrew Bailey, the designate head of the Prudential Regulation Authority, rather plaintively noted that to bring a criminal action against a bank would be a 'very destabilizing issue. It's another version of too big to fail.'¹⁶ The limitation is that absent substantive requirements to change not only compliance practice but also

12 Lanny Breuer, 'Address to the New York City Bar Association', Speech delivered to the NYC Bar Association, New York, 13 September 2012.

13 The use of deferred prosecutions is also under consideration the UK, with particular references to violations of that jurisdiction's Bribery Act of 2010. A consultation process, now under review advocated its expansion, see Ministry of Justice, *Deferred Prosecution Agreements* Cm 8348 (2012).

14 For review, see Brandon Garrett, 'Globalized Corporate Prosecutions' (2011) 97 Virginia Law Review 1776. In the UK there has long been considerable interest in introducing the measure. The government had already signalled its strong support for the introduction of the measure, see Caroline Binham, 'Garnier Eyes US Style Fines and Bargains', *Financial Times*, 28 September 2011. A consultation process highlighted one critical difference from practice in the United States. There is to be judicial involvement in the initial decision as to deploy the mechanism and the parameters of the proposed terms, see Ministry of Justice (n 13). On 23 October 2012, the Ministry of Justice announced its introduction, see Ministry of Justice, 'New Tool to Fight Economic Crime', Press Release, London, 23 October 2012, quoting Justice Minister, Damian Green that: 'Deferred Prosecution Agreements will give prosecutors an effective new tool to tackle what has become an increasingly complex issue. This will ensure that more unacceptable corporate behaviour is dealt with including through substantial penalties, proper reparation to victims, and measures to prevent future wrongdoing.'

15 See Larry Thompson, 'Principles of Federal Prosecution of Business Organizations' (Washington DC, US Department of Justice, 20 January 2003). Following criticism of requirements that organizations under investigation should waive client-attorney privilege and withhold payment of legal fees to individuals prosecuted, most notably in the prosecution of KPMG, these components were subsequently dropped.

16 Harry Wilson, 'Banks are too Big to Prosecute Says FSA's Andrew Bailey', *Daily Telegraph*, 14 December 2012 <www.telegraph.co.uk/finance/newsbysector/banksandfinance/9743839/Banks-are-too-big-to-prosecute-says-FSAs-Andrew-Bailey.html>. Although the introduction of the Deferred Prosecution mechanism is designed primarily for the dealing with corruption, there is no doubt it could be applied by the Serious Fraud Office in relation to the Libor scandal in the event that a criminal prosecution eventuates. On 11 December 2012, British authorities announced that three people had been arrested, see Jill Traynor, 'Bleak Day for British Banking as Libor Arrests Follow Record HSBC Fine', *The Guardian*, 11 December 2012 <www.guardian.co.uk/business/2012/dec/11/banking-libor-fine-hsbc>. On 19 December, the Department of Justice in the United States announced that two traders within UBS are to face criminal prosecution in relation to the manipulation of Libor, see Department of Justice, 'Attorney General Eric Holder Speaks at the UBS Press Conference', Press Release, Washington DC, 19 December 2012. The press conference revealed that UBS would face a combined fine of just over \$1.5bn, to be shared disproportionately between the Department of Justice (\$500m), the Commodity and Futures Trading Commission (\$700m), the UK's Financial Services Authority (\$260m) and the Swiss Financial Regulator, which while unable to levy a fine recouped \$69m in improper profits, see Kara Scanell et al, 'UBS in \$1.5bn Libor Settlement', *Financial Times*, 20 December 2012, at 1. For details of the scale of the deception, which spanned three continents see Financial Services Authority, 'Final Notice for UBS AG', London, 19 December 2012 <www.fsa.gov.uk/static/pubs/final/ubs.pdf>.

broader risk and corporate governance reporting frameworks, the financial penalties, while substantial, could be and often are written off as part of the cost of doing business.¹⁷

There is, therefore, a triangulated policy dilemma. If major banks are too big to fail, too big to prosecute and too big to manage, how does one secure substantive warranted commitment to ethical restraint and pro-social rather than anti-social behaviour from such entities? One way of offsetting that limitation is to ensure that cultural change is ongoing through the imposition of an external monitor.¹⁸ It is the application of this component of the regulatory toolbox that differentiates the Department of Justice's approach to HSBC. In section 2 the article examines the charges themselves, which draw heavily from the damning Senate Permanent Sub-Committee on Investigations report.¹⁹ Section 3 details the remedial action taken by HSBC to date and that mandated by the deferred prosecution agreement. Given this cooperation, section 4 then examines how and why the department did not accept remedial action at face value but instead imposed an external monitor with granular terms of reference. Section 5 of the article assesses the implications of that decision on regulatory design. It argues that the deal represents not the Department's weakness, as broadly reported. Instead it reflects its growing strength. This strength has far-reaching consequences, not just in the United States but internationally. Section 6 concludes.

2 A flawed business model

HSBC was found on 11 December 2012 in the Federal Court in Brooklyn of being responsible for systematic sanctions violations and the facilitation of money laundering on an industrial scale. It was held accountable for threatening national security by providing financing facilities to a Saudi Arabian bank with links to terrorist groups.²⁰ The four-count charge found that the bank had wilfully failed to develop, implement and maintain an effective anti-money laundering programme in contravention of the Bank Secrecy Act of

17 This is made manifest in two highly influential if disparate sources, Lex, 'HSBC/StanChart—Rap on the Knuckles', *Financial Times*, 11 December 2012 (noting 'transgressions normally only become public a long time after the fact. Markets seem happy to view these latest as one-off episodes of ancient history.').; see also Matt Taibbi, 'Outrageous HSBC Settlement Proves the Drug War is a Joke', *Rolling Stone*, 11 December 2012 <www.rollingstone.com/politics/blogs/taibblog/outrageous-hsbc-settlement-proves-the-drug-war-is-a-joke-20121213>.

18 Christie Ford and David Hess, 'Can Corporate Monitorship's Improve Corporate Compliance' (2009) 34 *Journal of Corporation Law* 679 (noting the danger that these are exercises in symbolism with 'monitors not conducting deep dives into the corporation's culture': at 737); see also Vikramaditya Khanna and Timothy Dickinson, 'The Corporate Monitor: The New Corporate Czar' (2007) 105 *Michigan Law Review* 1713 (noting the de facto creation of a new professional class of advisors and advocating allocation of fiduciary duty to shareholders: at 1727). The critical issue, therefore, pivots on willingness to use nascent power and to whom accountability is owed. There can be no mistaking the potential to gain effective control of corporate strategy. In 2006, for example, the corporate monitor installed at Bristol-Meyer Squibb advocated the sacking of the chief executive officer and the general counsel, recommendations accepted by the board, see Brooke Masters, 'Bristol-Meyers Ousts its Chief at Monitor's Urging', *Washington Post*, 13 September 2006, D1.

19 *US Vulnerabilities to Money Laundering, Drugs, and Terrorist Financing: HSBC Case History* (n 1); see also Carl Levin, 'Levin Statement on HSBC Settlement', Press Release, Washington DC, 11 December 2012, noting: 'In an age of international terrorism, drug violence, and organized crime, stopping illicit money flows is a national security imperative. Global banks have global responsibilities to prevent participation in illicit or suspect transactions. The HSBC settlement sends a powerful wakeup call to multinational banks about the consequences of disregarding their anti-money laundering obligations. It also shows the value of congressional oversight in exposing wrongdoing and the ongoing need to hold banks accountable.'

20 *United States of America v HSBC Bank USA NA and HSBC Holdings Inc* 12 Cr 763 (EDNY, 11 December 2012) <http://lib.law.virginia.edu/Garrett/prosecution_agreements/pdf/HSBC.pdf>. The online resource contains the charges, the text of the deferred prosecution, an agreed statement of facts and the terms governing the appointment of an external monitor. Hereafter these are referenced as 'United States Charges', 'Deferred Prosecution Terms', 'Attachment A: Statement of Facts' and 'Attachment B: Corporate Compliance Monitor'.

1970.²¹ The legislation, progressively extended in both granularity and geographic scope over the years to address an increase in criminal money-laundering activities utilising financial institutions, requires regulated entities to detect and report suspicious activity. Furthermore they are required to maintain records that could be used in criminal, tax or regulatory investigations or court proceedings. The bank's failure to comply with anti-money laundering legislation was not historical. Rather the deficiencies encompassed the period January 2006–December 2010; a period that straddled the global financial crisis. In the same period it was charged that HSBC wilfully failed to conduct due diligence on correspondent bank accounts for non-United-States persons. Correspondent accounts are set up to make or receive payments from individuals or organisations with which the US-based bank has no direct relationship. Under the terms of the Bank Secrecy Act, HSBC Bank USA was required to conduct extensive due diligence on the financial institutions for which it held these correspondent accounts. Inexplicably, HSBC Bank USA failed to do so in relation to accounts held by its affiliate in Mexico, notwithstanding the fact that there is no exception for foreign financial institutions within the same holding company. This, the count charged, inhibited the collection of material, which would have reasonably allowed for the detection and reporting of instances of money laundering and other suspicious activity.²²

The risk posed by initial failure to conduct the due diligence on the establishment of the accounts was magnified by an ongoing failure to monitor wire transfers within and between them. It was further compounded by the absence of anti-money laundering protocols in the HSBC Mexico operation itself. The combination was rendered catastrophic for the parent company by its use of vertical reporting lines. This meant that HSBC Bank USA was not directly informed of growing unease of regulatory, diplomatic and law enforcement agencies on both sides of the Rio Grande about a rapid expansion of money laundering across the Mexican banking sector and in which HSBC Mexico played a pivotal if unwitting role. The money-laundering charges were conjoined with two counts dealing with sanctions violation. The third count charged violations of the International Emergency Economic Powers Act of 1977.²³ Between 2001 to 2006 HSBC 'knowingly, intentionally and willfully facilitated prohibited transactions for sanctioned entities in Iran, Libya, Sudan and Burma'. HSBC knowingly and willingly circumvented government safeguards designed to block terrorist funding, allowing, for example, affiliates to shield the fact that thousands of transactions involved links to Iran. The Senate investigation suggested the problem was even more widespread. An independent audit paid for by HSBC found the bank facilitated 25,000 questionable transactions with Iran between 2001 and 2007.²⁴ The report also detailed that HSBC worked extensively with Saudi Arabia's Al Rajhi Bank, some owners of which have been linked to terrorism financing. HSBC's US affiliate supplied Al Rajhi with nearly \$1bn-worth of US banknotes until 2010, and worked with two banks in Bangladesh linked to terrorism financing. The fourth count charged that HSBC had engaged in similar activity in relation to Cuba in violation of the Trading with the Enemy Act of 1917.²⁵

The reputational damage to HSBC comes primarily, however, from the first two counts, not least because of the immediate cost of the drugs war on American society.²⁶ Astonishingly, the failure of the compliance policies and procedures is estimated to have caused at least \$881m in drug proceeds to filter through the United States financial

21 31 USC 5311–32.

22 The United States Charges (n 20) 11.

23 50 USC 1702.

24 *US Vulnerabilities to Money Laundering, Drugs and Terrorist Financing: HSBC Case Study* (n 1) 6.

25 12 USC §95aff.

26 It is this aspect of the case that has dominated media coverage, see, for example, Taibi (n 17).

system.²⁷ This was primarily achieved through the preference of drug cartels to use HSBC Mexico as a conduit for what was termed the 'Black Market Peso Exchange',²⁸ going as far as designing special containers that fit precisely under the teller windows installed across the bank's branch network. Building on an investigation launched in 2008 by the Department of Homeland Security into how HSBC Bank USA had been compromised, the litigation paints a dismal picture of wilful neglect at national and international levels within the bank.²⁹ When read in conjunction with the detailed congressional investigation, the Statement of Facts reveals how the cartels operated with apparent impunity.³⁰ Together, as highlighted above, they point to significant flaws in the entire HSBC business model.

The inability of HSBC, the London-headquartered self-styled world's local bank, to know how its affiliates were operating in critical markets does more than puncture a marketing myth. It also demonstrates the limited power that compliance departments at both national and broader group levels had to influence strategic direction. In operating on a franchise basis, rebranding foreign acquisitions without necessarily changing their culture or integrating them fully into global template, the bank institutionalised a silo approach to corporate governance and risk management. A senior London-based compliance officer noted the risk and likely result in discussions with a counterpart in Mexico as late as 2008,³¹ six years after HSBC's acquisition of Grupo Financiero Bital in 2002, at the time the

27 The United States Charges (n 20) 9–12.

28 Attachment A: Statement of Facts (n 20) para 49: 'In the BMPE, middlemen, often referred to as peso brokers, transform bulk cash from the sale of illegal drugs into revenue from the sale of legitimate goods. In this process, the peso brokers purchase bulk cash in United States dollars from drug cartels at a discounted rate, in return for Colombian pesos that belong to Colombian businessmen. The peso brokers then use the U.S. dollars to purchase legitimate goods from businesses in the United States and other foreign countries, on behalf of the Colombian businessmen. These goods are then sent to the Colombian businessmen, who sell the goods for Colombian pesos to recoup their original investment. In the end, the Colombian businessmen obtain U.S. dollars at a lower exchange rate than otherwise available in Colombia, the Colombian cartel leaders receive Colombian pesos while avoiding the costs associated with depositing U.S. dollars directly into Colombian financial institutions, and the peso brokers receive fees for their services as middlemen.'

29 Ibid, para. 50: 'The Department alleges, and HSBC Bank USA and HSBC Holdings do not contest, that, beginning in 2008, an investigation conducted by HSI's El Dorado Task Force, in conjunction with the U.S. Attorney's Office for the Eastern District of New York, identified multiple HSBC Mexico accounts associated with BMPE activity. The investigation further revealed that drug traffickers were depositing hundreds of thousands of dollars in bulk U.S. currency each day into HSBC Mexico accounts. In order to efficiently move this volume of cash through the teller windows at HSBC Mexico branches, drug traffickers designed specially shaped boxes that fit the precise dimensions of the teller windows. The drug traffickers would send numerous boxes filled with cash through the teller windows for deposit into HSBC Mexico accounts. After the cash was deposited in the accounts, peso brokers then wire transferred the U.S. dollars to various exporters located in New York City and other locations throughout the United States to purchase goods for Colombian businesses. The U.S. exporters then sent the goods directly to the businesses in Colombia . . . The investigation further revealed that, because of its lax AML controls, HSBC Mexico was the preferred financial institution for drug cartels and money launderers. The drug trafficking proceeds (in physical U.S. dollars) deposited at HSBC Mexico as part of the BMPE were sold to HSBC Bank USA through Banknotes. In addition, many of the BMPE wire transfers to exporters in the United States passed through HSBC Mexico's correspondent account with HSBC Bank USA.'

30 For analysis of the Congressional investigation and HSBC's response, see Justin O'Brien, 'Where the Buck Stops', *Australian Financial Review*, 27 July 2012, R1, 14–15.

31 Ibid para 34, noting that in July 2007, a senior compliance officer at HSBC Group told HSBC Mexico's Chief Compliance Officer that: '[t]he AML committee just can't keep rubber-stamping unacceptable risks merely because someone on the business side writes a nice letter. It needs to take a firmer stand. It needs some cojones. We have seen this movie before, and it ends badly.'

country's fifth biggest bank, with 1400 branches and 6m customers.³² The catalogue of failure within and between the disparate components of HSBC as outlined in the agreed Statement of Facts is as extensive as it is shocking.

Specifically, HSBC Bank USA ignored the money laundering risks associated with doing business with certain Mexican customers and failed to implement a BSA/AML program that was adequate to monitor suspicious transactions from Mexico. At the same time, Grupo Financiero HSBC, S.A. de C.V. ('HSBC Mexico'), one of HSBC Bank USA's largest Mexican customers, had its own significant AML problems. As a result of these concurrent AML failures, at least \$881 million in drug trafficking proceeds, including proceeds of drug trafficking by the Sinaloa Cartel in Mexico and the Norte del Valle Cartel in Colombia, were laundered through HSBC Bank USA without being detected. HSBC Group was aware of the significant AML compliance problems at HSBC Mexico, yet did not inform HSBC Bank USA of these problems and their potential impact on HSBC Bank USA's AML program.³³

The identified problems started within the Mexican operation. Despite the fact that the Mexican financial regulatory authority, the Comision Nacional Bancaria y Valores (the CNBV), had flagged its concerns in external reviews, which were, in turn, escalated to the chief executive officer of HSBC Holdings, no integrated approach on how to rank country risk was initiated.³⁴ Notwithstanding growing national and international concern about the rise of drug trafficking in and through Mexico, HSBC Bank USA maintained a risk ranking of 'standard'. This was the lowest rated risk. It meant that the accounts were given only cursory examination.³⁵

Given the critical financial relationship between HSBC Mexico and its counterpart in the United States and awareness in both jurisdictions as well as headquarters in London of how the Mexican financial system was used as a global money-laundering gateway, this amounted to a reckless disregard towards risk management. Over \$200trn in wire transfers passed between HSBC Bank USA and its global affiliates, with \$659bn coming from Mexico alone. The risk was not confined, however, to the retail bank operation. The systemic risk was magnified by the fact that HSBC's global banknotes operation, headquartered in New York, is the largest volume trader of physical currency in the world, controlling 60 per cent of the market.³⁶ \$9.4bn in physical banknotes were purchased from accounts linked to the Mexican operation in the period July 2006–July 2009 alone. The bank derived its revenue from commissions on the sale or purchase of physical dollars and its transportation and storage at the Federal Reserve. The Statement of Facts notes, however, that the banknotes compliance operation was not only almost ludicrously understaffed. It also lacked an automated monitoring function.³⁷ Throughout this period the bank, while aware of the risk, failed to 'provide adequate staffing and other resources to maintain an effective anti-money laundering program'.

32 Ibid: 'At the time of the acquisition, HSBC Group's Head of Compliance acknowledged there was "no recognizable compliance or money laundering function in Bital at present." HSBC Group Compliance believed it would take one to four years to achieve its required AML standards at HSBC Mexico. However, until at least 2010, HSBC Mexico's AML programme was not fully up to HSBC Group's required AML standards for HSBC Group Affiliates.' para 30.

33 O'Brien (n 30) para 9.

34 Ibid para 31.

35 Ibid para 18. The Statement of Facts further notes that 'from 2006 until May 2009, when HSBC Bank USA raised Mexico's risk rating to high, over 316,000 transactions worth over \$670 billion from HSBC Mexico alone were excluded from monitoring in the CAMP system': at para 19.

36 O'Brien (n 30) para 20.

37 Ibid para 22.

The clear inference is that such were the profits deriving from the operation it was not in the interests of HSBC, at any level, to investigate much less close suspicious accounts. According to the Statement of Facts the problems were addressed only on receipt of a cease and desist order issued by the Federal Reserve and the Office of the Comptroller of the Currency in October 2010.³⁸ Although the filing of the criminal charges and their subsequent deferral relate only to anti-money laundering control violations in relation to Mexico, the statement of facts makes clear that the nature of HSBC and its geographic exposure constituted an inherent risk. It sets out that 'HSBC Group Affiliates conducted business in many high-risk international locations, including regions of the world presenting a high vulnerability to the laundering of drug trafficking proceeds'.³⁹ This speaks directly to the possibility of broader systemic risks. Unstated in the report but clearly inferred is that HSBC's failure in relation to Mexico may well be only the tip of the iceberg.

In this regard two factors in relation to the HSBC settlement warrant significant attention. First, the scale of the HSBC disgorgement and civil penalties fine sends an unambiguous message that materiality is increasing. As the *Financial Times* has noted, 'a billion here, a billion there and pretty soon you are talking about serious money'.⁴⁰ Second, the fine is the least of HSBC's concerns in relation to its ongoing corporate governance and risk evaluation. The imposition of an external monitor sends an unambiguous message that the bank's commitment to reform should not be taken at face value. Before exploring the rationale, terms and implications, it is essential to highlight the extent to which HSBC has already transformed the compliance function.

3 Remedial action

HSBC has done much to improve the quality of its internal governance, including recruiting former heavyweights from the Department of Justice, Treasury and the Department of Homeland Security to pivotal management positions. The newly appointed chief legal officer, Stuart Levey, in particular, was an inspired choice. He was recruited to the bank direct from the US Department of Treasury, where he had developed a formidable reputation as Under Secretary for Terrorism and Financial Intelligence. As HSBC's chief legal officer, Levey flagged many of the remedial actions taken by the bank in an assured performance to the Senate Sub-Committee on Investigations in July:

While our old model served us well historically, it does not work in an interconnected world where transactions cross borders instantaneously and where weaknesses in one jurisdiction can be quickly exported to others . . . We have learned that our approach to compliance – and AML in particular – was not adequate to address the risks we face as a global institution. And we have learned that we did not share information effectively enough across our affiliates, with

38 O'Brien (n 30) para 11. Although the OCC is the recipient of a \$500m fine it is important to note significant unease over its monitoring operations, see *US Vulnerabilities to Money Laundering, Drugs, and Terrorist Financing: HSBC Case History* (n 1) 316. Specifically, the Senate report into the Mexican operation is as critical of the OCC as it is of HSBC itself: 'For more than six years, from July 2004 until April 2010, despite compiling a litany of AML deficiencies, the OCC never cited HBUS for a violation of law, never took a formal or informal enforcement action, and turned down recommendations to issue Cease and Desist Orders targeting particularly egregious AML problems, even though the same problems surfaced again and again. The OCC's failure to compel HBUS to remedy the AML deficiencies repeatedly identified by its examiners over a six-year period indicates that systemic weaknesses in the OCC's AML oversight model require correction.'

39 Attachment A: Statement of Facts (n 20) para 12.

40 Editorial, 'Paying the Price for the Banks' Mistakes', *Financial Times*, 11 December 2012 <www.ft.com/intl/cms/s/0/7e97873a-43a1-11e2-a48c-00144feabdc0.html#axzz2F4BqZ9tM>.

serious consequences . . . We must implement a global strategy to tackle the root causes of our identified deficiencies.⁴¹

These deficiencies centred on the fact that at a global level compliance served an advisory rather than a control function, which had neither the resources or empowerment to provide a monitoring function. Responsibility for ensuring that standards were being implemented was delegated to country level. As the bank now acknowledges ‘this led to inconsistency and in some cases confusion about ownership and escalation responsibility’. In candid testimony to Congress, Levy detailed how he negotiated the job parameters: ‘In our conversations, the Chairman of the Board and the new CEO were candid with me about the problems HSBC faced, the reforms they wanted me to help them implement, and the empowerment that I would need and have’, he said. It appears he has now that power.⁴² Group Compliance is empowered to set standards across the organization and now has the necessary authority to reach down into affiliates and ensure that those standards are being met . . . The work we have undertaken is ambitious and complicated given our size and our global footprint, but we all recognize that it must be done’, he told a receptive audience on Capitol Hill.⁴³ The work plan centres on the creation of four core business units – global markets and banking; commercial banking; private banking and retail banking; and wealth management.

We gave the heads of each business and function the authority over all personnel in their respective organizations all over the world, thus creating the ability to manage their business or function on a global basis, making it easier to implement consistent policies, standards, and processes [he said]. What that means is that the most senior people responsible for managing HSBC globally sit around a table every month, look at our risks, and make decisions . . . Better global integration makes us better situated today to manage our risk on a global basis, better able to see where risk in one part of HSBC may impact another part, and better able for the first time to ensure that consistent compliance standards and practices are implemented across all of our affiliates.⁴⁴

It is a laudable vision but one that cannot be vouchsafed without external review and validation. The Department of Justice has itself praised the level of cooperation. Significantly, however, in sharp contrast to prior cases involving sanctions violations (or the approach taken by the Securities and Exchange Commission in its non-prosecution deal with Goldman Sachs to trust the bank to reform), the Department of Justice has not taken HSBC’s word for it. It is the fear that symbolism will trump substance that underpins the decision to appoint an external monitor. The terms governing the appointment are exceptionally revealing of the level of distrust. It is abundantly clear that the Department of Justice is, at best, sceptical of self-regulation. That scepticism has an explicit extra-territorial dimension and extends beyond the governance of the bank to the global markets in which it operates.

4 The imposition of an external monitor

In December 2012 two different approaches to embedding restraint began to take shape as London-headquartered banks reflect on the exceptional power of the United States Department of Justice to shift cultural mores through the flexing of its prosecutorial discretion. Both provide tangible evidence of the Department’s renewed interest in the

41 Stuart Levy, ‘Written Testimony for Senate Permanent Subcommittee on Investigations’ (US Congress, Washington DC, 17 July, 2012).

42 Ibid.

43 Ibid.

44 Ibid.

financial sector. HSBC is in the process of submitting to the Department of Justice a pool of three suitably qualified candidates to the position of independent compliance monitor, a pool that the Department can unilaterally reject.⁴⁵ Meanwhile Barclays, which reached a financial settlement in relation to its role in the Libor scandal in August without the imposition of an external monitor is also ruminating over its future. In December 2012 Barclays announced that it had recruited Hector Sants, the former chief executive of the Financial Services Authority as group head of compliance and government and regulatory relations. Given Sants' previous stated interest in and support for the necessity of regulating culture, the appointment serves as a litmus test for both the bank and his own credibility.⁴⁶

The critical but unresolved question for the banks and regulatory authorities on both sides of the Atlantic as well as here in Australia is to what extent the imposition of an external monitor who reports to the regulator rather than the board reflects 'the new normal' – the theme of the upcoming Australian Securities and Investments Commission Annual Forum on international regulatory developments.⁴⁷ At its core this involves an adjudication of what constitutes the appropriate level of external oversight over ongoing corporate practice. As such it extends far beyond narrow issues of capitalisation. It focuses attention instead on the critical questions of how to ensure warranted trust in the operation of free markets while balancing more intrusive supervision with requisite levels of both expertise and accountability.⁴⁸ It also underscores the critical importance of evaluating when and on what basis these decisions are made. Notwithstanding their prevalence, there is a remarkable lack of consistency in the application of the use of a deferred prosecution, the size of the fine and whether an external monitor is imposed. A database compiled by the University of Virginia Law School reveals that of 258 negotiated prosecutions, involving either a deferred or non-prosecution 79 have imposed an external monitor: 21 in the case of non-prosecution deals and 58 in which the prosecution is deferred.⁴⁹ Of the total, 56 have involved firms in

45 Attachment B: Corporate Compliance Monitor (n 20) para 1.

46 Barclays Bank, 'Barclays Appoints Sants As Head of Compliance and Government and Regulatory Relations', Press Release, London, 13 December 2012 <<http://group.barclays.com/news/news-article/1329927766649>>. In the period 2009–2010, Sants made three influential speeches on how to design, legitimate and implement regulatory initiatives surrounding the embedding of cultural restraint; see Hector Sants, 'Delivering Intensive Supervision and Credible Deterrence', Speech delivered at the Reuters Newsmaker Event, London, 12 March 2009, at 2, noting: 'The limitation of a pure principles-based regime have to be recognized. I continue to believe the majority of market participants are decent people; however a principles-based approach does not work with people who have no principles.'; Hector Sants, 'Annual Lubbock Lecture in Management Studies', Speech delivered at Said Business School, University of Oxford, 12 March 2010, noting: 'We need to answer the question of whether a regulator has a legitimate focus to intervene on the question of culture. This arguably requires both a view on the right culture and a mechanism for intervention . . . My personal view is that if we really do wish to learn lessons from the past, we need to change not just the regulatory rules and supervisory approach, but also the culture and attitudes of both society as a whole, and the management of major financial firms. This will not be easy. A cultural trend can be very widespread and resilient – as has been seen by a return to a "business as usual" mentality. Nevertheless, no culture is inevitable.'; and Hector Sants, 'Can Culture Be Regulated', Speech delivered at the Ethics and Values in the City Conference, London, 5 October 2010, noting: 'The regulator must focus on the actions a firm takes and whether the board has a compelling story to tell about how it ensures it has the right culture that rings true and is consistent with what the firm does.'

47 In the interests of full disclosure, this author is a keynote speaker (although his address centres on the historical underpinnings of the rationale for intervention).

48 See Morford (n 2) 5: the Department of Justice clearly differentiates the role of the monitor, arguing that the 'monitor is not responsible to the corporation's shareholders. Therefore, from a corporate governance standpoint, responsibility for designing an ethics and compliance program that will prevent misconduct should remain with the corporation, subject to the monitor's input, evaluation and recommendations.'

49 All data sourced from Brandon L Garrett and Jon Ashley, *Federal Organizational Prosecution Agreements*, University of Virginia School of Law <http://lib.law.virginia.edu/Garrett/prosecution_agreements/home.suphp>.

the financial sector.⁵⁰ Of these an independent monitor has been imposed 14 times, equally split between both non-prosecution⁵¹ and deferred cases.⁵² When the decision is made to impose a monitor, however, the terms and conditions follow a generic template.⁵³ The Department of Justice has also sought to impose consistency in how the monitors operate with the public release of guidance to individual prosecutorial units.⁵⁴ The HSBC requirement to subject itself to an external monitor flows precisely within those guidelines.

- 50 Salomon Brothers (jurisdiction not specified; 1/5/92; NPA); John Hancock Mutual Life (Massachusetts; 22/3/94; NPA); Prudential Securities (NY-Southern; 27/10/94; DPA); Lazard Freres (Massachusetts; 26/10/95; NPA); Arthur Andersen (Connecticut; 17/4/96; DPA); Coopers & Lybrand (jurisdiction not specified; 1/10/96; NPA); Credit Lyonnais (California—Central; 7/6/99; NPA); JB Oxford Holdings, Inc. (California—Central; 14/2/00; NPA); HSBC (NY—Southern; 1/12/01; NPA); BDO Seidman (Illinois—Southern & USDOJ Criminal; 12/4/02; DPA); Banco Popular de Puerto Rico (Puerto Rico; 16/1/03 DPA); Bank of New York (NY—Southern and Eastern; 27/5/03; NPA); PNC Financial (Pennsylvania—Western & USDOJ—Criminal; 1/6/03; DPA); Merrill Lynch (USDOJ—Enron; 17/10/03; NPA); Canadian Imperial Bank of Commerce (USDOJ—Enron; 22/12/03; DPA); AmSouth Bancorp (Mississippi—Southern; 12/10/04; DPA); American International Group (AIG-IP PAGIC Equity Holding Company & AIG Financial Products) (Pennsylvania—Western & USDOJ—Criminal; 1/11/04; DPA); Edward D Jones (Missouri—Eastern; 1/12/04; DPA); KPMG (NY-Southern; 26/10/05; DPA); HVB (NY-Southern; 1/2/06); American International Group (USDOJ—Criminal; 7/2/06; NPA); BankAtlantic (Florida—Southern; 25/4/06); BAWAG psk (NY-Southern; 2/6/06; NPA); Mellon Bank, NA (Pennsylvania—Western; 14/8/06; NPA); Prudential Equity Group (Massachusetts; 28/8/06; NPA); Electronic Clearing House (ECHO) Inc. (NY—Southern; 27/3/07; NPA); Omega Advisors (USDOJ—Criminal Division & NY Southern; 5/7/07; NPA); United Bank for Africa (NY-Southern; 6/7/07; NPA); NETeller plc (NY-Southern; 17/7/07; DPA); American Express Bank Intl (USDOJ—Criminal; 6/8/07; DPA); Union Bank of California (USDOJ—Criminal; 17/10/07; DPA); Sigue (USDOJ—Criminal; 23/1/08; DPA); Unum Group (California—Southern; 1/6/08; NPA); Lloyds TSB Bank plc (USDOJ—Criminal; 22/12/08; DPA); UBS AG (Florida—Southern; 18/2/09; DPA); Credit Suisse AG (USDOJ—Criminal; 16/12/09; DPA); General Reinsurance (USDOJ—Criminal; 18/1/10; NPA); Wachovia (Florida—Southern & USDOJ—Criminal; 16/3/10; DPA); Metropolitan Life Insurance Company (California—Southern; 15/4/10; NPA); AllianceOne (USDOJ—Criminal; 6/8/10; DPA); Barclays Bank (USDOJ—Criminal; 16/8/10; DPA); Deutsche Bank AG (NY-Southern and USDOJ—Tax; 21/12/10; NPA); Baystar Capital Management LLC (California—Northern; 1/3/11; DPA); Community One Bank (North Carolina—Western and USDOJ—Criminal; 1/5/11; NPA); UBS AG (USDOJ—Antitrust; 4/5/11; NPA); JPMorgan Chase & Co (USDOJ—Criminal; 6/7/11; NPA); Ocean Bank (Florida—Southern; 12/8/11; DPA); Islamic Investment Cos. Of the Gulf (Bahamas) Ltd (USDOJ—Tax; 12/8/11; NPA); Wachovia Bank NA (USDOJ—Antitrust; 6/12/11; NPA); Aon Corp (USDOJ—Criminal; 20/12/11; NPA); GE Funding Capital Markets Services Inc (USDOJ—Antitrust; 23/12/11; NPA); Diamondback Capital Management LLC (NY-Southern; 20/1/12; NPA); Imperial Holdings Inc. (New Hampshire; 30/4/12; NPA); BDO USA LLP (NY—Southern & USDOJ Tax; 6/6/12; DPA); ING Bank NV (USDOJ—National Security & Criminal Division; 12/6/12; DPA); Barclays Bank (USDOJ—Criminal; 26/6/12; NPA).
- 51 Coopers & Lybrand (obtaining confidential bid information and lying to grand jury); JB Oxford Holdings, Inc (securities fraud; failure to disclose activities and beneficial ownership); Bank of New York (money laundering; unlicensed money transfers; no anti-money laundering programme); Merrill Lynch (false statements; aided and abetted Enron); American International Group (misstatements in periodic financial reports; Bank Secrecy Act; failure to maintain effective anti-money laundering programme); Mellon Bank NA (theft of government property; theft of mail matter; conspiracy); Deutsche Bank AG (tax evasion).
- 52 The deferred prosecutions requiring an external monitor comprise Prudential Securities (fraud in sale of partnership interests – \$330m settlement with SEC; 3 years); Canadian Imperial Bank of Commerce (aided and abetted accounting fraud by Enron – \$80m settlement with SEC; 3 years); American International Group (violations of antifraud provisions and aiding and abetting violations of reporting and record keeping – \$80m settlement with Department of Justice); KPMG (tax fraud; conspiracy to defraud IRS; tax evasion – \$466m settlement with Department of Justice (\$128m disgorgement of fees; \$228m restitution to IRS; \$100m fines to IRS); 3 years); NETeller plc (conspiracy to conduct an illegal gambling business; failure to maintain an anti-money laundering programme – \$136m forfeiture: 2 Years); Lloyds TSB Bank plc (knowing and wilful violations of International Emergency Economic Powers Act – \$175m forfeiture to Department of Justice; 2 years); AllianceOne (violation of Foreign corrupt Practices Act – no fine; 3 years).
- 53 The requirements and language used to describe those requirements are almost identical to those used in cases against Bionet and Smith & Nephew.
- 54 See Morford (n 2).

'To the extent that HSBC Holdings' compliance with obligations as set forth below requires it, HSBC Holdings agrees to require that its wholly-owned subsidiaries comply with the requirements and obligations set forth below, to the extent permissible under locally applicable laws and regulations, and the instructions of local regulatory agencies', runs the opening paragraph of the job description for the position of corporate compliance monitor.⁵⁵ The position is a fixed term for five years, at the end of which HSBC must sever ties with the monitor for at least one year. The role is to evaluate the effectiveness of the internal controls, policies and procedures of the holding company and its subsidiaries in relation to both anti-money-laundering legislation and the remedial action taken in response to the identified failures. An initial report is required within 90 calendar days of the appointment, which itself is mandated within 60 days of the agreement. Four additional reviews are to be conducted on an annual basis, unless the agreement is either terminated or rendered moot because a further material breach triggers immediate indictment.

The reports are to be contemporaneously submitted to the Board of Directors of HSBC Holdings and the Chief of the Asset Forfeiture and Anti-Money Laundering Section of the Criminal Division, the address of which is helpfully provided, as well as to the Federal Reserve and the Financial Services Authority in London. Interestingly, however, the Financial Services Authority is not given any defined right to engage with the monitor, nor are any of the other parties to the agreement.⁵⁶ This is the Department of Justice's show. Although HSBC can identify and propose the candidate, the Department of Justice retains a veto over the appointment and the procedures governing the production of her reports. The arms-length terms as they relate to HSBC are explicit. The appointee cannot have had a material association with the bank. They are less clear-cut in relation to the Department of Justice itself. It does not have to justify its preference beyond ensuring that the appointee is regarded as having requisite if generically explained expertise.⁵⁷ Once appointed, the independent monitor has the capacity to utilise enormous leverage from the Department of Justice. At stake here, therefore, is not just the credibility of the monitor but also the Department.^{57a}

55 Attachment B: Corporate Compliance Monitor (n 20) para 1.

56 The Financial Services Authority has separately agreed that HSBC should establish an anti-money-laundering/sanctions compliance board level committee, review policies and procedures and notes the employment of an independent monitor who is to communicate to the board and to regulators, see Financial Services Authority, 'FSA Requires Action of the HSBC Group', Press Release, London, 13 December 2012.

57 Deferred Prosecution Agreement (n 20) para 9: 'demonstrated expertise with regards to the Bank Secrecy Act; demonstrated expertise in the design and review of corporate compliance policies, procedures and internal controls; the ability to access and deploy resources as necessary to discharge duties and sufficient independence from HSNB Holdings to ensure effective and impartial performance'. For examination of how monitors carry out their roles, see Khanna and Dickinson (n 18) 1725–31 (noting that most tend to be former prosecutors); see also David Hess and Cristie Ford, 'Corporate Corruption and Reform Undertakings: A New Approach to an Old Problem' (2008) 41 Cornell International Law Journal 307, at 341 (noting the importance of industry experience, the necessity of being 'structurally and psychologically independent from the corporation' and having 'own reputational capital at stake'); see also Cristie Ford, 'Towards a New Model for Securities Law Enforcement' (2005) 57 Administrative Law Review 757, at 797–802 (noting the emergence of the monitor as an agent of behavioural change).

57a In December 2012 the American Bar Association announced the formation of a working group to draw up a set of best-practice principles governing how monitorship should operate, see 'ABA Launches Taskforce on Corporate Monitors', *Corporate Crime Monitor*, 5 December 2012 <www.corporatecrimereporter.com/news/200/abacorporatemonitortaskforce12052012/>. The taskforce includes Larry Thompson, the former Deputy Attorney General, who rapidly expanded the use of deferred prosecution (see n 15) and Mary Jo White, the former District Attorney Southern District of New York, who pioneered the extension of the mechanism in the prosecution of Prudential Securities in 1994 (n 50).

The symbiotic nature of the relationship is explicitly spelt out in the terms of the negotiated settlement. The monitor has the right to report any difficulties associated with gaining access to sensitive material, with the Department having the right to make a final determination on what should be disclosed without reference to further external adjudication. The monitor, although ostensibly independent, is unquestionably, therefore, an agent of the Department. On an ongoing basis the work plan for conducting the evaluations of policies, procedures and remedial action, must be submitted to and approved in advance by the Department. Moreover, 'any disputes between HSBC Holdings and the Monitor with respect to the work plan shall be decided by the Department in its sole discretion'.

Although the monitor is encouraged to work closely with HSBC in the preparation of the reports, the bank itself lacks the discretion on whether to implement any recommendation unless considered 'unduly burdensome, inconsistent with local or other applicable law or regulation, impractical, costly or otherwise inadvisable'. In such an event the bank has to provide reasons for the objections 'and shall propose in writing an alternative policy, procedure or system designed to achieve the same objective or purpose'. The parties are then given 30 days to reach an agreement.

In the event HSBC Holdings and the Monitor are unable to agree on an acceptable alternative proposal, HSBC Holdings shall promptly consult with the Department, which will make a determination as to whether HSBC Holdings should adopt the Monitor's recommendation or an alternative proposal, and HSBC Holdings shall abide by that determination.⁵⁸

Moreover, the Department is to be informed if in the course of the monitor's investigation of the efficacy of internal controls, policies and procedures improper conduct or a material violation of the law is uncovered as well as reporting such activity directly to the bank's chief legal officer. This can be bypassed if deemed appropriate by the monitor. The whistle-blowing protection is further embedded in the contractual terms as 'HSBC Holdings shall not take any action to retaliate against the Monitor for any such disclosures or any other reason'.⁵⁹ The Department of Justice recognising that the information contained in the compliance monitors reports may include 'proprietary, financial, confidential, and business information' has agreed, in principle, to keep the reports classified.⁶⁰ Public disclosure 'could discourage cooperation, impede impending or potential government investigations and thus undermine the objectives of the Monitorship'.⁶¹ Even here, however, the Department can override the commitment to confidentiality if it 'determines in its sole discretion that disclosure would be in furtherance of the Department's discharge of its duties and responsibilities or is otherwise required by law'.⁶²

Taken together the provisions governing the appointment and ongoing work of the monitor reflect an unparalleled extension of external oversight. As such they allay judicial

58 Attachment B: Corporate Compliance Monitor (n 20) para 5.

59 Ibid para 8

60 Ibid.

61 Ibid.

62 Ibid para 9.

suspicion about limited exercise of discretion.⁶³ Just as significantly they transfer knowledge directly to the Criminal Division of the Department of Justice, whose remit is governed by very different imperatives than prudential or market conduct regulators. A new cop is on the beat and making its presence felt. Those drinking in the last chance saloon are on notice that anti-social behaviour orders have been written and will be applied in the event of further infractions. It is not before time. The challenge for the Department of Justice, however, is to exercise its enhanced power with restraint and within accountable boundaries. If not, the regulatory cycle will turn once more, with accusations of overreach and unconstitutionality replacing quiescence in the creation of robust external oversight.

5 The policy implications

The external monitor at HSBC holds what Lanny Breuer describes as a ‘Sword of Damocles’ over the bank. It also applies to the Department of Justice itself. Future violations will automatically trigger the criminal conviction and could produce the very outcome the settlement is designed to avoid.⁶⁴ Equally, the application of external stewardship can have far-reaching consequences. In the aftermath of the settlement, the HSBC share price rose marginally, reflecting a degree of closure. If anything, however, the sword is even more delicately poised. As with the global media industry in the aftermath of the Leveson Inquiry in the UK⁶⁵ and its facsimile in Australia,⁶⁶ however, the banking sector is drinking in the last chance saloon as a consequence of the burgeoning Libor scandal.

Compliance or cultural problems within a single bank, no matter how serious, can be contained by one of three methods. First, the company can adopt voluntary structural reform, an approach initially favoured by HSBC’s Stuart Levey but ultimately rejected by the

63 *SEC v Bank of America* 09 Civ. 6829 (SDNY, 14 September 2009) 8. Judge Jed Rakoff held ‘the proposed settlement in relation to claim that Bank of America had misled investors over the payment of bonuses to executives within Merrill Lynch is described as “a contrivance designed to provide the SEC with the facade of enforcement and the management of the Bank with a quick resolution of an embarrassing inquiry”’. Judge Rakoff reluctantly signed off on the settlement, citing judicial restraint but stating that the settlement was ‘half baked justice at best’, see *SEC v Bank of America* 09 Civ 6829 (SDNY, 22 February 2010) 14. Similar frustration has been voiced by Judge Ellen Segal Huvelle, who refused to endorse a \$75m fine agreed by Citigroup to settle charges that the bank had misled investors over its sub-prime exposure, see Kara Scannell, ‘Judge Won’t Approve Citi-SEC Pact’, *Wall Street Journal*, 17 August 2010, B1: ‘I look at this and say, “Why would I find this fair and reasonable” . . . You expect the court to rubber stamp, but we can’t.’ See generally, Binyamin Appelbaum, ‘US Judges Sound Off on Bank Settlements’, *New York Times*, 23 August 2010, B1 (noting broader opposition to recent settlements proposed with Barclays, Citigroup and Bank of America). In a subsequent case taken against Citigroup, a firm Judge Rakoff described as a ‘recidivist’ offender, the District Court Judge refused to endorse the agreement. The judgment is currently under appeal, with the Securities and Exchange Commission describing it as unwarranted judicial interference on its discretion, see Securities and Exchange Commission, ‘Enforcement Director Statement on Citigroup Case’, Press Release, Washington DC, 15 December 2011. The UK intends to ensure that ‘a prosecutor is not entering into a “cosy deal” with a commercial organization behind “closed doors” by ensuring judicial oversight of the entire process’, Ministry of Justice (n 13) 21.

64 See Masters (n 18). For review of how individual corporations have fared post corporate prosecution, see Gabriel Markoff, ‘Arthur Andersen and the Myth of the Corporate Death Penalty: Corporate Criminal Prosecution in the Twenty First Century’ (2013 forthcoming) University of Pennsylvania Journal of Business Law, working paper available at <http://ssrn.com/abstract=2132242>. The paper notes that no publicly traded corporation convicted in the period 2001–2010 has failed, which suggests that corporate prosecutions should be privileged because of its inherently stronger demonstration effect: at 7. He does accept, however, that the deferred prosecution should be used in situations where ‘a prosecution might actually threaten a company’s survival’: at 44. Arguably banks are in this position.

65 Lord Justice Brian Leveson, *An Inquiry into the Culture, Practices and Ethics of the Press* (The Stationery Office 2012) <www.official-documents.gov.uk/document/hc1213/hc07/0780/0780.asp>.

66 Justice Raymond Finkelstein, *Independent Inquiry into Media and Media Regulation* (Australian Government, 2 March 2012).

Department of Justice. Secondly, one can, as Lanny Breuer of the Department of Justice has advocated, use a deferred prosecution to facilitate 'a truly transformative effect on particular companies and, more generally, on corporate culture across the globe'.⁶⁷ Thirdly, if necessary, closure, an option advanced by Senator Carl Levin is available. When the identified problems extend to allegations of collusion between banks, however, the entire social construction of the market itself comes under scrutiny. The corruption of core stated values has reached an inflection point with the multifaceted international investigation now underway into price-fixing within Libor. As the influential UK Treasury Select Committee reported in August: 'the standards and culture of Barclays, and banking more widely, are in a poor state. Urgent reform, by both regulators and banks, is needed to prevent such misconduct flourishing'.⁶⁸

The now emboldened Department of Justice and, in particular, its Criminal Division under the direction of Lanny Breuer, is playing a pivotal role in these discussions. Its leveraging power in this and other cases is further strengthened by enhanced whistle-blowing legislation in the United States. In particular, the expansion of a bounty system for those willing to report improper, unethical conduct significantly increases the possibility that such conduct will be reported to external agencies.⁶⁹ The critical question then will not be on the strength of the legal claim but the calculation on whether the complained of conduct can be defended in the court of public opinion.

The reality of complex litigation is that when taking enforcement action regulatory agencies balance the effect of conviction with the political costs associated with bringing uncertain cases to trial.⁷⁰ Beyond the merits of an individual action, wider demonstration effect requires changing both the content and context of the underpinning regulatory regime.⁷¹ First, the preparation of the case and its subsequent staging – including the critical initial presentation of the evidential base – needs to reconfigure media representations of what constitutes acceptable conduct, irrespective of the strength at law of the material claim. Precisely because trial strategies tend to bifurcate between competing (if partially understood) narratives that subsequently gain media traction, it is essential to 'own' the

67 Breuer (n 12).

68 Treasury Select Committee, *Fixing Libor: Some Preliminary Findings* (HM Parliament, London, 22 August 2012); see also O'Brien (n 30); and Editorial, 'Banks Must Learn From Past Scandals', *Financial Times*, 16 December 2012 <www.ft.com/intl/cms/s/0/deb826ea-4600-11e2-b7ba-00144feabdc0.html#axzz2FRxBaPIj> noting: 'a new culture is required at the top of financial institutions' precisely because 'the desire to reinvent banking as a high-growth, high-return business has belied its true social function as a utility . . . This is not something that can be changed by a few rule-tweaks. It requires new direction and leadership.' The problem, however, extends far beyond British banking as the investigation to UBS's involvement and the \$1.5bn settlement makes clear (n 16).

69 Wall Street Reform and Consumer Protection Act of 2010 (Dodd-Frank), s 922; see also Securities and Exchange Commission, *Implementation of the Whistleblower Provisions of Section 21F of the Securities Exchange Act of 1934* (Washington DC, 25 May 2011). Dodd-Frank also provides new enforcement tools to deal with fraud and manipulation in the futures, swaps and broader commodities markets by introducing a reckless standard (s 753), which reduces the scienter threshold from deliberate intent.

70 In an interview conducted in the aftermath of the Enron and WorldCom accounting scandals, Steve Cutler, then Director of Enforcement at the Securities and Exchange Commission, noted, the 'reluctance on the part of federal prosecutors to take on complicated accounting fraud cases. These are very difficult cases and require lots of resources, lots of time, [are] difficult to explain to juries and that makes for a less than ideal track record as far as a prosecutor is concerned.' Interview with Steve Cutler, Director of Enforcement, Securities and Exchange Commission (Washington DC, 11 May 2005).

71 A regulatory regime can be defined as the 'complex of institutional [physical and social] geography, rules, practice and animating ideas that are associated with the regulation of a particular risk or hazard,' see Christopher Hood, Henry Rothstein and Robert Baldwin, *The Government of Risk* (Oxford University Press 2001) 8.

media agenda.⁷² Second, the litigation needs to be capable of recalibrating – without credible dissension – the broader policy reform agenda.⁷³

This coupling is essential to ensure that neither judicial failure nor premature settlement will translate into an incremental erosion of wider support for the legitimacy of the regulator's operational imperatives.⁷⁴ As a consequence of Dodd–Frank along with public and judicial disquiet at the weakness of settlements, however, the calculation has changed. The agencies most poised to take advantage of looser *scienter* standards include the Commodity and Futures Trading Commission, which has spearheaded the investigation into Libor scandal. According to its head of enforcement, David Meister, the agency 'is looking to bring high impact cases that influence market behavior'.⁷⁵ The HSBC settlement is arguably, therefore, the most important and likely to be most influential on both regulator and regulated communities alike as a bargaining chip in this complex negotiation. The unresolved question is whether it is an outlier or reflects a determination to ensure ongoing substantive monitoring in order to prevent what the Assistant Attorney General Breuer terms the 'Sword of Damocles' now hanging over the banking sector from falling.⁷⁶

6 Conclusion

All too often in the past banks have made empty promises at congressional hearings before going on to commit further violations, with monetary fines written off as the cost of doing business. In part HSBC's apparent conversion can be traced to narrow self-interest. Senator Carl Levin had warned that regulators must consider the ultimate sanction of bank charter revocation in the US if international banks fail to internally police deviance, the primary reason he endorsed the muscular action taken by the Department of Financial Services in New York.⁷⁷ In part, also, however, the HSBC response reflects an awareness of custodian and broader gatekeeper obligation, which if monitored effectively offers a potential model to transform.

The Department of Justice has recognised the value of such an approach but has made it clear that self-regulation can only work effectively if enforced. It feeds into a crisis that

72 See Janet Malcolm, 'Anatomy of a Murder Trial', *New Yorker*, 3 May 2010, 36. For application to financial crisis, see John Cassidy, 'Scandals', *New Yorker*, 3 May 2010, 21: 'Few things excite the public as much as financial scandals . . . the result is a barrage of news stories that most people do not fully understand but which create a widespread sense that some unprecedented skullduggery has been revealed and that villainous investment bankers will finally be held to account.'

73 Jonathan Nash, 'Framing Effects and Regulatory Choice' (2006) 82 *Notre Dame Law Review* 314.

74 For trenchant critique of the deferred prosecution as an abuse of process, see Richard Epstein, 'The Deferred Prosecution Racket', *Wall Street Journal*, 28 November 2006, A14 (arguing that the agreed statement of claims 'often read like the confessions of a Stalinist purge trial').

75 Stuart Gittleman, 'US Regulators Caution Financial Firms to Stay Alert for New Priorities, "Game Changers"', Thomson Reuters Accelus, 14 December 2012, <www.complinet.com/global/news/news/article.html?ref=160823>.

76 Dominic Ruse and Jill Treaner, 'HSBC's Record \$1.9bn Fine Preferable to Prosecution, US Authorities Insist', *The Guardian*, 11 December 2012 <www.guardian.co.uk/business/2012/dec/11/hsbc-fine-prosecution-money-laundering>. The Department of Justice imposed a non-prosecution agreement on UBS in large part because of the scale of cooperation and extent of management change, see Department of Justice, 'Assistant Attorney General Lanny Breuer Speaks at UBS Press Conference', Washington DC, 19 December 2012. It also, however, secured a guilty plea from UBS Securities Japan, the subsidiary at the heart of the deception, which will not invalidate UBS's US banking licence.

77 Carl Levin, 'Levin Statement on Standard Chartered Bank Settlement', Press Release, Washington DC, 15 August 2012). Levin argued that the settlement 'showed that holding a bank accountable for past misconduct doesn't need to take years of negotiation over the size of the penalty; it simply requires a regulator with backbone to act. New York's regulatory action sends a strong message that the United States will not tolerate foreign banks giving rogue nations like Iran hidden access to the US financial system.'

calls into question as never before both the activities of the banks and their regulators. Globally, the practical and conceptual underpinnings of financial regulation are being questioned as never before. The legitimacy problem is serious, pressing and structural. It is one we ignore at our peril. Following the banking scandals of 2012, it is unsustainable for regulation to be decided and implemented and monitored at a national level. As HSBC has acknowledged, global oversight has become an imperative to reduce the conflicts of interest that may create profitable industries, but not socially beneficial ones. The monitor, as custodian of that purpose, will play an essential validating role. As such, the Department of Justice has taken a first, if uncertain, step towards recognition of globalised agendas. It is an exploration to be welcomed, as much in New York and Washington as in London.

Leading the charge? Payments for single use carrier bags in Wales

LORI FRATER AND ROBERT G LEE

ESRC Research Centre for Business Relationships, Accountability, Sustainability and Society, Cardiff University

Introduction

On 1 October 2011, the Welsh government (WG), employing powers conferred by ss 77 and 90 of and Schedule 6 to the Climate Change Act 2008 (the Act),¹ introduced the first mandatory charge for single use carrier bags in the UK through the implementation of the Single Use Carrier Bag Charge (Wales) Regulations 2010² (the Regulations). Under the Regulations, retailers are required to charge a minimum of 5p for all single use carrier bags, which meet the definitional requirements laid down in the Regulations. This definition includes plastic, paper, biodegradable and recyclable carrier bags. All retailers, not only those retailers that sell food and groceries, are affected by the Regulations, which apply equally to sales in store and online. The Regulations extend to goods purchased in Wales and goods delivered in a single use carrier bag to someone in Wales. Wales was the first of the devolved administrations to introduce such a scheme, though the Republic of Ireland introduced a nationwide charge for plastic carrier bags as long ago as 2002.³ At the time of writing, consultations on similar schemes are underway in Scotland and recently concluded in Northern Ireland. This paper considers this activity on the part of devolved administrations in the UK. It then draws on our two empirical studies to analyse the experiences of the implementation process in Wales of both the general public and of retailers whose behaviour is effectively regulated by the change. It concludes with some reflections on introducing legislative change which will affect the everyday behaviour of the population as a whole. It opens, now, with a consideration of why one might want to reduce the number of single use carrier bags in circulation.

It is estimated that 500bn plastic carrier bags are used worldwide each year.⁴ In the UK alone, in 2010 there were approximately 6.8bn plastic bags handed out by retail shops. In 2009, an estimated 445m carrier bags were used by shoppers from the major supermarkets

1 The Act extends to both England and Wales.

2 2010 No 2880 (W 238) as amended by Single Use Carrier Bags Charge (Wales) (Amendment) Regulations 2011 (2011 No 2184 (W 236)).

3 SI No 605/2001 Waste Management (Environmental Levy) (Plastic Bag) Regulations, 2001 made under the Waste Management Acts 1996 and 2001 and see F Convery, F McDonnell and S Ferreira, 'The Most Popular Tax in Europe? Lessons from the Irish Plastic Bags Levy' (2007) 38 *Environment and Resource Economics* 1–11.

4 K A Spokas, 'Plastics: Still Young, but Having a Mature Impact' (2007) 28(3) *Waste Management* 473–4 and 'Waste: An Overview' (2005) 77(9) *Geographical* 34–35.

in Wales⁵ which equates to 273 bags per household. On average, every kilogram of litter collected contains 3.4 plastic bags,⁶ constituting around 2.7 per cent by weight of all litter and costing Welsh local authorities an estimated £1m to clean up plastic bag litter every year.⁷ Many single use carrier bags are made of oil-based plastic, which is a non-renewable resource and can take up to 500–1000 years to decompose. The United Nations Environment Programme estimates that there are 46,000 pieces of plastic litter floating in every square mile of ocean.⁸

It is no accident that the powers to tackle carrier bag use in England and Wales are contained in the Climate Change Act 2008. A large amount of harmful emissions are produced during manufacture, shipment and recycling of plastic bags. Paper carrier bags, which are often considered a more environmentally friendly alternative, also have environmental consequences as much of the pulp used for paper shopping bags is virgin pulp because it is considered stronger. Paper production requires hundreds of thousands of gallons of water as well as toxic chemicals like sulphurous acid, which can lead to acid rain and water pollution. Even paper bags that come from a renewable source and are biodegradable require more energy than plastic bags to manufacture and transport. A report published by the Environment Agency estimates that rather than use a plastic carrier bag just once, we need to use a paper bag three times to match the global-warming potential.⁹

Enabling provisions and devolved government

The Climate Change Act 2008 introduces a long-term legally binding framework to tackle the dangers of climate change. The central pillars of the legislation are legally binding targets for reducing emissions of greenhouse gases (GHG) by 2020 and 2050.¹⁰ The Act aims for a built-in series of duties, actions and reports, which, whilst not necessarily enforceable in their own right,¹¹ create the transparency, accountability and political pressure necessary to achieve the purpose of the legislation.

Powers to introduce a charge for carrier bags are contained in Part 5 of the Act, alongside supporting provisions aimed at reporting on, measuring and reducing carbon emissions in different sectors of the economy and society. Schedule 6 of the Act enabled the WG to introduce a minimum charge for carrier bags but not a maximum charge.¹² Schedule 6 is silent as to where the proceeds of the charge should be directed. Between the passage of the 2008 Act and the time that Wales began to plan to act on carrier bags, it had acquired additional legislative competence. The Government of Wales Act 2006 had

5 WG (2010) *Proposal for a Charge on Single Use Carrier Bags: A Consultation on the Draft Single Use Carrier Bag Charge (Wales) Regulations 2010*, June 2010.

6 Waste Awareness Wales, 2011 <www.wasteawarenesswales.org.uk/recycle/plastic_bags.html>.

7 WG (n 5).

8 United Nations Environment Programme (2006), UNEP-IUCN, *Ecosystems and Biodiversity in Deep Waters and High Seas*, UNEP Regional Seas Report and Studies No 178.

9 Environment Agency (2011), *Evidence: Life Cycle Assessment of Supermarket Carrier Bags: A Review of the Bags Available in 2006 Report* SC030148.

10 Client Earth (CE), *The UK Climate Change Act 2008: Lessons for National Climate Laws* (CE November 2009).

11 Cf *Friends of the Earth and Ors v Secretary of State for Energy* [2009] EWCA Civ 810 and see M Stallworthy, 'Legislating against Climate Change: A UK Perspective on a Sisyphean Challenge' (2009) 72 *Modern Law Review* 412–36.

12 See Schedule 6 para 4.

introduced a certain competence to legislate in environmental matters¹³ and, following the National Assembly for Wales (Legislative Competence) (Environment) Order 2010, the National Assembly was able to make provision (inter alia) relating to preventing, reducing, collecting, managing, treating or disposing of waste. This it did by the Waste (Wales) Measure 2010 and that measure amended Schedule 6 of the 2008 Act in relation to the destination of proceeds from the charge in Wales by inserting a new para 4A. That paragraph allows the Welsh ministers to provide for the application of the net proceeds of the charge to specified purposes.

In spite of the passage of the 2008 Act, the Westminster government has at no point sought to use the powers conferred under this Act to introduce a carrier bag charge in England.¹⁴ Part of the reason might be that retailers have attempted to persuade a government committed to deregulation that much could be achieved by voluntary action. Coinciding with the Climate Change Act, six major retailers, the British Retail Consortium, the Department of the Environment, Food and Regional Affairs (DEFRA) and the Scottish, Welsh and Northern Irish governments entered into a voluntary agreement to cut the number of carrier bags distributed by the end of May 2009 by 50 per cent (against 2006 levels) and to achieve a 70 per cent cut 'in the longer term'.¹⁵ This was in fact working with the trend since between 2006 and 2009, plastic bag use in the UK declined by about 40 per cent to under 6.5bn¹⁶ as against 2006 when shoppers in the UK were taking home approximately 11bn plastic bags, equating to more than 400 bags per household or 475m plastic bags a month. By May 2009, this agreement had resulted in a 48 per cent reduction across the UK, with a 49 per cent reduction in Wales on the number of single use carrier bags given out.¹⁷ However, since the recession, plastic bag use is once again on the increase. The British Retail Consortium has accredited this increase to shoppers changing their method of shopping, now making a number of short trips rather than a single big weekly shop.¹⁸ In Wales, it was felt that progress was insufficient and that further action was required in order to adapt shopping habits to a level of improved sustainable consumption.

Scotland consulted on a mandatory plastic bag charge of 10p in 2005 and had proposed to introduce the Environmental Levy on Plastic Bags (Scotland) Bill but decided not to proceed to legislate amidst doubts about the level of support. As it now revisits the issue, the Scottish government is proposing regulations under s 88 of the Climate Change (Scotland) Act 2009 which provides specific powers to make regulations requiring retailers to charge for carrier bags with the proceeds from the charge going to the advancement of environmental protection or improvement.

In Northern Ireland an initial consultation process closed in October 2011. Following the Welsh example, government in Northern Ireland consulted on the types of retailers to be covered by the charge, what type of bag should be included, the level of the charge/levy

13 The National Assembly's competence to legislate on these topics is found in Matters 6.1, 6.3 and 6.4 contained in Field 6 of Part 1 of Schedule 5 to the Government of Wales Act 2006 as extended by the National Assembly for Wales (Legislative Competence) (Environment) Order 2010.

14 Though in September 2011, the Prime Minister, David Cameron, did warn retailers that they had to reduce the number of plastic bags they distributed or in the future a statutory ban on plastic bags or a mandatory charge may be introduced. See, 'Cameron Threatens New Law Following "unacceptable" Rise in Plastic Bag Use', *The Guardian*, 29 September 2011.

15 ENDS Report 428, *Carrier Bag Cuts behind Target*, 28 September 2010, 19–20.

16 BBC, 8 September 2011, 'Plastic bag ban "could be considered", minister says' <www.bbc.co.uk/news/uk-politics-14841492>.

17 WG (n 5).

18 'Plastic Bag Use on the Rise after Years of Decline', *The Guardian*, 28 July 2011 <www.guardian.co.uk/environment/2011/jul/28/plastic-bag-rise>.

and enforcement and sanction provisions. The consultation proposal, which remains, is that the funds raised from the charge will be forwarded to the government¹⁹ whereas it is recommended that the funds raised by the charge in Wales are distributed to charities. Consultation on the draft Single Use Carrier Bags Charge Regulations (Northern Ireland) 2012 has recently concluded and a two-phase programme of charging is recommended with a 5p levy on single use carrier bags, effective from April 2013 in phase 1. From April 2014, in phase 2 the levy will be increased to 10p and extended to lower-cost reusable bags.²⁰ In Northern Ireland, in an independent survey of consumers, carried out for the Department of the Environment, 66 per cent of those surveyed said they would welcome a bag levy. In contrast, 42 per cent of businesses identified additional charges for their customers during a period of recession, while 46 per cent of consumers expressed concern at the prospect of having to buy bags.²¹

The experience of the Republic of Ireland suggests that a levy can quickly reduce plastic bag use, which fell from an estimated 328 per person to 21 per person with the introduction of the measure in 2002. However, the charge rose rapidly following its introduction. Since 2002, the Irish government has increased the charge from €0.15 to €0.22. Under the Waste Management (Landfill Levy) Regulations 2011,²² the plastic bag levy can be amended once in any financial year by the application of the consumer price index plus an additional 10 per cent at the discretion of the Minister for the Environment. The ceiling for the plastic bag levy is set at €0.70. The levy was increased as a rise in bag distribution had occurred from 21 bags per capita to 31 bags per capita and following the heightened levy, per capita bag usage fell back to 21. However, an unintended consequence of the levy in Ireland was an increase of around 75 per cent in sales of bin liners and refuse sacks after the levy was introduced, as fewer people used carrier bags as bin liners.²³

One notable feature of the move to limit bag use is that it is an increasingly rare example of environmental regulation not mandated by the EU. However, the issue is now on the EU environmental agenda. Janez Potočnik, European Commissioner for Environment, introducing an EU consultation has argued that:

Fifty years ago, the single-use plastic bag was almost unheard of – now we use them for a few minutes and they pollute our environment for decades. But social attitudes are evolving and there is a widespread desire for change.²⁴

From May until August 2011, the European Commission held the consultation on whether to tax or simply ban plastic carrier bags. The consultation, open to the public, organisations and public authorities, asked if charging and taxation would be effective, or if other options such as an EU-level ban on plastic carrier bags would be a better regulatory option.²⁵ The consultation also aimed to gather views on the adequacy of current requirements on compostability and biodegradability of plastic carrier bags as provided in the EU Packaging

19 Department of the Environment, Northern Ireland, *Proposals for a Charge on Single Use Carrier Bags*, 20 July 2011, at 7.

20 Department of the Environment Northern Ireland, *Consultation Document: The Draft Single Use Carrier Bags Charge Regulations (Northern Ireland) 2012*, 16 April 2012.

21 'Consumers Give Green Light to Plastic Bag Tax', *Belfast Telegraph*, 9 December 2011 <www.belfasttelegraph.co.uk/business/business-news/consumers-give-green-light-to-plastic-bag-tax-16088785.html>.

22 SI No 434/2011.

23 M Brennan, 'Shoppers Still Bagging Plastic Sales', *Irish Examiner*, 29 January 2003.

24 European Commission, 'Commission Seeks Views on Reducing Plastic Bag Use', Press Release IP/11/580, 18 May 2011.

25 Ibid.

Directive.²⁶ That Directive makes no clear distinction between biodegradable products that should biodegrade in natural conditions in the environment and compostable products that only biodegrade in industrial composting facilities. Therefore, the consultation proposed that each plastic bag should carry a label clearly defining whether it is 'biodegradable', meaning that it will biodegrade in natural conditions in the environment; or whether it is 'compostable', meaning that it will break down only in industrial composting facilities.

The European Commission acknowledged that some member states had already taken action to reduce the use of plastic carrier bags through: pricing measures (Republic of Ireland, Germany, Denmark); agreements with the retail sector (UK); and bans on certain types of bags (Italy, France). However, no specific measures exist at the EU level, though in March 2011 EU Environment Ministers discussed the environmental impact of plastic carrier bags and the concerns they raised indicated that effective EU action is needed. At the end of the consultation, over 15,500 responses had been received and approximately 53 per cent of respondents 'strongly agreed' with the most stringent measure proposed, namely an EU ban on plastic bags. In addition, a total of 65 per cent of people strongly agreed that some form of measure needed to be adopted at EU level to reduce the use of plastic bags.²⁷

It may be then that UK legislation in this area needs to be adapted at some future point to meet a harmonised EU solution. For the moment, however, three of the four UK administrations look likely to tackle the issue of retail bag use in some way. However, the outcome may be four different carrier bag schemes across the UK. As many retailers operate across the UK, they may have to introduce different systems for their shops in Wales, Scotland, Northern Ireland and England. One national retailer interviewed during our survey, which is outlined below, commented: '[We are] not going to change our bags, we are a massive organisation with only a couple of stores in Wales so not going to change just for Wales.' We now turn to review the Welsh experience.

The carrier bag charge in Wales

As from 1 October 2011, retailers located in Wales or delivering goods to someone in Wales by single use carrier bags are required to charge a minimum of 5p for all single use carrier bags (reg 6) whether in store and online sales (reg 4). The charge applies to a wide range of retailers from high street to local shops, from market stalls to charities, from takeaway restaurants to opticians.²⁸ Under the Regulations, a single use carrier bag is defined as one made from paper, plant-based material or natural starch and is not manufactured for multiple use (reg 3(2)) or is made from plastic, not intended for multiple use and is not classified as a 'bag for life' (reg (3)(3)). Under reg 7 and Schedule 1, specific bags are exempt from the charge including: cloth, jute, cotton, hessian, hemp, wicker and heavy duty plastic. Also exempt are bags for unwrapped food items such as fruit and vegetables, bags for uncooked raw meat and fish or small flat paper bags for greeting cards. One particular exemption is for items provided on a prescription or as part of another NHS service

The exemptions are a source of misunderstanding and confusion. Under the Regulations, an exempt bag may lose its exempt status if goods, which do not qualify under Schedule 1, are also put into the exempt bag. The Regulations are aimed at the retailer who distributes a single use carrier bag and therefore, the exemption is lost when, in a single transaction, goods which do not fall under Schedule 1 are put into a bag with the Schedule 1

26 European Parliament and Council Directive 94/62/EC of 20 December 1994 on packaging and packaging waste.

27 E Gyekye, 'Public Strongly Agree with EU Bag Ban Plans', *Packaging News*, 12 January 2012.

28 See Annex 1 for the full list.

goods. In a pharmacy, for example, if prescription goods (identified under Schedule 1) are put into a bag, the bag is exempt and no charge applies. However, if the customer is also buying other items, for example, shampoo or moisturiser (goods not included under Schedule 1), and these items are also put into the bag, the bag is no longer exempt and a charge ought to apply. This type of problem can occur in other situations. Bags provided where a shop provides a service, for example, shoe repair, are exempt and not subject to a charge, but if the business also provides goods such as shoe polish it will need to charge for the bag.²⁹ The Regulations only allow for a bag to be charged for once. Consequently, where charities give out bags that have already been used, the shoppers would not need to pay the charge on those bags.

The record-keeping requirements on retailers are quite onerous as a record must be kept of: the number of single use carrier bags supplied which meet the requirements of the charge; the gross amount collected from the charge; the net proceeds, minus VAT liability and any other reasonable costs; and how the net proceeds have been distributed. However, by amending regulations, retailers with fewer than 10 employees are exempt from the requirement to keep records. Nonetheless, if they are VAT registered they are still required to pay all VAT liabilities arising from the charge and consequently will need to keep records for this purpose. Any retailer who does not charge for a non-exempt single use carrier bag, or fails to keep, retain, supply and publish records (where required) is in breach of the regulation (reg 11) and could face civil sanctions of a fixed penalty of up to £200 (Schedule 2) or a discretionary penalty up to £5000 or up to £20,000 if they provide false or misleading information (Schedule 3). Local Authorities are responsible for administering and enforcing the charge (reg 5). They are permitted to make test purchases of goods for the purposes of ascertaining whether the retailer is complying with the Regulations (reg 14(2)(a)). The Regulations do not provide any requirement on the final destination of the charge, but the WG has recommended that the proceeds are passed on to good causes in Wales,³⁰ though there is no requirement on the retailers to disclose the decisions they make about how they use the additional funds generated by the obligation to charge (reg 8).

Public attitudes to charging

In mid-September 2011, a team of researchers from the ESRC BRASS³¹ Research Centre at Cardiff University investigated the introduction of the new charge in Wales in two distinct but inter-related phases. Phase I was an analysis of the attitudes, awareness and acceptance of the charge by the public, whilst phase II was an assessment of the attitudes of retail companies, their level of understanding of the regulations and WG guidance, the practicalities of introducing the charge (including staff-training and record-keeping) and questions relating to the proceeds from the charge. The first phase of this work is now considered.

Face-to-face surveys of 600 members of the public were conducted at four different sites across Cardiff in mid-September 2011. The interviewees were randomly selected by a group of surveyors. The four sites were chosen to capture respondents within an area which offered as wide as possible selection of retail outlets within the city. The four sites captured potential and/or actual shoppers from brand-name stores, supermarkets, fast-food and takeaway outlets, clothing and charity shops as well as market stalls in both local and city-centre locations. Demographic information relating to gender, age group and employment status/sector was collected in addition to a further 10 questions on the level

29 WG, *Guidance on the Single Use Carrier Bag Charge (Wales) Regulations 2010*, September 2011, para 21.

30 WG, *Retailer Information Pack, Single Use Carrier Bag Charge: Key Points* (WG 2011).

31 ESRC Research Centre for Business Relationships, Accountability, Sustainability and Society (BRASS) at Cardiff University of which both authors are members <www.brass.cf.ac.uk>.

of awareness of the introduction of the charge,³² whether interviewees knew the extent of the charge in relation to shops and types of bag included. Respondents were further asked whether and why they agreed or disagreed with the charge and why they believed the WG had introduced such a charge. They were also asked what maximum charge they would be willing to pay for carrier bags.

The survey sample was not a proportional representation of the population of either Wales or Cardiff. It was based on a random selection of 600 people in Cardiff who were willing to complete the survey questionnaire. The sample does represent, however, a close proportional representation of the Wales gender demographics, which is 51 per cent female and 49 per cent male as against the survey sample which was 54 per cent female and 46 per cent male.³³ In Cardiff by June 2010, the population between the age of 16 and 64 (excluding students) was 196,800. Of this figure 79 per cent was employed (the survey captured 43 per cent) and 8 per cent unemployed (the survey captured 6 per cent). Given the time of day of the survey (mid-morning to mid-afternoon), the number of employed respondents is likely to be less than the Cardiff employment percentage. In 2009, the WG estimated that in Cardiff 16 per cent of the population was of retirement age (SDR 40/2009, Statistical Focus of Age in Wales, 2009). Of the people in the survey sample, 19 per cent classified themselves as retired.

There was a high level of awareness of the forthcoming charge with 85 per cent of all respondents stating that they were aware that a carrier bag charge was to be introduced. However, of those who were aware, only 41 per cent knew that the charge would commence on 1 October 2011 (within two weeks of the survey taking place). Moreover, more detailed knowledge of the charge was often lacking. 60 per cent of respondents (wrongly) believed that the charge applied only to plastic carrier bags and although 51 per cent identified (correctly) that the charge would apply to all types of shops, 27 per cent believed that it applied to supermarkets only. There was widespread support for the charge, with 70 per cent of the respondents agreeing with the introduction of the charge, and only 26 per cent disagreeing. The remaining 4 per cent had either no view, were undecided or were not interested.

Of the 70 per cent who agreed with the charge,³⁴ they cited as their main reasons for support: the positive impact on the environment; improved litter control; and a reduction in the use of plastic bags as well as a general reduction in waste generated. Many within this group defined environment as their local surroundings and linked environment to improved litter control. There was little appreciation of wider environmental issues such as resource use or indeed that the move might be linked to such issues as emissions' reductions and wider issues of climate change as reflected in the 2008 Act.³⁵

Of the 26 per cent who disagreed with the charge, the main reason given was the additional cost on already expensive food bills. There was also a feeling that the charge should not apply to items such as clothes or expensive goods as a carrier bag was a part of the retail service experience. Interestingly, 38 per cent of respondents said that they were

32 On the whole attitudes based on gender, age, etc are not examined in the analysis presented here, which deals with broad questions of attitudes to the charge.

33 All statistics here taken from StatsWales 2010 <<http://statswales.wales.gov.uk/index.htm>>.

34 This figure seems remarkably resilient as it matches that in a study published in June 2012 of attitudes to the charge: see W Poortinga, L Whitmarsh and C Suffolk, *Evaluation of the Introduction of the Single-Use Carrier Bag Charge In Wales: Attitude Change and Behavioural Spillover*, Report to Welsh Government by Cardiff University, June 2012.

35 Cf the comment by one respondent: "No need to introduce [it], climate change is happening and we can't stop it." (Interviewee 242).

willing to pay more than 5p for a carrier bag, though for some this did depend on the quality of the bag, and 29 per cent said 5p was the maximum amount they would pay. A further 24 per cent said that they would be unwilling to pay at all for a bag but this did not necessarily indicate distaste for the policy but meant that it was their firm intention to shop with their own reusable bag and therefore never make a payment. Only 11 per cent of respondents said they were likely to buy bags,³⁶ with 67 per cent indicating an intention to use their own bags. There was an identifiable link between willingness to pay and the use to which payments would be put. This might well influence behaviour in Northern Ireland where revenues will return to the government rather than the charitable sector, which is often the case in Wales. In the survey, 13 per cent believed that the reason the charge had been introduced was to raise money for the WG. It was viewed (wrongly) as a tax or as a means of dealing with the economic climate by bringing in additional public revenue. Few respondents knew that the funds would not be returned to government, but respondents did feel that they would be more willing to pay the charge if they were assured that the funds would be distributed to charities.³⁷

In summary, whilst the majority of the public surveyed knew about the charge and were in agreement with it, many of the respondents were hazy on the detail, for example, believing that the charge applied solely to plastic carrier bags. In addition, respondents in favour of the charge still expressed concerns about the reasons why the WG was introducing it and were unaware that the funds would not be returned to the government but that the government had recommended that the funds be given to charities. Respondents aware that the funds were to go to charity were generally supportive of the charge. Local charities in particular seem to have been beneficiaries of the revenues generated by the Welsh scheme and recognition of this, following the introduction of the scheme, appears to have helped public acceptance of the charge. In spite of the fact that there is no legal requirement to forward the proceeds of the charge to charity, the majority of retailers do so and make something of a virtue of this. Again it will be interesting to compare the experience in Northern Ireland, but one might expect that there may be less support by both retailers and consumers for what could be seen as a revenue-raising measure.

Retailer reactions

Phase II of the work was a business survey conducted one month after the public survey (Phase I) in mid-October 2011 and was spread over a two-week period. A database of over 300 companies located in Cardiff was developed from which participating companies were sought. Companies were selected randomly, although a small sub-sample was also identified based on their size to ensure a sample that represented as wide a possible spectrum of the different types of retailers covered by the charge. Retail outlets included bookshops, hairdressers, jewellers, stationers, electronics suppliers and the like as well as departmental stores and food outlets. The latter category itself was wide ranging covering takeaways, bakers, butchers, fruit and vegetable shops, greengrocers, health food shops and off-licences.

The sample also included a broad spectrum of company sizes but did pay particular attention to small to micro-sized organisations as they represent the largest percentage of

36 In the event the food retail sector has seen reductions of between 95 per cent and 70 per cent according to the British Retail Consortium: see John Griffiths AM, Written Statement: An Update on the Single Use Carrier Bag Charge in Wales (WG, 4 July 2012).

37 Reported figures of charitable donations from supermarkets in Wales included £136,352 from Asda (as at end April 2012, and £170,642 from Morrisons (as at 6 April 2012) <<http://greenroom.asda.com/2012/6/20/wales-air-ambulance-benefit-from-carrier-bags>> and <www.morrisons.co.uk/corporate/CR/Our-carrier-bag-savings-in-Wales/>.

companies in Wales.³⁸ The size of company was measured by the full-time equivalent (FTE) staff for each site, however, 59 per cent of the retailers interviewed did belong to a group of shops, though some may have had only one or two other small shops in other locations in Wales; others, including the charities, belonged to larger organisations that spread across the UK. Overall, companies were willing to participate and the target of 50 companies was surpassed with survey data being collected from 90 retailers across Cardiff. The interviews were conducted either face to face or, where more convenient for the company, by telephone. Surveyors sought information on how the charge had been introduced by each company and also views of how the public appeared to be responding to the charge.

Once again, the survey sample was not a proportional representation of the population of retail stores in either Wales or Cardiff. It is a random selection of retail stores located in Cardiff based on type of store and size of store. The data was not grossed up to represent the findings of retail stores in general and is therefore only an indication of the 90 stores surveyed. Of the retailers surveyed, 59 per cent belonged to a chain of stores, with 52 per cent being small-scale chains based solely in Wales.

The data was collected using a standard questionnaire based on information that would reflect the key elements of the obligation placed on the companies by the introduction of the charge. This included data on: charging practices; identification and use of exempt bags; administration of the charge; level of information received and used; destination of proceeds; and customer reactions to the charge. All data collected from the retail stores was gathered in confidence and anonymised to conceal the identity of the companies who agreed to participate in the survey and to ensure appropriate data protection. The data was amalgamated to provide an overall percentage response from the 90 Cardiff retailers participating.

One immediately surprising finding was how little lead-in time many of the companies had had to prepare for the change with almost half of our respondents (48 per cent of companies) claiming to have learnt of the charge only one to two months prior to its introduction. The majority of the respondents (31 per cent) stated that their main source of information about the charge came from the media or news outlets. Just 18 per cent of respondent retailers had been informed about the charge by the WG either in the form of a letter or from the website, though this percentage may be a little artificial: 16 per cent were informed directly by their Head Office, which may well have obtained the information from the WG rather than from a trade association or the media. As many of the head offices of larger stores were based in England, many of the Welsh stores had not received the information directly from government. One consequence of this was that those responsible for introducing the charge in Wales only had partial information about how the charge would operate.

Only 7 per cent of retailers were charging for carrier bags prior to the 1 October 2011.³⁹ In contrast, 94 per cent of Welsh retailers are now charging the minimum price of 5p. Some retailers did state that this did not cover the purchase cost of the bag to the company, potentially indicating that they were unaware that the 5p charge was merely a minimum and that they were able to set a higher charge; and 3 per cent of retailers are charging more than

38 In the 2010 Size Analysis of Welsh Businesses, 60 per cent were identified as micro, small and medium-sized enterprises, SDR 180/2010, released 20 October 2010. In Cardiff, the size breakdown identified that 96.7 per cent of all businesses were identified as micro, small and medium. These figures apply to all business types and not just retail companies.

39 In evidence to government in Northern Ireland, Marks and Spencer disclosed that it introduced a 5p food carrier bag charge in May 2008 aiding a reduction in food bag usage by 80 per cent (or 1.7bn bags) as well as generating nearly £6m to charitable projects across the UK: see consultation response letter of 14 June 2012: <http://corporate.marksandspencer.com/documents/publications/consultations/ni_carrier_bags_2012.pdf>.

5p (between 6p and 10p) and 3 per cent admitted that they were not charging. The reasons provided by the latter group included that: the standard of the bag was of the same quality and description as a bag for life and therefore no charge was required under the legislation; the retailer no longer provided any kind of carrier bag; or customers were asked to make a donation to a charity. Only one company did state that they would not be charging, despite knowing that their bags fell within the regulations.

As explained above, the status of exempt bags is complex. Whilst 45 per cent of the sample stated that they did not have any exempt bags, 35 per cent of these were retailers of the type that one might expect to have some form of exempt bag either because they may sell loose food products (grocers), items that fit into small flat bags (card shops) and those that may reuse bags (charities). This therefore suggested doubts about the level of understanding amongst the companies as to what constituted an exempt bag. Some did state that they would charge for a bag irrespective of its status because they did not want to be prosecuted.⁴⁰ Of the 54 per cent of companies who stated that they distributed exempt bags, 32 per cent reported that customers were confused about how the exemption operated, in particular in relation to food items.

There was confusion also over what information had to be recorded, particularly in relation to 'reasonable costs' that could be deducted from the gross amount collected from the charge. Some retailers thought that they could deduct their purchase cost of the bag, which under the regulations is not permitted; and other retailers were merely operating a 'coin in box' system, which was then handed over to a charity. However, if these companies were VAT registered, they risk a breach of their obligations to pay VAT on the carrier bag charge since the sale of the bag attracts VAT. In fact 54 per cent of the companies were not aware of their VAT obligation, although for many this was administered centrally and therefore was not something they had themselves to consider. Only 7 per cent of companies stated that they would send the proceeds to environmental causes. The WG had stated that it hoped that the proceeds would be passed on to good causes in Wales, in particular environmental projects. However, whilst the majority of retailers were passing on the proceeds to charity, environmental causes were not the main recipients. In Scotland, legislation will allow for the Scottish ministers to direct net proceeds (i.e. profit after administrative costs have been deducted, estimated at £5–6m if an 80 per cent drop is achieved) to be used for 'environmental protection or improvement or to any other purposes that may be reasonably regarded as analogous'.⁴¹

Of those retailers who sold bags for life or other non-single use carrier bags (canvas, hessian or cotton), a minority of stores reported a significant (500 per cent-plus) increase in their distribution since 1 October 2011. The majority of the other stores either stated that there was no change or that sales had increased between 20 and 50 per cent. It should be noted that a distorting factor here is that a number of stores were handing out free bags for life prior to and during the first few weeks of the introduction of the charge. Of the retailers who provided an estimate of how many customers were bringing their own bag, the estimate varied between 20 and 95 per cent of customers coming prepared with their own bag. Those companies that reported significant increases between 60 and 95 per cent were mainly supermarkets and department stores in comparison to charities, book and clothing shops which reported an increase of less than 60 per cent. Nonetheless, 66 per cent of companies did report a reduction in carrier bags. The reduction varied from a slight

⁴⁰ Anecdotal experience of those living or working in Wales bears this out. It is not uncommon to be asked to pay for clearly exempt bags.

⁴¹ Scottish Government, *Consultation on a Programme to Encourage the Sustainable use of our Resources, to Support both the Environment and Economy*, 27 June 2012.

reduction up to approximately 95 per cent with those reporting a reduction of more than 80 per cent being the larger national retail chains. Those recording a reduction between 50 and less than 80 per cent were predominantly charities. Companies reporting a reduction of less than 50 per cent were a mixture of charities and clothes shops. Greengrocers reported only a slight decrease in carrier bags, as did pharmacies and electrical stores. Takeaways, however, did not record any reduction.

Conclusion: communicating change

In one sense the introduction of a charge for carrier bags was not the most difficult exercise in law reform. It had widespread support and there was a good deal of awareness of the impending change. Though an essentially simple regulatory measure in the form of a ban on freely available single use bags, much of the detail of the change was misunderstood in terms of essentials such as what type of bags and which shops would be subject to the charge. Moreover, in spite of public goodwill, very few people were aware that the money could go to charity, with a significant minority of people (13 per cent) believing that this charge was introduced for financial or economic reasons.

There was a widespread belief that the charge was restricted to plastic carrier bags but it was a rather unique feature of the Welsh policy that the charge would be levied on other sorts of bags, such as paper bags. Retailers later reported much stronger levels of consumer resistance when this was realised. Retailers themselves had poor understandings of the detail of the regulation, in particular regarding the use of bags exempt from the charge and, more seriously, in relation to the VAT implications of adding a charge for a bag. This was because, in spite of attempts by the WG to brief retailers by sending out 40,000 information packs, the majority of retailers surveyed in the research reported receiving no official information about the charge. They were reliant on the media or other retailers for the information. This was true also of the shoppers who had heard about the charge from either the media (television, radio and newspapers at 38 per cent) or from shops and supermarkets (33 per cent). Moreover, only 6 per cent had heard about the introduction of the charge from the government-led advertising campaigns with more people (8 per cent) reached by word of mouth. This suggests that, both in the case of retailers and shoppers, the best route to dissemination to those affected by regulation is likely to be via media briefings.

Policymakers may find it helpful to try and communicate some of the finer detail where a measure such as this affects the wider public. Although there was broad knowledge of the charge and good acceptance of it, support improved where the environmental benefits of the scheme and understandings of its charitable nature were better understood. In communicating this sort of change, policymakers may need to understand that government itself is unlikely to be the first source of information either for those directly regulated (in this case retailers) or otherwise affected (as with the shoppers). If other forms of media are the main source of information, policymakers may need to take this into account. For example, the survey shows that knowledge of the change was much higher among older shoppers (over 55s) and much lower among young people. Policymakers may need to consider the types of media accessed by certain age (or other types of) groups.

The uncertainty among retailers, regarding the detail of the regulation, did give rise to difficulty. It was never the intention of the WG to immediately deploy the penalties and sanctions in the legislation for failure to adequately implement the charge. However, misunderstandings concerning potential liability for VAT might not be dealt with only by the WG and, with VAT at 20 per cent, a failure to hold back a penny when making a charge for a carrier bag might have serious consequences for the regulated community when faced with a VAT bill. It would seem that of the £400,000 said to have been spent on

communicating the measure, money may have been better spent securing media coverage of the change than in trying to reach the retail community directly with information packs. Our research suggests that 41 per cent of the retailers surveyed had become aware of the measure in the two months prior to its introduction. This probably followed heightened media discussion of the issue but it meant a short lead-in time to respond to the regulation for those retailers. Earlier media briefings would have certainly been helpful to this body of respondents (which were generally small and medium-sized enterprises).

Nonetheless, the policy should be considered a success and has been assessed as popular and effective in a survey of attitudes towards the charge.⁴² The habit of taking one's own bag to the shop has increased across all age and gender groups. The carrier bag charge attracted widespread support before its introduction and there is no sign that this has diminished.⁴³ Indeed, rather than the public resenting the non-availability of free bags, the policy has attracted support because the general public appear to have responded by retaining their own bags particularly for supermarket shopping. The charitable underpinnings of the scheme have helped maintain popular support.

The ban on carrier bags has been described as a cuddle blanket that was briefly, and horribly, fashionable. This is not quite true since, with measures still to come in Northern Ireland and Scotland, it seems still to be in fashion. But the same commentator⁴⁴ also described the measure as a deadly distraction, presumably because there are more serious environmental threats about which we should worry and which might more fittingly be the subject of legislation. This is undoubtedly so, but it ignores the rhetorical quality of the measure in Wales as a sign of an intent to begin tackling some of these wider problems. However, if the ban on single use bags is more important for its symbolism than its environmental impact, then this might be considered all the more reason why communication of the measure to both retail organisations and the wider public mattered so much and, judging by the results of the surveys, an opportunity to engage on a green agenda was poorly exploited.

42 Poortinga et al (n 34).

43 Ibid.

44 T Gold, 'This Plastic Bag Conspiracy is a Truly Deadly Distraction', *The Guardian*, 3 August 2012.