Northern Ireland Legal Quarterly

Volume 60 Number 4

EDITOR
PROFESSOR SALLY WHEELER

Queen's University Belfast
School of Law

Contents

Representation, bargaining and the law: where next for the unions?
Michael Doherty 383

SOX as a window on transference of corporate governance norms across jurisdictions
George Peter Gilligan 403

A hierarchy of disability rights? A comparative examination of the regulation of digital television in the United States of America and the United Kingdom
Eliza Varney 421

Keeping bad company: building societies – a case study
Lorraine Talbot 443

The limits of legal ideologies
Sean Coyle 471

NILQ 60(4): 383–401

Representation, bargaining and the law: where next for the unions?*

DR MICHAEL DOHERTY

Lecturer in Law, School of Law and Government/Socio-legal Research Centre, Dublin City University

Introduction

The right to freedom of association and the right of trade unions to bargain on behalf of workers are rights that have long been internationally recognised. This can be seen in international law instruments such as the International Labour Organisation's (ILO) Convention No 87 of 1948, the Freedom of Association and Protection of the Right to Organise Convention and, under European law, in the shape of Article 28 of the European Union's Charter of Fundamental Rights and Article 11 of the European Convention on Human Rights and Fundamental Freedoms (ECHR). The right to freedom of association is also protected by Article 40.1.6(iii) of the Irish Constitution. However, as Hogan and Whyte point out, as interpreted by the Irish Supreme Court, this does not imply any duty on the employer beyond respecting the right in itself; in particular, "it does not oblige him to negotiate with any association which employees may form".[1]

This constitutional position reflects the fact that trade union bargaining rights in Ireland, as in the UK, have traditionally been premised on the notion of "voluntarism", i.e. the avoidance of statutory regulation.[2] The position in Ireland remains, however, unique in the Western world in that Irish trade unions have no legislative right to be recognised in the workplace for collective bargaining purposes.[3] This article considers a recent attempt, in the form of the Industrial Relations (Amendment) Acts 2001–2004 (hereinafter referred to as "the Acts"), to address this situation. This research looked at the perceived effectiveness of the legislation from the point of view of the various employment relations stakeholders (unions, employees, employers and the state's industrial relations bodies), by way of a detailed examination of the day-to-day operation of the Acts by the Labour

* I would like to express profound thanks both to the anonymous reviewer and to Anthony Kerr of University College Dublin for the insightful and helpful comments received. I would also like to thank my colleagues in the Socio-Legal Research Centre at DCU, particularly Dr Noelle Higgins. Many thanks, too, to Sinead Ring (UCC) for her work in helping compile the data.

1 G Hogan and G Whyte, *Kelly: The Irish Constitution* 4th edn (Dublin: Butterworth 2003), p. 1803.

2 T Kerr and G Whyte, *Irish Trade Union Law* (Dublin: Butterworth 1985).

3 Twice, in 1967 and again in 1996, groups charged with reviewing the Constitution have considered whether a constitutional amendment was necessary in the area of freedom of association. In both cases, the view expressed was that this matter would be best dealt with by legislation; see *Report of the Committee on the Constitution* (1967) and *Report of the Constitution Review Group* (1996). A statutory procedure for gaining recognition rights now exists in the UK; see the Employment Relations Act 1999.

Court, from their coming into operation in June 2001 to the immediate aftermath of the decision of the Irish Supreme Court in the *Ryanair* case,[4] which, as we will see, led many commentators to proclaim the legislation effectively neutered.[5] The article questions whether the Acts can, despite the pessimism of many commentators, still play a role in Irish industrial relations (IR).

However, more broadly, the Acts represented an innovative attempt to steer a "half-way house" course between legally mandated collective bargaining and voluntarism. As a result, the article will also consider some possible implications of the "Irish model" for IR in other European jurisdictions, particularly in the light of recent developments concerning collective bargaining rights at EU level.

The article is laid out as follows. In the next section, the legislation and the decision in *Ryanair* will be outlined in more detail. Then, the methodology used in the analysis will be explained. The implications of the *Ryanair* decision will then be considered in light of the empirical evidence relating to the interpretation of the Acts. Finally, some implications of the Acts for industrial relations practice and collective bargaining more generally (in Ireland and elsewhere) will be teased out.

The Industrial Relations Acts 2001–2004

The absence of any statutory protection for collective bargaining became progressively more of a concern for the Irish trade union movement in the 1990s. This was partly because the decline in union density (the proportion of workers who are members of a trade union) became increasingly pronounced and partly as a result of a change in state policy in the 1980s, where state industrial development agencies began "marketing" Ireland (particularly to US multinational corporations (MNCs)) as a non-union environment.[6] The unions' dissatisfaction was arguably compounded by the participation since 1987 of the Irish Congress of Trade Unions (ICTU) in the social partnership process, which has given the union movement a key institutional role at national level in relation to socio-economic policy, while at workplace level unions have no such institutional security and must fight recalcitrant employers for recognition rights.

As a result, the union movement pushed for progress on recognition rights for unions at workplace level.[7] Under the fourth of the social partnership agreements, *Partnership 2000*, a high-level group comprising union and employer representatives and industrial relations experts was set up to examine the issue of trade union recognition rights.[8] The result was the drawing up of the Code of Practice on Voluntary Dispute Resolution[9] and the Industrial Relations (Amendment) Act 2001. The Code of Practice and the 2001 Act explicitly exclude the imposition of any "arrangements for collective bargaining", on the grounds of protecting Ireland's voluntarist tradition. The general philosophy behind both

4 *Ryanair* v *The Labour Court* [2007] 4 IR 199.

5 M Connolly, "Industrial Relations (Miscellaneous Provisions) Act 2004 – implications for industrial relations law and practice of the Supreme Court Decision in *Ryanair* v *Labour Court and IMPACT*" (2007) 4 *IELJ* 37; M Doherty, "Union sundown? The future of collective representation rights in Irish law" (2007) 4 *IELJ* 96; cf. A Kerr, "Industrial relations law" in M Regan (ed.), *Employment Law* (Dublin: Tottel 2008).

6 See J Wallace et al., *Industrial Relations in Ireland* (Dublin: Gill & Macmillan 2004); T V Roche, "Pay determination, the state and the politics of industrial relations" in W K Murphy and T V Roche (eds), *Irish Industrial Relations in Practice* (Cork: Oak Tree Press 1997).

7 Particularly in the light of a bitter and high profile 1998 union recognition dispute at the airline company, Ryanair.

8 See para. 9.22 of *Partnership 2000*.

9 Industrial Relations Act 1990 (Code of Practice on Voluntary Dispute Resolution) (Declaration) Order 2000 (SI No 145 of 2000).

is that disputes relating to union recognition should be dealt with within the context of voluntary collective bargaining (with parties offered recourse to the advisory and conciliation services of the Labour Relations Commission (LRC)). Thus, the 2001 Act does not provide for union recognition, but for a range of procedures to allow unions to seek to have specific disputes with regard to pay, terms and conditions of employment and dispute resolution procedures addressed.[10] The provisions of the Act are used as a fallback measure whereby, in a situation where the parties cannot come to agreement under the "voluntary leg" of the process, a union or excepted body[11] may request a further investigation by the Labour Court, which can issue a Recommendation.[12] Should the issue remain unresolved, the court has the power to issue a legally binding Determination on pay and terms of employment. If the employer does not comply with a Labour Court Determination, the trade union may apply to the Circuit Court for an order directing the employer to carry out the Determination in accordance with its terms.

Changes to the legislation were agreed under the *Sustaining Progress* partnership agreement[13] and resulted in the passing of the Industrial Relations (Miscellaneous Provisions) Act 2004. SI No 145 was repealed and replaced by the Industrial Relations Act 1990 (Enhanced Code of Practice on Voluntary Dispute Resolution) (Declaration) Order 2004.[14] The 2004 Act provided that the processing of disputes under the Voluntary Dispute Resolution Code should take place within an indicative overall time frame of 26 weeks, with the possibility of extending it to a maximum of 34 weeks. Under the legislation, therefore, an employer may be compelled to grant union representatives the right to represent unionised employees on workplace issues relating to pay and terms and conditions of employment, but cannot be forced to make arrangements for collective bargaining.

Fasten your seatbelts – the *Ryanair* case

Trade unions had clearly hoped to use this legislation as a "springboard" to greater recognition rights,[15] but such aspirations appear to have been dashed given the decision in *Ryanair* v *The Labour Court*.[16] There, the Supreme Court upheld the company's complaints against the Labour Court's operation of the legislation, basing the decision on two key factors. First, the Supreme Court was highly critical of the *procedures* adopted by the Labour Court in hearing claims under the legislation. In particular, the Supreme Court felt that employees on behalf of whom claims were taken should ideally give oral evidence. The court held that the Labour Court did not adopt fair procedures by permitting complete non-disclosure of the identity of the persons on whose behalf the union was purporting to act. Furthermore, and most controversially, the Supreme Court criticised what it referred to as the Labour Court's "mindset", which favoured the way particular expressions are used

10 C Ryan, "Leaving it to the experts – in the matter of the Industrial Relations (Amendment) Act 2001" (2006) 3 *IELJ* 118.

11 "Excepted body" is defined by s. 6(3)(h) of the Trade Union Act 1941 (as inserted by s. 2 of the Trade Union Act 1942) and refers to "a body all the members of which are employed by the same employer and which carries on negotiations for the fixing of wages or other conditions of employment of its own members (but no other employees)".

12 Note the "Labour Court" is not part of the formal courts system but is a specially established industrial relations tribunal set up under the Industrial Relations Act 1946 to provide fair, informal and inexpensive arrangements for the adjudication and resolution of trade disputes.

13 See Art. 8.9 of *Sustaining Progress*.

14 SI No 176 of 2004.

15 B Sheehan, "Employers and the traditional industrial relations system: how the bonds have been loosened' in T Hastings (ed.), *The State of the Unions* (Dublin: Liffey Press 2008).

16 [2007] 4 IR 199.

and particular activities are carried out by trade unions and which hinted that collective bargaining in a non-unionised company must take the same form and adopt the same procedures as would apply in collective bargaining with a trade union. This is somewhat surprising, as the Superior Courts have traditionally been quite deferential to the Labour Court's expertise in relation to industrial affairs. In the *Ryanair* case itself, Hanna J in the High Court[17] had endorsed this view:

> The Labour Court is very much in charge of its own procedures. It has provided over many years a vital and invaluable service to the State in the often fraught area of industrial relations. It is not a court of law and the practice and procedure which it has evolved over the years, understandably, necessarily involved pursuing a less ritualistic and formalistic path than might be the case before these courts.

In fact, Hanna J pointed out, the present case was somewhat noteworthy in that lawyers *did* attend on behalf of the applicant. The Supreme Court's criticism of the procedures adopted by the Labour Court will likely have the effect of encouraging a greater formality in respect of Labour Court hearings and perhaps encourage a further "juridificiation" of the process.[18]

Secondly, the Supreme Court ruled that the Labour Court had erred in law (due partially to this "mindset") in its interpretation of the legislation; in particular, in deciding what was meant by the "practice" of "collective bargaining", in assessing when "internal procedures" had been exhausted, and in deciding what constituted an "excepted body". Essentially, the Supreme Court was not satisfied that the Labour Court was correctly confirming its jurisdiction to hear claims under s. 2 of the 2001 Act. As we will see, the number of claims processed under the 2001–2004 Acts has fallen dramatically in the wake of the decision.

The sample and method

This study looked at 103 Labour Court hearings up to, and including, the end of 2007, which were reported on the Labour Court website.[19] Two hearings were reported in 2002; 10 in 2003; 20 in 2004; 31 in 2005; 31 in 2006; and 9 in 2007. It should be noted that 89 of these hearings (86 per cent of the total) were heard after the amendments introduced by the 2004 Act came into effect in April 2004. Eighty-nine different employers featured in the 103 hearings.[20] If one looks at the occupational characteristics of the groups of employees involved in the cases, five sectors (using the Central Statistics Office (CSO) classification)[21] accounted for the bulk of the cases; manufacturing industry; transport, storage and communication; wholesale and retail; real estate, renting and business activities (including security services); and community social and personal service activities (excluding health and education). When looked at in detail, the cases are almost exclusively taken in respect of relatively low-pay, low-skill groups of workers.[22]

Sixty-four of the employers involved in the hearings (72 per cent) were indigenous organisations, while 25 (28 per cent) were Irish subsidiaries of foreign-owned MNCs. It was not possible in all cases to establish the size of the workforces involved. However, of the 76 cases in which this information was ascertained, in 15 of the cases (20 per cent) the company employed fewer than 20 persons, in 35 cases (46 per cent) between 20 and 100

17 [2006] ELR 1, at p. 17. See also the judgment of Clarke J in *Ashford Castle* v *SIPTU* [2007] 4 IR 70.

18 J Browne, *The Juridification of the Employment Relationship* (Aldershot: Avebury 1994).

19 www.labourcourt.ie. Just two hearings were held in 2008. There were no hearings in 2009.

20 Some employers participated in multiple hearings (discussed further below).

21 See www.cso.ie/statistics/LabourForce.htm.

22 Although there are a limited number of exceptions; the airline pilots in the *Ryanair* case, for example.

persons and in 26 cases (34 per cent) over 100 persons. Thus, the majority of cases involved relatively large employers, most of which were indigenous.

Of the 93 hearings where substantive issues were considered (i.e. excluding preliminary hearings under s. 2(1) of the 2001 Act discussed in the next section), binding Determinations under s. 6(1) were made in 26 cases (28 per cent). In 19 cases the Determination confirmed the original Recommendation in full and in seven cases it confirmed the Recommendation in part.[23]

A random sample of 48 hearings (47 per cent of the total) was chosen for more in-depth analysis.[24] Remuneration[25] was at issue in 41 of the hearings. Grievance and disciplinary issues were at issue in 30 hearings. Of the latter, 21 specifically involved the situation where an employee or group of employees were seeking to be represented by a trade union representative in respect of grievance and/or disciplinary proceedings. Non-pay benefits (e.g. canteen facilities) were at issue in 11 cases, leave entitlements in 10, working hours and employer failure to comply with statutory obligations both featured in six cases with assorted "other" issues featuring in a further 17 cases.[26] Thus, the issues raised involve traditional "core" union issues; pay and conditions and protection/representation in respect of grievances and disciplinary matters. Very few issues raised related to more "qualitative" issues (for example, family-friendly working).

The *Ryanair* decision: procedures

Confirming preliminary jurisdiction

Under s. 2(1)(a) of the Industrial Relations (Amendment) Act 2001 (as amended) the Labour Court may investigate a trade dispute where it is satisfied that:

> it is not the practice of the employer to engage in collective bargaining negotiations in respect of the grade, group or category of workers who are party to the trade dispute and the internal dispute resolution procedures (if any) normally used by the parties concerned have failed to resolve the dispute.

Under s. 3 of the 2001 Act (as amended), any question as to whether the requirements specified in s. 2 have been met may, as the court considers appropriate, be determined either by way of a hearing preliminary to the court's investigation under that section, or as part of that investigation.

Of the 103 Labour Court hearings held prior to 2008, just 10 were preliminary hearings under s. 2, where the court did not discuss the substantive issues but considered only its jurisdiction to hear the union's claim. Of the 10, the court confirmed its jurisdiction in six cases and found it did not have jurisdiction to hear the case in the remaining four cases.

Of the 48 cases chosen for closer analysis, the employer made arguments contesting the Labour Court's jurisdiction to hear the claim in 20 cases. In four cases, the employer made submissions questioning the union's claim to represent employees; in six cases the company

23 S. 6 specifies that Determinations should be in the same terms as any Recommendation made under s. 5, unless the court agrees a variation with the parties, or the court decides that the Recommendation, in whole or in part, was grounded on unsound or incomplete information.

24 Of these, in three cases there was no discussion of the substantive issues, as only the court's jurisdiction to hear the case was considered. As three of the hearings featuring Ashford Castle all involved the same issues, a sample of 43 cases involving distinct substantive issues was examined. The sample was random, except in the sense that, where a particular employer was involved in multiple hearings, all of these were examined.

25 Which here is taken to mean basic salary as well as issues like shift pay, overtime and sick pay.

26 For example, bullying and harassment, dignity at work; one case of victimisation on the grounds of trade union membership was raised.

argued that internal procedures had not been exhausted; in three cases the employer argued that no trade dispute existed; in four cases the employer argued that it did, in fact, engage in collective bargaining with the relevant workers; in two cases the company argued that it engaged in collective bargaining with a different union to that taking the claim; and in one case the company claimed the union had engaged in industrial action (in breach of the Acts).

Overall, in the sample of 48 cases, the court confirmed its jurisdiction to hear claims in 42 (including three preliminary hearings). In the six cases where the court decided it did *not* have jurisdiction to hear the claim, it found in three cases that the employer did engage in collective bargaining, but with a union other than that taking the claim.[27] In three further cases, the court found that internal procedures had not been exhausted.[28]

In *Ryanair*, the Supreme Court was critical of the Labour Court in respect of its failure adequately to address the issue of preliminary jurisdiction. The small number of stand-alone preliminary hearings held demonstrates that, in the vast majority of cases where jurisdiction was contested, the issue was decided by the Labour Court as part of the substantive investigation. In the light of *Ryanair*, it is likely more preliminary hearings will be necessary in future cases, as the preliminary examination would need to be more thorough and more formal procedures would need to be employed. This, of course, will have time and cost implications. Also, given the restrictive interpretations applied to key elements of s 2. of the 2001 Act in *Ryanair* (analysed below), it is likely more cases would emerge in which jurisdiction was not confirmed. This can already be seen in *Bell Security*.[29]

Identity of union members and representation

The Supreme Court in *Ryanair* held that the Labour Court did not adopt fair procedures by permitting complete non-disclosure of the identity of the persons on whose behalf the union was purporting to act. This aspect of the decision has led to a concern that, fearing victimisation from employers displaying a Ryanair-like aversion to trade unions, many employees would be likely to be discouraged from pursuing claims under the legislation.[30] The issue had arisen before the Labour Court, which decided in *Genesis Group*[31] that the legislation:

> does not require a trade union or excepted body to disclose the names or other identifying details of its members as a condition precedent to the making of an application.

The Supreme Court, however, seemed to view the issue of disclosure as a fundamental aspect of fair procedures. It is interesting that, in the immediate aftermath of the *Ryanair* decision, the LRC issued a press release, stating that, in its view, the verification of union

27 *Fernley Airport Services* (Case LCR18845 issued on 26/2/2007); *Federal Security Services* (Case LCR18621 issued on 4/07/2006); and *MCM Security* (Case LCR18206 issued on 26/05/2005).

28 See, for example, *Banta Global Turnkey* (Case DECP041 issued on 13/07/2004).

29 Case LCR19188 issued on 11/04/2008 discussed below at p. 395.

30 This point is particularly pertinent in the light of Geary's research, which showed that the propensity to unionise in non-union workplaces is especially manifest in situations where employers offer their support for union representation; the propensity to unionise drops markedly in situations where employers are not prepared to support union organisation. See J Geary, "Employee voice in the Irish workplace: status and prospect" in P Boxall et al. (eds), *What Workers Say: Employee voice in the Anglo-American Workplace* (Ithaca: Cornell UP 2007). It should be noted that there are anti-victimisation procedures under the Acts (ss. 8 and 9 of the 2004 Act), which have rarely been used.

31 Case DIR0511 issued on 22/9/2005. Note, however, that the Labour Court restricted the application of its decision to the members of the union. Therefore, although the union did not have to disclose its membership as a condition of making the application, it would have to disclose in order to benefit from the decision.

members at the LRC stage of the process (at the request of employers) was in accordance with the Supreme Court's ruling:

> this was done in a very straightforward way by the Advisory Officer obtaining a list of members from the trade union and cross checking this against the employer's own data such as payroll . . . our advice is that, while we must satisfy ourselves that the trade union has members in the employment concerned, the judgement does not oblige us to compel a trade union to disclose to employers the names of those taken into membership.[32]

A related area of controversy is the extent to which unions taking claims under the legislation are, in fact, representative. In 78 of the 103 hearings (76 per cent) no information can be gleaned from the Labour Court report as to the number of members the union in question claimed to have in membership. In the other 24 cases, the number of members claimed varies significantly. In *Castle T Furniture*,[33] for example, SIPTU (Services Industrial Professional and Technical Union) claimed to represent half of the 18 employees in the company, while in *Schering Plough*[34] the union claimed to represent 306 of the 700 employees at the company plant. At the other end of the scale, in *QK Cold Stores*[35] the union's claim was on behalf of just seven employees out of a workforce of over 100, while, in *Finlay Breton*,[36] BATU's (Building and Allied Trade Union) claim was in respect of three members out of a workforce of 300. Thus, in some cases the unions involved pursued a claim where they declared to have a considerable existing presence, while in others claims were taken on behalf of a handful of employees only.

In three of the 48 cases chosen for closer analysis, the employer explicitly challenged before the Labour Court how representative the unions in question were.[37] In *Goode Concrete*,[38] the company contended that the union was not actually representative of *any* of its employees and that there could not, therefore, be a dispute between it and the union. The union claimed to represent 30 drivers and offered to provide the court, on a confidential basis, with details concerning the number and identity of company employees who had joined the union, as the members did not consent to their identity being disclosed to their employer for reasons outlined to the court. The court noted that the Acts do not prescribe any membership threshold which a trade union must meet before it can bring an application under s. 2(1) and accepted the assurances of the union that it was representative of employees in dispute with the company.

The issue of verifying to what extent a union is representative of workers in dispute is now paramount in the light of *Ryanair*. Unions are, and have been, reluctant publicly to divulge information about membership levels in non-union companies for fear this may lead to employees being identified by employers and, potentially, being victimised or disadvantaged. The Supreme Court decision goes beyond this by explicitly requiring some identification of *individual* employees in dispute. This is so even though, as the Labour Court

32 www.lrc.ie/viewdoc.asp?Docid=555&Catid=28&StartDate=1+January+2008&m=n.

33 Case LCR19002 issued on 01/10/2007.

34 Case LCR18226 issued on 15/06/2005.

35 Case LCR18556 issued on 10/05/2006.

36 Case LCR 062 issued on 6/04/2006.

37 *Hillview Nursing Home* (Case LCR 18271 issued on 29/07/2005); *Analog Devices* (Case LCR18137 issued on 21/03/2005); and *Goode Concrete* (Case LCR18037 issued on 09/12/2004).

38 *Goode Concrete* (Case LCR18037 issued on 09/12/2004).

has pointed out, there is no requirement under the legislation for unions to meet any representation threshold prior to taking a claim.[39]

This is a real problem for unions, not only as many (most?) members will likely be unwilling to put their heads above the parapet in pursuing a claim under the legislation, but also more broadly in the sense that evidence has shown that, in the absence of employer support, many Irish workers are fearful of the consequences of joining unions at all lest union membership damage their career prospects.[40] In this sense, there may be some conflict between the position following the Supreme Court decision and Ireland's obligations under Article 11 of the ECHR. It should be remembered that in the *Wilson and Palmer* case[41] the European Court of Human Rights found that "employees *should be free* to instruct or permit their union to make representations to their employer or to take action in support of their interests".[42] The court went on to say that if workers were prevented from so doing, their freedom to belong to a trade union became illusory and that it was the "role of the State to ensure that trade union members were not prevented or restrained from using their union to represent them in attempts to regulate their relations with their employers".[43]

In relation to the issue of identifiable union members giving evidence, one solution might be for some number of employees, on whose behalf the union is acting, to swear affidavits which would be supplied to the Labour Court. The court, having verified that the employees are union members and work for the company (in a manner similar to that outlined by the LRC above), could then supply anonymous versions of these to the employer. This, however, would run into another procedural issue identified by the Supreme Court that relates to oral evidence (discussed in the next section).

A second way out of the identification issue might be to re-fashion s. 2(1) altogether. Instead of fulfilling the criteria presently laid down in order to take a claim, a union could be required to meet some specified threshold of membership (which, again, could be verified by the LRC or Labour Court). This would have the advantage for employers of their not being subject to the legislation where only a handful of workers are union members.[44] There is some recent precedence for this type of arrangement, too, in the Employee (Provision of Information and Consultation) Act 2006, which requires employers to inform and consult with employee representatives on a range of issues, where such a request is made by at least 10 per cent of the workforce.[45] For the unions, where such a threshold was met, this approach would give "moral legitimacy" to their claims. This,

39 This idea of pressure on unions to disclose to courts and tribunals sensitive information regarding membership has interesting parallels with the view posited by some that the result of the ECJ's decisions in cases like *Viking* could well be that unions, in order to defend the proportionality of industrial action in claims relating to EU free movement rights, could be forced to disclose to national courts and tribunals potentially oppressive volumes of materials on internal union strategy, tactics and policy; see K Ewing and J Hendy, "The ECJ decisions and trade union freedom: lessons from the UK" in K Ewing and J Hendy (eds), *The New Spectre Haunting Europe – The ECJ, trade union rights and the British government* (Liverpool: Institute of Employment Rights 2009). Case C-438/05 *International Transport Workers' Federation and Finnish Seamen's Union* v *Viking Line ABP* [2008] IRLR 14. See further below, at 398.

40 Geary, "Employee voice", n. 30 above.

41 *Wilson & the NUJ, Palmer, & Others* v *The UK* [2002] IRLR 128.

42 Ibid., at para. 46 (emphasis added).

43 Ibid.

44 As in *Finlay Breton*, n. 36 above.

45 The 2006 Act implements Directive 2002/14/EC of 11 March 2002 establishing a general framework for informing and consulting employees in the European Community OJ L80/29; see M Doherty, "It's good to talk . . . isn't it? Legislating for information and consultation in the Irish workplace" (2008) 15 *DULJ* 120.

in fact, would not be a million miles from the position pre-2001, where unions took recognition claims under s. 20(1) of the Industrial Relations Act 1969. Recommendations under the 1969 Act that the employer should recognise the union in respect of those workers it had in membership were not binding on the employer and were often ignored.[46] However, such a Recommendation still allowed a union taking industrial action in support of recognition to show that it had done its best to abide by procedures. Under the 2001–2004 Acts, of course, any Recommendation would not relate to union recognition, but would be binding.

Whether this would satisfy the unions is another question. As noted, the unions may have seen the legislation as a potential "springboard" to full recognition rights in companies. To meet a threshold unions would need to build up substantial support in a company prior to making a claim (a problem that has bedevilled British unions attempting to invoke the statutory recognition procedure there),[47] whereas, under the present s. 2(1) they can invoke the legislation on behalf of any number of members, however small. Success can then result in a beneficial "demonstration" effect to other employees and help unions to garner more members. Similarly, from an employee's point of view, the purpose of the legislation, as outlined by Geoghegan J in *Ryanair*, was to protect employees in non-unionised companies from the obvious danger that: "employees may be exploited and may have to submit to what most reasonable people would consider to be grossly unfair terms and conditions of employment".[48]

One of the objectives of the Acts was to permit a right to representation for individual union members in the face of their employer's opposition. Because such employees, as individuals, normally have little bargaining power, the Acts accorded with arguably the main object of labour law; to be a countervailing force to counteract the inequality of bargaining inherent in the relationship, concealed by that "indispensable figment of the legal mind, the contract of employment".[49]

Oral evidence

A final procedural issue raised by the Supreme Court related to oral evidence. The court felt that factual issues in dispute should be resolved on oral evidence from parties who participated in the process or who could give first-hand evidence on how the employer's procedures operated.[50] Therefore, direct evidence on any issue is generally to be preferred to a legal submission, or an opinion or references to documents unsupported by direct evidence.

The reference here to an "opinion" is particularly worrying for the unions, as it relates to the issues discussed in the previous section around employees giving direct evidence. In order to protect employees' anonymity, the usual practice under the Acts has been for a union official to outline the employee case. Indeed, of the sample of 48 cases looked at here, in only one did an employee who was party to the dispute appear to give oral evidence.[51] If direct evidence from those involved is to be preferred, this option will no longer be enough. It is interesting to note in *Ryanair* that Hanna J, in the High Court,[52]

46 C Higgins, "The right to bargain law: is it working?" (2001) 45 *IRN*.

47 See further below at n. 98.

48 [2007] 5 IR 99, at p. 215.

49 Kahn-Freund, *Labour and the Law* (London: Stevens & Sons 1977), p. 6.

50 Therefore, the suggestion re "anonymous affidavits" mooted above would seem to fall foul of this aspect of the Supreme Court decision.

51 *Little Rascal Crèche* (Case LCR18648 issued on 24/07/2006).

52 [2006] ELR 1.

pointed out that whether or not oral evidence is offered in a case is a call made on a daily basis by advocates before the ordinary courts, where parties are free to offer viva voce evidence or not as the case may be. He went on to observe that, while there might be circumstances in which the Labour Court might take a more "activist role" in determining what oral evidence it might wish to hear (for example, where there is a marked imbalance of "firepower" in the representation of the parties before it), this was not such a case.

The Supreme Court decision: interpretation

The second limb of the Supreme Court's criticism in *Ryanair* related to the interpretation given to key elements of the amended s. 2(1) of the 2001 Act by the Labour Court. For the Labour Court to assert jurisdiction in such cases it must be satisfied that it is not the "practice of the employer to engage in collective bargaining" with a trade union or an excepted body. The view accepted by the Labour Court was that if a group of employees unilaterally withdraws from the internal negotiating procedures, it could not thereafter be said that the employer had a practice of engaging in collective bargaining with them. The Labour Court laid down its definition of collective bargaining in the *Ashford Castle*[53] case, noting that the expression is not defined in industrial relations legislation and that it is not a legal term of art:

> Collective bargaining comprehends more than mere negotiation or consultation on individual employment related issues, including the processes of individual grievances in relation to pay or conditions of employment. In the industrial relations context in which the term is commonly used, it connotes a process by which employers or their representatives negotiate with representatives of a group or body of workers for the purpose of concluding a collective agreement fixing the pay and other conditions of employment applicable to the group of workers on whose behalf the negotiations are conducted.
>
> Normally the process is characterised by the involvement of a trade union representing workers, but it may also be conducted by a staff association, which is an excepted body within the meaning of the Trade Union Act 1941, as amended. However an essential characteristic of collective bargaining, properly so called, is that it is conducted between parties of equal standing, who are independent in the sense that one is not controlled by the other.

The Supreme Court, however, objected to the view "arguably hinted at" in the definition that collective bargaining in a non-unionised company must take the same form and adopt the same procedures as would apply in collective bargaining with a trade union. The Supreme Court criticised the Labour Court for acknowledging a special, trade union meaning of the expression "collective bargaining negotiations" and held that the phrase should be given simply an ordinary meaning and not any distinctive meaning as understood in trade union negotiations. According to Geoghegan J:

> if there is a machinery in Ryanair whereby the pilots may have their own independent representatives who sit around the table with representatives of Ryanair with a view to reaching agreement, if possible, that would seem to be "collective bargaining".[54]

Furthermore, the unilateral withdrawal by employees from machinery put in place by the employer would not of itself entitle the employees to assert that there was no collective bargaining process in being; ultimately, where an employer has an internal non-union

53 Case DECP032 issued on 19/11/2003.

54 [2007] 4 IR 99, at p. 218.

collective bargaining unit in place, this might constitute an excepted body under the legislation and satisfy the requirements of s. 2.

The definition of collective bargaining provided by the Supreme Court and its indication that a collective bargaining unit can, it seems, amount to any group of employees as long as the group is recognised for this purpose by the company concerned have provoked much comment. In *Ashford Castle* v *SIPTU*, Clarke J, in the High Court, noted that the legislation:

> only applies in circumstances where there is no collective bargaining. The only reasonable inference to draw from that provision is that the intention of the Oireachtas was to confer upon employees, who did not have the benefit of collective bargaining, a means of attempting to achieve terms and conditions comparable to those who had the benefit of collective bargaining.[55]

This was also the view underpinning the Labour Court's approach to the four cases examined in this research where the employer argued that it did, in fact, engage in collective bargaining.[56]

In *Exel Technologies*,[57] the company claimed that "monthly communications meetings" were held to discuss all matters relating to employment, including pay and non-pay terms and conditions of employment, and were attended by elected staff representatives from each department. Therefore, the employer submitted, this process was no different from that which occurs in a company that engages with a trade union. SIPTU contended that the monthly communications meeting system was under the control of the employer and, consequently, employees had no appropriate means of processing claims with independent representation, no means of appeal, and no opportunity to refer to third parties. The court was satisfied that no details had been submitted by the employer to show that disputes concerning terms and conditions of employment were normally or routinely dealt with through this process. Consequently, the court found that the company's procedure was not of the type envisaged by s. 2(1)(a) of the Act.

In *Ryanair*,[58] the company outlined a system (which it contended amounted to collective bargaining) whereby employees, including pilots, elect employee representatives to Employee Representative Committees (ERCs). The various ERCs then negotiate directly with the company on an ongoing basis in relation to all terms and conditions of employment. It was accepted that the Dublin pilot representatives had withdrawn from the ERC in August 2004 and no new representatives had been appointed. The court found that the ERCs were established by Ryanair and that the company had organised and controlled the election of employee representatives to them, including specifying the criteria of eligibility for election (e.g. no representative could serve more than one term). Employees were informed of the outcome of ERC discussions by Ryanair in a newsletter which it published and in respect of which it retained copyright. As a result (and by reference also to company documents) the Labour Court found that collective bargaining did not take place within the company.

The Labour Court will now have to rethink its underlying approach in the light of the Supreme Court decision. While the latter, significantly, did not set down precise rules or offer guidelines for the operation of a non-union internal bargaining unit, it seems from the

55 [2007] 4 IR 70, at p. 75.

56 In addition to the examples discussed in the text, see *Little Rascal Crèche* (Case LCR18648 issued on 24/07/2006) and *Ashford Castle* (Case DECP032 issued on 19/11/2003).

57 Case LCR18274 issued on 25/07/2005.

58 Case DECP051 issued on 25/01/2005.

judgment that employers would be free to determine the form, structure and organisation of any internal collective bargaining units, as long as these have a degree of permanency and are not ad hoc. Thus, if an employer were to set up such a unit, it could presumably decide on issues such as how employees would be elected or chosen to be members, the remit of the unit, the terms of office of its members, and the rules and procedures of its operation.[59]

Disquiet has been expressed that the Supreme Court's definition of collective bargaining tends to ignore not only Ireland's industrial relations traditions, but also ILO Conventions and Declarations to which Ireland is a signatory.[60] Ireland has ratified a number of ILO instruments, which explicitly require that the framework within which collective bargaining must take place if it is to be viable and effective be based on the principles of the independence and autonomy of the parties and the free and voluntary nature of the negotiations.[61] Furthermore, ILO principles require that all legislation establishing machinery and procedures for arbitration and conciliation designed to facilitate bargaining between both sides of industry should guarantee the autonomy of parties to collective bargaining; explicitly excluded is the notion of employer-dominated bodies or company unions being considered as mechanisms for collective bargaining.[62]

Moreover, in addition to its international law obligations, there must be a real concern as to the compatibility of the Irish position with Article 28 of the Charter of Fundamental Rights, which guarantees the "right to negotiate and conclude collective agreements". With the entry into force of the Lisbon Treaty in December 2009, the Charter has attained the status of primary EU law.[63] The provisions of the Charter, according to Title VII, are addressed to the institutions and bodies of the EU with due regard for the principle of subsidiarity and to the member states *only* when they are implementing EU law. However, there is an obligation on Community Institutions and Member States to *promote* the rights in the Charter. Finally, given that Ireland has now incorporated into domestic law the ECHR,[64] there is an obligation on the Irish Courts to interpret and apply any statutory provision on rule of law in a manner that is compatible with the country's obligations under the Convention.[65] In this respect, it is important to note the recent decision of the European Court of Human Rights in *Demir and Baykara* v *Turkey*[66] where the court declared that Article 11 of the ECHR includes a *right* to bargain collectively and precludes a blanket ban on a right to strike.[67]

There must be a real concern, therefore, that the Supreme Court, by not indicating more precisely what a non-union internal bargaining unit would look like, opens the possibility that employers will set up units that may not be genuine bargaining fora and thus

59 Doherty, "Union sundown", n. 5 above.

60 D D'Art and T Turner, "Ireland in breach of ILO conventions on freedom of association, claim academics" (2007) 11 *IRN*.

61 B Gernigon et al., "ILO principles concerning collective bargaining" (2000) 139 *International Labour Review* 33.

62 Ibid. at p. 44.

63 New Art. 6 Treaty on European Union.

64 By virtue of the European Convention on Human Rights Act 2003.

65 Ibid., s. 2.

66 Application No 34503/97, 12 November 2008. See also *Enerji Yapi-Yol Sen* v *Turkey* (Application No 68959/01, 21 April 2009).

67 Rulings which potentially bring the court's jurisprudence into conflict with that of the ECJ.

abuse the process. At the very least, it seems, legislative intervention will be required.[68] Most likely this will require a statutory definition of "collective bargaining". The ILO defines collective bargaining as:

> all negotiations which take place between an employer, a group of employers or one or more employers' organisations, on the one hand, and one or more workers' organisations, on the other, for—
>
> (a) determining working conditions and terms of employment; and/or
>
> (b) regulating relations between employers and workers; and/or
>
> (c) regulating relations between employers or their organisations and a workers' organisation or workers' organisations.[69]

Where national law or practice recognises the existence of other forms of workers' representatives, national law or practice can determine the extent to which the term collective bargaining shall also extend to negotiations with these representatives.[70] Thus, although ILO definitions emphasise the role of *trade unions* in collective bargaining, they do recognise the potential role of non-union representatives. Crucially, however, they also insist that such representatives (and bargaining units) must be genuinely *independent* of employers. It can be argued that worker representatives (whether elected or appointed) who are employees of the undertaking, and therefore dependent on the goodwill of management for their employment and prospects of promotion, can never be genuinely independent.[71] However, at a minimum, Irish legislation could be amended to, first, lay down conditions to be met regarding the establishment and operation of internal non-union bargaining units.[72] Secondly, and in contrast to the Supreme Court's view, it might be provided that collective bargaining could not take place in a context where an employer refuses to engage with the trade union or excepted body which employees indicate (by way, perhaps, of ballot or independent verification) they wish to represent them.

The effects of the Supreme Court's decision can already be clearly seen. First, as noted above, the number of claims pursued under the Act has dropped dramatically and many union officials have proclaimed the Acts effectively dead. Secondly, the decision in *Bell Security*[73] illustrates the new approach taken by the Labour Court. Here, two employees gave oral evidence to the court in support of the union's claim (both were union shop stewards). The Labour Court found that a trade dispute was in existence. It further found that the internal dispute resolution procedures had been exhausted. However, on the facts of the case the court decided that it was the company's practice to engage in collective bargaining with employees. The Labour Court's concluding remarks are worth quoting at this point:

---

68 One option would be to go down the UK route of statutory union recognition; this approach, though, would be strongly resisted by employer groups and does not seem to be favoured by the government; J Lavelle et al., "Unions on the edge? Industrial relations in multinational companies" in T Hastings (ed.), *The State of the Unions* (Dublin: Liffey Press 2008).

69 Art. 2 of ILO Convention (154) concerning the Promotion of Collective Bargaining (1981).

70 The ILO defines non-union worker representatives as "elected representatives, namely, representatives who are freely elected by the workers of the undertaking in accordance with provisions of national laws or regulations or of collective agreements and whose functions do not include activities which are recognised as the exclusive prerogative of trade unions in the country concerned"– Art. 3 of the Convention Concerning Protection and Facilities to be Afforded to Workers' Representatives in the Undertaking (1971).

71 D'Art and Turner, "Ireland in breach", n. 60.

72 A template of sorts can be found in Sched. 2 of the Employee (Provision of Information and Consultation) Act 2006, which lays down the standard rules for the establishment of an employee information and consultation forum.

73 Case LCR19188 issued on 11/04/2008.

> The Court accepts the union's submission that on the facts of this case there was an inherent and manifest inequality of negotiating capacity between the employee and management representatives. A professional HR specialist and senior managers represented the Company. Electricians who had no training or skills in negotiation or bargaining represented the employees. It is clear on the evidence that because of this the employee representatives came to see their role as involving little more than carrying messages back and forth and considered the process to be a waste of time and going nowhere.
>
> It is nonetheless clear on the evidence that there was a practice whereby representatives of the employees, including representatives of the Dublin engineers, sat around a table with representatives of the Company with a view to reaching agreement if possible. That is collective bargaining negotiations for the purpose of the Act.
>
> If the Court were considering the factual matrix of this case in an industrial relations context it might take a different view. However it must apply the law as it finds it and following the decision in *Ryanair* there can be no doubt as to the correct legal approach to the questions arising in this case.

Thus, the court accepted that the Supreme Court decision requires it to take an approach outside of the "industrial relations context". Of course, the court's unease with this, as well as its comments on the inequality of firepower between the parties, should be noted with concern. However, it does seem that it is the definition of "collective bargaining" that is key here. The union quoted both ILO and dictionary definitions. Legislative movement on this issue in line with these may well, therefore, have the effect of resuscitating the Acts.

An "Irish model" for representation rights?

The study did not seek to investigate in depth the extent to which union claims were successful or unsuccessful under the legislation. However, a number of points are noteworthy in terms of the Recommendations made by the Labour Court. This section will assess some of the key impacts of the legislation, focusing particularly on the extent to which the acts plugged the "representation gap" that exists where employers refuse to recognise employees' union(s). Moreover, potential implications of this "Irish model" for IR practice elsewhere are considered.

Mandating the "model employer" – an Irish approach to the "Laval question"?

In 34 of the 48 cases looked at in depth, the Labour Court made a Recommendation on remuneration. Two types of Recommendation are noteworthy. First, in eight of the cases, the court felt the company should pay the terms of the *national pay agreement* in place at the time. For example, in *Creagh Transport*,[74] it was established that the company had no fixed or formal system of reviewing the pay of its employees. The company claimed that, due to the economic and commercial circumstances of the business, it had been unable to pay any increases for over four years. The Labour Court was of the view that the company's failure to provide for any increases in pay over such an extended period could not be justified. The court went on to say that, whilst the increases provided by national partnership agreements were not an automatic statutory or contractual entitlement, in the absence of any other established or agreed method of pay determination, they represented an "appropriate reference point" for establishing a fair and reasonable level of pay adjustment. The court rejected the submission that it was precluded from recommending increases in line with

74 Case LCR17933 issued on 18/08/2004.

those provided by national agreements as any recommendation made by the court to that effect could be implemented *without* the necessity for collective bargaining at the level of the enterprise. The court recommended that, in future, pay should be adjusted by reference to the increases provided by national agreements subject to the right of the company to plead inability to pay through the mechanisms provided by those agreements.

Secondly, in 18 cases where a Recommendation on remuneration was made, this was on the basis of *pay norms* in the given industry.[75] In *Bank of Ireland*,[76] the court pointed out:

> The powers which are given to the Court by the Act are a far reaching departure from the normal approach to the resolution of industrial relations disputes. They provided, in effect, that the Court may arbitrate in a dispute on the unilateral application of one party and in circumstances where the other party may not consent to the process. It seems to the Court that, having regard to the voluntary nature of our industrial relations system, such an intervention is only appropriate where it is necessary in order to provide protection to workers whose terms and conditions of employment, when viewed in their totality, are significantly out of line with appropriate standards.

Thus, the court, in its Recommendations on pay, has sought to introduce the idea of the "model employer"; in other words, it has effectively benchmarked respondent companies against others in the sector. This can be seen in *Fournier Laboratories*,[77] where the Court found that the company's pay determination system was out of line with accepted standards in that it was based solely on performance assessment, rather than by reference to a basic "rate for the job", the predominant practice in the sector. Similarly, in *Cooley Distillery*,[78] the court accepted pay rates agreed by the union (through collective bargaining) with other employers both locally and nationally as indicative of the industry norm. It recommended the respondent increase its pay rates to this more "appropriate standard".[79]

Therefore, where companies fall below the general, prevailing industry standards (as located by the court) they have been told to raise standards to that level (frequently identified as those set down by national pay agreements). From an employer's point of view, this can be seen as unwarranted interference with the right to operate a business and with property rights (as it effectively forces up industry norms in terms of pay). Furthermore, using the "unionised" standard of national agreements seems to threaten the right (recognised in *Ryanair*) to operate a non-union company. This most likely feeds into the Supreme Court's criticism of the Labour Court's "union mindset".

However, it has been frequently suggested (although evidence is somewhat sketchy)[80] that in many industries the national agreements *do* act as a benchmark for non-union firms. In addition, in pay claims before the Labour Court under any legislation, the national agreements are arguably a useful benchmark to use. The alternative is for the court to

---

75 See, for example, *Galway Clinic* (Case LCR18815 issued on 18/01/2007).

76 Case LCR17745 issued on 28/01/2004.

77 Case LCR18582 issued on 24/05/2006. This decision of the Labour Court was quashed by way of a consent order, following the institution of Judicial Review proceedings, made by the High Court on 13 November 2007.

78 Case LCR17908 issued on 19/07/2004.

79 In *Sterile Technologies* (Case LCR17906 issued on 19/07/2004), the court, in holding pay rates were *not* out of line with industry norms, took account of the employer's provision of recently published survey information on comparable rates of pay compiled by the IBEC and the CSO. The court also noted that no industry comparators were put forward by the union.

80 Cf. D G Collings et al., "Between Boston and Berlin: American MNCs and the shifting contours of industrial relations in Ireland" (2008) 19(2) *International Journal of HRM* 242.

depend on the parties' submissions on local and sectoral pay comparators. In such cases, both sides will have clear agendas, which can distort the true picture, information of this nature may be difficult to obtain and partial, and there are concerns about confidentiality. Furthermore, intervention in companies' wage-setting is nothing new; all employers are bound by the national minimum wage and some industries are subject to levels of pay set by Joint Industrial Councils (JICs) and Joint Labour Committees (JLCs).[81] Clearly, though, there is a distinction between legally binding *minimum* standards and the setting of an industry wide "fair rate for the job" (particularly where the latter is set, not by the parties themselves in collective agreements, but by a state industrial relations tribunal).

It is this issue of binding *minimum* standards, as distinct from collectively agreed *norms*, that has been at the heart of controversial recent European Court of Justice (ECJ) decisions on collective rights. In a series of cases, *Laval*, *Viking*, *Rüffert* and *Luxembourg*,[82] the ECJ has severely affected the rights of trade unions and member states to protect collective agreements in cases where the rights of free movement of services or establishment are involved. The court ruled in *Laval* that, in accordance with the free movement of services provisions of the EC Treaty and the terms of the Posted Workers Directive (PWD),[83] Swedish trade unions could not take industrial action to compel a Latvian builder operating in Stockholm, and "posting" Latvian workers there, to observe the terms and conditions of collective agreements operating in Sweden. Similarly, in *Rüffert* a Polish contractor could not be compelled to observe collective agreements that were locally, but not nationally, applicable and in *Luxembourg* posting employers could not be forced to observe collectively agreed minimum terms and conditions of employment beyond the mandatory matters listed in Article 3 of the PWD. At the heart of all these rulings is the view that where collective agreements are not declared universally applicable, extended *erga omnes* to non-union workplaces, or their provisions protected, in some way, by member state legislation, they cannot be imposed on service providers from other EU jurisdictions operating in the member state in question.[84] All that can be required of such service providers is that they observe statutory minima terms and conditions of employment.

In this respect, the Irish legislation, ironically, given that it falls short of union demands for strengthened collective bargaining rights, might offer a mechanism to protect "prevailing rates", rather than minimum standards, that would withstand ECJ scrutiny. In terms of fairness and social equity, the Labour Court approach, in attempting to benchmark and mandate good practice, has much to commend it. The court, after all, is explicitly set

81 See Parts IV and V of the Industrial Relations Act 1946. JLCs provide for the fixing of legally binding minimum rates of pay and the regulation of employment in certain sectors where there is little or no collective bargaining and where significant numbers of vulnerable workers are employed (e.g. the hotels sector). Collective agreements made by JICs (voluntary negotiating bodies for an industry or part of an industry, designed to facilitate collective bargaining at industry level in certain sectors) are also registered with the Labour Court and are legally binding. They generally exist in sectors with a relatively high level of unionisation (e.g. the construction sector). The Industrial Relations Amendment Bill 2009 (just published at the time of writing) seeks to strengthen both mechanisms.

82 Case C-341/05 *Laval* v *Svenska Byggnadsarbetareförbundet* [2007] ECR I-11767; Case C-438/05 *International Transport Workers' Federation and Finnish Seamen's Union* v *Viking Line ABP* [2008] IRLR 14; Case C-346/06 *Rüffert* v *Land Niedersachsen* [2008] IRLR 467; and Case C-319/06 *European Commission* v *Luxembourg* [2009] IRLR 388.

83 Directive 96/71/EC of 16 December 1996 concerning the posting of workers in the framework of the provision of services OJ L18/01.

84 See S Deakin, "The labour law perspective: the economic implications of the decisions" (2007–08) 10 CYELS 463; C Barnard, "*Viking* and *Laval*: a single market perspective" in K Ewing and J Hendy (eds), *The New Spectre Haunting Europe – The ECJ, trade union rights and the British government* (Liverpool: Institute of Employment Rights 2009); A C L Davies, "One step forward, two steps back? The *Viking* and *Laval* Cases in the ECJ" (2008) 37(2) *Industrial Law Journal* 126.

up to reflect and accommodate the different, often conflicting, interests of labour and business in the interest of employment relations harmony. However, it has been persuasively argued[85] that the Irish judiciary in a number of major areas of constitutional interpretation frequently tends to defer to individual values, and particularly individual property rights, over those of the community or collective groups. This approach of the Superior Courts is, perhaps, visible again in the Ryanair decision. It would require a legislative (or Supreme Court) rethinking of what "bargaining" entails in the Irish context (in line with ILO and ECHR formulations) for the Acts to truly offer a way out of the Laval dilemma. We will return to this question in the concluding section.

A Trojan horse? Union recognition under the Acts

A key concern of employers has been that the Acts effectively promote a form of "back door" union recognition; that essentially the Acts allow unions to get their "foot in the door" and force employers to deal with them on some level.[86] Undoubtedly, the unions hoped that this would be the case. It should be noted, however, that in only two cases, *Federal Security Services Ltd*[87] and *Hillview Nursing Home*[88] did employers actually concede bargaining recognition to a union (and in both cases, the employer decided to recognise unions *other than* those taking the respective claims). Therefore, neither the fears of employers nor the hopes of the unions seem to be borne out empirically.

Nevertheless, the Acts have had some interesting effects worthy of comment. In a limited number of rulings, the Labour Court seemed to indicate that, on certain issues, unions and their members should have a collective influence. This can be seen in *Carlingford Nursing Home*[89] where it was recommended that the company put in place a new harassment and equal treatment policy. In drawing this up, the union (SIPTU) was to be permitted to make submissions on the content of such policies and these submissions were to be taken into account. More starkly, perhaps, in the long-running saga involving *Ashford Castle*[90] the company (while not conceding union recognition) eventually indicated to the Court that it would be prepared to facilitate the union in providing paid training for its shop stewards, who could then represent members in grievance procedures.

Most worryingly of all for non-union employers is the case of Esker Lodge nursing home, where it was reported that the Irish Business and Employers' Confederation (IBEC) may have advised the company to recognise the union on pragmatic grounds; that is, that it would be easier to simply concede recognition than to become caught up in the procedure under the Acts.[91] It should be stressed, though, that there seems little further evidence of cases where unions have been able to use the spectre of the Acts to persuade employers to engage directly in a traditional collective bargaining relationship.

In this respect, the Acts throw up an intriguing dilemma for unions. Taking claims under the Acts involves considerable time and expense. Often claims are taken on behalf of non-members, and frequently in respect of small groups of workers. Although unions may have been willing to pursue these claims initially in an attempt to demonstrate the

85 G Morgan, *A Judgment Too Far? Judicial activism and the Constitution* (Cork: Cork UP 2001); in relation to trade union rights, see p. 38ff.

86 See, for example, T Dobbins, "Union recognition law used for benchmarking private sector pay" (2005) 36 *IRN*.

87 See n. 27 above.

88 Case LCR18440 issued on 22/12/2005.

89 Case LCR17932 issued on 17/08/2004.

90 Case LCR18820 issued on 22/01/2007.

91 T Dobbins, "'Fear' of 2004 Act persuades nursing home to concede full recognition" (2006) 36 *IRN*.

value of union representation to non-union workforces, the sustainability of such a strategy is questionable. It certainly is not a traditional form of union representation nor does it fit within many of the prescriptions offered to halt union decline in the literature on union "revitalisation".[92]

Furthermore, the fact that the legislation explicitly excludes Recommendations on collective bargaining and empowers the court to deal only with specific, defined issues, means that unions may face the prospect of being forced to take multiple claims against a particular employer. The study showed that 89 different employers featured in the 103 hearings, but some employers featured in multiple hearings; Ashford Castle, for example, was involved in four separate hearings.[93] Again, this type of action probably involves an unsustainable commitment of union time and resources. Clearly, the unions will take such cases in order to try and persuade employers to recognise them in a traditional bargaining relationship. Where employers refuse, however, multiple claims seem to be the only option under the legislation. Outside of the impacts on unions and employers (who will have the threat of Labour Court imposed pay and conditions hanging over them), there are important public policy considerations here also, not least the implications in terms of the growing demands on the time and resources of the LRC and the Labour Court.

A final issue needs to be addressed here in relation to union recognition. As noted above, one of the key objections to any attempts to introduce a statutory recognition procedure (along the lines of the UK) has been the alleged threat this would pose to foreign direct investment (primarily from US MNCs) on which the Irish economy is so dependent. Similar concerns have been voiced about the impact of the 2001–2004 Acts on MNC activity in Ireland.[94] In this regard, it should be noted that the study reveals the majority of cases involved indigenous employers (72 per cent). Although the number of cases involving MNCs is perhaps higher than had been previously thought,[95] the main impact of the legislation, as we have seen, has been on employers whose pay and conditions fall below the industry norm. The larger, "flagship" MNCs, where HR practices would generally be relatively sophisticated and where pay and conditions of employment would generally be at the higher end of the spectrum, barely feature at all.[96] Thus, the focus on the "threat to investment" argument seems to be misplaced.

Conclusion

This article considered the operation of the 2001–2004 right to bargain legislation in the light of the seminal Supreme Court decision in the *Ryanair* case. Throughout, the concern was to identify, first, whether the legislation has any future role to play and, secondly, to consider the impacts of the Acts and the *Ryanair* decision on the various IR stakeholders for whom the rights (or lack thereof) to union representation is of paramount concern. On the first issue, various suggestions were made that could have the effect of "resuscitating" the legislation, in particular a re-formulation of the definition of "collective bargaining".[97]

92 C Frege and J Kelly, *Union Revitalisation Strategies in Comparative Perspective* (Sage: London 2003).

93 Case DECP032 issued on 19/11/2003; Case LCR17760 issued on 23/03/2004; Case LCR17914 issued on 22/07/2004; and Case LCR188220 issued on 22/01/2007; as well as a High Court hearing: [2007] 4 IR 70.

94 See n. 6 above.

95 Sheehan, "Employers", n. 15 above.

96 And, for example, in *GE Healthcare* (Case LCR18013 issued on 22/11/2004) SIPTU's pay claim was rejected by the court as employees had actually received pay increases in excess of the national pay deals.

97 See also E Gilvarry and B Hunt, "Trade union recognition and the Labour Court: picking up the pieces after *Ryanair*" in T Hastings (ed.), *The State of the Unions* (Dublin: Liffey Press 2008).

On the second issue, there must be a real concern that Irish law remains unique in the Western world in not offering legally guaranteed rights to union representation.

However, the innovative nature of the legislation is worthy of some concluding comments in the light of recent developments in other EU jurisdictions. First, it has been noted that the UK, with a similar voluntarist tradition to Ireland (albeit without the corporatist tendencies evident in the social partnership era), introduced a mandatory union recognition law under the 1999 Employment Relations Act. The legislation has certainly not been an unqualified success from the point of view of bolstering collective bargaining there, as, while there has been a sharp decline in the *derecognition* of unions, concerns have been expressed that recognition has been primarily achieved in "core" union constituencies (where existing membership levels are relatively high); that the scope of collective bargaining is relatively restricted and more often represents consultative, rather than joint, regulation; and that the delays and costs of organising recognition ballots frustrate efforts to recruit and retain members.[98] Recognition laws, clearly, are not a panacea for union ills.[99]

Secondly, we have discussed above the impact of recent ECJ decisions regarding the interaction between economic rights and the right to uphold collective agreements. The ECJ has effectively made voluntary collective agreements that require service providers to comply with terms and conditions that go beyond statutory minima impossible to enforce against providers from other EU jurisdictions. The Irish legislation and the Labour Court's operation thereof, at least before its emasculation in the *Ryanair* case, did provide an opportunity for prevailing norms in an industry or (through voluntary national pay deals) the economy as a whole to be enforced, *even where* these exceeded statutory minimum standards. Where legal protection for collective bargaining rights under EU law appears to be under threat, it may be that the Irish approach offers an alternative model worthy of exploration.

98 See, for example, R Dukes, "The statutory recognition procedure 1999: no bias in favour of recognition?" (2008) 37(2) *Industrial Law Journal* 236; W Brown and D Nash, "What has been happening to collective bargaining under New Labour? Interpreting WERS 2004" (2008) 39(2) *Industrial Relations Journal* 91; B McArthur, "The efficacy of statutory union recognition under New Labour: a comparative review" (2004) 20(3) *International Journal of Comparative Labour Law and Industrial Relations* 399; and S Oxenbridge et al., "Initial responses to the statutory recognition provisions of the Employment Relations Act 1999" (2003) 41(2) *British Journal of Industrial Relations* 315.

99 As Kahn-Freund put it "a healthy union movement can take a great deal of legal intervention whilst weak unions may be its victim"; *Labour and the Law*, n. 49 above, at p. 121.

NILQ 60(4): 403–19

SOX as a window on transference of corporate governance norms across jurisdictions

DR GEORGE PETER GILLIGAN[1]

Monash University, Melbourne, Australia

Abstract

This paper considers the issue of the transference of norms across jurisdictions in corporate governance contexts through the lens of an Australian case study. The paper focuses on the impacts of the United States of America (US) legislation the Sarbanes-Oxley Act 2002 (SOX)[2] from an Australian perspective. The paper draws on a series of semi-structured interviews (n=14), with senior personnel of: accounting firms; business organisations; consumers; financial exchanges; government; institutional investors; investment banks; law firms; private investors; professional associations; and regulators. The findings from the study are that key stakeholders in Australia have taken notice of SOX and its effects in the US, but that the influence of SOX in specifically Australian contexts has been limited. The general perception in Australia seems to be that SOX has had some flaws in its inception and in its subsequent delivery in the US, but also that it has produced some positive outcomes. However, domestic factors and influences are overwhelmingly more important in shaping how financial regulation and corporate governance evolve in Australia. Therefore, it seems that SOX does not signify in any substantive way a regulatory hegemony emanating from the US that determines financial market regulation or the evolution of corporate governance in Australia.

Globalisation and norms transfer

Private, corporate and state forces are continually at work in the worlds of governance and regulation. Regulatory control is a reciprocal arrangement, shaped by negotiated and symbiotic relationships and as such is a social and political process of continuing adaptation within a regulatory setting, not value-neutral and/or pre-destined. These forces can be observed in how regulatory regimes evolve because norms and standards in regulation can be local, national or international phenomena, with regulatory norms and standards interacting at a number of levels in different ways. These interactive effects have

1 Dr George Gilligan is Senior Research Fellow in the Department of Business Law and Taxation at Monash University, Melbourne, Australia. His research interests centre on governance and regulatory theory and practice, especially in relation to the financial services sector. Email address: george.gilligan@buseco.monash.edu.au.

2 Details of the enactment process and formal title of SOX are listed below. The common usage of the Sarbanes-Oxley Act refers to its two main promoters. Paul Sarbanes was a Democratic senator from Maryland from 1967–2007, who in 2002 was chair of the Senate Banking, Housing and Urban Affairs Committee. Michael G Oxley was a Republican member of the House of Representatives from 1981–2007, who in 2002 was chair of the House Financial Services Committee.

become more complex in current conditions of globalisation.[3] Globalisation has been a widely used term since the 1980s as technological innovation, especially in the information technology sector has surged,[4] more easily connecting people and organisations around the world and reducing the effects of time-space distantiation in a host of economic, social, cultural and environmental contexts. The ongoing processes of globalisation reflect this increasing interaction between human activity with closer economic integration across jurisdictional borders, movement of people across borders and sectors, and similarly technology transfer across borders and sectors. The cumulative effect of these forces prompts greater interaction and interdependence between national economies. One of the substantive implications of globalisation is that information, knowledge and normative behaviours (in particular for the purposes of this article, business norms and protocols) may be transmitted and dispersed across national, sectoral and cultural boundaries. These normative transmissions and dispersals can become constituent factors in the *game* between jurisdictions as they seek to attract and/or retain capital investment. These processes of regulatory competition are more pronounced in this current era of globalisation, when economic and political ties between many jurisdictions are deepening and jurisdictions increasingly are playing a mediating role regarding the interests of much business that may be conducted within and around their spheres of influence.[5]

However, transference of norms, including legal norms is, of course, not a novel phenomenon that has been occurring merely in the last 30 years or so.[6] Rather, norms, whether, social, economic, legal or cultural have been transferred across national and societal boundaries throughout history. Some commentators have argued that social norms can be more influential than legal rules in directing how corporate governance is made operational within and between firms.[7] Almost inevitably, the debate and various studies focusing on the relative influence of norms and rules on governance contexts become murky and riddled with subjective evaluation – this is perhaps unavoidable given the scale and nature of the subject matter. Nevertheless, there are myriad instances, in life in general and in business in particular, when people will behave in certain ways, because they are the individual and/or collective normative standards to which they are committed. For example, Tyler's research in the US, based on citizen experiences of the criminal justice system, concluded that citizens' levels of commitment to the system were directed by how fairly

3 Globalisation can be a contested concept and this article will not discuss the utility of various definitions of globalisation or how best to evaluate its impacts. There are many texts that analyse globalisation including: M Bordo, A M Taylor and J G Williamson, *Globalization in Historical Perspective* (Chicago: University of Chicago Press 2003); J A Frieden, *Global Capitalism: Its fall and rise in the twentieth century* (New York: W W Norton & Company 2006); J Gray, *False Dawn: The delusion of global capitalism* (London: Granta Books 1998); P Hirst and G Thompson, *Globalization in Question: The international economy and the possibilities of governance* (Cambridge: Polity Press 1998); K Ohme, *The End of the Nation State* (New York: Free Press 1995); J A Scholte, *Globalization: A critical introduction* (Houndmills: Palgrave Macmillan 2005); and J Stiglitz, *Globalization and its Discontents* (New York: W W Norton & Company 2003).

4 For example, Internet World Stats estimates that, by the end of 2008, 1.6bn or 24% of the world's population were internet users and this rate is rising rapidly, see www.internetworldstats.com/stats.htm.

5 G. Gilligan, "Whither or wither the European Union Savings Tax Directive? – A case study in the political economy of taxation" (2003) 11(1) *Journal of Financial Crime* 56.

6 There is a substantial literature which considers issues of norm transference. See, for example, J N Drobak (ed.), *Norms and the Law* (New York: Cambridge UP 2006); J Elster, "Social norms and economic theory" (1989) 3(4) *Journal of Economic Perspectives* 99–117; R A Posner, "Social norms and the law: an economics approach" (1997) 87 *American Economics Review* 365–9; E Posner, *Law and Social Norms* (Cambridge, Mass: Harvard UP 2000); J F Scott, *Internalization of Norms: A sociological theory of moral commitment* (Englewood Cliffs, NJ: Prentice-Hall 1971); and O Williamson, *The Economic Institutions of Capitalism* (New York: Free Press 1985).

7 For example, M Eisenberg, "Corporate law and social norms" (1999) 99 *Columbia Law Review* 1253; G P Gilligan, *Regulating the Financial Services Sector* (London: Kluwer Law International 1999).

they perceived the system treated them.[8] Coffee comes to a similar conclusion with regard to the corporate arena, noting that managers and controlling shareholders are often constrained in their behaviour by social forces which are independent of underpinning legal sanctions, and that the relative capacity of social norms to influence controlling shareholders can vary significantly across jurisdictions.[9]

It can be difficult to know sometimes where the influence of social norms begins and the influence of legal rules ends because in so many cases they can meld or be mutually sustaining. This is because there is not always a clear distinction between when norms are internalised by actors and the extent to which compliance with, or acceptance of, norms is motivated by the potential impact of third-party enforcement through formal mechanisms such as the justice system, or more informal ones, such as a firm's internal business culture or philosophy. So, when considering relative influences in the production of corporate governance norms, there is merit in emphasising behavioural approaches to how governance standards are produced and how they are evaluated by those who are involved in their application, what Stout describes as "the reality of socially contingent, other-regarding preferences".[10] Similarly, Rock and Wachter posit that firms are largely governed by norms which they interpret as non-legally enforceable rules and standards.[11]

Despite the difficulties of assessing the impacts of norms on business practice, there has been no shortage over the last 15 years or so of international organisations generating codes of preferred behaviour which they hope will gain traction as influential corporate governance norms across jurisdictions. Examples include the International Monetary Fund (IMF),[12] the World Bank[13] and the Organisation for Economic Co-operation and Development (OECD).[14] These initiatives do not bind jurisdictions but instead are preferred benchmarks to stimulate improved corporate governance in both national and international contexts. What is of special interest for this article is not a general appraisal of international codes of governance and how well they may travel in terms of their impact on different countries.[15] Rather, the article's purpose is to examine how influential those corporate governance norms, emanating from the world's most powerful nation the US, may be in terms of shaping prevailing corporate governance standards in other jurisdictions, in this case Australia. In particular, whether US corporate governance norms produced in the US are more privileged in terms of their capacity to shape regulatory initiatives in other jurisdictions.

---

8 T R Tyler, *Why People Obey the Law: Procedural justice, legitimacy and compliance* (New Haven: Yale UP 1990).

9 J C Coffee Jnr, "Do norms matter?: A cross-country examination of the private benefits of control", *Columbia Law and Economics Working Paper* No 183 (2001), pp. 3 and 29: http://papers.ssrn.com/sol3/papers.cfm?abstract_id=257613.

10 L A Stout, "Other-regarding preferences and social norms", *Georgetown Law and Economics Research Paper* No 265902, *Georgetown Public Law Research Paper* No 265902 (2001), p. 32, http://papers.ssrn.com/sol3/papers.cfm?abstract_id=265902.

11 E B Rock and M L Wachter, "Islands of conscious power: law, norms and the self-governing corporation" (Spring 2001) *Symposium, Norms and Corporate Law, University of Pennsylvania Review*, http://papers.ssrn.com/sol3/papers.cfm?abstract_id=269328.

12 IMF, *Good Governance: The IMF's role* (Washington, DC: IMF Publication Services, August 1997).

13 International Bank for Reconstruction and Development/World Bank and OECD, *Global Corporate Governance Forum: 1 Focus: Corporate Governance and Development* (Washington, DC 2003).

14 OECD, *OECD Principles of Corporate Governance* (Paris: OECD Publications Service 2004).

15 For an analysis of the capabilities of initiatives by international organisations to promote convergence in national corporate governance schemes, see F de Zwart and G Gilligan, "Comparative corporate governance schemes and their relevance for the sporting sector", *Monash University Department of Business Law and Taxation Research Paper* No 16 (2008): http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1295682.

Specifically, the Australian case study is designed to explore the role that the characteristics of Australia's political economy play in shaping a national regulatory regime and scoping its responses to international changes in governance of financial markets. In particular, the causes and consequences of the enactment of stringent US legislation on corporate liability, the Public Company Accounting Oversight and Investor Protection Act (2002) which is more commonly referred to as Sarbanes-Oxley (SOX).[16] SOX has emerged as a piece of domestic legislation that has implications for many jurisdictions other than the US, in particular for overseas firms that elect to list on US exchanges and thus must meet US listing requirements, including the more stringent innovations (discussed below) that have arrived as part of the SOX regime. Before discussing these provisions and Australian reactions to them, it is necessary to consider why SOX emerged in the way and shape that it did.

Enron, dirty washing and SOX

SOX was a specific US response at a particular time to what was widely perceived as a crisis in the governance standards of contemporary corporations following the collapse of Enron and WorldCom. As the Enron saga unfolded, or (should that be?) imploded, it was a moment in time when the "dirty washing" of US corporate life was hung out for all to see. The demise of Enron in December 2001 was sudden, a year earlier it had been the seventh largest company in the US with assets valued at almost $US70bn.[17] As late as September 2001, Axiss Australia[18] published an interview with Paul Quilkey, managing director of Enron Australia, in which it stated that "Enron is succeeding in its primary task – to always make money."[19] Less than three months later, Enron was in bankruptcy amid allegations of fraud for its use of special purpose entities in dubious off-balance-sheet accounting practices. These practices later led to convictions for its chief executive Jeffrey Skilling (imprisoned for 24 years), its chief accountant Richard Causey (seven years), its chief financial officer Andrew Fastow (10 years), and its external auditor, global accounting behemoth Arthur Anderson, ceased to exist as a substantial commercial entity.[20] These were severe sentences and prompted some academic analysis of whether SOX had altered sentencing contexts in relation to white-collar crime in the US.[21] There is a large literature that has dissected the Enron and WorldCom scandals[22] and considered not only what the core causes were, but also the likely value of regulatory responses – especially SOX – to

16 Sarbanes-Oxley Act 2002, Public Law No 107-204, 116 Stat. 745.

17 L Collins, "Death of a titan – biggest collapse in US corporate history", *Australian Financial Review*, 30 November 2001, p. 20.

18 Axiss Australia (now part of the Australian Trade Commission – Austrade) was established by the Australian Commonwealth government to act as a one-stop reference point for information on Australia's financial services industry, see www.austrade.gov.au.

19 Axiss Australia, "The energiser – Enron the commodity and financial futures dynamo" (September 2001) *Markets Perspective* 4–8, at p. 4.

20 A Barrionuevo, "Enron scandal stalks Skilling", *Australian Financial Review*, 25 October 2006, p. 68.

21 For example, P J Henning, "The changing atmospherics of corporate crime sentencing in the post Sarbanes-Oxley Act Era" (2008) 3(2) *Journal of Business and Technology Law*; C S Lerner and M A Yahya, "'Left behind' after Sarbanes-Oxley" (2007) 44(4) *American Criminal Law Review*.

22 See, for example, F Partnoy, *Infectious Greed: How deceit and risk corrupted the financial markets* (New York: Henry Holt & Company 2004); K Vasatka, "WorldCom scandal: a look back at one of the biggest corporate scandals in US history", *Associated Content*, 8 March 2007.

them.[23] Bainbridge offers a five-year overview of SOX since its inception and a pragmatic description of how SOX impacts upon the directors and managers of publicly held corporations that are subject to SOX provisions.[24]

The substantial academic literature which has grown around the advent of SOX and its associated academic discourse regarding the relative legitimacy and efficacy of SOX has in some case been quite heated. For example, Ribstein and Butler perceive SOX as a costly mistake, in fact a debacle, which in their opinion has been a colossal failure, poorly conceived and hastily enacted during a regulatory panic.[25] Fisch and Rosen argue that the legislative reaction to Enron and its associated US business culture problems, as epitomised by SOX-driven corporate governance reform, has been flawed and that there should be a greater emphasis on demand-side approaches which better align incentives with promoting good corporate governance.[26] Perino also is critical of SOX arguing that it adds little to the deterrence capabilities of US laws and regulators.[27] In one of the most strident and early critiques of SOX, Romano labels the legislation as *quack corporate governance* and a policy blunder which should be stripped of its mandatory provisions.[28]

However, not all the reaction by academe to SOX was negative, especially, and interestingly so given the focus of this article, from those commentators who placed a greater emphasis on the capacity of norms transfer to be affected by regulatory intervention. For example, Brown refutes Romano's claim that SOX is *quack corporate governance* and, instead, argues that it has stimulated improvements in US corporate governance processes.[29] Fanto believes that SOX was a symbolic expression of outrage by the community against misconduct by some members of the elite in the 1990s, which sought to reassert the need for the social value of professionalism in capital markets, and thus was "fundamentally a reassertion of social values against the socially destructive aspects of the self-interest ideology in the US which views individuals solely as profit-maximisers".[30] Frankel also sees positive potential in SOX and argues that it presents an opportunity to reward truthful corporations and their management.[31]

---

23 See W W Bratton, "Enron, Sarbanes-Oxley and accounting: rules versus principles versus rents" (2003) 48(4) *Villanova Law Review* 1023; D M Branson, "Enron – when all systems fail: creative destruction or roadmap to corporate governance reform?"(2003) 48(4) *Villanova Law Review* 989; J N Gordon, "Governance failures of the Enron board and the new information order of Sarbanes-Oxley", *Columbia Law and Economics Working Paper* No 216, *Harvard Law and Economics Discussion Paper* No 416 (2003), http://papers.ssrn.com/sol3/papers.cfm?abstract_id=391363.

24 S M Bainbridge, *The Complete Guide to Sarbanes-Oxley* (Avon, Mass: Adams Media 2007).

25 H N Butler and L E Ribstein, *The Sarbanes-Oxley Debacle – What we've learned: how to fix it* (Washington: American Enterprise Institute Press 2006).

26 J E Fisch and K M Rosen, "Is there a role for lawyers in preventing future Enrons?", *Fordham School of Law Pub-Law Research paper* No 21, *Villanova Law/Public Policy Research Paper* No 2003-15, http://papers.ssrn.com/sol3/papers.cfm?abstract_id=367661.

27 M A Perino, "Enron's legislative aftermath: some reflections on the deterrence aspects of the Sarbanes-Oxley Act of 2002", *Columbia Law and Economics Working Paper* No 212 (2002), http://papers.ssrn.com/sol3/papers.cfm?abstract_id=350540.

28 R Romano, "The Sarbanes-Oxley Act and the making of quack corporate governance", *NYU Law and Econ Research Paper* 04-032; *Yale Law and Econ Research Paper* 297; *Yale ICG Working Paper* 04-37; *ECGI Finance Working Paper*, 52/2004; http://papers.ssrn.com/sol3/papers.cfm?abstract_id=596101.

29 J R Brown Jr, "Criticising the critics: Sarbanes-Oxley and quack corporate governance", *U Denver Legal Studies Research Paper* No 07-11; (2006) 90 *Marquette Law Review* 209, http://papers.ssrn.com/sol3/papers.cfm?abstract_id=959443.

30 J A Fanto, "A social defense of Sarbanes-Oxley" (2007/08) 52 *New York Law School Review* 517–32, at p. 532.

31 T Frankel, *Trust and Honesty: America's business culture at a crossroad* (Oxford: Oxford UP 2006).

Despite these disagreements about the merits or otherwise of SOX, and whether it deserves such a derogatory label as *quack corporate governance*, both supporters and critics of SOX agree that it was enacted extremely swiftly by US standards. Romano believes that the SOX legislative process proceeded in an atmosphere of crisis following spectacular accounting frauds at several large public corporations. These frauds fuelled what Romano terms:

> a media frenzy . . . From January to July 2002, 471 of 613 business-related stories on the major network evening news were on corporate scandals, compared to 52 of 489 business stories in the same period the previous year . . . over eighty per cent of those news stories looked to government to solve the perceived problem . . . The confluence of spectacular financial scandals, a declining stock market, waning public confidence in business, and a media frenzy in an election year resulted in a restrictive legislative debate and progressively more lopsided votes in support of greater regulation.[32]

The demand in the media for a swift and significant legislative response was driven partly by the political connections that Enron and some of its senior executives had to the White House administration of President George Bush. For example, over a period of 10 years, Enron and its CEO had contributed more than $US1bn to various Bush political campaigns.[33] A combination of intense media scrutiny and the need for Congress to be seen to be making specific and strong responses to widespread community disquiet about prevailing ethical standards in US corporate life ensured that the votes in the US Congress certainly were overwhelming numerically, 423:3 in the House of Representatives[34] and 99:0 in the Senate.[35]

So, SOX was passed in mid-2002 and there has been disagreement ever since about whether its provisions and its application have been beneficial to US interests, and in particular whether the additional SOX-related costs that have had to be borne by US firms have been justified. There is uncertainty about the cost-effectiveness of the SOX reforms. For example, the 2008 survey of US financial executives by Oversight Systems Inc. found that only 29 per cent of its sample had confidence that the SOX reforms would help to reduce financial fraud (a drop from 40 per cent in 2005).[36] However, there is no uncertainty on the issue of whether SOX has led to increases in the costs of regulatory compliance which have raised the costs bottom-line for US firms. Bainbridge notes that meeting SOX requirements has resulted in audit costs across the board for US firms being estimated to have risen by up to 30 per cent and that some of the largest firms pay annual fees as high as $2m to help fund the SOX-established Public Company Accounting Oversight Board (PCAOB).[37] Similarly, the 2007 Financial Executives International (FEI) Survey found that for companies with a market capitalisation above $75m the annual costs for compliance with SOX provisions was $1.7m.[38]

---

32 R Romano, "Does the Sarbanes-Oxley Act have a future?", *Yale Law School, John M Olin Center for Studies in Law, Economics and Public Policy Research Paper* No 385 (May 2009), pp. 11 and 13, http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1404967.

33 L Collins, "Enron: the fallout, UBS lands energy trading deal", *Australian Financial Review*, 14 January 2002, p. 7.

34 148 Cong. Rec. 5544-48, H4580, daily edn, 25 July 2002.

35 148 Cong. Rec. S7365, daily edn, 25 July 2002.

36 Oversight Systems, *The 2008 Oversight Systems Financial Executive Report on Sarbanes-Oxley* (2008), www.oversightsystems.com/collateral/2008OversightSOXReport4.pdf.

37 S M Bainbridge, *The Complete Guide to Sarbanes-Oxley* (Avon, Mass: Adams Media 2007), p. 3.

38 Financial Executives International, *FEI Survey: Average 2007 SOX Compliance Cost $1.7 Million* (2007), http://fei.mediaroom.com/index.php?s=43&item=204.

There are numerous provisions of SOX that have sought to bring about transformative change in the norms of US business culture. Section 301 requires all listed companies to have audit committees comprised entirely of independent directors. Section 304 requires the forfeiture of bonus, incentive and equity compensation for a firm's chief executive officer and chief financial officer if there is a material restatement of the company's financials due to misconduct. Section 406 requires a company to disclose if it has a code of ethics for its senior financial officers and if not, why not. Section 802 provides for criminal penalties for violation of SOX that can be as high as 20 years' imprisonment. Section 1107 provides for penalties including possible imprisonment of up to 10 years for retaliation against whistleblowers.

However, it is s. 404 which has created the most controversy and attracted the most ire from critics of SOX. Section 404 requires each company's annual report to contain: 1) a statement of the management's responsibility for internal control structures and financial reporting; 2) management's assessment each year of the effectiveness of these control structures and reporting processes; and 3) a report by the auditors regarding the above assessments by management. With the potential sanction of substantial gaol terms, SOX in general, and s. 404 in particular, sought to sheet accountability for any misstatement of company financials at the apex of corporate structures in efforts to avoid future Enron-type scandals. Unsurprisingly, when those at the top of an organisation may face punishment for mistakes/deceptive behaviour lower down the corporate food chain, a stronger and more costly organisational emphasis on regulatory compliance emerges. So, something of a growth industry around s. 404 has emerged in the US and this has inevitably increased operational costs to companies subject to SOX provisions. Maher and Weiss found that the annual SOX compliance costs range on average from 0.298 per cent to 0.618 per cent of sales in each of the first four years after SOX was enacted.[39] In a survey of Fortune 500 firms in the US (n = 206), Kipperman et al. found general support for the proposition that the specific effects of SOX had reduced the perceived value of the firm.[40]

One of the more contentious and contestable issues surrounding the effects of SOX and its perceived impact on US economic interests is whether it has made firms less likely to opt for a listing in the US, especially in relation to the compliance costs associated with s. 404, which is generally accepted as being the most onerous section of the Sarbanes-Oxley Act. For example, one strident critic of SOX has been the Committee on Capital Markets Regulation, a group of 22 US experts from the investment community, business, finance, law, accounting and academia which was formed due to concerns about US financial actors becoming less competitive. The committee claimed that the reductions in global financial market share suffered in recent years by US financial markets, for example only seven per cent of the value of global Initial Public Offerings in 2007, compared to 58 per cent 10 years earlier was due largely to the effects of an increasing regulatory burden, especially SOX. In its report the committee made a series of recommendations which included scaling back some of the SOX reforms, especially in relation to s. 404.[41] The committee was specifically concerned about the US losing business to the UK and the chair of the London Stock Exchange had been quoted previously as saying that "the hard rules of Sarbanes-Oxley

39 M W Maher and D Weiss, "Costs of complying with the Sarbanes-Oxley Act", *UC Davis School of Management Research Paper* No 10-08 (2008), http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1313214.

40 F Kipperman, R P McAfee and N Vakkur, "Non-pecuniary costs of Sarbanes Oxley", *RAND Working Paper* No WR-554-CCEG (2008), http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1312090.

41 Committee on Capital Markets Regulation, *The Competitive Position of the US Public Equity Market* (2007), www.capmktsreg.org/pdfs/The_Competitive_Position_of_the_US_Public_Equity_Market.pdf.

made the US far less attractive and welcoming to foreign issuers".[42] Kamar et al found that SOX induced small firms to exit the US public capital market.[43] However, Bartlett's study suggested that non-SOX factors were the major motivation for buyouts commonly evoked as evidence that SOX has harmed the competitiveness of US capital markets.[44]

So, it would seem that the evidence on this issue is mixed and, as yet, there is no compelling case that Sarbanes-Oxley definitely has led to a reduction in the numbers of companies choosing to list in the US. It is under this canopy of uncertainty regarding the utility and effects of SOX that the Australian case study was conducted as part of a broader project, managed by Queen's University Belfast and funded by the ESRC, which sought to gauge the influence of SOX in a number of jurisdictions around the world.[45]

Did SOX travel well? Evidence from Australia

The Australian case study draws on a series of semi-structured interviews (n=14)[46] with senior personnel of: accounting firms; business organisations; consumers; financial exchanges; government; institutional investors; investment banks; law firms; private investors; professional associations; and regulators.

The first question posed to respondents focused on their perceptions of SOX's effect in the US from the perspective of an Australian observer. None of the respondents felt that the label *quack corporate governance* was appropriate, but all saw SOX as a political reaction (over-reaction in the minds of some), to business scandals. Rather, the general view was that SOX should be seen as "reactive governance" – a particular type of response to a particular set of corporate collapses such as Enron and WorldCom – "a very US solution" as one corporate governance professional put it. Respondents agreed that attitudes towards SOX in the US were mixed, with business groups by and large being more negative towards SOX and regulators having more positive views. One should not assume from this that Australian regulators in the study were absolutely positive about SOX because they did feel that some of the US criticism of SOX "had traction". In particular, a sense that although the principles underlying SOX were largely sound its provisions were too prescriptive, engendering excessive compliance costs. This reflects the Australian context in which, compared to the US, there is less emphasis on black letter law and more reliance on principles/best practice guidelines emphasising an "if not, why not" culture. However, as one regulatory professional noted, it is difficult to argue against some of the SOX reforms which addressed obvious conflicts of interest such as looking at auditors doing a variety of non-audit work for clients, establishing audit committees, requiring top executives to certify

42 L Taylor, "Sarbanes-Oxley gives UK an edge", *Australian Financial Review*, 24 December 2002, p. 13.

43 E Kumar, P Karaca-Mandic and E Talley, "Going-private decisions and the Sarbanes-Oxley Act of 2002: a cross-country analysis" (2009) 25(1) *Journal of Law, Economics and Organization* 107–33.

44 R P Bartlett, "Going private but staying public: reexamining the effect of Sarbanes-Oxley on firms' going-private decisions" (2009) 76(1) *University of Chicago Law Review* 7–44.

45 The interviews for this Australian case study were conducted in 2007/2008 by Gilligan and his Monash colleague, Associate Professor Ken Coghill. Similar case studies took place in Ireland, South Africa, South Korea, Turkey, the United States and the United Kingdom. All the case studies were part of the project Regulatory Regime Change in Financial Markets, which was coordinated by the Institute of Governance at Queen's University Belfast and funded by the Economic and Social Research Council (UK), with additional funding to help support the Australian case study supplied by the Department of Business Law and Taxation and Department of Management at Monash University. The author would like to acknowledge the support not only of all those mentioned above, but also the valuable contributions of the respondents who participated in the interviews.

46 Prior to the interviews the respondents were sent the discussion topic questions contained in the Appendix to this article, thus heightening the level of focus for the individual interviews and permitting a greater level of consistency across them as a grouping.

accounts, and giving whistleblowers more protection if they reported any suspicions of fraud. In particular, several of the respondents stressed that repeated failures by regulators and business to deal with the issue of conflicts of interest were crucial in explaining why SOX emerged in the way that it did.

The specific knowledge of SOX amongst the respondents was high. Most emphasised that a number of the SOX provisions are similar to existing Australian regulatory requirements under the Corporations Act and ASX Corporate Governance Principles and Recommendations of 2007.[47] For example, the certification under SOX s. 302 is comparable in many ways to certifications made under s. 295 of the Corporations Act and in response to ASX Recommendations 7.2 and 7.3. Also, the disclosures regarding the code of ethics required by SOX s. 406 are comparable to ASX Recommendation 3.1. Several respondents were of the view that to a large extent, other than for SOX s. 404, the corporate governance practices in SOX are consistent with leading corporate governance practices and requirements in Australia. The provisions of s. 404 did draw specific attention from respondents, not only because of both the absence of an equivalent requirement under the Corporations Act or the ASX Corporate Governance Principles and recommendations, but also because of the relatively high cost associated with compliance. Respondents felt that the underlying intent of s. 404 – that a company should report on the effectiveness of its internal control over financial reporting – is sound. However, the prevailing view was that the costs of meeting this obligation have been unnecessarily inflated because of insufficient upfront guidance from the Securities Exchange Commission (SEC), and the influence of PCAOB Auditing Standard No 2 on both registrants and their external auditors. However, several respondents felt that this issue had been redressed by subsequent guidance from the SEC and the issue of Auditing Standard No 5, which have helped to produce more efficient and effective assessment programmes.[48]

When asked whether SOX has been a positive innovation for the US or not, respondents perceived SOX as generally not a bad thing, but as one said, "it was a very American response to problems in the US". A notable exception was the fund manager interviewed who felt that SOX was neither positive nor innovative, perhaps reflecting the views of some of his US business counterparts. The broad view held was that SOX's main effect has been to "functionalise" codes of conduct processes and build more structure into the US corporate governance context in its efforts to counter those firms operating in the greyer zones of compliance.

Unsurprisingly there was a belief across the respondents that a major part of the problem in evaluating SOX is the inability to specify the major evaluative criteria. Respondents felt that instances of financial statement fraud or restatements – or the absence thereof – should not be directly attributed only to SOX. Any criteria for evaluating SOX must be derived in the impact that SOX has had on attitudes and behaviours – within the business community, among the members of audit committees, and of stakeholders. There were assorted prickly variables thrown up by respondents that might be added into the SOX evaluative mix. For example, how does one judge and measure market confidence? Or, has SOX mitigated the harmful effects of the current "bear market"? Or does SOX contribute to a more precise capability to predict corporate failure? Or have

47 Corporations Act 2001; ASX Corporate Governance Council, *Corporate Governance Principles and Recommendations* 2nd edn, (Sydney: Australian Stock Exchange, 2 August 2007).

48 PCAOB, Auditing Standard No 5: "An audit of internal control over financial reporting that is integrated with an audit of financial statements and related independence rule and conforming amendments", 12 June 2007, Washington DC, www.pcaobus.org/Rules/Docket_021/2007-06-12_Release_No_2007-005A.pdf.

improved SOX internal controls and disclosure brought gains in terms of costs to shareholders or cost of capital? These are thorny and complex issues that merit substantive research in their own right.

Respondents were unanimous that there has been something of a push-back against SOX in the US, which, given prevailing political realities and the worsening economic climate both domestically and globally, has acted as a restraint on further regulatory innovation by US politicians and regulators in relation to business activity. However, the views of the respondents diverged regarding the intensity of any backlash or the degree to which this acted as a brake on regulatory innovation. Some respondents felt that "backlash" was too strong a term, viewing SOX more as an expected "human nature response" to poor behaviour by corporate actors and a sense of "crisis" in the values of business in the US. A longer-term issue mentioned by several respondents and which has implications for non-SOX jurisdictions is the "brain drain of directors" because the risk–reward matrix regarding governance expectation and potential liability may be moving out of kilter, so that successful directors are particularly pressured in this regard. This issue is a growing concern in many jurisdictions, not just a SOX-affected US.

Respondents felt that resentment against SOX in the US has largely been directed at s. 404. The volume of this resentment has been somewhat reduced subsequent to the recent SEC guidance and issue of PCAOB Auditing Standard No 5, particularly from larger US corporate organisations that have now established more efficient and effective assessment programmes. However, respondents felt that there is still discontent among smaller US corporate organisations that are now implementing and operating assessment programmes without the resources available to larger companies.

Whether SOX had an effect on their own organisation's activities obviously varied according to occupational role, with some respondents affected directly and others only indirectly. Certain Australian regulatory actors were directly affected as they have to be able to exchange information and work with US regulatory agencies such as the SOX-created PCAOB, and, unsurprisingly, it has been s. 404 which has had the most significant effect. Overall, the effects of SOX have been limited, with those responsible for the supervision of markets and industry professionals always conscious of the need for the highest possible levels of integrity regardless of reforms in the US or elsewhere. Interestingly, some of the private-sector respondents have felt the impacts of SOX in particular ways. For example, it has made tasks such as corporate restructuring much more difficult because, under SOX rules (especially on conflicts), different people in organisations will have different levels of knowledge and this can make managing a restructure quite cumbersome.

When asked about the most significant changes in the regulatory regime of Australia in the period immediately coincidental with the advent of SOX, the general view was that the CLERP (Corporate Law Economic Reform Program) 9 legislation[49] and the ASX Corporate Governance Council's Corporate Governance Principles and Recommendations are the most significant changes in Australia following SOX. The Corporate Governance Principles also draw on the UK Combined Code,[50] particularly the "if not, why not" approach, as well as the OECD Corporate Governance Principles and represent a very different approach to SOX. As CLERP 9 emanated largely from the HIH Royal

49 The CLERP (Audit Reform and Corporate Disclosure Act) 2004, which came into effect on 1 July 2004.

50 Financial Reporting Council (FRC), *The Combined Code on Corporate Governance*, June 2006, London (the UK Combined Code).

Commission,[51] timing of these changes with SOX was felt by respondents to be largely coincidental, rather than a specific regulatory reflex action across borders. Nevertheless a couple of the respondents were of the view that there is generally an increasing influence of US law and regulation in Australia, with Australian courts increasingly been persuaded by some of the developments in US law.[52] Also, some respondents pointed to some of the unexpected consequences of SOX, such as the conflicts of interest and large chartered firms issue, prompting a substantial growth in small firms specialising in corporate restructures being established by ex-partners of the major global accounting firms, for example, KordaMentha.

Question 3 focused on the influence of Australian domestic factors relative to SOX's capacity to transfer governance norms. There was universal agreement that the spate of collapses and fraud instances in the late 1990s and early 2000s in Australia (e.g. HIH, One-Tel) confirmed that the factors leading to SOX were not unique to the US and that a local response was necessary. Also, there was a broad view among respondents that these factors were situational rather than structural in Australia. Certainly, these influenced the types of reforms included in the CLERP legislation and were also an important factor that led to the formation of the ASX Corporate Governance Council in August 2002. All concerned were keen to avoid the "black letter" approach taken in SOX and for this reason the council worked on putting out a series of flexible principles and recommendations which are backed up by the Listing Rules 4.10.3 and 12.7. The first set of Corporate Governance Principles was developed in six months and before the CLERP 9 legislation. Respondents directly involved in that process confirmed that the council was conscious at the time that, if it did not arrive at a market-based solution quickly, legislation was highly likely. The second set of principles was developed more slowly with the benefit of broad public consultation – and involved much more of an evolutionary journey. This is an important point from the perspective of transference of legal norms because there was a strong sense amongst respondents that corporate governance in Australia has been on an evolutionary path, so that, for most companies, implementation of corporate governance regulation has been a slower process of formalisation and improvement rather than an outright transformation.

All respondents agreed that the Australian response was consistent with a more principles-based approach: for example, the use of "if not, why not" disclosure in the ASX Corporate Governance Principles rather than mandating compliance. Continuous disclosure is a very Australian regulatory response and SOX interacts partially with Australian regulatory praxis, but its influence is limited. As one respondent put it, the strength of the development of corporate governance in Australia has been its emphasis on "functionality of governance". Certainly, as far as companies are concerned, respondents felt that Australia has struck the right balance between self-regulation and black letter law, and it was emphasised repeatedly that there was stakeholder input into the regulatory change process, for example, with regard to CLERP, which helped to reduce some of the angst levels of business in relation to the changes that were happening. This melded with a very strong view across all those interviewed that the standard of corporate governance in

51 Justice Neville Owen was appointed in August 2001 by the Commonwealth Government of Australia to lead a Royal Commission of Inquiry into the reasons for and the circumstances of the collapse of HIH Insurance. Justice Owen's report was tabled in April 2003 and was influential in shaping policy directions for corporate governance in Australia. See *Report of the HIH Royal Commission* (2003), www.hihroyalcom.gov.au/finalreport/index.htm.

52 For example, the Sons of Gwalia case which drew upon US case law and whose decision now allows the possibility under Australian law for the claims of shareholders to be ranked equally with unsecured creditors: *Sons of Gwalia Ltd* v *Margaretic*; *ING Investment Management LLC* v *Margaretic* [2007] HCA 1.

Australia appears very high, with Australian corporate law being firmer in some of the areas that SOX sought to strengthen in the US. This in itself reduced the need for specific SOX-type reform but, regardless, overwhelmingly in the view of respondents, local context will impact significantly on how a regulatory innovation like SOX travels.

Question 4 sought comment on how influential have changes in the global regulatory environment been upon Australian regulatory innovations. Respondents agreed that changes in international or regional regulatory regimes will always be influential in the sense that Australia looks to the way in which other jurisdictions deal with particular issues, such as corporate collapses or the role of the auditor. However, respondents all agreed that there is increased sensitivity towards corporate governance issues in general and that it is important that Australia responds to these changes in a way that is appropriate for Australian conditions, including appropriate levels of regulatory burden and compliance cost. On the issue of Company Boards, respondents perceived them as being increasingly sensitive towards education and compliance issues regarding corporate governance. For example, more Australian companies are heeding calls for increasing numbers of independent directors and there are a decreasing number of recalcitrant companies on this issue. Respondents noted that listed companies in particular are very sensitive to what investors want to hear, with board composition seen as especially important. Respondents delivered a comprehensive "no" on whether Australian regulatory innovation was a direct and coordinated response to the implementation of US regulatory regime policies. All believe that Australian factors have been more significant regarding corporate governance reforms in Australia, which highlights again the importance of subsidiarity factors when considering the issue of transference of legal norms across borders.

Question 5 considered how the legal approach to, and enforcement of, corporate governance may have changed. Respondents perceived the issues raised by this question as not so much how has the legal approach changed but more how "enforcement" of corporate governance has changed. For example, change can be seen in the original governance ASX Listing Rules which asked companies to describe their practices in fairly general terms as opposed to the more recent ASX Corporate Governance Principles which are quite detailed. Respondents noted that CLERP 9 has enacted certain specific corporate governance principles as law in the Corporations Act, thereby enabling increased regulatory enforcement. This has dovetailed with the approach of the ASX Corporate Governance Principles and Recommendations which, while not having the force of the Corporations Act, have driven a more consistent view of, and approach to, corporate governance practices. One respondent noted this synergy with evolution at the political level, in that in Australia there is now a Minister for Corporate Governance and Regulation (as part of the Ministerial Team at the Treasury), perhaps signalling a concerted move away from the "boom, bust, investigate, regulate cycle" and an increased emphasis from governments on market dynamics and corporate governance. A number of respondents also commented on the fact that there has been in Australia a growing media influence in the prosecution of corporate governance (following US trends in this regard, for example, a sharper focus on insider trading). This trend has been accompanied by an increasing use of the media by regulators themselves and this is an area of burgeoning influence which merits focused research.

Question 6 asked whether the government, the "governance profession" and business have different goals and, unsurprisingly, not only did respondents have mixed views on this issue, they also found it tricky to answer. As one respondent noted, co-objectives will be held by any entity or grouping whether it is labelled as "governance" or as some other category. Some felt that ultimately all these groups will have similar goals: deep liquid

markets that are internationally competitive and a market that has the highest integrity, the right balance of regulation and which provides good returns to all investors. Other respondents felt that there were significant differences. For example, the goals of government include regulating the corporate governance practices of businesses in order to protect the integrity of capital markets and their participants; the goals of businesses include assuring shareholders and other stakeholders that the company is well governed; and the goals of the governance profession including assisting businesses to provide assurance to shareholders and other stakeholders about corporate governance practices (including compliance with relevant government regulation). Nevertheless, as one respondent (a senior commercial lawyer in private practice) stressed, there has been a "sea change" in relation to Corporate Social Responsibility (CSR) in Australia. No longer is lip service paid to this issue by Australian corporations, it is now a very important reputational risk management issue, with the corporate governance profession actively educating business in this area. He often uses the "Page 1" test when advising boards by asking: "How would you feel if this was on page 1 of a broadsheet newspaper?" If the board members are uneasy about this, then they are probably handling the issue incorrectly and so the issue of avoiding conflicts, for example, is now a commercial priority for many businesses in Australia. The combined effects of domestic factors mean that it has not required a dramatic SOX-type legislative bombshell to engineer this change in Australia, but a more evolutionary process.

All respondents answered yes to question 7 on whether corporate governance should concern itself with stakeholder interests, but there was variance in what the intensity of concern should be. Some respondents emphasised the revised ASX Corporate Governance Principles which talk about stakeholders in two places in the context of Principle 3 (to "promote ethical and responsible decision-making") and the commentary to Principle 7 which talks about the need to consider risks relating to some matters characterised as CSR/sustainability when looking at the issue of risk. Other respondents commented on the almost inevitable tensions raised in a particular area of stakeholder interest, corporate governance regulation relationship, and it remains a source of debate in contemporary discourse on corporate governance in Australia.[53]

With regard to question 8 and whether there is a trade-off between shareholder and stakeholder interests, generally respondents felt that there may be competing objectives between shareholders and stakeholders, but these do not necessarily translate to differences in corporate governance objectives. For example, objectives relating to the reliability of financial reporting and effectiveness of related controls are relevant to all stakeholders, including shareholders. Of course, respondents recognised that in some contexts trade-offs are likely between shareholders (who emphasise returns/profits) and stakeholders (who want to ensure that they can get their money back and can operate in a relatively secure investment environment), but these may well vary across and between industries.

The interview's final question sought summary assessment of both the beneficial and adverse effects of the changes made by SOX and the contemporary Australian reforms. In terms of its positive effects, all respondents agreed that SOX has helped increase awareness of corporate governance in general and avoidance of conflicts in particular, and that these changes have been coupled with a rise in regulatory activity in the area of corporate

53 This can be seen most recently in the July 2009 report of the Corporations and Advisory Markets Committee (CAMAC), *Aspects of Market Integrity*. The report focused on: dealing by company directors in the shares of their company; spreading of false and misleading information – *rumourtrage*; and disclosure of information in the briefing of analysts. The report is available at www.camac.gov.au/camac/camac.nsf/byHeadline/Whats+NewMarket+Integrity+Report+June+2009?openDocument.

governance. Specific beneficial effects noted by respondents included: the increased rigour in relation to issues such as the independence of auditors; assurance in relation to financial statements; and a greater general consciousness of the importance of these issues. Listed companies do take these issues seriously, as is evidenced by the overall improvements in reporting levels in relation to corporate governance disclosures. There was consensus amongst the respondents that the main benefit of SOX was that it focused increased attention on areas that should have been the focus of attention anyway, especially for business organisations themselves. These areas included: improved business ethics; better disclosure; clearer relationships between risk management and financial management; and increased truth in auditing. As one respondent commented, corporate managers and executives in Australia benefit from the SOX reforms, because they can help to underpin some of their assertions about how a company should be run and increase the confidence of external actors in the financial reporting and other governance aspects of firms. However, several respondents pointed out that there is real difficulty in quantifying the benefits of the SOX reforms.

Respondents agreed that the negative effects of SOX largely centred on costs and the scale and industry of an organisation seems to be crucial in any decision-making matrix. One respondent estimated that the Big Four Australian banks would have had to invest $25m per year in order to be compliant with the initial s. 404 regime. Following those initial sunk costs and given the revised s. 404 regime as a result of the SEC changes annual costs of s. 404 compliance were probably less than $10m – probably about $8m per annum. Interestingly, that figure was viewed as tolerable to a large organisation like a Big Four bank because of the reputational protection that s. 404-type compliance can bring to a firm of such organisational scale which still maintains much of its revised governance arrangements (introduced as a direct result of s. 404 in SOX), and because it felt that the firm benefits from the changes. Big bank audit committees remain in favour of retaining many of the s. 404 changes. In the view of this particular respondent that is the "proof of the pudding" really – none of the Big Four banks in Australia are significantly scaling back their s. 404 changes, although some do tweak their internal governance systems to suit the firm's objectives. The banks are making rational actor cost–benefit and risk management decisions about the intrinsic value of an internal s. 404-type oversight and accountability system. However, other respondents were less well-disposed to the cost-benefit impacts of SOX, feeling that SOX flatters to deliver increased probity but in reality makes no difference and may even make the situation worse. One cryptically noted that just because there are more pages about corporate governance in an annual report does not mean that in reality there will be improved governance within a company or in general. As one put it, "SOX draws a very long bow between costs and complexity of impacts versus the problems being addressed." Several respondents felt that SOX was poorly drafted and that CLERP 9 in Australia was not only better drafted but better targeted. Another referred to SOX as "an unfinished symphony, too black-letter, in need of review in order to produce more codification, especially in the area of shareholder rights".

Conclusion

Overall, the findings from the study are that key stakeholders in Australia have taken notice of SOX and its effects in the US, but that the influence of SOX in specifically Australian contexts has been limited. The general perception in Australia seems to be that SOX has had some flaws in its inception and in its subsequent delivery in the US, but also that it has produced some positive outcomes. However, domestic factors and influences appear to be overwhelmingly more important in shaping how financial regulation and corporate

governance evolve in Australia. All respondents agreed that SOX has helped produce an increased awareness of corporate governance in general and avoidance of conflicts of interest in particular, and that these changes have been coupled with a perceived rise in regulatory activism in the broader area of contemporary corporate governance.

There were specific SOX-related beneficial effects noted by respondents. For example, the increased rigour in relation to issues such as the independence of auditors. Similarly, respondents drew comfort from what they saw as raised levels of assurance in relation to financial statements and the greater general consciousness of the importance of these issues. All the respondents agreed that listed companies take these issues seriously, as evidenced by the overall improvements in reporting levels in relation to corporate governance disclosures. Also, there was consensus amongst the respondents that the main benefit of SOX was that it focused increased attention on areas that should have been the focus of attention anyway, especially for business organisations themselves. These areas included: improved business ethics; better disclosure; clearer relationships between risk management and financial management; and increased truth in auditing. However, it is important to emphasise that several respondents pointed out that there remains a real difficulty in quantifying the benefits of the SOX reforms.

The US regulatory influence is obviously most strong upon those Australian firms who have chosen to remain in the US as Foreign Private Issuers. Nevertheless, there was widespread agreement amongst the respondents that, on balance, the level of influence that US regulatory activity has on Australian regulatory structures and processes is appropriate. Australia notes and takes into account US trends but does not slavishly take the lead from the US – the CLERP 9 process and the ASX Corporate Governance Council activities demonstrate this regulatory reality. Respondents also agreed that, given that the US is the dominant market globally, international investors expect Australia to have comparable governance standards to the US, or they would consider Australia to be a sub-par regulatory environment. So, SOX is seen as merely accelerating an ongoing process across the world which had been kick-started by the Cadbury reforms in the UK. One respondent commented specifically that some people can be too quick to say that Australia is on a US bandwagon – CLERP 9 shows that Australia has been very much following its own way in terms of corporate governance reform.

The results of the Australian case study confirm the need to take into account the importance of local cultural influences in governance which is an issue that is stressed by Kirkbride et al.[54] Similarly, Cunningham notes in what he perceives as the clumsy global reach of SOX, that national conceptions of corporate purpose differ around the world and inevitably this will shape the choices a jurisdiction makes regarding its corporate governance tools, how these tools are described and how they are applied.[55] This seems to be a sensible and legitimate summary about how legal norms may travel across national and cultural boundaries. This perspective appears to have been borne out in this Australian case study as there was universal agreement amongst the respondents that Australia does strive to achieve "global best practice" and so inevitably is influenced by regulatory developments in the US in particular, but also Europe. However, the degree of US regulatory influence was seen as largely benign and SOX was not perceived as signifying in any substantive way a regulatory hegemony emanating from the US that determines financial market regulation or the evolution of corporate governance in Australia. Different jurisdictions have different

54 J Kirkbride, S Letza, X Sun and C Smallman, "The boundaries of governance in the post-modern world" (2008) 59(2) *Northern Ireland Legal Quarterly* 161–75, at p. 174.

55 L A Cunningham, "From convergence to comity in corporate law: lessons from the inauspicious case of SOX" (2004) 1(3) *International Journal of Disclosure and Governance* 269–98, at p. 272.

perceptions about what are their respective legitimate interests and appropriate regulatory strategies. This political reality reflects in our current era of networked governance, as in previous eras, that successful regulation is at heart an issue of balance.

Appendix – Topics used to guide the semi-structured interviews

Regulatory regime change in world financial markets – Australia

Guide to semi-structured interview questions.

1. Perceptions of Sarbanes-Oxley's (SOX's) effect in the US from an Australian observer
   a. SOX has had mixed reviews in the US amongst academe, business, media and politics. One prominent academic (Romano: 2004) has labelled SOX "quack corporate governance". Do you think such a view is justified?
   b. Do you think SOX has been a positive innovation for the US or not, and in your view what should be the major criteria for evaluating SOX?
   c. Do you perceive that currently there is a backlash against SOX in the US, and if so, do you think that this is acting as a restraint on further regulatory innovation by US politicians and regulators in relation to business activity?
2. SOX & you.
   a. Has SOX had an effect on your/your organisation's activities and if so, in what way?
   b. What do you see as the most significant changes that have occurred in the regulatory regime of Australia in the period immediately coincidental with and following changes in the US regulatory regime (i.e., Sarbanes-Oxley)?
3. What domestic factors have determined and shaped regulatory regime changes in Australia relative to SOX?

   In your view to what extent were those changes attributed to local structural conditions (e.g., historical regulatory policy patterns) or situational factors (e.g., local scandals or political upheaval/turmoil)?
4. How influential have changes in the global regulatory environment been upon Australian regulatory innovations such as CLERP 9? Specifically, to what extent are the changes related to:
   a. Changes in international or regional regulatory regimes
   b. Changes in prevailing dominant corporate behaviour norms
   c. A direct and coordinated response to the implementation of US regulatory regime policies?
5. How has legal approach to, and enforcement of, corporate governance changed?
6. Do government, the "governance profession" and business have different goals?
7. Should corporate governance concern itself with stakeholder interests?

8. Is there a trade-off between shareholder and stakeholder interests?
9. Assessment of effects
    a. In summary then, what is your assessment of,
        i. firstly, the beneficial effects and
        ii. secondly, the adverse effects of the changes made by
            a. SOX
            b. the subsequent Australian reforms?
    b. Do you feel that the SOX story to date shows that the levels of influence that US regulatory activity has on Australian regulatory structures and processes are appropriate or not?

NILQ 60(4): 421–42

A hierarchy of disability rights? A comparative examination of the regulation of digital television in the United States of America and the United Kingdom

Dr Eliza Varney*

Keele University

Introduction

Information and communication technologies (ICTs) play an increasingly important role in everyday life.[1] For disabled people, technological advances are particularly important, as they can contribute towards facilitating independent living.[2] Technological developments include the proliferation of new services for accessing digital television such as audio description (video description), closed signing[3] and the availability of subtitles (captions) in live broadcasts, enabled by new speech-to-text technologies.[4] Effective access to information can be crucial in enabling participation in society as citizens.[5] Unfortunately, disabled people still face significant barriers in accessing digital television, including, among other things, the use of multiple remote controls and the provision of difficult to navigate on-screen displays.[6] Although these access barriers have the potential to affect a considerable number of disabled people, legislative responses tend to concentrate on assisting people with sensory disabilities, often overlooking the needs of people with dexterity or cognitive disabilities.[7] Whilst ensuring an effective protection for people with sensory disabilities is important, it is disappointing that people with cognitive and dexterity disabilities are not given a stronger degree of protection. It can, therefore, be argued that we are faced with a hierarchy of disability rights in the digital television sector.

---

* Lecturer in Law, Keele University. I would like to thank Mike Varney, Ruth Fletcher, Jim Chen, Ani Satz and the anonymous referees for constructive comments on an earlier draft of this article. Any errors and omissions are, of course, my own. The issues discussed in this article will be developed in a forthcoming book: *Disability and Information Technology: A comparative study in media regulation* (Cambridge: CUP).

1 Measuring Progress of eAccessibility in Europe (MeAC), *Assessment of the Status of eAccessibility in Europe: Main Report* (2007).

2 E L Myers III, "Disability and technology" (2004) 65 *Montana Law Review* 289–307, at p. 290.

3 Closed signing allows the public to switch the sign interpreter on and off. See M George and L Lennard, "Ease of use issues with domestic electronic communications equipment" (2007), para. 4.17, www.ofcom.org.uk/research/tv/reports/easeofuse/easeofuse.pdf.

4 MeAC, *Main Report*, n. 1 above, p. 49.

5 M Feintuck and M Varney, *Media Regulation, Public Interest and the Law* 2nd edn (Edinburgh: Edinburgh UP 2006), p. 250.

6 George and Lennard, "Closed signing", n. 3 above.

7 MeAC, *Main Report*, n. 1 above, p. 49.

This article aims to explore the protection of disability rights in the regulation of digital television in the United States of America and the United Kingdom, in view of assessing whether a hierarchy of disability rights exists in these jurisdictions. The study adopts a comparative approach to determine whether regulators across the Atlantic are confronted with similar challenges and whether similar solutions are adopted to address these challenges. Part one highlights the importance of ICTs in the lives of disabled people and focuses on the idea of access to information as a tool for enabling people to participate in society as citizens. This article relies on a definition of disability focused on the social barriers faced by disabled people rather than on their medical conditions.[8] Parts two and three explore the protection of disability rights in the regulation of digital television in the United States of America and the United Kingdom respectively. Part four discusses the importance of a regulatory framework for the electronic communications sector focused on safeguarding citizenship values such as equality and dignity and stresses that in a regulatory framework genuinely committed to protecting equality of citizenship, there is no room for any hierarchy in the protection of disability rights.

I Rights of access to ICTs for people with disabilities

ICTs have become central to people's lives.[9] As Goggin and Newell suggest, "technology is present in everyday life in ways that we often do not notice once we become accustomed to and reliant upon it".[10] Developments such as closed captioning, closed signing and audio description have the potential to improve the lives of disabled people and to contribute towards independent living.[11] At the same time, advances in ICTs have brought an increased level of complexity due, inter alia, to the provision of difficult to navigate on-screen displays and the use of multiple remote controls.[12] The United States Department of Justice has stressed that "information technology can empower the lives of disabled people if it is accessible, or further segregate them from mainstream if it is not".[13] A number of accessibility provisions need to be put in place to ensure that disabled people can have equal access to digital television. For example, people with visual disabilities could benefit from the provision of audio description, which consists of an additional channel that describes the video content transmitted.[14] People with hearing disabilities could benefit from the provision of sign language interpretation or from the closed captioning of digital television programming.[15] For people with dexterity disabilities, accessibility solutions could include the design of keypads with better-spaced buttons.[16] People with cognitive disabilities could benefit, for instance, from the design of easy to use remote controls and easy to navigate on-screen displays.[17]

Such measures are designed to address the barriers faced by disabled people in accessing ICTs. Equality of access to information plays an important role in enabling the participation

8 M Connolly, *Discrimination Law* (London: Sweet & Maxwell 2006), p. 299.

9 G Goggin and C Newell, *Digital Disability: The social construction of disability in new media* (Lanham, MD: Rowman & Littlefield 2003), p. 4.

10 Ibid., p. 3.

11 B M Sullenger, "Comment: Telecommuting: a reasonable accommodation under the Americans with Disabilities Act as technology advances" (2007) 19 *Regent University Law Review* 537–60, at p. 537.

12 George and Lennard "Closed signing", n. 3 above.

13 USDJ, "Section 508 of the Rehabilitation Act: Accessibility for people with disabilities in the information age" (2008).

14 MeAC, *Main Report*, n. 1 above, p. 4.

15 Ibid., p. 5.

16 Ibid., p. 6.

17 Ibid.

of disabled people in civil society as citizens.[18] Gregg stresses that "access to technology is one determinant of who can participate in the social, cultural, political and economic facets of a society",[19] while Feintuck and Varney suggest that effective access to information is "a prerequisite for any meaningful concept of citizenship".[20] It is important that regulatory measures for the digital television sector reflect a perception of disabled people not just as consumers but also as citizens. The notion of consumer reflects a narrow perception of the public as economic actors in the pursuit of self-interest, who would benefit from lower prices, increased choice and increased quality of products.[21] On the other hand, the citizenship notion is much wider and comprises democratic values such as the protection of human dignity and equality between all members of society.[22] When acting as citizens, people participate within the wider social and political sphere and tend to take into account the interests of others.[23] The citizenship values of equality and dignity should, therefore, play an important role in any regulatory framework for the digital television sector.

Measures designed to tackle access barriers in the digital television sector reflect a social definition of disability. Unlike the medical approach, which focuses on the physical or mental impairments of disabled people, the social model is concerned with the "social barriers" faced by disabled people.[24] While the medical model assumes that one of the "primary problem[s]" faced by disabled people is "[their] incapacity to participate in society",[25] the social model points towards the failure of society to construct environments which take into account disabilities.[26] The American National Council on Disability (NCD) highlights that "people with almost every type of disability encounter barriers" to accessing ICTs.[27] These access barriers are likely to affect a considerable number of people, as it is estimated that "approximately 20 per cent of all people live with a disability".[28] These barriers can be traced, inter alia, to general challenges to equal opportunity and independent living (given that, statistically, disabled people tend to have fewer resources than others and this perpetuates disparities in opportunity)[29] as well as to the poor design of ICTs.[30] Ultimately, these accessibility barriers can result in isolation[31] and a sense of dependency and frustration.[32]

Even when provisions are put into place to facilitate the access of disabled people to ICTs, these tend to reflect a hierarchy of disabilities. Regulatory efforts are mainly focused

18 European Commission, "Challenges for European information society beyond 2005", COM (2004) 757, p. 6.

19 J L Gregg, "Policy-making in the public interests: a contextual analysis in the passage of closed-captioning policy" (2006) 21(5) *Disability and Society* 537–50, at p. 537.

20 Feintuck and Varney, *Media Regulation*, n. 5 above, p. 250.

21 M Feintuck, *"The Public Interest" in Regulation* (Oxford: OUP 2004).

22 Ibid.

23 C R Sunstein, *After the Rights Revolution: Reconceiving the regulatory state* (Cambridge, Mass: Harvard UP 1990), p. 58.

24 D Chalmers et al., *European Union Law* (Cambridge: CUP 2006), p. 891.

25 R Scotch, "Models of disability and the Americans with Disabilities Act" (2000) 21 *Berkeley Journal of Employment and Labor Law* 213–22. at p. 214.

26 K Monaghan, *Equality Law* (Oxford: OUP 2007), para 1.21.

27 NCD, "When the Americans with Disabilities Act goes online: application of the ADA to the internet and the worldwide web" (2003) Position Paper.

28 Goggin and Newell, *Digital Disability*, n. 9 above, p. xv.

29 NCD, "When the ADA goes online", n. 27 above.

30 Ibid.

31 Scotch, "Models of Disability", n. 25 above, pp. 43–4.

32 European Commission, "Ageing well in the information society: an i2010 initiative", COM (2007) 332, p. 5.

towards the protection of people with sensory disabilities (e.g. visual and/or hearing), through the adoption of measures concerned with accessibility services such as subtitling, signing and audio description. At the same time, the interests of people with dexterity disabilities and cognitive disabilities, such as learning disabilities, are often overlooked in the regulation of the ICT sector. As Campbell suggests, "the disability community is not immune from the same stereotypical attitudes about different disabilities that affect the non disabled community".[33] In fact, "there is a hierarchy of disabilities within the community itself" and people with cognitive disabilities are "among the most stigmatised, even amongst those with disabilities".[34] The following two sections explore the extent to which this hierarchy of disabilities is reflected in the regulation of digital television in the USA and the UK.

II The protection of disability rights in the regulation of digital television in the United States of America

In the American system, the protection of disability rights in the digital television sector tends to concentrate on safeguarding the rights of people with sensory disabilities, often overlooking the rights of people with dexterity and cognitive disabilities. It is regrettable that people with cognitive and dexterity disabilities are not given a stronger degree of protection, especially given the commitment of broadcasters to comply with "public interest" obligations.[35] Broadcast licensees are regarded as "public trustees" and, in exchange for the right to broadcast over a channel of publicly owned radiofrequency spectrum, they are required to "broadcast in furtherance of the 'public interest, convenience and necessity'".[36] What exactly constitutes "public interest" programming was left to be defined by the regulators.[37] However, as Varona indicates, the American communications regulator, the Federal Communications Commission (FCC) has failed to put forward a coherent definition of the "public interest" in the broadcasting context.[38] In the specific "public interest" obligations imposed on broadcasters, Congress has drawn a link between these obligations and the protection of disability rights in the communications sector.[39] Yet, these provisions are limited to enhancing television access for people with sensory disabilities through the use of closed captions and other assistive technologies.[40] This limited level of protection conferred upon disabled people is disappointing, given the strong link between the "public interest" concept and the protection of citizenship rights.[41] As Feintuck suggests, the value base for interpreting the "public interest" notion should comprise the "democratic imperatives that underlie our society", including the protection of equality and human dignity.[42]

33 J Campbell, "Unintended consequences in public policy: persons with psychiatric disabilities and the Americans with Disabilities Act" (1994) 22(1) *Policy Studies Journal* 133–46, at p. 134.

34 Ibid.

35 Communications Act 1934, as amended, 47 USC, s. 301ff.

36 Ibid. See A E Varona, "Changing channels and bridging divides: the failure and redemption of American broadcast television regulation" (2004) 6(1) *Minnesota Journal of Law, Science and Technology* 1–116, at pp. 3–4.

37 A E Varona, "Out of thin air: using First Amendment public forum analysis to redeem American broadcasting regulation" (2006) 39(2) *University of Michigan Journal of Law Reform* 149–98, at p. 151.

38 Ibid.

39 47 USC, s. 613 (Supp. V. 1999). See D P Graham, "Public interest regulation in the digital age" (2003) 11 *Communications Law Conspectus* 97–144, at p. 107.

40 Ibid.

41 Feintuck, *The Public Interest*, n. 21 above, p. 28.

42 Ibid., p. 58.

The key provision safeguarding the rights of people with hearing and visual disabilities in the digital television sector is s. 713 of the Telecommunications Act 1996.[43] For people with hearing disabilities, s. 713(a–e) empowers the FCC to adopt regulations requiring closed captions for television programmes.[44] Closed captions enable people with hearing disabilities to access television programming "by displaying the audio portion of a television program as text on the television screen".[45] Under s. 79(1) of the Telecommunications Act, by the beginning of 2006, all new programming (i.e. programming broadcast on or after 1 January 1998) is expected to be broadcast with closed captions.[46] Furthermore, 75 per cent of pre-rule programming (i.e. programming broadcast before 1 January 1998) is expected to be broadcast with closed captions by the beginning of 2008.[47] Certain categories of programming (such as programming broadcast between 2am and 6am local time) are exempted from these provisions.[48] Furthermore, the FCC can waive the closed captioning requirement if this imposes an "undue burden" on the programming provider[49] (i.e. if the captioning requirement would cause "significant difficulty or expense").[50] Although there is no evidence to suggest that this provision has had a detrimental effect on the availability of closed-captioned programming,[51] the "undue burden" defence illustrates how cost-related considerations constitute potential barriers for an increased level of access for disabled people to ICTs.

Closed captions are displayed by television receivers through either the use of a set-top box (STB) decoder or an integrated decoder circuitry.[52] The integrated decoder circuitry is preferable to a separate decoder, due to the cost implications and the installation requirements of the latter.[53] In 1990, Congress adopted the Television Decoder Circuitry Act (TDCA),[54] which requires all television receivers with screens of 13 inches or larger to contain built-in decoder circuitry for the display of closed captions.[55] The TDCA has been praised for enabling millions of people with hearing disabilities to access broadcast information[56] and for acknowledging that "for many disabled people, access to technological media lies at the heart of economic advancement and social integration".[57] Organisations representing the interests of people with hearing disabilities have called for a stricter enforcement of the TDCA, arguing that captioned text is "often filled with errors and terminated before the end of the program" and that this has a negative effect on the ability of people with hearing disabilities to access information effectively.[58]

43 47 USC, s. 613 (2004).

44 Ibid. See Myers, "Disability and technology", n. 2 above, p. 291.

45 FCC, "Closed captioning" (2008) Consumer Fact Sheet, 29 October.

46 47 CFR, s. 79(1)(a) (2004).

47 Ibid.

48 Ibid., s. 79(1)(d); MeAC, *Assessment of the Status of eAccessibility in Europe: Policy inventory* (2007), p. 263.

49 47 CFR, s. 79(1)(d)(2).

50 Ibid., s. 79(1)(f)(2).

51 Myers, "Disability and technology", n. 2 above, p. 291.

52 FCC, "Closed captioning", n. 45 above.

53 Gregg, "Policy making", n. 19 above, p. 547.

54 Public Law 101–31, 104 stat. 960 (1990).

55 MeAC, *Policy Inventory*, n. 48 above, p. 264.

56 E J Markey, "Electronic oases take root in Mr Minow's vast wasteland" (2003) 55(3) *Federal Communications Law Journal* 545–52, at p. 547.

57 A M Schloss, "Web-sight for visually-disabled people: does Title III of the Americans with Disabilities Act apply to internet websites?" (2001) 35(1) *Columbia Journal of Law and Social Problems*, pp. 35–59, at p. 48.

58 Gregg, "Policy-making", n. 19 above, p. 546.

Unlike the provisions for closed captioning under s. 713(a–e) of the Telecommunications Act 1996,[59] the provisions for video description under s. 713(f–g) designed to assist people with visual disabilities, are much weaker.[60] Video description technology provides people with visual disabilities with aural descriptions of the key visual elements in a television programme.[61] While Congress required the FCC to adopt closed-captioning regulations,[62] the requirements imposed with regards to video description are limited to the preparation of a report by the FCC to Congress.[63] Despite these limitations, the FCC proceeded to adopt rules on video description, with the aim of facilitating the access of people with visual disabilities to television programming.[64] These rules required major network and cable channels in the top 25 to broadcast a minimum of four hours per week of described programming.[65] These rules were challenged by the industry (on grounds that Congress did not entrust the FCC with the adoption of video description rules) and were ultimately struck down by a ruling of the United States Court of Appeals for the District of Columbia Circuit in *Motion Picture Association of America* v *FCC*.[66]

All of these difficulties could have been avoided if s. 713 of the Telecommunications Act 1996 had been adopted in its original format, as presented in the House of Representatives, which included a provision empowering the FCC to adopt video description rules.[67] The ultimate version of s. 713 was watered down, removing this authority from the FCC.[68] It is regrettable that Congress has treated closed captioning and video description differently, providing people with visual disabilities with a lesser degree of protection than people with hearing disabilities. A Bill to reinstate the FCC requirements on video description[69] has been put forward before Congress.[70] However, even if the FCC rules on video description are ultimately reinstated, it is doubtful that four hours of described programming per week constitutes a genuine commitment to ensuring that people with visual disabilities benefit from equal access to programming. Some organisations representing the interests of people with visual disabilities have expressed opposition to this Bill, unless an amendment is introduced "to require a process that will lead to the voicing of text printed on the screen".[71] Ensuring effective access to programming for people with visual disabilities is seen as "a matter of . . . fairness", as "video description is for the blind people what closed captioning is for those who are deaf".[72] Such provisions would avoid the risk of a hierarchy of protection for people with sensory disabilities.

According to Myers, an increased level of protection for people with visual disabilities in accessing digital television could be provided if Congress amended s. 255 of the

59 47 USC, s. 613(a–e) (2004).

60 Ibid., s. 613(f–g).

61 *Motion Picture Association of America, Inc. et al.* v *Federal Communications Commission et al.* 309 F3d 796, 8 November 2002.

62 47 USC, s. 613(b–c) (2004).

63 Ibid., s. 613(f–g). See *Motion Picture Association*, n. 61 above.

64 FCC, "Implementation of video description of video programming, report and order" (2000) 15 *FCCR* 230.

65 MeAC, *Policy Inventory*, n. 48 above, p. 264.

66 *Motion Picture Association*, n. 61 above.

67 Ibid.

68 Ibid.

69 HR 951 Video Description Restoration Act 2005.

70 MeAC, *Policy Inventory*, n. 48 above, p. 264.

71 S Maneki, "Imagining a brighter future for blind Americans: a report on the 2005 convention resolutions" (2005 August/September) *The Braille Monitor*.

72 American Council of the Blind, "The American Council of the Blind answers your questions about video description and the Video Description Restoration Act".

Telecommunications Act 1996 to include video description services.[73] Section 255 requires manufacturers of telecommunications equipment and service providers to ensure that their equipment is accessible and usable by disabled people.[74] However, this requirement applies only if accessibility is "readily achievable"[75] (i.e. if accessibility can be ensured without much difficulty or expense).[76] In determining whether accessibility is "readily achievable", the FCC makes assessments on a case-by-case basis, balancing accessibility costs with the available resources of a particular market player.[77] In implementing s. 255, the FCC seems to be more favourable to industry interests, often at the expense of the interests of disabled people.[78] As Kanayama points out, the FCC regulations implementing s. 255 rely on voluntary efforts by the telecommunications industry limited to encouraging it to adopt accessibility solutions.[79] Therefore, while an extension of s. 255 to cover video description services would be beneficial in facilitating the access of people with visual disabilities to digital television,[80] such a provision will only have a genuine impact in assisting disabled people if the FCC moves away from the "pro-industry regulatory bias" affecting the telecommunications industry.[81]

The specific provisions adopted for the television sector seem to be focused on assisting people with hearing disabilities and to a lesser extent people with visual disabilities, overlooking the needs of people with dexterity and cognitive disabilities. The remainder of this section focuses on the general equality provisions, particularly the Americans with Disabilities Act (ADA),[82] assessing the extent to which these provisions assist disabled people in accessing digital television. The ADA is the main piece of legislation adopted in the United States of America for combating disability discrimination. The Act aims to guarantee equality of opportunities for disabled people in a number of sectors, including employment (Title I),[83] public services (Title II),[84] places of public accommodation (Title III)[85] and telecommunications (Title IV).[86] The ADA is designed, inter alia, "to provide a clear and comprehensive national mandate for the elimination of discrimination against individuals with disabilities" and "to provide clear, strong, consistent, enforceable standards addressing discrimination against individuals with disabilities".[87] The enactment

73 Myers, "Disability and technology", n. 2 above, p. 307.

74 47 USC, s. 255(c) (2004).

75 Ibid., s. 255(a)(2).

76 47 CFR, s. 6.3(g) (2004). See Myers, "Disability and technology", n. 2 above, p. 293.

77 Myers, "Disability and technology", n. 2 above, p. 293; FCC, "Section 255 telecommunications access for people with disabilities" (2008b) Consumer Fact Sheet, 8 October.

78 T Kanayama, "Leaving it all up to industry: people with disabilities and the Telecommunications Act of 1996" (2003) 19 *Information Society* 185–94, at p. 185.

79 Ibid.

80 Myers, "Disability and technology", n. 2 above, p. 307.

81 Kanayama, "Leaving it", n. 78 above, p. 185.

82 42 USC, s. 12101ff.

83 Ibid., ss. 12111–17.

84 Ibid., ss. 12131–65.

85 Ibid., ss. 12181–9.

86 47 USC, ss. 225 and 611.

87 42 USC, s 12101(b)(1) and (2). The provisions under the ADA are supplemented by statutes under state law. However, an examination of the provisions under state law is outside the scope of the present research.

of the ADA has led to both "physical changes" (such as accessible public facilities) and "fundamental changes in public attitudes" towards disabled people.[88]

While an earlier version of the ADA included provisions applicable to the television sector (including the development and operation under "appropriate circumstances" of assistive technologies such as captioning and audio information),[89] these provisions were subsequently left out of the final version of the Act[90] and the provisions dealing specifically with the communications sector were limited to telecommunications relay services for people with hearing and speech disabilities[91] and closed captioning on public service announcements.[92] The adoption of a watered-down version of the communications provisions in the ADA is believed to be the result of concerns that the Act would face opposition in Congress, as a result of lobbying from the broadcasting industry.[93] This is unsatisfactory, as it illustrates the influence exercised by the industry over policy making,[94] often at the expense of the interests of the public.[95] Although the ADA was ultimately complemented by the adoption of the TDCA, the inclusion of access provisions for the broadcasting sector in the ADA would have sent the message that access to this sector is an integral part of the process of combating the social barriers faced by disabled people and for promoting equality between all citizens. The incorporation of provisions regarding the television sector in the ADA would have also benefited from the social model reflected in this instrument. The ADA seeks to eliminate discrimination and to tackle the barriers faced by disabled people[96] and reflects the "emergence of a new socio-political consciousness about disability".[97] It is indeed regrettable that the provisions facilitating the access of disabled people to the television sector have been left out of the final version of the ADA, as the social model of disability reflected in this instrument could have led to a stronger commitment to tackling the barriers faced by disabled people in the communications sector.

However, even if the ADA were applicable to the digital television sector, the limitations in Title IV of the Act render questionable the effectiveness of these provisions in tackling the access barriers faced by disabled people. Unlike the provisions under Titles I–III, which are concerned with safeguarding the interests of all disabled people, the protection conferred by Title IV is confined to people with hearing and speech disabilities.[98] This limitation is unsatisfactory, given that the ADA relies on a wide definition of disability which includes physical or mental impairments, a record of such impairments and perceived disability.[99] The limited application of Title IV is in strong contrast to the objectives of the ADA, which are, among others, to ensure the "equality of opportunity, full participation,

88 J Bick, "Americans with Disabilities Act and the internet" (2000) 10(2) *Albany Law Journal of Science and Technology* 205–27, at p. 207.

89 HR 4498, s. 8(h)(3).

90 R Colker, "ADA Title III: a fragile compromise" (2000) 21 *Berkeley Journal of Employment and Labor Law* 377–412, at n. 39.

91 47 USC, s. 225. Relay services "enable hearing and speech impaired people who use devices that translate voice signals into written messages, to communicate with people who do not possess such special devices". See Schloss, "Web-sight", n. 57 above, p. 48.

92 47 USC, s. 611.

93 Gregg, "Policy-making", n. 19 above, p. 543.

94 Varona, "Changing channels", n. 36 above.

95 E Varney, "Regulating the digital television infrastructure in the EU. Room for citizenship interests?" (2006) 3(3) *SCRIPT-ed* 221–42.

96 Scotch, "Models of disability", n. above 25, p. 214.

97 J A Nelson, *The Disabled, the Media and the Information Age* (Westport: Greenwood Press 1994), at p. 4.

98 47 USC, s. 225.

99 42 USC, s. 12102.

independent living, and economic self-sufficiency" necessary to enable disabled people to access goods and services to the same extent as everyone else.[100] Finnigan et al. indicate that the provisions under Title IV are "narrowly tailored",[101] while Bowe points out that Title IV of the ADA is "the only section of the Act that protects only part, not all of the population of Americans with a disability".[102] In concentrating its protection only on people with hearing or speech disabilities, Title IV overlooks the needs of people with visual, dexterity and cognitive disabilities in accessing communication technologies.[103] Once again, the final version of the ADA represents a watered-down version of the original provisions. Rather than classifying the protection conferred upon disabled people under separate titles, previous drafts were divided into sections concentrated on prohibiting discrimination in a number of environments, including in "access to services or programs".[104] The application of a broad definition of disability to the telecommunications sector would have assisted all disabled people. From this perspective, the ADA represents a missed opportunity for safeguarding the interests of all disabled people to access effectively the communications sector.

Even if the ADA provided disabled people with a stronger degree of protection for accessing the communications sector, disabled people would still be faced with considerable challenges in their search for justice. Although the ADA reflects the social model of disability, the courts conditioned the application of the ADA on the provision of medical evidence, reinforcing the medical approach to disability.[105] According to Smith, the approach based on medical evidence reflects "a deep-seated scepticism" of disabled people who seek to enforce their rights under equality legislation.[106] For these reasons, the ADA is seen as providing disabled people with "a hope for equality and access that has not been fulfilled".[107] The courts' reliance on the medical model of disability when dealing with claims under the ADA could be due to the use of the term "impairment" in the statute.[108] As Oliver suggests, the term "impairment" concentrates on the physical or mental condition of individuals and is associated with the medical model of disability, while the term "disability" concentrates on the social barriers faced by people and is, therefore, associated with the social model of disability.[109] Smith concludes that the use of the term "impairment" in the ADA creates a tension between social and medical concepts of disability[110] and that the courts' insistence on medical evidence for individuals wishing to rely on the ADA places disabled people "in a position of being pathologized, pitied and therefore disempowered" and overlooks the perception of disabled people as citizens.[111]

100 42 USC, s. 12101(a)(8). See Bick, "Americans with Disabilities Act", n. 88 above, p. 213.

101 M O Finnigan Jr et al., "Accommodating cyberspace: application of the Americans with Disabilities Act to the internet" (2007) 75 *University of Cincinnati Law Review* 1795–826, at p. 1802.

102 F G Bowe, "Access to the information age: fundamental decisions in telecommunications policy" (1993) 21(4) *Policy Studies Journal* 765–74, at p. 766.

103 Ibid.

104 Colker, "ADA Title III", n 90 above, p. 382, referring to HR 448 and S 2345 (1988).

105 Scotch, "Models of disability", n. 25 above, p. 218. See, for example, *Sutton* v *United Air Lines, Inc.* 527 US 471 (1999) but note that the narrow definition of disability adopted in *Sutton* (which relied on the "ameliorative effect of mitigating measures" to determine "whether an impairment substantially limits a major life activity") has been rejected in the ADA Amendments Act of 2008, Public Law 110-325, 25 September 2008.

106 D M Smith, "Who says you're disabled? The role of medical evidence in the ADA definition of disability" (2007) 82(1) *Tulane Law Review* 1–76, at p. 1.

107 Ibid.

108 Ibid., p. 7.

109 M Oliver, *The Politics of Disablement* (Basingstoke: Palgrave Macmillan 1990).

110 Smith, "Who says", n. 106 above, p. 11.

111 Ibid. pp. 71–2.

While the ADA Amendments Act of 2008 relies on a broad definition of disability "constructed . . . to the maximum extent permitted by the terms of this Act",[112] this does not necessarily mean that this instrument reflects a social approach to disability, given its continued reliance on the notion of physical and mental "impairment".[113]

This section has illustrated that, despite the number of provisions adopted in order to facilitate the access of disabled people to digital television, there is still room for improvement. The measures adopted in the digital television sector tend to focus on the provision of assistance for people with sensory disabilities (particularly hearing disabilities), overlooking the needs of people with other disabilities, such as cognitive or dexterity disabilities. Furthermore, the provision of an effective level of protection for disabled people seeking to enforce their rights under the ADA is affected by the continuing reliance of the courts on the medical definition of disability, which is in strong contrast to the social focus of this Act. The approach adopted by the courts tends to focus on disabled people as disempowered victims and fails to perceive disabled people as citizens.[114] The final section of this paper calls for a regulatory approach for the communications sector which is focused on citizenship values such as equality and dignity and which aims to protect the rights of disabled people to access digital television as a citizenship right. Before that, however, the following section aims to explore the extent to which the regulatory framework for the communications sector in the UK safeguards the rights of disabled people.

III The protection of disability rights in the regulation of digital television in the United Kingdom

Similar to the American system, the approach adopted in the UK for safeguarding disability rights in the digital television sector tends to concentrate on assisting people with sensory disabilities and provides only a limited level of assistance for people with cognitive or dexterity disabilities.[115] This limited level of protection is regrettable, particularly given the general duty imposed by the Communications Act 2003[116] on the British communications regulator Ofcom to have regard to the needs of disabled people[117] when fulfilling its principal duty "to further the interests of citizens in relation to communications matters" and "to further the interests of consumers in relevant markets, where appropriate by promoting competition".[118] The key provisions for safeguarding the rights of disabled people in the digital television context are s. 303 (television services for deaf and visually impaired people), s. 310 (Code of Practice for Electronic Programme Guides (EPGs)) and s. 10 (use of apparatus with ease and without modifications) of the Communications Act. According to s. 303, Ofcom is required to adopt and maintain a code providing guidance on accessibility services for people with hearing disabilities and/or visual disabilities.[119] Such accessibility

112 Public Law 110-325, 25 September 2008. See 42 USC, s. 3(3)(A).

113 Ibid.

114 Smith, "Who says", n. 106 above, pp. 71–2.

115 For a more detailed discussion of these issues see E Varney, "Disability rights in the communications sector: an examination of digital television regulation in the United Kingdom" (2008) 13(6) *Communications Law* 187–96.

116 Communications Act 2003, c. 21.

117 Ibid., s. 3(4)(j).

118 Ibid., s. 3(1).

119 Ibid., s. 303(1).

services include subtitling, signing and audio description. In 2004, Ofcom adopted a Code on Television Access Services (subject to subsequent reviews)[120] which applies to

> licensed public service channels, [digital television] programme services, television licensable content services and restricted television services, as well as any [digital television] programme services provided by the Welsh Authority.[121]

The code also applies to the public television services provided by the BBC.[122] The Communications Act makes provision for targets with respect to subtitling, signing and audio description, which must be fulfilled from the fifth and the 10th anniversary of the "relevant date" (set in accordance to when a service commenced).[123] These targets are based on the size of the audience.[124] However, Ofcom can exempt certain programmes from the obligations to comply with these targets.[125] The 10-year targets for broadcasters are generally set at 80 per cent for subtitling, 5 per cent for signing and 10 per cent for audio description.[126] While the application of the code to both public service broadcasters and commercial broadcasters has been praised in a report on measuring eAccessibility in Europe,[127] organisations representing the interests of disabled people, such as the Royal National Institute of Blind People (RNIB), have expressed disappointment that the 10 per cent target for audio description is too low[128] and, therefore, is unlikely to have a significant impact in facilitating the access of people with visual disabilities to digital television. Similar to the American context, the provision on accessibility services for people with hearing disabilities seem much stronger than the level of assistance provided to people with visual disabilities. Despite the existence of guidelines for broadcasters on subtitling, signing and audio description,[129] the Code on Television Access Services makes no provisions for any technical standards regarding the way in which these accessibility services are made available.[130] Service providers are merely expected to use "reasonable endeavours" to ensure that the greatest number of people can access television services.[131] This seems to allow a considerable degree of flexibility in the hands of the industry. The fact that market players often lack the economic incentives to search for interoperable solutions[132] makes it unlikely that this approach is the most effective for safeguarding the interests of disabled people in the digital television sector.

Similar to the Code on Television Access Services, the Code of Practice on EPGs adopted under s. 310 of the Communications Act is also limited to providing assistance to "people with hearing and/or visual impairments".[133] This code of practice requires EPG

120 Ofcom (2006) "Television access services review", at para. 2.1, www.ofcom.org.uk/tv/ifi/codes/statement/.

121 Ofcom (2008) "Code on Television Access Services", at para. 3, www.ofcom.org.uk/tv/ifi/codes/ctas/.

122 Cl. 59, "Agreement between Her Majesty's Secretary of State for Culture, Media and Sport and the British Broadcasting Corporation", July 2006, Cm 6872.

123 Ss. 303(4) and (5) and 305, Communications Act.

124 MeAC, *Policy Inventory*, n. 48 above, p. 220.

125 S. 303(7), Communications Act.

126 Ofcom, "Code on Television", n. 121 above, para. 8.

127 MeAC, *Main Report*, n. 1 above, p. 44.

128 "Report of the Joint Committee on the Draft Communications Bill" (*Puttnam Report*), July 2002, HL Paper 169-1, HC 876-1, para. 313.

129 Ofcom, "Code on Television", n. 121 above, Annex 2.

130 Ibid, para. 31.

131 Ibid.

132 Communications Committee (2004) "Report from the Inclusive Communications Subgroup" COCOM 04-08, p. 42.

133 Ofcom (2004) "Code of Practice on EPGs", www.ofcom.org.uk/tv/ifi/codes/EPGcode/epgcode.pdf.

providers to make adjustments to their EPGs, so far as practicable, in order to ensure the access of people with sensory disabilities to these services and to promote awareness on accessibility information on the EPGs.[134] The failure to include any form of assistance for people with other disabilities, such as cognitive and dexterity disabilities, is disappointing. Organisations such as Age Concern[135] and Ricability[136] have called for a wider definition of disability in the regulation of EPGs and for addressing issues such as the use of remote controls in accessing EPGs. As Ricability points out, poorly designed remote controls can constitute a barrier for many disabled people in accessing digital television,[137] including people with dexterity disabilities.

While the Code of Practice on EPGs makes no provisions for assisting people with dexterity and cognitive disabilities, the assistance conferred to people with sensory disabilities is far from satisfactory. The requirement to provide accessibility solutions only as far as practicable is vague and seems to provide a significant degree of flexibility in the hands of the industry. Furthermore, the code states that "at present, there is limited scope to reconfigure EPGs so as to facilitate their use" by people with sensory disabilities, while the provision of accessibility solutions is reliant on "the future development of EPGs".[138] This approach is regrettable, as the code should have provided an increased level of incentives for EPG providers to seek accessibility solutions in order to assist disabled people to take advantage of these facilities.[139] When justifying its refusal to require EPG providers to take "rapid steps" in improving the accessibility of EPGs "through changes on the set-top box", Ofcom has pointed towards concerns that the hasty implementation of new features on the STB could affect the operation of EPGs.[140] It is likely that Ofcom's decision was influenced by cost-related concerns raised by EPG providers.[141] In balancing the interests of disabled people against the interests of the industry, Ofcom should have placed more emphasis on the urgency in the need for accessibility solutions for disabled people in the digital television sector.[142]

While the assistance provided by the Code on Television Access Services and the Code of Practice on EPGs is limited to facilitating the access of people with sensory disabilities to digital television, s. 10 of the Communications Act has the potential to assist all disabled people. This section requires Ofcom to encourage the development of electronic communications apparatus "which is capable of being used with ease and without modification by the widest possible range of individuals (including those with disabilities)".[143] Ofcom is also under a duty to encourage the wider availability of electronic communications apparatus.[144] As defined under s. 10(4) of the Communications Act, such

134 Ofcom, "Code of Practice", n. 133 above, para. 6.

135 Age Concern, "Response to the Ofcom Consultation on EPGs" (2004). para. 3.1, www.ofcom.org.uk/consult/condocs/epg/responses/age_concern.pdf/.

136 Ricability, "Response to the Ofcom Consultation on EPGs" (2004), para. 6.0, www.ofcom.org.uk/consult/condocs/epg/responses/ricability.pdf.

137 Ibid.

138 RNIB, "Response to the Ofcom Consultation on EPGs", (2004), p. 8, www.ofcom.org.uk/consult/condocs/epg/responses/rnib2.pdf.

139 Disability Rights Commission, "Response to the Ofcom Consultation on EPGs" (2004), p. 3, www.ofcom.org.uk/consult/condocs/epg/responses/drc.pdf.

140 Ofcom, "Statement on Code on Electronic Programme Guides" (2004) p. 18, www.ofcom.org.uk/consult/condocs/epg/statement_archived/statement.pdf.

141 Ibid. 19.

142 MeAC, *Main Report*, n. 1 above, 13.

143 S. 10(1)(a), Communications Act.

144 Ibid., s. 10(1)(b).

apparatus sends or receives information transmitted by means of an electronic communications network. The Explanatory Notes to s. 10 stresses that under this provision, Ofcom has the potential to encourage the development of easy to use remote controls.[145] This could represent a step forward in facilitating the access to digital television for people with dexterity disabilities. Nevertheless, the potential impact of this provision in assisting disabled people is limited by the fact that s. 10 merely requires Ofcom to "encourage" the development and availability of easy to use apparatus, without providing the communications regulator with formal regulatory powers to enforce such provisions.[146] Once again, the reliance on voluntary efforts by the industry overlooks the sense of urgency in developing accessibility solutions for the digital television sector, particularly in light of the switchover to digital television planned to be completed by 2012.[147]

As the specific provisions for facilitating the access of disabled people to digital television seem focused on assisting people with sensory disabilities, it would be interesting to see the extent to which people with other disabilities could rely on the general equality legislation, in order to challenge the barriers they face in accessing digital television. The Disability Discrimination Act (DDA) 1995 (as amended in 2005)[148] is the main piece of legislation adopted in the UK for combating disability discrimination. The DDA combats disability discrimination in a number of sectors, including access to information and communication services.[149] As one of the categories of service providers covered by the Act, broadcasting organisations have the responsibility to ensure that people accessing their services do not face unjustified discrimination on grounds of disability.[150] Discriminatory treatment includes the refusal to provide a service or the provision of a lower standard of service to a disabled person.[151] According to s. 21 of the DDA, service providers are placed under a duty to take reasonable steps to ensure that disabled people do not face discriminatory treatment. In complying with this duty, service providers are required, inter alia, to alter particular physical features in their services in order to enable access.[152] This provision could prove particularly useful for pursuing equality objectives in the digital television sector, for example, by requiring EPG controllers to make their services available to disabled people.[153] Nevertheless, the duty imposed on service providers by s. 21 is subject to limitations based on financial considerations.[154]

Public authorities, such as the communications regulator Ofcom, are subject to a disability equality duty (DED), requiring them to ensure that disabled people enjoy equal access to services.[155] They must be proactive in ensuring equal access to services for disabled people and must involve disabled people in the search for solutions.[156] Furthermore, public authorities are expected to publish yearly reports highlighting the key

145 Explanatory Notes to the Communications Act, para. 42.

146 Ofcom, "Usability in the communications sector" (2008), p. 3, speech by Huw Irranca-Davies MP, Minister for Digital Inclusion, www.ofcom.org.uk/research/usability/usability08/report.

147 Culture, Media and Sports Committee (CMS), "Analogue switch-off: a signal change in television" (March 2006), HC Paper 650-I.

148 Disability Discrimination Act 1995, c. 50 (as amended by the Disability Discrimination Act 2005 c. 13).

149 Ibid., s. 19(3)(b)(c).

150 Ibid., s. 19; DRC, "Revised Code of Practice on Part 3 DDA" (2006), para. 3.3.

151 Ss. 19 and 20, DDA.

152 Ibid., s. 21(2).

153 Ofcom, "Code on Television", n. 121 above, p. 35.

154 S. 21(6) and (7), DDA.

155 Ibid., s. 49A.

156 MeAC, *Policy Inventory*, n. 48 above, p. 228.

steps taken for achieving the equality objectives.[157] The DED has the potential to act as "a powerful tool" in promoting accessibility for disabled people in the communications sector.[158] As indicated in the first disability equality scheme issued by Ofcom in 2006, the issue of access to electronic communications is perceived as a priority area for promoting disability equality.[159] It is interesting to note that the Equality Bill[160] aims to introduce a streamlined equality duty placed on public authorities, covering not only the existing duties regarding disability, gender and race but extending to age, gender reassignment, religion or belief, and sexual orientation.[161] A number of organisations representing disabled people have stressed the need to ensure that this extended duty does not weaken the existing DED[162] and does not lead to the situation where public authorities "cherry-pick which activities they needed to undertake . . . leav[ing] disabled people at the bottom of the political agenda".[163]

While, at least at a theoretical level, the DDA has the potential to assist disabled people in challenging access barriers to the digital television sector, the practical benefits of this Act for the digital television sector are yet to be seen. Disabled people wishing to rely on the application of the DDA to the digital television sector are not aided by the absence of case law covering this issue.[164] Furthermore, as Carmichael et al. suggest, the provisions designed to ensure that disabled people do not face discrimination in accessing services are seen as "secondary" to the definition of circumstances which allow service providers to justify discriminatory treatment.[165] Another potential difficulty could be posed by the fact that this Act adopts a medical approach to disability. Section 1(1) of the DDA defines disability as "a physical or mental impairment which has a substantial and long-term adverse effect" on the ability of a person "to carry out normal day to day activities". Cases brought on the basis of the DDA can become focused on whether a claimant has an impairment which satisfies the definition of disability put forward by the Act, rather than on removing the social barriers to participation faced by disabled people.[166] Unfortunately, there is no indication that future legislative developments will depart from the medical model of disability in favour of the social model. The government response to the consultation on the Equality Bill (which was subsequently introduced in Parliament in the 2008–2009 session)[167] has stressed that disability discrimination law should protect "only those people who are disabled in the generally recognised sense of the term i.e. because they have a long-term or permanent impairment, with substantial adverse effect".[168] The definition of disability adopted in cl. 6 of the Equality Bill continues to place emphasis on the link between disability and impairment.[169] This continued emphasis on impairment and the

157 MeAC, *Policy Inventory*, n. 48 above, p. 228.

158 Ibid., p. 227.

159 Ofcom, "Disability equality scheme" (2006), para. 2.3, www.ofcom.org.uk/consult/condocs/des/statement/desstatement.pdf.

160 Equality Bill 2008–2009, Bill 85 08–09, introduced in the House of Commons on the 27 April 2009.

161 Cl. 143(6), Equality Bill.

162 Government Equalities Office, "The Equality Bill – government response to the consultation" (July 2008), Cm 7454, p. 18.

163 Ibid., p. 19 (referring to the submission by the Association of Disabled Professionals).

164 A Carmichael et al., "Digital switchover or digital divide: a prognosis for usable and accessible interactive digital television in the UK" (2006) 4 *Universal Access in the Information Society* 404–16, at p. 408.

165 Ibid.

166 Monaghan, *Equality Law*, n. 26 above, para. 5.76.

167 Equality Bill, n. 160 above.

168 Government Equalities Office, "The Equality Bill", n. 162 above.

169 Cl. 6(1), Equality Bill.

reliance on medical evidence for individuals wishing to rely on equality legislation before the courts reflects a perception of disabled people as disempowered victims and overlooks the perception of disabled people as citizens.[170] It is very doubtful that the emphasis of the legislative framework on individual litigants who have to prove that their "impairment" fits the "recognised" definition of disability[171] can achieve equality in practice to the same extent as the social model would. As discussed in the following section, a preferable alternative to individual litigation for tackling the barriers faced by disabled people in accessing digital television would be the provision of a regulatory framework which reflects citizenship values such as equality and dignity.

IV Lessons to be learned?

The regulatory frameworks for electronic communications in both the United States of America and the United Kingdom have already taken important steps for facilitating the access of disabled people to digital television. However, as illustrated in parts two and three of this study, more needs to be done to ensure that all disabled people enjoy equal access to digital television. In the absence of positive steps towards this aim, technological developments will have the undesirable effect that they "disempower rather than empower and to isolate rather than integrate" disabled people.[172] Advances in digital technology could, therefore, bear the risk of creating barriers for disabled people in accessing information and lead to "the creation of disability and the construction of dependency relationships".[173] This section calls for a regulatory framework focused on tackling the social barriers faced by disabled people in accessing digital television and on safeguarding citizenship values such as equality. The discussion reflects on the key challenges associated with the accessibility provisions in the USA and the UK and calls for the need to tackle the hierarchy of disabilities in the regulation of digital television, in order to ensure that all disabled people enjoy effective access to information. The analysis also stresses the importance of the social model of disability in tackling the access barriers faced by disabled people in the digital television context and calls for a regulatory framework which emphasises the rights of disabled people to access digital television as citizenship rights, and which prioritises these rights over the economic interests of the industry.

The regulatory provisions for digital television in both the USA and the UK seem focused on addressing the needs of people with sensory disabilities, overlooking the needs of people with other disabilities such as cognitive or dexterity disabilities. Whilst ensuring an effective level of protection for people with sensory disabilities is important, it is also important that the regulatory framework for digital television addresses the needs of all disabled people. It is essential that all disabled people have effective access to information, which would enable them to participate effectively in society as citizens. Unfortunately, the regulatory framework for the digital television sector is affected by a hierarchy of disabilities. Deal indicates that non-disabled people's attitudes towards disabled people have been "predominately negative" and that "behaviours appear to vary" according to the nature of disability.[174] People with cognitive disabilities such as learning disabilities

170 Smith, "Who says", n. 106 above, p. 72.

171 Monaghan, *Equality Law*, n. 26 above, para. 5.76.

172 L Johnson and E Moxon, "In whose service? Technology, care and disabled people: the case for a disability politics perspective" (1998) 13(2) *Disability and Society* 241–58, at p. 246 (quoting Oliver).

173 B Sapey, "Disablement in the informational age" (2000) 15(4) *Disability and Society* 619–36, at p. 619 (referring to works by Finkelstein and by Oliver).

174 M Deal, "Disabled people's attitudes towards other impairment groups: a hierarchy of impairments" (2003) 18(7) *Disability and Society* 897–910, at p. 898.

encounter an increased level of difficulty towards acceptance in society.[175] According to Deal, the hierarchy of disabilities is also reflected in the attitudes of disabled people, as they "do not always wish to be associated with other impairment groups for a variety of complex reasons", including stigma and the perception that people with other disabilities have got a different agenda with regards to issues such as access to services.[176] Similarly, Fox and Kim argue that certain groups within the disability population "appear to lag behind their peers in important indicators of independent living".[177] This could be due to the fact that, despite the growing perception of disability as being attributed to social barriers, learning disability continues to be associated with the medical approach.[178] Goodley stresses the need to ensure that cognitive disabilities such as learning disabilities are perceived from a social rather than a medical perspective.[179] In a study about the citizenship of people with learning disabilities, Walmsley indicates that these citizens "have often been marginalised" but that their active involvement in society is far more than just "an empty dream".[180] As Walmsley suggests,

> despite ideological and structural barriers . . . some people with learning difficulties are learning skills which enable them to participate in a meaningful way in decisions about their own lives and in activities which might be categorised as pertaining to citizenship.[181]

It is, therefore, important to ensure that their interests are effectively taken into account by any regulatory framework designed to facilitate the access of all citizens to information.

Both the United Kingdom and the United States of America need to move away from the hierarchy of disabilities in the regulation of digital television and need to ensure that accessibility solutions reflect a wide range of disabilities. The solution rests with the adoption of a regulatory framework for the digital television sector based on citizenship values such as equality and human dignity.[182] In such a regulatory framework, there is no room for a hierarchy of disabilities. As Bynoe et al. point out, in safeguarding the rights of disabled people, facilities must be organised around "a framework of enforceable citizenship rights" and emphasis must be placed on facilitating the active participation of all citizens in society.[183] The citizenship values of equality and human dignity are inherent in every human being and any legislative framework committed to safeguarding these values aims to ensure that "everybody is treated as having value or worth".[184] A formal approach to equality seeks to ensure that individuals are not treated less favourably because of a specified characteristic[185] and adopts a reactive approach to discrimination, providing individuals with the right to sue if they have been treated in a discriminatory manner.[186] On the other hand, the substantive vision of equality is concerned with tackling "systematic"

175 Deal, "Disabled people's attitudes", n. 174 above.

176 Ibid., pp. 903–4.

177 F H Fox and K Kim, "Understanding emerging disabilities" (2004) 19(4) *Disability and Society* 323–37, at p. 326.

178 D Goodley, "'Learning difficulties', the social model of disability and impairment: challenging epistemologies" (2001) 16(2) *Disability and Society* 207–31, at p. 210.

179 Ibid.

180 J Walmsley, "'Talking to top people': some issues relating to the citizenship of people with learning difficulties" (1991) 6(3) *Disability and Society* 219–31, at p. 220.

181 Ibid.

182 Feintuck, *The Public Interest*, n. 21 above.

183 J Bynoe et al., *Equal Rights for Disabled People. The case for a new law* (London: IPPR 1991), p. 3.

184 S Fredman, "Equality: a new generation?" (2001) 30(2) *Industrial Law Journal* 145–68, at p. 155.

185 Ibid., p. 154.

186 Ibid., p. 164.

forms of discrimination rooted in society and with ensuring social inclusion for under-represented groups.[187] Substantive equality adopts a proactive approach to tackling discrimination and to promoting equality, stressing the need to alter practices and structures in order to "bring about real change".[188] The concept of human dignity plays an important role in ensuring substantive equality.[189] Despite a certain degree of reluctance from some critics who stress that the term is too ambiguous, the concept of human dignity can act as a powerful tool in human rights discourse[190] and in enhancing the protection of fundamental values such as equality of citizenship.[191] This concept is often linked with the notion of "autonomy" or "the freedom of the individual to choose according to his view of the 'good life'"[192] and has been defined as "the right not to be disadvantaged or humiliated by virtue of one's subjective characteristics".[193] The adoption of a regulatory framework for the digital television context based on the notions of substantive equality and human dignity would ensure that such measures do not differentiate between the levels of protection conferred to different disabilities.

The social model of disability should prevail in the regulatory framework for dealing with disability discrimination in both the USA and the UK. Scotch warns against the risks associated with the medical approach to disability, stressing that a perception of disabled people as having "pathological attributes typically linked to incapacity and dependency" places the onus on the need for disabled people to "overcome" their disabilities in order to fit into mainstream society.[194] Rather than perceiving disabled people as citizens who face social barriers to participation, disabled people are perceived as "victims" who fail to fit into mainstream society.[195] There is a significant difference between the perception of disabled people as victims (associated with the medical model) and their perception as citizens entitled to equal treatment (associated with the social model).[196] Measures which perceive disabled people as victims tend to adopt a reactive approach to disability discrimination, relying on individual litigation whenever discrimination has occurred. Nevertheless, an approach based on individual litigation relies on action by an individual "with the energy and resources to bring a claim" and depends on the potential to trace inequalities to a particular perpetrator.[197] Furthermore, if a claimant is successful, the impact of the case is limited to providing compensation for that particular claimant, without any obligation to tackle the social barriers that gave rise to disability discrimination in the first place and which are likely to affect disabled people as a group.[198] On the other hand, measures which perceive disabled people as citizens are rooted in a framework of principles based on values, such as equality and the protection of human dignity,[199] and adopt a proactive approach to

187 Chalmers et al., *European Union Law*, n. 24 above, pp. 897 and 916.

188 Fredman, "Equality", n. 184 above, p. 163.

189 Ibid., p. 155.

190 C McCrudden (ed.), *Anti-Discrimination Law* (Dartmouth: Ashgate 2004).

191 D G Réaume, "Discrimination and dignity" (2003) 63 *Louisiana Law Review* 645–95.

192 Fredman, "Equality", n. 184 above, p. 155.

193 Z Apostolopoulou, "Equal treatment of people with disabilities in the EC: what does 'equal' mean?" (2004) Jean Monnet Working Paper 09/04.

194 Scotch, "Models of disability", n. 25 above, p. 219.

195 Gregg, "Policy-making", n. 19 above, p. 538.

196 Fredman, "Equality", n. 184 above, p. 164.

197 S Fredman, *Human Rights Transformed: Positive rights and positive duties* (Oxford: OUP 2008), p. 189.

198 Ibid., p. 190.

199 Feintuck, *The Public Interest*, n. 21 above.

achieving equality in practice.[200] Such measures focus on systematic cases of discrimination and are designed to protect the rights of groups rather than just individuals.[201] The proactive approach calls for cooperation amongst all stakeholders, including policy makers and service providers.[202] This approach advocates social change and concentrates on the removal of barriers which hinder participation in society.[203] In the digital television sector, the social model of disability provides an effective approach for tackling the access barriers faced by disabled people and calls for positive action in order to ensure that all disabled people enjoy effective access to information. This may require the need to "reshape society" in order to "include the whole spectrum of abilities"[204] but, as Gooding suggests, this is not an impossible task if it is rooted in a broad framework of rights for disabled people.[205]

The regulatory frameworks in both the USA and the UK do not place sufficient emphasis on the rights of disabled people to access digital television as citizenship rights. In fact, in the UK, the absence of "a coherent concept of the public interest" in the Communications Act 2003 has led to criticism that the British system lacks "solid foundations" for regulatory intervention in the pursuit of citizenship values.[206] While the Act imposes on Ofcom the general duty to have regard to the needs of disabled people[207] when fulfilling its principal duty to safeguard the interests of citizens and consumers in the communications sector, it is not clear from the language of the Act whether disabled people are perceived as citizens or as consumers. Yet, these interests are not always synonymous and regulatory intervention to meet consumer interests could overlook wider citizenship concerns such as the need to safeguard equality and human dignity. In a recent consultation where Ofcom aims to clarify its dual responsibility to protect citizenship and consumer interests, the key provisions designed to facilitate the access of disabled people to digital television are classified under the heading "duties to further citizenship interests".[208] Nevertheless, Ofcom stresses that citizenship and consumer interests can often overlap and that having access to services as consumers is often a prerequisite for having access to information in order to participate in society as citizens.[209] Potential difficulties are, however, likely to arise in the case of conflict between citizenship and consumer interests. At the moment, Ofcom is entrusted with "wide-ranging and largely unstructured discretion"[210] to resolve the conflict in the manner "they think best in the circumstances".[211] As Feintuck indicates, a preferable approach would have been to prioritise citizenship interests over the interests of consumers, as suggested by the Puttnam Report,[212] as failure to do so risks leaving citizenship interests "vulnerable to defeat by

200 Fredman, "Equality", n. 184 above, p. 164.

201 Ibid.

202 Fredman, *Human Rights Transformed*, n. 197 above, p. 190.

203 Fredman, "Equality", n. 184 above, p. 150.

204 C Gooding, *Disabling Laws, Enabling Acts. Disability rights in Britain and America* (London: Pluto Press 1994), p. 3.

205 Ibid., p. 30.

206 M Feintuck, "Walking the high-wire: the UK's Draft Communications Bill" (2003) *European Public Law* 9(1), pp. 105–24, at p. 107.

207 S. 3(4)(j), Communications Act 2003.

208 Ofcom, "Citizens, communications and convergence" (2008), para. 3.9, www.ofcom.org.uk/consult/condocs/citizens/discussionpaper.pdf.

209 Ibid., para. 2.22.

210 M Feintuck, "Protecting non-commodity values in 'the public interest'" in T Prosser et al. (eds), *Law, Economic Incentives and Public Service Culture*, Working Paper 05/129 (CMPO: Bristol 2005), pp. 70–80.

211 S. 3(7), Communications Act 2003.

212 *Puttnam Report*, n. 128 above.

other factors".[213] Furthermore, by prioritising citizenship interests over the interests of consumers, the Communications Act could have avoided any ambiguity as to the range of values protected in any particular provisions adopted under the Act.[214] Due to the absence in the Communications Act of any specific reference to the rights of disabled people as citizenship rights, the provisions adopted under ss. 10, 303 and 310 of the Act tend to reflect a perception of disabled people as "users" rather than citizens. Therefore, these provisions fail to emphasise the citizenship values of equality, dignity and individual autonomy as the basis for regulatory intervention designed to facilitate the access of disabled people to information. A positive step towards acknowledging the need to safeguard citizenship values in the communications sector is reflected in the recent House of Lords Select Committee report on the *Ownership of the News*,[215] which recommended in the context of media mergers that "when Ofcom considers . . . public interest . . . considerations . . . it should be required to put the needs of the citizen ahead of the needs of the consumer".[216] While this development, limited to the context of media mergers, provides a certain degree of optimism, more needs to be done in order to ensure that citizenship values are prioritised over consumer interests in all areas of the communications sector, including in the provision of access for disabled people to digital television.

The absence of a clearly defined framework of principles to assist the communications regulator in the provision of access for disabled people to digital television has also affected the American system. While the FCC is entrusted with acting in pursuit of the "public interest"[217] and licensed broadcasters are required to broadcast "in furtherance of the 'public interest'",[218] the public interest notion is ambiguous.[219] Given this ambiguity, there is a significant risk that this concept can be misused by political or commercial forces in the pursuit of self-interest.[220] According to Feintuck, this risk can be avoided in a regulatory framework that puts forward a clear definition of the public interest, reflecting democratic values such as equality of citizenship.[221] Furthermore, the citizenship values inherent in the public interest concept can play an important role in interpreting regulatory provisions in the communications sector, which deal with the issue of access for disabled people.

The regulatory frameworks in both the USA and the UK reveal a tendency of the regulators to side with industry interests, often at the expense of those of citizens. This tendency is reflected, for example, in the watered-down measures adopted for the provision of access for disabled people to digital television. These provisions are likely to be the result of financial considerations put forward by the industry with regards to the costs of accessibility provisions. Walker and Townsend suggest that the integration of disabled people is likely to face not only "environmental and psychological barriers" but also considerable financial barriers.[222] Nevertheless, as Shakespeare indicates, society must be prepared to make additional investments in order to address the barriers faced by disabled people.[223]

213 Feintuck, "Protecting non-commodity values", n. 210 above.

214 Feintuck, "Walking the high-wire", n. 206 above.

215 House of Lords, Select Committee on Communications, report on the *Ownership of the News*, June 2008, HL Paper 122-I.

216 Ibid., para. 275.

217 Feintuck, "Walking the high-wire", n. 206 above, p. 122.

218 47 USC, s. 301ff. See Varona, "Changing channels", n. 36 above, p. 4.

219 D Freedman, *The Politics of Media Policy* (Cambridge: Polity Press 2008), at p. 64.

220 Feintuck, *The Public Interest*, n. 21 above, p. 28.

221 Ibid., p. 58.

222 A Walker and P Townsend, *Disability in Britain: A manifesto of rights* (Oxford: Martin Robertson 1981), p. 16.

223 T Shakespeare, *Disability Rights and Wrongs* (London: Routledge 2006), p. 66.

There is a danger that citizenship interests are overlooked in favour of the economic interests of the industry. This could be due to strong lobbying from the industry, influencing policy making. According to Varona, the influence exercised by the broadcasting industry in the United States of America is manifested in relation to both the FCC and Congress, arguing that the FCC has been "captured" by industry interests and that Congress is "so beholden to broadcast interests that its link to broadcasters has been characterised as that of an 'umbilical cord'".[224] As the legislative history of the Telecommunications Act and of the ADA illustrates, while Congress had the chance to go beyond closed captions in order to authorise the FCC to pursue the interests of all disabled people in accessing digital television, it chose not to do so. The influence exercised by commercial players over Congress and the FCC renders the chances of adopting "heightened 'public interest' requirements" for the communications sector which would benefit, disabled people, and others, to be "remote at best".[225] Similar concerns regarding the influence of the industry over the regulatory realm have been expressed in relation to the UK,[226] illustrated, for example, by the deregulatory nature of the Communications Act and the "unwillingness" of policy makers "to take tough decisions"[227] in favour of citizenship interests.[228] There is a significant discrepancy between the influence exercised over the regulatory realm by commercial players on the one hand and the public on the other. Kanayama stresses that "typically, the regulated business and industry groups have the loudest voices among the interest groups".[229] Industry groups are well organised and benefit from sufficient resources which enable them to influence policy making.[230] On the other hand, citizenship voices are much weaker, due, in part, to high participation costs, such as legal fees, limited resources and lack of expertise in technical proceedings.[231]

The regulatory framework for the communications sector needs to facilitate participation in policy making from groups representing the interests of disabled people. Campbell highlights the importance of ensuring that disabled people are powerful enough to influence their own future,[232] while Evans indicates that "those who know best the needs of disabled people and how to meet those needs are disabled people themselves".[233] Furthermore, Oliver acknowledges that new technologies have the potential either to empower or confine disabled people and that the disability movement can play a crucial role in ensuring that disabled people are empowered by technological advances.[234] The adoption of the TDCA in the United States of America represents a success story for the involvement of disability groups in policy making in the communications sector.[235]

224 Varona, "Changing channels", n. 36 above, pp. 82 and 84.

225 Ibid., p. 91.

226 E Varney, "Winners and losers in the communications sector: an examination of digital television in the United Kingdom" (2005) 6(2) *Minnesota Journal of Law, Science and Technology* 645–85.

227 T Crane, "OFCOM – a new order for communications regulation or a bureaucratic nightmare" (2003) 9(2) *Computer and Telecommunications Law Review* 37–40, at p. 37.

228 Feintuck, "Walking the high-wire", n. 206 above, pp. 107–08.

229 Kanayama, "Leaving it", n. 78 above, p. 186.

230 Ibid.

231 Ibid.

232 J Campbell, "Valuing diversity: the disability agenda – we've only just begun" (2002) 17(4) *Disability and Society* 471–8, at p. 472.

233 J Evans, "How disabled people are excluded from independent living" (2002) presentation for Madrid Conference on European Disabled People, 21 March.

234 Oliver, *Politics of Disablement*, n. 109 above. See Johnson and Moxon, "In whose service?", n. 172 above, p. 255.

235 Bowe, "Access", n. 102 above, p. 766.

Organisations representing people with hearing disabilities played a very active role in the debate about closed captioning, advancing their "direct experience" in using this technology and resisting calls from the industry for the imposition of a lower level of obligations on television manufacturers.[236] Unfortunately, the success story for social forces in the adoption of the TDCA represents the exception rather than the rule.

In order to ensure that the citizenship interests of disabled people are not sacrificed in favour of the economic interests in the industry, it is important that the regulatory framework for the communications sector facilitates the involvement of disabled people in policy making. Positive steps towards this aim have already been taken in the United Kingdom in the Communications Act. For example, in issuing and reviewing the Code on Television Access Services, s. 304 of the Act requires Ofcom to consult, among others, persons representing the interests of people with sensory disabilities. It is, however, necessary that the consultation process has a real impact and that the views of disabled people are adequately taken into account in policy making for the electronic communications sector. It is also important to ensure that the consultation process involves organisations representing the needs of all disabled people, including people with cognitive and dexterity disabilities.

The regulatory framework should also ensure that citizenship rights are prioritised over the interests of the industry in policy making. The success of such provisions may depend on broader social changes. As Gooding suggests, we should "not ask too much of individual pieces of legislation", given that a genuine commitment to equality can be achieved only within a "context of broader economic and social change".[237] Similarly, Kanayama notes that the promotion of accessibility solutions for disabled people may require a change in people's perceptions of these issues, as a strong level of public support for accessibility solutions is an effective way to "overcome organised opposition like industry groups".[238] A regulatory framework which ensures that disabled people have a strong voice in policy making is an important step towards achieving these objectives.[239]

V Conclusion

The regulatory frameworks for electronic communications in both the USA and the UK reflect a hierarchy of protection for disability rights of access to digital television. The key accessibility provisions adopted for the digital television sector tend to concentrate on assisting people with sensory disabilities, failing to address the needs of people with other disabilities, such as dexterity or cognitive disabilities. Whist ensuring an effective degree of protection for people with sensory disabilities is important, it is disappointing that the interests of people with other disabilities are overlooked, especially given the growing importance that ICTs play in people's everyday life.[240] This article has stressed the need for a regulatory framework for the electronic communications sector based on democratic values such as equality of citizenship,[241] arguing that under such a framework there is no room for a hierarchy of disability rights. Such a regulatory framework would also perceive disabled people as empowered citizens rather than vulnerable consumers and would adopt a proactive approach for tackling the barriers faced by disabled people in accessing digital

236 Gregg, "Policy-making", n. 19 above, p. 545.

237 Gooding, *Disabling Laws*, n. 204 above, p. 29.

238 Kanayama, "Leaving it", n. 78 above, p. 193.

239 C Barnes and M Oliver, "Disability rights: rhetoric and reality in the United Kingdom" (1995) 10(1) *Disability and Society*, 111–16.

240 Goggin and Newell, *Digital Disability*, n. 9 above, p. 4.

241 Feintuck, *The Public Interest*, n. 21 above.

television, reflecting therefore a social approach to disability. Despite the positive measures adopted in both the USA and the UK towards ensuring an inclusive society, the current regulatory frameworks in both jurisdictions do not go far enough in safeguarding the rights of disabled people to access digital television as citizenship rights. In fact, both jurisdictions reflect a considerable degree of influence exercised by the telecommunications industry over the regulatory realm, often at the expense of citizenship interests. In order to redress the balance in favour of citizenship rights and to ensure that the interests of disabled people are effectively protected in the electronic communications sector, it is essential that the regulatory framework for this sector facilitates the involvement of disabled people in policy making and prioritises citizenship rights over the economic interests of the industry.

NILQ 60(4): 443–70

Keeping bad company: building societies – a case study

Lorraine Talbot

University of Warwick

Introduction

The title, "Keeping bad company", is intended to convey a number of assertions made throughout this paper. When building societies converted from mutual building societies to public companies, the erstwhile building societies became bad companies. They were bad because, necessarily, they followed the corporate model dictated by the neo-liberal ideology of the time. So, to put it another way, by incorporating, they fell into bad company. And when you keep bad company, you pick up bad habits.

To substantiate these assertions, I compare two organisational forms – or at any rate, two ideological perspectives on organisational form. These are the management controlled organisation (MCO); and the shareholder value organisation (SVO). I will attempt to assess their relative strengths and to find in that contrast an explanation for the fact that converted building societies, as SVOs, have been part of – indeed central to – the current financial crisis, whereas their MCO counterparts in what one can still call the mutual *movement* have survived largely unscathed. I will draw upon diverse disciplines in order to make this argument, including institutional economics, organisational sociology, labour process theory, political economics and the law.

I begin by examining how the original small terminating societies were usurped by larger permanent societies and the economic developments which led societies to become MCOs.[1] I go on to assess MCO theory and the normative implications of this approach which is both supportive and critical of the MCO. Critics cite MCOs' hierarchical and oligarchic nature while proponents of MCOs characterise them as a rational mechanism for enhancing production and ensuring economic stability. I then show how theory critical of MCOs dovetailed with the neo-liberal economic and political project, resulting in the reform of building society law. This legislative reform and ideological shift resulted in many societies becoming market organisations; companies governed by shareholder value concerns. I then assess the problems with shareholder value and SVOs and argue

1 It can also be noted here that, given the one-member-one-vote basis of member democracy in mutual building societies, the managerial model better describes building societies than it does the large US public companies it was mainly developed in respect of. This is because, when voting entitlement is based on share ownership, there is always a possibility that wealthy investors, acting individually or as a group, may exercise control over their company's management, whereas there is no such possibility in a mutual building society.

that they enslave governance to the capital markets with destructive consequences. Finally, I compare the approach of MCO building societies to the current crisis with that of SVO converted societies.

Early mutual building societies and the emergence of the management controlled building society

The first recorded building societies were formed in the last quarter of the 18th century, in newly industrialising cities,[2] by people whom Hobsbawn might describe as "labour aristocracy"; artisans, highly skilled and relatively well-paid workers. These early mutual organisations had a small membership which saved collectively, making monthly contributions at a set amount with the aim of acquiring privately owned property and escaping high rents. The usual system involved saving sufficient funds to purchase one dwelling place which could be allocated to one member through a lottery. Members would continue to save until all the membership had purchased a property after which the society would come to end. Because of the finite nature of their operation, these societies were known as "terminating societies".

Early societies operated for over 60 years without any legislation specifically designed for their usage. This left their internal organisation to be prescribed by the members rather than by legislation. Yet, despite the absence of coordination between societies or cohering regulation, most building societies organised themselves in a fairly similar manner.[3] There were minor differences in the cost of subscriptions, methods of allocating property and penalties for defaulting members. However, all societies operated under a system whereby members received benefits upon making contributions. That is, all members were both borrowers and savers; beginning as savers, once allocated property they became borrowers until all the members had been allocated property. Likewise, in terms of governance, all members had a single vote and decisions in respect of the society were made collectively. Indeed, this governance approach was noted in an early case which turned on the very issue of member democracy.[4]

After legislation was eventually passed in the form of the Benefit Building Society Act 1836 building societies rapidly grew in both number and size. For 70 years building societies had been small in size. Their membership never exceeded 80, and there were less than 100 known societies.[5] By 1850, over 2000 societies had registered under the 1836 Act, some with thousands of members.[6] In the context of the large societies at least, this meant that the traditional terminating societies had become increasingly inappropriate, designed as they were for a small, non-fluctuating group of saver/borrower members.

2 For example, the first known building society was formed in Birmingham in 1775 and took its name from the local hostelry in which it was formed, Ketleys: E J Cleary, *The Building Society Movement* (London: Elek Books Ltd 1965), p. 11.

3 Ibid., p. 13.

4 *Pratt* v *Hutchinson* (1812) 15 East KB 511. In this case, the guarantor of a member who had failed to keep up his payments sought to avoid liability on the basis that building societies were illegal organisations under the Bubble Act 1720. Building societies contravened the Act, he argued, because, they allocated shares to their members through subscription (shares that could be transferred in many cases) and they did so without obtaining a charter or an Act of Parliament. The court did not agree and instead found that membership-based restrictions on the transferability of shares meant that building societies fell outside of the ambit of the Bubble Act. This Act only applied to organisations which sold "freely" transferable shares and building societies were member-based democracies because in building societies members voted on share transfers.

5 Cleary, *Building Society Movement*, n. 2 above, p. 11.

6 A Scratchley, *Industrial Investment and Emigration* (West Strand, London: John W Parker 1858, 1st edn 1849), p. 50.

The organisational form which became effective in these larger societies was the permanent form. In permanent societies, members would have an individual contract with the society designating them as either investors receiving interest or borrowers paying interest. Investors would be able to withdraw their investment with relative ease, while borrowers would make periodic repayments (usually monthly) over a fixed period of years and would only withdraw their membership after repaying the agreed amount. In this way the society could extend its borrowing according to the amount invested in it and borrowers could join without catching up on the payments made by founder members. In 1849, actuary and key player in the building society movement, Arthur Scratchley published a book describing the permanent system.[7] Later when the Royal Commission on Friendly Societies set up in 1870 focused on Building Societies, it recommended that the permanent form replace the terminating form.[8] The resulting 1874 Building Societies Act therefore required that all new societies be registered as permanents, allowing the terminating societies to complete their lifecycle and gracefully depart the scene.[9] So, although in 1874 most of the estimated 1500 societies were terminating societies,[10] time would soon recalibrate this.

So the 1874 Act provided a framework for the regulation of building societies as distinct financial institutions and to facilitate the monitoring of their activities whilst allowing them certain freedoms in their internal organisation.[11] It was intended to maintain building societies as socially useful organisations which facilitated the respectable activities of the "industrious classes".[12]

However, whilst the legislative framework tightly controlled building societies' business activities, it could not control the governance issues that arose as a result of their growth. The emergence of the permanent system facilitating a large membership resulted in a dilution of member control and the emergence of management control. The one-member-one-vote approach to governance which reflected the equal nature of members' financial interests and commitment in small terminating societies, continued to operate in the permanent society. When permanent societies grew in size, the huge numbers of members they represented were unable to exercise control over their management. In societies which had a large membership, voting on a one-member-one-vote basis naturally dispersed power. By the 1930s, almost all societies had a large membership which left management in effective control. Successive annual reports from the Registrar of Friendly Societies show the huge rate of merger activity which further dispersed member control.[13] In 1890 there

7 Cleary attributes this innovation to actuary James Henry James whose pamphlet on the subject was published in 1845. "Benefit Building Societies, their Advantages if Legitimately Constituted, the defects of the principles generally Applied to their establishment, with Suggestions for their Improvement and their Profitable Union with Life Assurance Societies practically considered": Cleary, *Building Society Movement*, n. 2 above, p. 47.

8 Royal Commission on Friendly Societies, *First Report from the Commissioners on Friendly Societies* (1870). It thought it probable that permanent societies grew organically from the tendency of larger terminating societies to incorporate members throughout their existence, creating indefinitely existing terminating societies so that "a permanent society is a terminating society to every individual from the date at which he enters": p. 14.

9 The Building Societies Act 1874, s. 9.

10 Cleary, *Building Society Movement*, p. 119, n. 2 above.

11 The Act established liability for officers of registered societies who allowed loans and deposits which exceeded amounts prescribed by the Act and made them personally liable for the amount received or loaned in excess of this sum. Furthermore, the falsification of accounts could result in their summary conviction upon complaint by the Registrar of Friendly Societies. The Benefit Building Societies Act 1836 established a regulator, the Certifying Barrister for England, Scotland and Ireland who was renamed the Registrar by the Friendly Societies Act 1846.

12 A phrase used in the Act and by the Royal Commission.

13 Reports of the Chief Registrar of Friendly Societies of the Building Society Commission of each year noted.

were 2286 registered societies with 600,000 members but by 1920 there were only 1271 societies. In 1950 there were only 819 societies with a membership of 1.5m and by 1970 the number of societies had nearly halved (to 481) although their assets and membership had increased tenfold (to £10.8bn and 10.9m members). By 1988, just before the first demutualisation there were 188 societies holding assets of £188.844bn and having 48.1m members. Today, most of the major societies have converted, leaving just 59 societies which hold total assets of £369bn.[14] The Building Society Association (BSA) reported over 23m investing members and over 2.9m borrowing members in 2007.[15]

Another key factor leading to the rise of management control was the sheer complexity presented by the permanent system. Arthur Scratchley had argued vigorously for a trained management in large permanent societies whose growth necessitated professional managers and risk assessors drawn from the educated middle class.[16] This clearly raised management costs in large permanents in contrast to the small terminating societies whose management was largely voluntary. As statistics from Cleary demonstrate, terminating societies, whose management was drawn from the common membership, had low management costs whereas the large permanents had management costs which were twice as much on average.[17]

This layer of professionals became increasingly oligarchic as societies merged into ever larger permanent entities. This expressed itself in the rise of a trade association for building societies which became ever more proactive in building society policy. This organisation, originally called the Building Societies Protection Association was established in 1869 to inform and advise members on such issues as economic and legislative changes. It campaigned for legislative reform and was instrumental in the passage of the 1874 Act. Its journal, *The Building Society Gazette*, began publication in 1869 and provided a central, popular forum and mode of communication to those in the building society movement.

The Building Societies Protection Association (which later became the Building Society Association (BSA)) gradually took a much more interventionist approach to the coordination of building societies and policy responses to crises. During the 1930s' economic depression the lack of demand for money led to lower and lower interest rates and thus increased competition between building societies for borrowing members on whom they relied in order to pay high interest rates to lending members. A market-based resolution to this would have inevitably resulted in the economic ruin of many of the less competitive societies and so the BSA took action to protect its members as a collective. The solution was to annihilate competition through a policy of "recommended interest rates". The rate recommended by the BSA would be the one to which all societies, outside the few that were not members, would be obliged to adhere. As late as 1983, a member who did not

14 Data up to August 2008 at www.bsa.org.uk/keystats/index.htm (last visited 28 November 2008).

15 www.bsa.org.uk/publications/newsbite/june_08.htm (last visited 28 November 2008).

16 Scratchley, *Industrial Investment*, n. 6 above.

17 Cleary, *Building Society Movement*, p. 69, n. 2 above.

adhere to the advised rates,[18] New Cross Building Society, was forced to close following a decision by the Chief Registrar of Friendly Societies.[19]

The oligarchy in control of the BSA was concentrated on the largest societies as those societies became dominant in the building society sector as a whole. In 1983, for example, the BSA's policy-making body, the council, consisted of representatives drawn from the 10 largest societies. Moreover, the chairman of the BSA was generally drawn from the largest building society in the context of the five largest societies holding 50 per cent of all the capital of the entire industry and the two largest holding 33 per cent.[20] Thus, a large part of managerial decision making for all societies was removed from the executives of individual societies and centralised under the BSA, in an elite management group.

From the 1930s through to the 1970s, scholarship and government policy was orientated around MCOs, of which building societies were a classic example. Some of this scholarship provides support for the MCO, which has a great deal of resonance for building societies. However, much of it is highly critical of MCOs and as we shall see these theories laid the foundation for the subsequent neo-liberal approaches in both scholarship and policy that transformed most large building societies into SVOs.

Theory: the management controlled organisation

1 Origins

The notion of management controlled organisations is generally attributed to Berle and Means' famous thesis in which they argued that the largest corporations in the United Sates were controlled by a salaried non-stock-holding professional management.[21] Furthermore, they argued, the tendency for corporations to grow meant that the managerial control model would soon subsume other forms of ownership and control.[22] In these corporations the owners or stockholders were too widely dispersed and held too small a proportion of corporate stock to be able to assert control over the managers. In Marxist terms, stockholders were money capitalists, contributing funds in order to expand them, not entrepreneurs whose claim on profits was based on their entrepreneurial activity. In this

18 The BSA replaced the "recommended system" with one of advised rates although its affect was much the same.

19 New Cross's chief executive embarked on some radical departures from the standard practices "advised" by the BSA. It operated with a small personnel, few branches, it raised three-quarters of its funds through a mortgage marketing team and by insurance salesmen rather than through its branches. Most significantly, in March 1983 it put up interest rates by 2 per cent for its larger and longer-term investors and extended quick higher-interest loans. Clearly, these two policies ran contrary to the BSA's advised rate. Furthermore, New Cross's customers tended to be of a wealthier variety and not the traditional middle-income family favoured by building societies. The BSA contacted the auditors of New Cross, Dearden Farrow, and advised them to scrutinise more carefully the 1982/83 accounts. On doing so they found certain essentially technical discrepancies relating to "special advances" and, without offering the society the opportunity to redress the problem, the chief registrar issued orders to close the New Cross: C Wolman, "Downfall of a heretic: behind the New Cross closure", *Financial Times*, 21 January 1984, p. 24.

20 P Barnes, *The Myth of Mutuality* (London: Pluto Press 1984).

21 A A Berle and G Means, *The Modern Corporation and Private Property* rev'd edn (New York: Harcourt, Brace and World 1967, 1st edn 1932), p. 108. They further noted that in those corporations categorised as management controlled the dominant shareholding was never over 5 per cent.

22 Ibid., at p. 109. Berle and Means calculated, or in their words "carefully guessed" that 21 per cent of corporations were controlled by legal device, such as pyramiding, and 23 per cent by minority control when assessed as a percentage of corporations, but these figures dropped to 22 per cent and 14 per cent respectively when assessed by wealth.

respect, therefore, shareholders were no different from *rentiers*.[23] Berle and Means viewed the corporation as a productive entity which is not controlled by shareholders or primarily motivated by their self-interest. The normative implications of this radical thesis, therefore, have been vigorously debated.

II Normative implications of the separation thesis: positive

Berle's own initial position was that the law should vigorously protect shareholders against potential abuses of power by managers.[24] In particular, he argued, the law should impose strict fiduciary standards on directors in respect of issuing stock, declaring or withholding stock, acquiring stock in other corporations, powers in respect of mergers and powers to amend the corporate charter. In short, the governance implication of this thesis, as far as Berle was concerned, was the rigorous construction and enforcement of directors' fiduciary duties. In a later revised version of the modern corporation, Berle adopted a more identifiably institutional economic position which emphasised the shift of the corporation away from profit expansion goals into an entity no longer primarily subject to market forces. Accordingly, the corporation was no longer comprehensible according to classical economics. It was not driven by the self-interest of individual entrepreneurs making favourable bargains in a free market. Instead, the corporation was driven by managerial goals such as power, prestige and the personal job security of the manager. In particular these organisations were concerned with growth rather than profit. In the 1967 edition of *Modern Corporation* he argued that the corporation was so socialised into a plurality of goals and the state so involved in its operations that the corporation could not sensibly be understood as a private institution. The profit motive was now just one of the goals of the quasi-public corporation. In MCOs, managers were the risk takers, replacing the risk-taking role of the entrepreneur. In an entrepreneur controlled business the risk taker could personally benefit from risk taking, however, in the MCO there were no such benefits for the manager. Indeed, any loss resulting from risk taking would merely jeopardise the manager's professional position. Because of this, Berle argued, in an MCO, high profit is replaced by stability and good labour standards as the primary managerial goals.[25]

Following a similar perspective, Galbraith described the waning importance of shareholders in the governance of corporations.[26] Like Berle he argued that there was little point in managers pursuing risky high returns when they would suffer greatly for failures but would gain little from success. Galbraith argued that a manager's self-interest was served by steering a steady course for the corporation. Managers would pursue a minimum not a maximum profit and they would seek goals such as sales growth and prestige. Absent powerful shareholders, the large corporation had become subservient to the state, hence to society's needs. The state, thereby, had replaced market relations and steady growth, rather than high returns for shareholders, characterised modern capitalism.

Many theorists considered the MCO to be a progressive development replacing the less efficient market-based entrepreneurial capitalism. In his classic text *The Visible Hand* Alfred Chandler argued that MCOs were a historical response to the problems of entrepreneurial

23 P Ireland, "Defending the rentier: corporate theory and the reprivatisation of the public company", in J Parkinson, A Gamble and G Kelly (eds), *The Political Economy of the Company* (Oxford: Hart 2001); R Pollin, "Resurrection of the rentier" (2007) *New Left Review* 46.

24 A Berle, "Corporate powers as powers in trust" (1930–1) 44 HLR 1049.

25 Berle and Means, *Modern Corporation*, n. 21 above.

26 J K Galbraith, *The New Industrial State* (New York: Houghton Mifflin Co. 1967).

capitalism which additionally enhanced efficiency.[27] His history of American business shows how the free market was replaced by connected industries drawn together in huge management-led organisations. This shift originated with the construction of the railroads where logistical difficulties and high capital requirements necessitated the emergence of a trained managerial elite to oversee this massive infrastructural development. Many of the methods developed in that context were then used in an industrial context, particularly when industry faced recession and financial crisis. Managing that crisis became the job of trained professionals who gradually came to usurp the control of the original entrepreneurs and their families. For example, the "inventory crisis" of 1920–22 forced companies to write down the value of their stock and led to many companies establishing sophisticated management structures in order to anticipate problems and synchronise activities.[28] At General Motors, they developed a multidivisional structure where "autonomous divisions continued to integrate production and distribution by coordinating flows from suppliers to consumers in different, clearly defined markets".[29] These divisions were headed by middle managers, with top management positions taking overarching strategic decisions on their financial and market performance. Recession, argues Chandler, "transformed General Motors from an entrepreneurial to a managerial enterprise".[30] The management structure developed in General Motors was adopted by other large industrial organisation in the 1920s and 1930s.

Chandler also shows that the necessity for trained management was anticipated and catered for by the rapid emergence of professional bodies, journals and business schools at the beginning of the 20th century.[31] When Harvard opened its Graduate School of Business Administration in 1908 (nearly 20 years before the crash which initiated Berle and Means' thesis), it did so with a clear emphasis on training managers for large business entities.[32] The "professionalisation of managers" was itself a business.

Chandler's history shows how management enterprises continued to grow between the two world wars as large firms expanded by "adding new units and by internalizing their activities and the market transactions involved".[33] These large industrial enterprises encompassed research and development activities and diversified into many new product lines, often because of mergers. By 1929 over two-thirds of the personnel employed in industrial research were concentrated in five groups (of industries).[34] In retail, the chain store developed, with enterprises often producing their own brands because in-house production reduced the expense of transacting outside of the business; size meant efficiency. Indeed, argues Chandler, even when share dispersal did not lead to management control, so efficient was this model that family-owned businesses would choose to put their business under professional management. For example, he notes, in retail, internally generated profits allowed expansion to occur without recourse to bank loans or dispersed ownership so that, mass retailers could "remain entrepreneurial enterprises much longer

27 A D Chandler, *The Visible Hand: The managerial revolution in American business* (Cambridge, Mass. and London: Belknap Press of Harvard UP 2002, 1st edn 1977).

28 At General Motors, for example, the write-down between 1921–22 was over $83m: ibid., p. 457.

29 Ibid.

30 Ibid., p. 459.

31 Ibid., p. 465.

32 Ibid., p. 467.

33 Ibid., p. 469.

34 Ibid., p. 474.

than did the integrated industrials".[35] However, whilst they could continue to operate as owner-controlled enterprises they did not choose to do so. In Chandler's words,

> by 1917 representatives of an entrepreneurial family … almost never took part in middle management decision on prices, output, deliveries, wages, and employment required in the coordinating of currents flows. Even in top management decisions concerning the allocation of resources, their power remained essentially negative.[36]

Indeed, he argues, unless a family member was actually trained as a professional manager they would not take part in management.

Likewise, in the banking sector, efficiency was the driving force behind the emergence of large banks which usurped the numerous local banks which characterised the 19th century. By

> . . . 1919, 464 banks operated 1,082 branches, and by 1929, 816 had 3,603 branches. The share of bank resources held by the multi-unit enterprises rose from 16 per cent in 1919 to 46 per cent in 1929.[37]

So, by the 1930s, the multi-unit management-led organisation dominated. This is particularly significant for our understanding of building societies in England. Although they were not the fully commercial organisations that banks were, they experienced similar shifts in forms. The number of building societies decreased, the number of branches increased as did the number of investors and the resources held. Efficiency is claimed by Chandler as the source of these developments, though building societies were accused of the opposite.[38]

Whilst Chandler's account is a history of business, its theoretical integrity and its underlying notion of efficiency was posited on the earlier work of Ronald Coase in *The Nature of the Firm.*[39] Just a few years after the publication of *Modern Corporation*, Coase's seminal piece sought to explain why commercial organisations or "firms" existed and why they had become the massive entities that characterised American capitalism. Coase's central thesis here was that hierarchical firms exist because they are more economically efficient than discrete market transactions undertaken by small entrepreneurs. Coase argued that markets are the most efficient way to transact *if* the cost of transacting is nil. However, as these costs are in fact high (including the cost of information, contracting and enforcement of contracts) it becomes more efficient to transact within the firm. Price mechanisms directed resource allocation outside the firm and in the market, whilst in the firm "market transactions are eliminated"[40] and replaced by "the entrepreneurial co-ordinator, who directs production".[41] He concluded that firms would emerge in an exchange economy and when the cost of negotiating the price of each transaction in the market was more costly than if those transactions took place within an organisation under the guidance of the skilled and risk assuming entrepreneur.[42]

35 Chandler, *Visible Hand*, n. 27 above, p. 471.

36 Ibid., p. 491.

37 Ibid., p. 472.

38 Specifically, in the *Wilson Report* examined later.

39 R Coase, *The Nature of the Firm* (1937) reproduced in *The Nature of the Firm, Origins, Evolution, and Development* (Oxford: OUP 1993).

40 Ibid., p. 24.

41 Ibid.

42 This is particularly the case when considering on-going transactions which are service related. That is, if a person's skill and labour is required on a regular basis it is much cheaper to have that person employed within the firm than to enter into numerous contracts with him or her on the market.

Accordingly, to Coase, the size of a business organisation would be determined by the relative effectiveness of these two strategies, firm versus market. Where it was more efficient to organise in-house and expand into many different products or services, the size of a firm would increase. Where it was more efficient to transact on the market, firms would be smaller. Firms might also increase in size under a competent entrepreneur who had possession of good information. However, if the entrepreneur's subjective understanding of existing information was poor, or objectively there was an information deficit, market transacting might become more efficient given that it is the best provider of information (although as noted above there are costs involved). Correspondingly, improvements in managerial technique would tend to increase the size of firms by improving the entrepreneurs' ability to accumulate good information.

Coase's subsequent attractiveness to law and economics theorists is his insistence, in contrast to the managerial theories which dominated from the early 1930s up until the end of the 1970s, on the retention of market forces.[43] Coase did not accept the underlying premise of managerialism that the market was subverted by the emergence of the large corporation so that competitive forces no longer determined economic relations. Instead, he described a refinement of competitive forces in which the firm could contract (reduce in size) or even disappear if ordinary market forces proved more efficient. In accordance with this theory, an organisation is large when and because it is more efficient to be so. A Panglossian conclusion perhaps, but one which could be used to defend the emergence of large building societies as an efficient outcome.

Similarly, for Chandler, the MCO emerged only partly because of the shifts in ownership described by Berle and Means which diluted large stock ownership and thus owner control. Chandler maintains that this dilution chiefly arises as an effect of the rational growth of enterprises, where, following Coase's theory, economic efficiency is better achieved by internalising previously individualised market transactions. The demands of a complicated, dynamic and evolving economy with correspondingly complicated business enterprises, necessitated the rise of professional management, rather than entrepreneurial management. Trained management overseeing multi-unit enterprises was key to financial success by promoting efficient methods of production and by preparing for and responding to economic crises. It is a perspective with which Arthur Scratchley would have been very comfortable. Likewise, Chandler's claim that professional managers constituted a coherent group joined by shared knowledge and training, attendance at business schools and engagement with similar literature, echoes the orientation of the BSA.

Chandler's detailed analysis of a history of business which develops in response to social, political and economic change from small concerns into a rational oligopoly managed by a professional oligarchy echoes developments in the building society movement. He demonstrates that the emergence of these forms is progressive and desirable because it allows business better to achieve its goals (production in the case of industry, mortgage lending and savings in the case of building societies) whilst simultaneously protecting these businesses from crises.

Similarly, Marxists Sweezy and Baran, though highly critical of large corporations generally, argued that their management maintained such a high degree of group coherence that it was able to operate as a group and manage financial crises. Capitalism had developed into "monopoly capitalism" where effective control of business resided within a

43 Coase's theory does not expressly associate efficiency with shareholder value and is equally likely to emphasise production, distribution, growth and product development. Furthermore, it does not promote shareholder value.

coordinated small elite.[44] Indeed, they argued, the dominance of the MCO meant that financial crises could become a thing of the past, a creature of competitive entrepreneurial capitalism, not of today's professionally driven monopoly capitalism. Thus, if we accept the argument that with the emergence of large corporations emerges a small group of elites who have various and often informal mechanisms which allow them to work as a coordinated self-supporting group and thus avoid financial crises, we must accept the following. The correlating developments in the building society movement coupled with the formal institutionalisation of an oligarchy under the auspices of the BSA must be at least as equally adept at avoiding crises.

Likewise, the meritocratic dominance of the professional manager operating in a management bureaucracy must, according to Chandler, be more efficient than power which derives from ownership alone. He applauds an oligarchic economy as one which exceeds the efficiency of pure market transactions. In this respect he reflects the views of Weber's earlier work which emphasised the superior nature of the bureaucratic organisation in all aspects of the social world. Bureaucracy, for Chandler, as it was for early Weberianism, is a neutral good.[45]

The efficiency of large organisations resulting from such factors as trained management, rational hierarchies, rational diffusion of information and so on are evident in building societies to an even greater degree than large corporations. Building society policy has been coordinated by a small group of professionals who have reached the top of their profession by becoming the top managers of the largest societies. The BSA runs annual conferences, coordinates society meetings and has an informative website among many other cohering activities. It liaises with the building society regulator and it has been historically influential in building society legislative reform.[46]

III Normative implications of the separation thesis: negative

For many theorists, the MCO perpetuated and assisted the "sectional interests" of management.[47] It was not justifiable on the grounds of efficiency or on the grounds of effective and just allocation of resources. A small number of MCO-premised critics argued from a Marxist perspective. In *Monopoly Capital* (1976), Baran and Sweezy acceded to the existence of a managerial revolution but did so within a class-based Marxist framework. They argued that although management organisation dominated capitalism in the United States at least, management is neither benign nor neutral. Instead, they argued, management exists to represent a capitalist class interest and will do so with the same degree of commitment as the capitalist class had previously done for itself. They also noted the existence of large investors represented on the board who may directly intervene in the pursuit of their class interest. As previously noted, Baran and Sweezy also viewed monopoly capitalism as an economic form which is capable of transcending the economic cycles and crises which characterise entrepreneurial capitalism. Monopoly capitalism manages crises because of management's expertise, experience, and planning and because management forms a self-perpetuating oligarchy that may act in sufficient concert to avoid severe

44 S Sweezy and P Baran, *Monopoly Capital* (New York: Monthly Review Press 1966).

45 W Weber, *The Protestant Ethic and the Spirit of Capitalism* (New York: Courier Dover Publications 2003, 1st edn 1905).

46 Furthermore, having interviewed many building society executives and attended a number of conferences, I can personally testify to their social homogeny – they are invariably white, middle-aged (plus) and male.

47 "Sectional managerialists", as Nichols termed those who accepted the existence of MCOs but were highly critical of them, theorised from a number of different perspectives: N Nichols, *Ownership Control and Ideology* (London: Allen and Unwin 1969).

economic collapse. Monopoly capitalism, in other words, sidesteps the destructive nature of the competitive entrepreneurial capitalism which preceded it. In the final analysis though, it is not a benign form of organisation.

Weber's later work bristles with criticisms of the large bureaucratic organisations whose structures, hierarchies and rules are determined by the internal logic of the organisation rather than any substantive rationality. For Weber, large organisations are malignant bureaucracies which infect all areas of society.[48] Similarly, Harry Braverman demonstrated the destructive nature of managerial power upon labour.[49] Specifically, he argued that the scientific managerialism promoted by "taylorism" enhanced the social division of labour in these respects. It separated labour from skills, promoting monotonous low-skill repetitive tasks; it separated manual from mental labour; and finally management took possession of information of the whole process thereby separating labour from knowledge. The purpose of the MCO therefore was to secure absolute power for the manager by a conscious disempowering of other actors within the organisation.

Some years before Braverman's work, Burnham too had argued that managers' expropriation of knowledge and skills and their development of bureaucratic organisations had simultaneously deskilled labour and disempowered investors.[50] According to Burnham, the objective of management in developing these systems is to secure its position by ensuring that investors or other employees cannot unravel its complicated organisational structures. Managers had, thereby, secured a "managerial revolution" wherein their control over society's economic resources had resulted in them becoming the new ruling class. The era of bureaucratic organisations had therefore succeeded both capitalism and socialism. Unlike the former, it did not exist to enrich capitalists nor, like the latter, did it exist to redistribute society's resources. Instead, its rationale was the protection and perpetuation of management itself. All other outcomes were secondary to this goal.

Organisational sociologist Charles Perrow also voiced criticism of the large organisation, much as Louis Brandeis has some decades before, essentially on the grounds of bigness.[51] According to Perrow, networks of small organisations bring greater economic and social benefits than large organisations which either intentionally or by dint of their operation (size) affect society in socially undesirable ways. To briefly characterise his argument, the absence of big public governmental organisations in America (in contrast to European countries such as France and Germany) meant that the state was not equipped to counter the unabated growth of big economic, and private, organisations. In contrast, "big government" European countries where the state is much more interventionist are characterised by smaller economic organisations which are traditionally understood to be "closely held" by family groups. This echoes much of Roe's work on the political determinates of corporate governance.[52] Like the later Weber, Perrow viewed the large organisation as being a power in itself, of having a "life of its own".

> Bureaucratic organizations are the most effective means of unobtrusive control human society has produced, and once large bureaucracies are loosed upon the world, much of what we think of as causal in shaping our society – class, politics,

48 W Weber, *The Theory of Social and Economic Organisation* (New York: Free Press 1947).

49 H Braverman, *Labour and Monopoly Capital* (New York: Monthly Review Press 1974).

50 J Burnham, *A Managerial Revolution: What is happening in the world?* (New York: John Day Co. 1941).

51 C Perrow, *Organising America: Wealth power, and the origins of corporate capitalism* (New Jersey: Princetown UP 2002).

52 M Roe, *The Political Determinates of Corporate Governance* (Oxford: OUP 2003).

> religion, socialisation and self conceptions, technology, entrepreneurship – became to some degree, shaped by organisations.[53]

Accordingly, the large organisation creates a number of negative effects on society. Among these he cited wage dependency, urban crowding, socialisation of members into the organisation's norms and the concentration of power in the economy as a whole. Furthermore, he argued, large organisations are slow to respond to economic change as they cannot, "engage in smooth production" which results in "boom-and-bust cycles and layoffs".[54] In contrast, an economy with many small organisations will disperse both wealth and power. And, if these small organisations operate cooperatively, they will be productive and responsive to economic shifts.

Within the broad spectrum of MCO theorists, it is those on the left that are most critical of its operations. An oligopoly is variously criticised for being socially oppressive, determined by its own (irrational) internal logic, inefficient (on that account and on others) and perpetuating the interests of its managers. The normative implications of these MCO theorists informed the criticisms of building societies as MCOs. However, the criticisms might not have developed into policy were it not for the emergence of a critique based on the very different premises of neo-liberalism.

In the next section I will assess the MCO-based criticisms of building societies and show that they dovetailed with the neo-liberal approach of the Thatcher government which legislatively facilitated demutualisation. In so doing it will assess the neo-liberal theories which informed this pro-market approach as against the organisational approach. I will conclude this section by looking at some of the problems of a neo-liberal shareholder value approach.

Building society reform, neo-liberalism and shareholder value

I The drive for building society reform

In response to mounting criticisms of the building societies, such as the lack of accountability of building society managers and the huge influence asserted by these large organisations operating as a connected group under the BSA, the *Wilson Report* made investigations and recommendations.[55] The report noted that building societies had engaged in a number of policy decisions for reasons of stability or even ideology but that these had had the effect of crushing competition and institutionalising discrimination. For example, it noted that as a bulwark against fluctuating interest rates the BSA had required its member societies to increase their quantities of reserves over and above that required by statute and had sustained the policy of recommended rates since 1939. The BSA's strict prudential policies, it maintained, "discriminate against those who find it difficult to raise the necessary deposit", and, "because of their tradition of encouraging self-help and thrift, they have frowned upon the concept of 'low start' mortgages".[56] The *Wilson Report* recommended abandoning the policy of protecting weaker societies through the BSA's recommended rates system (set at a rate all societies could afford) and subjecting all societies to a competitive market as a way of opening up the market to more borrowers. Furthermore, the loosening of such anti-competitive practices would facilitate

53 Roe, *Political Determinates*, n. 52 above, p. 3.

54 Perrow, *Organising America*, n. 51 above, p. 14. It is an exact contradiction to the efficiency argument of Chandler, Sweezy and Baran.

55 Committee to Review the Functioning of Financial Institutions, *Wilson Report* (HMSO: London 1980) Cmnd 7937. The committee, which started work in 1977, presented its report to Parliament in June 1980.

56 Ibid., p. 85, para. 289.

the kind of innovation which could turn a profit from the traditionally less desirable customers noted above. Competitive market practice would thereby succeed the oligarchic practices of the BSA.

> The only sure way of providing a competitive spur to building societies is in our view to end the recommended rate system, that is, to allow societies to set their own rates according to their circumstances and to break the present automatic link between the rates paid to depositors and those charged to borrowers.[57]
>
> . . . The dismantling of the recommended rate system would give building societies greater freedom to compete with other deposit taking institutions as well as with themselves.[58]

Drawing on the *Wilson Report* and reflecting the broad spread of critics of the MCO, Paul Barnes argued that building societies were inefficient at everything bar the provision of security for their managers. Building society managers were concerned to pursue managerial goals such as personal prestige, job security, status symbols and empire building. Utilising Williamson's "expense preferences" theory, Barnes argued that in the absence of the restraining influence of the profit required by both shareholders and the financial market, managers will make spending decisions in ways that are most satisfying to themselves. When the profit goal is low priority, the preferred expenses are more likely to be "the expansion of staff, the expansion of physical plant and equipment".[59]

Barnes argued that the power, prestige and status which building society managers pursue in their "expenditure preferences" often expresses itself in the empire building activity of opening branches.[60] Citing statistics correlated by Davies and Davies,[61] he shows the huge increase in branch openings from 1970–78, arguing that this is an example of profit wasting in favour of managerial preferences.[62] Additionally, he noted, whilst in 1978 smaller building societies (with a total of £3711m-worth of assets in 1978) on average spent 76.7 pence per £100 of assets on operational costs, the operational costs of medium-sized companies (with a total of £14,278m-worth of assets in 1978) on average were 100.8 pence per £100.[63]

II The rise of neo-liberalism and the making of the Building Society Act 1986

As we have seen, the model for organisations until the later 1970s was that of the MCO. Management activity was orientated toward developing products within a hierarchical structure and gaining market position and (although a lesser concern) profits through the competitiveness of these products. Other managerial goals were cited as remuneration, job security and empire building. As we have observed, MCO theories covered a number of perspectives but what connected these diverse perspectives was an agreement that the MCO operated as a hierarchy; an institutional mechanism which ordered power within itself and

57 *Wilson Report*, n. 55 above, p. 113, para. 380.

58 Ibid., para. 384.

59 Barnes, *Myth of Mutuality*, n. 20 above, p. 44.

60 The *Wilson Report* showed that whilst in 1968 there were 1662 building society branches by 1978 this number had risen to 4595 whilst the average number of accounts per branch had only risen from 4921 to 5440 in the same period: *Wilson Report*, n. 55 above, p. 114.

61 G Davies and M J Davies, *Building Societies and their Branches: A regional economic survey* (London: Franey & Co. 1981), pp. 172–3. Cited in Barnes, *Myth*, n. 20 above, p. 50.

62 In 1970, the top five building societies had 746 branches but by 1978 this had risen to 1799. In medium-sized societies (those with assets over £100m), they began with 1041 branches and ended eight years later with 2310 whilst small societies began with 229 and ended with 486 in the same period.

63 Barnes, *Myth*, n. 20 above, p. 50.

which benefited people according to their role within it. Broadly speaking, the MCO was conceived as an organisation which was at least partly public in character and therefore justifiably regulated, at least to some degree, by the state. After recession in the 1970s, neo-liberalism spread throughout the academy and into mainstream political thought, in which context it was variously termed Thatcherism or Reaganomics. Neo-liberal ideology rejected notions of institutional and non-competitive practices and instead promoted the free market in which private property owners could contract as the single answer to economic prosperity. Within the academy this approach went under the banner of "law and economics". In the context of business organisations, law and economics theory cited its roots in the earlier works of Coase.

As noted earlier, Coase probably resonated for the law and economics constituency because he retained market forces in his theory even though he argued that large firms existed because they were better than the "invisible hand" (later followed up by Chandler with his conceit of the "visible hand"). Furthermore, Coase described relationships and resource allocation in contractual terms, whether inside or outside of the firm. Thus, in the law and economic reconfiguration of Coase, the firm could be conceived as operating as an *internalised* market.

Problematically, for neo-liberal theorists, Coase's theory envisaged the firm as dependent on the existence of the entrepreneur overseeing an organisation that was hierarchical in nature. Coase's theory promoted hierarchy, in contrast to "law and economics" contractualism and had thereby given the firm texture and validity as an organisation. The firm was an organisation because it was a structure in which people identified each other by their given roles.

It was these attributes of Coase's theory which the law and economics theorists sought to flush out. A hierarchical organisation necessarily ran counter to the normative values of contract such as freedom, equality and choice. It implied authority, discipline, inequality and involuntarily assumed tasks. Thus, when Alchian and Demsetz published the first *law and economics*-based article on the nature of the business organisation, they sought to deal with the unpalatable characteristics of Coase's hierarchical firm whilst retaining enough of the original theory that would bequeath their own with some heritage and authority. Tackling the hierarchical issue head on they argued:

> It is common to see the firm characterised by the power to settle issues by fiat, by authority, or by disciplinary action superior to that found on the conventional market. This is delusion. The firm does not own all its inputs. It has no power of fiat, no authority, no disciplinary action any different in the slightest degree from ordinary market contracting between any two people.[64]

In their reconstruction of Coase's theory all relationships in production are non-hierarchical and contractual. The firm is a "team productive process"[65] whereby "resource owners increase productivity through cooperative specialization".[66] The undeniable existence of managers is explained as a mechanism to ensure that the efficiency of all members of the team can be monitored and ensured. The team activity which constitutes the firm can obfuscate individual contributions or lack of them. This creates what they call a "metering problem" which is not present in individual entrepreneurial activity or ordinary market exchange. Proper metering, a mechanism to reward success and monitor shirking,

64 A A Alchian and H Demsetz, "Production, information costs, and economic organisation" (1972) *American Economic Review* 777, at p. 777.

65 Ibid., p. 778.

66 Ibid., p. 777.

an inevitable result of team activity, is undertaken by the monitor, the manager. The issue of who monitors the monitor is resolved in such a way as to justify shareholder claims to profits and other legal privileges. Shareholders monitor the monitor and the reward for their diligence is high dividends.

Having dealt a blow to the notion of hierarchies, Meckling and Jenson continued the project of annihilating the notion of institutions per se. In their argument the firm is a legal fiction. It is not an entity in its own right but a nexus of contracting individuals.[67] The firm is one form of legal fiction among many but is uniquely characterised by *some* contractors, the residual risk takers (shareholders) having residual claims on the assets and by those same individuals having the right unilaterally to end the contract, unencumbered by the views of other contractors. Accordingly the corporation should not be seen as an organisation formed around an entrepreneur. Indeed, "it makes little or no sense to try to distinguish those things that are 'inside' the firm (or any other organisation) from those things that are outside of it". Economic life was organised around contracts which were generally complete and fair arrangements. However, in the context of the contract between managers and owners the divergent interest between the two meant that costs would have to be assumed in order to bridge those divergences. These costs are famously known as agency costs and underpin modern corporate governance. In Jenson's analysis these costs include the cost of contracting, the cost of monitoring by the principal (the shareholder), such as rules to control the agent (the manager), the cost of the agent's bonding expenditures, such as incentives to profit maximise, for example, share options, the cost of residual losses and finally the cost of divergences that cannot be bridged.

Thus, in their analysis, the manager is the representative of the shareholder, engaging in activates and decisions which entrepreneurial shareholders *would* be doing had they not had the foresight to employ somebody to do these things on their behalf. However, by giving the agents the power to represent the principal there are issues that arise from the agent's possession of that power – these are agency issues. An empowered agent may find that his or her self-interest is better pursued by engaging in activities which do not represent the principal's interest. Thus, a key part of this ideological approach is to establish mechanisms which make it in the agent's self-interest to pursue the principal's best interest. As the shareholder/principal's interest is in dividends and share price, mechanisms are needed to make the agent's self-interest aligned to this. Accordingly, the mechanisms are share options and the "market for corporate control". Offering large share options to an agent means that as the agent will benefit from high dividends through his or her own shareholding it is in the agent's interest to make the pursuit of dividends and share price a managerial goal. The "market for corporate control" is the proposition that managers who fail to enhance share value, will find their reputation marred on the employment market, and, if shares fall in price, find that their corporation will be subject to hostile takeover bids which would see them out of a job.[68]

The normative implications of the contractualist approach in relation to the business organisation is brought out clearly in Easterbrook and Fischel's work in which they argue that not only is the corporation a nexus of contracts (in which agent/principal relations

67 M C Jenson and W H Meckling, "Theory of the firm: managerial behaviour, agency costs and ownership structure" (1976) 3 *Journal of Financial Economics* 305. Douglass North argues that all institutions are "contractual arrangement between principals": "Transaction costs, institutions and economic history" in E G Furubotn and R Richter (eds), *The New Institutional Economics* (Tübingen: J C B Mohr (Paul Siebeck) 1991).

68 H Manne, "Mergers and the market for corporate control" (1965) *Journal of Political Economy* 73.

are worked out) but, by so being, it is the most efficient mechanism for organising business.[69] Business is more profitable under this model, and jurisdictions which facilitate this model by reducing state inference into the business corporation will have the most successful economies.[70]

Neo-liberalism therefore conceives of MCOs as hierarchical organisations operating as oligopolies which have a negative impact on profitability because they reduce the positive affect of market forces such as free access to information. The normative implication of this conception is to reduce state interference in terms of regulation in business organisations so that individuals can negotiate freely and the invisible hand can function. By subjecting managers to a contractual arrangement with shareholders to enhance shareholder value, efficiency is enhanced and the economy can flourish. The agency problem inherent in this arrangement can be managed by offering share options and other incentives to managers. It can also be managed by subjecting managers to market forces, making their achievements visible through the mechanism of share price.[71] Thus, business organisations from all sectors would vastly improve their performance if subjected to markets forces and competition.

This new ideology impacted at every level of public life and in the context of building societies, whose oligarchic practices were already under great criticism, change was inevitable. The BSA's own *Spalding Report* recommended that societies should be able to convert into companies.[72] The Conservative government, with the support of actively anti-building society politicians on the left such as Ken Weetch, entered into a process of reform that would bring societies under the same shareholder value/agency costs ideology as corporations.[73] The proposals for building society reform were set out in a Green Paper, *Building Societies: A new framework*, which was presented to Parliament in July 1984. Its stated purpose was:

> To ensure that the building societies continue primarily in their traditional roles – holding people's savings securely and lending for house purchase – while loosening the legal restraints under which they have operated for a century or more so that they can develop in other fields.[74]

The package of proposals to liberalise building societies involved first allowing societies to develop business in other fields.[75] The Green Paper proposed that building societies

69 F Easterbrook and D Fischel, *The Economic Structure of Corporate Law* (Cambridge, Mass: Harvard UP 1998, 1st edn 1991).

70 Thus, in the context of corporate law in the United States they are one of the chief proponents of the "race to the top" theory.

71 M Jenson and R Ruback, "The market for corporate control: the scientific evidence" (1983) 11 *Journal of Financial Economics* 5–50.

72 "The future constitution and powers of building societies", the *Spalding Report* (chair John Spalding) (London: BSA 1983).

73 In the House of Commons oral answer session of 22 May 1980, Nigel Lawson was required to answer to politicians from right, left and centre (Budgen, Weetch and Grimond) on issues relating to building society inefficiency and oligarchic practices. *Hansard*, oral answer session, 22 May 1980.

74 *Building Societies: A new framework* (London: HMSO 1984) Cmnd 93L6, p. 2.

75 The introduction of market criteria had already begun through other legislation, which homogenised rules relating to taxation of banks and building societies. Tax-paying investors in building societies had previously paid a lower rate because they were paid interest net of a composite rate of tax. This privilege, the report stated, was already now extended to bank investors (from April 1985). In addition, the 1984 budget had already bought corporate tax down to 35 per cent for both banks and building societies, the previous position being corporate tax at 52 per cent for banks and 40 per cent for building societies.

should be relieved from their strict business objects.[76] It suggested that the business purposes section should now say that, "the *primary* purpose of a building society is to raise funds from individual members for lending on security by mortgage of owner-occupied residential property".[77] Other purposes would include the provision of other financial services such as unsecured loans and overdrafts.

The Green Paper also sought to reduce the oligarchy of the BSA thereby introducing more competitive practices. It noted that, in the previous year, the BSA had announced new arrangements whereby the "recommended" rates would be replaced by "advised" rates for ordinary shares and mortgage loans, in response to the criticism of the *Wilson Report*. However, the Green Paper sought to build on that by removing building societies' exemption from the Restrictive Trade Practices Act 1976. This would allow the Director General of Fair Trading to legally challenge even the "advisory" function of the BSA on the basis that it was contrary to the public interest. A successful challenge on these grounds would result in building societies setting their own interest rates without reference to the BSA, entirely subjecting rates to market criteria and, in the Green Paper's view, providing "a greater range of choice and a better service to building society members".[78]

However, the main liberalising proposal was that the conversion of building societies from mutual societies to companies should be possible.[79] If societies chose the conversion route, the new company would require a licence from the Bank of England under the Banking Act 1979. And, as a company, it would be subject to the Companies Act 1985.[80] Its directors would owe a fiduciary duty to the company which under case law (and now statute) equates with the interests of its shareholders, that is, profit.[81]

The Green Paper proposed that those societies which remained mutual would enjoy some financial liberalisations and could deal in a number of financial products. But their activity would be much more restricted than their corporate counterparts. Thus, the Green Paper proposed that, in addition to raising funds from investing members, societies could also raise moneys from non-retail sources, but such funding would be limited to 20 per cent of building society liabilities.

The central proposal to allow for conversion was duly enacted in the Building Societies Act 1986.[82]

---

76 *Building Societies*, n. 74 above, para 2.05. S. 1 of the Building Societies Act 1962, as previous Acts, restricted building society activities to "raising, by the subscriptions of the members, a stock or fund for making advances to members out of the funds of the society upon security by way of mortgage of freehold or leasehold".

77 Building Societies Act 1962, s. 1.

78 *Building Societies*, n. 74 above, para. 6.09.

79 Although, interestingly, it said that the government "does not accept that such a change is needed": ibid., para. 5.29.

80 Now the Companies Act 2006.

81 *Percival* v *Wright* [1902] 2 Ch 421; Companies Act 2006 s. 172.

82 Ss 97–103 in conjunction with Scheds 2 and 17 of the 1986 Act provide for the power of members to convert to plc status from mutual status. S. 97(1) states that: "A building society in accordance with this section and other applicable provisions of this Act may transfer the whole of its business to a company." As a constraint on conversion the Act provides that there must be an investors' special resolution and a borrowing members' resolution. In order to convert, 20 per cent of investing members entitled to vote must vote and, of that, 75 per cent of investors and 50 per cent of borrowers entitled to vote must vote in favour of conversion. When the transfer is to an existing company, as opposed to one specially created for the purpose, then the percentage is 50 per cent of those eligible to vote on the special resolution, or the holders of 90 per cent of share capital must vote in favour.

Since 1989, the vast majority of large building societies have converted, taking over 80 per cent of building society assets out of the mutual sector. As companies, the converted societies were subject to shareholder value ideology as the neo-liberal hegemony caused a rapid shift in management orientation from a more complicated palate of possibilities, such as the stakeholder approach, to a single goal; that of shareholder interests alone. Furthermore, company law reform proposed by both the Law Commission and the Company Law Review Steering Group proceeded on the premise of shareholder value, which is evident in the resulting legislation.[83] Thus, prevailing ideology, legislative reform and corporate governance orientation meant that building societies were converting in a context whereby they would become pure SVOs.

III The problem with SVOs

There is strong circumstantial evidence to support the argument set out here that SVOs tend to enhance economic instability. First, the SVO enslaves managers to the capital market and, secondly, SVOs seem to encourage economic bubbles.

In the SVO the manager is judged by share value alone and he or she is incentivised to pursue profits with little regard to risk and total regard for the share market. Enslaved to the capital market the manager's decisions are designed around its pacification. As even prime SVO proponent Jenson recently admitted, managers will manipulate outcomes to fit target-based corporate budgeting systems by such methods as delaying planned research and development until the next quarter, purely to appease the capital market.[84] The reason for doing this, he argued, was because the market heavily penalised companies which did not reach their targets by just a fraction but highly rewarded those which exceeded expectation by just a little.

The manager's tunnel-vision pursuit of shareholder value is further exacerbated by the rise of institutional investors as a constituent part of total investment. The role of institutional investors in such activities as monitoring corporate agents, has elicited mixed views.[85] However, there is general agreement on a number of features of the institutional investor. Most agree that institutional investors have limited expertise in the business in which they have invested funds and will have little interest in the operation of that business.[86] Furthermore, they are only interested in the businesses in which they have invested as profit maximisers. Institutions will pass their diverse portfolio of shares to their fund managers who will be rewarded according to the achievement of high, short-term returns.[87] This creates its own dynamic. Where institutional investors dominate ownership of shares in companies and their portfolio is managed by fund managers, managers of companies must undertake strategies which appease these interests. In contrast, it has been observed that managers of companies in Western Europe, characterised by insider ownership, will have diverse considerations when making business decisions.[88]

83 L E Talbot, *Critical Company Law* (London: Routledge-Cavendish 2008), ch. 4.

84 M Jenson, (2005) "Agency costs of overvalued equity" (2005 spring) *Financial Management* 5–19.

85 For an interesting appraisal of this material, see J C Coffee Jr, "Regulating the market for corporate control: a critical assessment of the tender offer's role in corporate governance" (1984) 84 *Columbia Law Review* 1145. The institutional investor formed a key part of the government's thinking on "enlightened shareholder value" (which described responsible share ownership) in which the former act as good monitors with positions on the board of directors.

86 L Roach, "CEOs, chairmen and fat cats: the institutions are watching you" (2006) *Company Lawyer* 27.

87 J Froud, *Financialisation and Strategy* (London: Routledge 2006).

88 H B Hansmann and R R Kraakman assert that this diversity (inter alia) accounts for the inferiority of these models compared with SVO models: "The end of history for corporate law" (2001) 98 *Geo LJ* 439.

The second and perhaps most important problem with SVOs is that the ideology of shareholder value assures investors that their interests are *the* corporate goal. An ideology of corporate governance which promotes shareholder value is therefore likely to promote an expectation that shareholders will receive high benefits from their shares. This expectation is reflected in share price, but what if those expectations are unrealistic, inflating the price and inadequately valuing the real economic strength of a company?

According to neo-liberal governance theorists, this couldn't happen. Shareholder value governance (which characterises Anglo-American governance) has been popularly claimed to be the most effective model in rewarding shareholders.[89] Part of the reason for the purported success of the SVO in achieving shareholder value is the effectiveness of the capital markets in communicating information about a company's economic performance through share price. Stock which carries more risk, that is, where the return is uncertain,[90] has the potential to achieve high returns – those that are substantially over the risk-free option of bank deposits or government bonds. There is, in other words, a premium for taking an investment with less certain outcomes. Correspondingly, stock in companies with steady returns will receive returns over the risk-free option, which are less than the risk-full stocks. It is uncertainty of returns coupled with the possibility of high returns and the prevailing rate of interest which determines share price. In law and economics' "efficient market theory", it is the ability of markets to reflect this information which makes markets efficient.[91] Markets, accordingly, are the best mechanisms for reflecting information.

An important assumption in this approach is that if shares are accurately priced, or at least as accurately as they can be, there is little opportunity to "buck the market". Risky investment may give high rewards but then may follow up with losses. Low-return, low-risk investments are less likely to surprise. So, as one of the earliest theorists on share prices and pioneer of financial economics Louis Bachelier illustrated in his 1900 doctoral dissertation, *Theorie de la Spéculation*, when calculated over a long period of time, the returns on investments are fairly even across the market.

However, in the recent period there appears to be a huge disjuncture between expectation and performance. Although dividends over the last 20 years on average only modestly exceed the prevailing rate of interest (at best), share prices have soared. Expectations on share performance seem to exceed actual performance so that the market does *not* seem efficiently to reflect all the available information on share values.[92] Statistics offered by Froud show that shares in large companies listed in the FTSE 100 and the S&P 500 consistently provide shareholder returns on the basis of share price rather than dividends.[93] On average, dividends make up around 25 per cent of total shareholder returns which on average are around 20 per cent. This makes the rate of profit, (the profit made from production) at a rough calculation, around 4–5 per cent. To take a typical example, from the years 1983–2002 companies listed in S&P 500 gave shareholders a healthy 20 per cent return on their investment on average. However, of this, 70.1 per cent was derived from its share price and only 29.9 per cent from dividends.[94]

If the value of shares are calculable according to the expected level of profitability (factoring in risk) above the prevailing rate of interest, the rise in share prices noted above

89 Hansmann and Kraakman, "End of history", n. 88 above.

90 The Sharpe-Lintner-Black model or capital asset pricing model.

91 According to proponents such as Jenson, Fama, Easterbrook and Fischel.

92 The view taken by Doug Henwood in *Wall Street* (London: Verso 1998).

93 Froud, *Financialisation*, n. 87 above, p. 78.

94 Ibid.

should be reflecting profitability (because in a shareholder value system profit should be immediately translated into dividend). To give an example, if expected dividends are 10 per cent when the prevailing rate of interest is 5 per cent, share price should double. However if dividends are, as above, around 4–5 per cent when the prevailing rate of interest is just below that, the rise in share price should be much more modest than it has been.

Work by Blair and Schary also backs the view that the profits from production are modest.[95] In critiquing Jenson's free cash-flow theory, in which he argues that takeovers, leveraged buy-outs and mergers are market controls over managers who hoard cash that should be returned to shareholders, they contend that such hoarding is rational, explicable and not an example of an agency problem. They show that the cost of capital (the return expected by shareholders) becomes too high when interest rates are high because further investment could not provide returns which exceed prevailing interest rates. They show that managerial hoarding peaked during the high interest rates of the 1980s because it was unwise to invest then, given that further investment would be unlikely to make returns that exceeded previous experience in an industry. They offer the cash-flow-rich nature of corporations during this period as the real reason for the financial restructuring seen in the 1980s, not, as Jenson suggests, market control in operation.

So, how can the disparity between market expectation/pricing and profitability be explained? Henwood argues that the market is simply irrational and inefficient.[96] Share prices reflect little more than the moods and ineptitude of inexperienced investors. Their disproportionate and unjustifiable effect on companies, he argues, has a correspondingly destructive effect on the productive economy. More moderately, Froud attributes this disparity to three factors; shareholder exuberance,[97] low interest rates and the injection of capital from institutional savings.[98] My own view is that there is something in all of the above explanations. The market is often incontinent and shareholder exuberance a very likely consequence of an SVO-based governance model so that rises in share prices may be said to result from ideologically (and often irrational) provoked expectations of share value. However, mass delusion, however hegemonic the causative ideology, seems insufficient to explain this disparity. A rational expectation that managers will pursue shareholder value activities, borne out by experience, must inform share value to some degree. Furthermore, this activity, which I will not examine here, must be grounded in activity other than dividend production.

The modest nature of dividends evidenced above takes us full circle to the institutional economics position held by Berle et al. in the "non"-shareholder value period of the 1930s–1970s. In their analysis, corporations make an adequate profit, but not a huge profit, and returns on investments, if accurately connected to profitability, are equally modest. However, when shareholder value ideology re-emerged with some force after this period and today, the economy did not (and indeed could not) radically change; ideology ran ahead of the objective limitations of the market. The 1970s saw a shift in ideology without a corresponding shift in the capacity of the market.

It has become commonplace to assert that the current financial crisis originates with greedy bankers taking huge risks. But those risks are at least in part attributable to the pursuit of shareholder value when profits are modest but expectation is high. Executives of financial institutions are treated as unconscionable scum. But they have just been rewarded

95 M Blair and M Schary, "Industry level indicators of free cash-flow", *The Deal Decade* (Washington: The Brooking Institution 1993), ch. 4.

96 Henwood, *Wall Street*, n. 92 above.

97 Based on R Shiller's book *Irrational Exuberance* (New Jersey: Princetown UP 2005).

98 Froud, *Financialisation*, n. 87 above, p. 78.

for doing what shareholder value required of them. Their pay is an agency cost properly incurred to bridge the divergent interests of principal and agent; incentivising executives by giving them a stake in profitability. These costs are borne precisely for the sake of shareholder value.[99] High executive pay therefore is a product of shareholder value ideology appeasing the capital market.[100]

Case Study: SVO (converted societies) versus MCO (mutual building societies)

In this section I will assess the affect of the current financial crisis on mutual building societies and on those societies that have been demutualised and are subject to a shareholder value model. In so doing, two areas of divergence between the mutual and demutualised society will be assessed; first their management's business model and second their response to crisis.

1 Business model

As a recent news article pointed out, building societies are not totally free from the charge of imprudent financing.[101] Many societies were exposed to subprime mortgages and a number offered 100 per cent mortgage deals. As late as January 2007, the Coventry Building Society offered a 125 per cent mortgage. And, whilst their losses have been significantly less than the demutualised societies, in large part because they relied less on the wholesale money markets and more on the retail market, it was the BSA which lobbied for the wholesale limits to be raised.[102] Under the Building Societies Act 1986 a building society could not raise more than 40 per cent of its funds from the wholesale market but this was later raised to 50 per cent[103] because of BSA campaigning as far back as 1992. Unsurprisingly, the societies who backed the BSA in seeking this reform were the very societies who would demutualise in the following years, the Cheltenham and Gloucester, the Bristol and West, the Bradford and Bingley, the National Provincial, the Alliance and Leicester and Northern Rock.[104]

However, in practice the societies that remained mutual did not make use of the new wholesale limits and they did not significantly alter their borrowing practices. With variations across the sector, on average, mutual societies raise less than 30 per cent of their funds from the wholesale market.[105] In contrast, two of the societies pushing for lower wholesale limits, the Bradford and Bingley and Northern Rock, once freed from the restrictions of the Building Societies Act 1997, took a huge slice of the wholesale market pie. At the time they collapsed they raised most of their funds from the wholesale market.[106]

Northern Rock's over-reliance on the wholesale market was certainly the cause of its collapse. One of the most notable characteristics of Northern Rock was that it was one of the first mortgage companies to engage in mortgage securitisations, collateralised debt

99 Jenson and Meckling, "Theory of the firm", n. 67 above.

100 Jenson recently blamed high executive pay on the target-base corporate budgeting systems noted above: Jenson, "Agency costs", n. 84 above.

101 www.timesonline.co.uk/yol/money/savings/article5859566.ece (last visited March 28 2009).

102 R Thompson, *The Independent*, "Building societies' cash flow squeezed", 22 July 1992.

103 HC Deb, 24 February 1995, vol. 255, cc. 339–41W; Building Societies Act 1997.

104 These societies demutualised in the few years following this debate. The Cheltenham and Gloucester in 1995, the Bristol and West, the National Provincial, the Alliance and Leicester, and Northern Rock in 1997, and the Bradford and Bingley in 2000.

105 www.timesonline, n. 101 above. As the article states, the BSA (now run by pro-mutual societies as opposed to the large pro-conversion societies noted above) wouldn't have allowed excessive reliance on the wholesale market.

106 ibid.

obligations (CDOs), that is, where mortgages are packaged up as bonds and sold to investors. The popularity of these securitisations meant that Northern Rock had access to cheap funds with which it could offer highly competitive rates to its customers. It rapidly raised its share of the mortgage market by relying on ever riskier, high-return strategies based on securitisation so that "roughly 10 per cent of Northern Rock's financing was coming from credit lines with a duration of less than one year"[107] and "more than 75 per cent of Northern Rock's funding came from wholesale markets, a greater proportion than any other UK lender".[108] This meant that £52.8bn of its funding was raised from securitisation and covered bonds.[109]

Northern Rock had never offered subprime mortgages but the American banks that had been securitising mortgage loans for decades had. It is generally accepted that is was investor realisation that many of these American mortgages were to risky investors who were defaulting on payments in a declining housing market that caused a rapid contraction in the securitisation market.[110] This meant that "when the credit markets seized on August 9, those financings (those relied on by Northern Rock) were moving on to maturities of less than a month."[111] Furthermore, Northern Rock was adversely affected by the "prevailing interest rate environment as well as its timing of transacting hedges for fixed rate mortgages".[112] By September 2007, Northern Rock was forced to borrow heavily from the Bank of England to maintain liquidity. However, following a mass withdrawal of funds by account-holders, famously queuing outside Northern Rock branches in droves, it could not recover. It was "temporarily" nationalised in February 2008.

In its report on the problems at Northern Rock, the Treasury Select Committee placed a major part of the blame on the bank's directors pursuing this "reckless business model which was excessively reliant on wholesale funding".[113] Another way of looking at this is surely to say that Northern Rock, a demutualised society, was no longer bound to the constraints of a building society and its particular construction of prudent behaviour. As a plc, subject to a shareholder value model which is short-termist by nature, its executives were bound to pursue profits and, thus, the risks that accompany high profit. Northern Rock delivered this for years and, when profits started to flatten in the early 2000s, its chief executive pursued this strategy even more aggressively. Ultimately, it failed to secure long-term stability, but was that its remit?

Likewise, management at the Bradford and Bingley pursued a risky business model strategy to gain high profits and fulfill shareholder value. At the Bradford and Bingley, it seems likely that this strategy expressed both the pressure to deliver shareholder value and the competitive inclinations of its management. Many executives of building societies expressed the view that the Bradford and Bingley's top managers were spoiling for the thrill of the market when it was a mutual society.[114] Although they had mounted a vocal defence of mutuality – with John Wrigglesworth stating in 1996, "our members would vote for

107 www.telegraph.co.uk/finance/markets/2815859/Why-Northern-Rock-was-doomed-to-fail.html (last visited 29 March 2009). But, by the end of 2008, roughly 50 per cent of the outstanding mortgages in the UK had been sold off in securitisation vehicles.

108 www.bsa.org.uk/policy/policyissues/consumerrelations/nrock.htm (last visited 30 March 2009).

109 www.companyinfo.northernrock.co.uk/downloads/Summary_of_Proposed_Business_Plan.pdf p2.

110 Ibid.

111 www.telegraph.co.uk, n. 107 above.

112 Business Plan, n. 109 above.

113 www.bsa.org.uk. Reported by the BSA (30 March 2009).

114 CEOs and chairs interviewed by Talbot at the 1998 BSA Conference (unpublished).

conversion for two quid. We have to show them the benefits of mutuality"[115] – they soon embraced the conversion route. In April 1999, the board of the Bradford and Bingley announced its intention to put a vote for conversion to its members after 62 per cent of voting members at an informal vote said they wanted their society to become a bank. In the run up to the conversion vote the management made some policy decisions which seemed to prepare the society for a "yes" vote. In October of 1999, the Bradford and Bingley announced plans to cut or move 385 jobs, which included closing two branches in Leamington Spa that employed 275 staff. Furthermore, the directors announced that all members of the building society would benefit from free shares as "equals". That is, all of its 3m members, including those who had recently joined would receive a windfall estimated at around £1000. Chief executive Christopher Rodrigous justified the scheme as impartial and not designed to achieve a "yes" vote by rewarding "carpetbaggers" (those who had joined precisely to gain the possible benefits from conversion).

The preparation continued with a rise in mortgage rates and a reduction in savings rates. Profitability was the goal. In the words of Mr Rodrigues, "as a mutual, we gave benefits through better rates. As a Plc, we will give benefits through greater product choice and dividends to shareholders."[116] And later, "the reason for our rate changes was simple: we need to be able to pay dividend to our shareholders and invest for the future. We intend to stay competitive as a Plc."[117] So despite stock market news of falling share prices in the largest converted building societies such as the Abbey National, Halifax, Northern Rock and the Woolwich, the motion in favour of conversion was overwhelming. With a high turnout (70 per cent of savers and 53.5 per cent of borrowers), 94.5 per cent of savers and 89.5 per cent of borrowers voted in favour.[118]

Bradford and Bingley plc quickly embraced the buy-to-let market as a mechanism to achieve high dividends. It was the biggest buy-to-let lender in the UK and it engaged in the highly risky practice of allowing self-certification of income by borrowers. Its buy-to-let lending rose from £748m in the first half of 2002 to £2.3bn in the first half of 2008. By June 2008 buy-to-let mortgages accounted for £24bn of its £49bn mortgage book.[119] Furthermore, 17 per cent of the Bradford and Bingley borrowers were self-certified.[120] Like Northern Rock it was heavily reliant on CDOs.

As early as 2003, even other banks not known for their cautious business plans were criticising buy-to-let lenders. Barclays stated that such lenders were taking "'suicidal' risks with their lending policies . . . lending very high amounts to the value of the property on 'dizzying' income multiples".[121] The Bradford and Bingley, which had increased its buy-to-let lending by 300 per cent that year, was, however, standing by this model. Mr Rodrigues responded to criticisms saying that "the buy-to-let market is not as risky as it is portrayed".[122] However, following a series of share collapses and unsuccessful attempts to raise funds through share issue the Bradford and Bingley was sold to Santander in 2008.

115 R Taylor, "Bradford & Bingley cuts mortgage rate to 7.24 per cent", *Financial Times*, 25 January 1996.

116 N Montagu-Smith, "Money go round: Bradford members weigh up windfalls", *Daily Telegraph*, 3 June 2000.

117 Ibid.

118 www.independent.co.uk/news/business/news/landslide-vote-for-conversion-at-bradford-amp-bingley-708027.html (last visited 9 December 2009).

119 www.dailymail.co.uk/news/article-1023533/Black-Monday-battered-banks (last visited 1 April 2009).

120 www.usatoday.com/money/economy/2008-09-28-2318700586_x.htm (last visited 30 March 2009).

121 Matt Barrett, the chief executive of Barclays, quoted in www.independent.co.uk/news/business/news/bradford-amp-bingley-defends-surge-in-buytolet-mortgages-589483.html (last visited 30 March 2009).

122 Ibid.

II Response to crisis

As argued earlier, one of the most significant attributes of an oligarchic MCO is its pursuit of stability over profitability. Building societies, as MCOs which operate as a group controlled by a small elite, pursue this stability goal in many ways. One of those ways is in the group's approach to financial crises.

Historically, building societies have responded to the financial problems of their members within the organisation of the BSA, typically with a financially strong society taking over the business of the failing society. Early on in building society history this approach had been additionally supported by strengthening the powers of the Registrar of Friendly Societies.[123] In 1892, the crash of the Liberator which was compulsorily wound up owing over £3m initiated the Building Societies Act 1894 which empowered the Registrar to take a proactive approach to a society's finances.[124] If it appeared after investigation that the society would be unable to meet its obligations to its members and that "it would be in their benefit that it should be dissolved", the registrar could "award that the society be dissolved".[125]

In the 1970s the fraudulent activities of the top executives of two building societies, which caused huge losses, were managed by the BSA and the registrar. In 1976, the accounts of the Wakefield Building Society revealed a shortfall of £633,000 due to the fraudulent activities of its general manager. The BSA was concerned to protect the Wakefield's business as well as the reputation of building societies per se so, whilst the society's reserves covered the shortfall, the Halifax took transfer of the Wakefield's engagements in order to circumvent any possible bad publicity.[126] In 1978, the Grays Building Society, whose total assets were stated as being £11m, was found to have deficiencies of over £7m, embezzled by its chairman and secretary Harold Jaggard over a 30-year period. Initially, the BSA arranged for the Woolwich Building Society to take over the Grays' engagements. However, when the full extent of the deficiencies became known, the BSA sought a more collective solution. They set up a rescue fund under s. 43 of the Building Societies Act 1962 to which all societies could contribute. Within three months, the investor members of Grays had access to their money and the Woolwich was able to take over the Grays business.[127] It was a collective solution to an individual problem made by organisations that operated as a collective.

Today's financial crisis has been dealt with in a similar fashion. Failing societies have merged with stronger societies and pre-emptive mergers have taken place between strong businesses which want to face the crisis as a larger organisation. The proposed merger of the Britannia Building Society and the Co-operative Financial Services (CFS) is an example of the pre-emptive approach. This merger, under the Building Societies (Funding) and Mutual Societies (Transfers) Act 2007, will be voted on by Britannia's members in April. The management of the Britannia is recommending the merger. In statements made on the society's website the merger is characterised as being the key to a brave new (financial world):

> The co-operative and mutual movements have never been more relevant. Owing to the damage done by the credit crunch, people have been crying out for a new way of doing business with a financial organisation of substance that truly has

123 Since 2001, the FSA has been the regulator.

124 For a long account of the Liberator's crash, see Cleary, *Building Society Movement*, n. 2 above.

125 Building Societies Act 1894, s. 7(2).

126 M Boleat, *The Building Society Industry* (London: Allen and Unwin 1986, first published 1982), p. 147.

127 Ibid., p. 149.

> their interests at heart - this merger will create that organisation and we'd hope to attract many thousands of new customers as a result.[128]

And later:

> Britannia members have an historic opportunity to help create a new way of doing business in British financial services by voting to bring together two leading customer-owned businesses with unrivalled reputations for social responsibility, customer satisfaction, employee engagement and member democracy. They can choose to be part of something good.[129]

This new "super co-operative" will join the Britannia with an organisation that boasts 109 retail and corporate centres, £40bn of assets and employs 8000 people. Unlike many contemporary building society mergers, this would seem to be a defensive move rather than a necessary one. There was, "'no question' of the companies being forced to merge because of funding difficulties".[130] To underline this, members would expect a substantial windfall payment following a vote in favour.

A more common form of merger activity in the recent crisis has been of stronger societies merging with weaker, smaller societies. Such mergers are falling into a typical pattern. First, there are no windfall payments for members and, second, the members of the larger society are typically not asked to vote on the merger. The reasons are clear. Windfalls are not appropriate for members whose society has suffered losses (and thus needs to merge) and transferee society members would probably not vote in favour of the merger because it involves taking on another society's losses. These are decisions taken by an elite management group, not members, characteristic of an MCO. Typical of this kind of merger were those involving the Skipton, Scarborough, Chelsea, Catholic, Nationwide, Cheshire, Derbyshire, Yorkshire, Barnsley and, more recently, the Dunfermline. The merger of the Skipton and Scarborough on the 20 of March 2009 was confirmed by the Financial Services Authority (FSA) under the Building Societies Act 1986, s. 95.[131] The Scarborough was financially vulnerable because of the fall in house prices but also because of the widely accepted problem of small societies being shut out of the wholesale capital market.[132] Under the Building Societies Act 1986, the FSA may direct a building society to transfer its engagements to another society if it considers it expedient to do so and, given the Scarborough's position, this was clearly the case.

Furthermore, in pursuance of this, a society that proposes to accept a transfer of engagements may do so by a resolution of directors, rather than a shareholder vote, if the FSA is satisfied that the transferee's assets are considerably higher than the transferors and the transferee's members would suffer no significant detriment.[133] The FSA was so satisfied and the transferences took place without a shareholder vote.[134] This was the act of an

128 www.britannia.co.uk/home/_site/corporate/news/press-releases/position-statement.html (last visited 30 March 2009).

129 Ibid.

130 www.telegraph.co.uk/finance/newsbysector/banksandfinance/4308247/Britannia-merges-with-Co-op-in-super-mutual.html (last visited 20 March 2009).

131 Decision of the FSA on 20 March 2009 following an application by the Scarborough Building Society and Skipton Building Society to confirm transfer of the Scarborough's engagements to the Skipton: www.fsa.gov.uk/pubs/other/scarborough_skipton.pdf (last visited 9 December 2008).

132 www.guardian.co.uk/money/2008/nov/03/banks-savings (last visited March 28 2009). Mutual societies, though not as reliant on the capital markets as converted societies, still draw a significant amount of funding from this source.

133 Building Societies Act 1986, s. 94(5)(b).

134 FSA decision, n. 131 above.

MCO, not a democratic decision-making body. Furthermore, no members received windfalls in order to preserve the merged society's capital base. This was a defensive action to preserve a preservable society.

Similarly, the Barnsley Building Society was taken over by the Yorkshire Building Society after it emerged it held £10m with the Icelandic banks Kaupthing Singer & Friedlander and Heritable – UK subsidiaries of Icelandic institutions which went into administration in September 2008. The two societies said that the merger followed "swift, pre-emptive action from the board of the Barnsley in approaching the Yorkshire to seek a merger after the identification of possible losses of deposits with Icelandic banks".[135] In kind with other similar mergers, there were no windfall payouts to Barnsley members but Iain Cornish, chief executive of the Yorkshire, was keen to put a positive spin on things saying that "this merger is a further step in demonstrating the strength of the mutual sector in the UK, with complete certainty to customers maintained at no cost to the taxpayer".[136]

In December 2008 the FSA confirmed transfer of the Catholic Building Society's engagements to the Chelsea Building Society[137] – 94.9 per cent of the Catholic's shareholding members and 94.83 per cent of its borrowing members voted in favour of the merger in a deal that was expected to give some windfall payments.[138] The Chelsea's members were not asked to vote and, instead, a director's resolution in favour of the merger was allowed as the Chelsea's assets were some 270 times larger than the Catholic's.[139]

Likewise, the merger of the Nationwide with the Cheshire and Derbyshire was in pursuance of a director's resolution and not a members' vote. This strategy is, of course, the faster route to merger. However, it runs contrary to the involved membership which the Nationwide, in particular, has been keen to foster over the last 10 years as a method of advertising the benefits of mutuality and as a bulwark against carpet-bagging.[140] It is at least possible that this merger would not have been supported by the Nationwide's members, as in 2008 the Derbyshire had made a loss of £17m in its subprime mortgages and the Cheshire had made a loss of £10.5m on one loan.[141] But as noted earlier, these building societies operate as MCOs with members exercising little control over their activities. It has been ideologically expedient for societies like the Nationwide to emphasise their member-based nature but such concerns will be set aside if they interfere with their practical operation.

The Nationwide mergers were announced in September 2008 and were swiftly confirmed by the FSA in November that year. It was described as a "prudent measure to ensure the financial strength of each Society" in the context of a financial crisis which threatens to destabilise the whole organisation.[142] Likewise, the announced collapse of the Dunfermline Building Society was soon followed by an announcement that the Nationwide would take over the accounts of its members and its prime mortgage lending book, with Dunfermline's head office, its branches and its employees becoming part of the Nationwide. The government and the FSA had ruled out an extensive rescue package for

135 www.guardian.co.uk/money/2008/oct/22/yorkshire-building-society-barnsley-building-society-merger (last visited 30 March 2009).

136 Ibid.

137 FSA's written statement at www.fsa.gov.uk/pubs/other/catholic_chelsea.pdf (last visited 30 March 2009).

138 Ibid., p. 5.

139 In accordance with s. 94(5) (b), ibid., p. 6.

140 This strategy dominated discussion in and out of the conference hall in the BSA Conference 1998.

141 Ibid. See also www.timesonline.co.uk/yol/money/savings/article5859566.ece (last visited 30 March 2009).

142 www.fairinvestment.co.uk/deals/news/mortgages-news-Nationwide-Cheshire-BS-merger-gets-the-FSA-go-ahead-2609.html (last visited 28 March 2009).

the Dunfermline because of its history of low profitability; estimated by Alistair Darling to be in the region of £5m a year, a sum "insufficient to even service the necessary loan from the government".[143] It would appear that the government is underwriting some of the Dunfermline's more "toxic" loans as a statement from the Nationwide's chief executive Graham Beale indicates:

> This transaction excludes high risk assets: commercial loans and some residential loans were not transferred, and the transaction will enhance the overall value to Nationwide's membership over the medium term.[144]

Conclusion

By the 1900s building societies had become MCOs, the organisational form which would dominate scholarly thought and government policy until the 1970s. Indeed, through the coordinating activities of the BSA, particularly after 1939, building societies had become MCOs par excellence. Many of the MCO theories during this time attributed MCOs with positive qualities and many of these qualities were possessed by building societies. For example, building societies were stable, saw steady but not excessive growth, and operated as hierarchies under trained professionals. Indeed, the power of top management was, under the auspices of the BSA, more centralised and arguably more able to plan and protect than the organisations assessed by Chandler. Like the modern corporations described by Berle and Galbraith, building societies were organisations that were connected to "social" business, they were tightly controlled by legislation and so were effectively public organisations. On the other hand, and in accordance with those theories critical of MCOs, building societies had many negative attributes. The one-member-one-vote governance was unlikely to deliver member control over management in societies with many hundreds of thousands of dispersed members. This enabled building society managers to become a self-perpetuating elite whose decision-making, centralised through the BSA, was often guided by managerial goals such as empire building and personal security. However, it is unlikely that these criticisms would have affected building society organisation without the rise of neo-liberalism. The neo-liberal approach which gathered pace throughout the economic failures in the United States and the United Kingdom galvanised a much stronger attack and promoted a much simpler solution. Building societies were economically inefficient and their managers were unaccountable. This could be solved through the legislative loosening of restrictions on building society business, reducing their privileges and enabling them to become truly market creatures, that is, companies. In this way a clear focus on shareholder value and market accountability would ensure efficient, dynamic profitability. Only it didn't. Or rather it seemed to for a while and then it didn't. Why? Because the high octane business of SVOs – pursuit of shareholder value, risk taking and high management rewards – leads to a dangerous enslavement to the capital market.

My argument then is critical of SVOs. However, it is not a call for the re-adoption of MCO forms. It is rather about a consideration of what it means to be a very large organisation, so large that today we call them too big or too complex to fail. MCO theories of the past cited public participation (as employees and consumers as well as investors), stability and state involvement, as reasons for abandoning the notion of business organisations as really private. Rather, these theories held that they should be reconceptualised as public institutions. Today, there is little disagreement about the need for tighter state regulation. However, there seems to be no correlation between state involvement and a consideration of business organisations as *public* institutions. The

143 Radio 4 news, 29 March 2009.

144 www.nationwide.co.uk/popup/dunfermline.htm (last visited 30 March 2009).

private property nature of corporations of all shades remains seriously unchallenged and with it the notion of shareholder value. But these are institutions which can only survive crisis by relying on public money. The society in which they operate can only survive by rescuing, and regulating, them. In these conditions, they cannot plausibly claim to be fundamentally private in nature. However, a new conceptualisation of large business organisations must not return to the inward-looking, self-serving elitism of the MCO. Rather, the challenge is to build a new, *public benefit* organisation to provide the structure for a truly progressive society.

NILQ 60(4): 471–86

The limits of legal ideologies

Sean Coyle

University of Exeter

In this essay I would like to suggest answers to two questions:

1. what is the role of law in determining human progress; and
2. can any understanding of the human person, or of human nature, be derived from the study of legal arrangements and institutions?

Modern jurisprudence has largely ignored the questions of human nature and human progress in its analysis of social institutions, but the present essay will suggest that these questions are, in fact, inseparable from such analysis and thus are central to jurisprudence.

Given the scope and centrality of these questions, the present essay can hope to do little more than illuminate certain dimensions of the problem only. The particular aspect I would like to explore is given concise expression by Lon Fuller. Fuller argued that the main purpose of law is to "rescue man from the blind play of chance and to put him safely on the road to purposeful and creative activity". In this way, law may be said to "create the conditions essential for a rational human existence".[1]

Fuller's claim implies the connection between the two questions I have raised: for in suggesting law as the pre-eminent means by which rationality is to be manifested and advanced in human affairs, Fuller of necessity presupposes certain assumptions about the nature of a "rational human existence" and the extent to which that mode of existence may be said to complete, rather than fight against, basic human impulses. Thus, the role and purpose of law are to be discovered in its rationality; and our understanding of this rationality gives substance to our beliefs concerning human life as a rational existence. My aim will be to take Fuller's suggestion as the starting point for discussion of my initial questions, and my argument will attempt to establish that, unless it is heavily qualified, Fuller's suggestion must be rejected.

The basis of the argument I wish to advance can be stated succinctly. It is that the notion of a rational human existence as something identifiable and intelligibly present within history is deeply problematic. This is not to deny that human action, or even human societies, may exhibit rationality; it is rather to suggest that it is impossible to derive from historical understanding anything which approximates or reduces to a definite "mode of life" that represents the necessary, unique or desired direction of human activity. Human

1 L Fuller, *The Morality of Law* (New Haven: Yale UP 1969), p. 9.

personality implies freedom, for the ability to differentiate "persons" as separate centres of thought and activity demands not only the possibility but also the reality of independence of each centre from all others. Freedom therefore implies creativity, in that each person (in manifesting genuine independence of thought and action) must be thought capable of originating modes of purpose and activity rather than simply manipulating pre-determined possibilities for action. This freedom is present within human history just insofar as historical explanation is incapable of furnishing explanatory historical laws: were such explanations of historic actions or events available, they could also be used for future prediction.[2] Creativity therefore entails the rejection of determinism and thus also the possibility of artificially determining the future form or direction of social action. Law may determine behaviour (by structuring choices), but it cannot determine the general mode of existence. Thus also, whatever rationality is present in law is not of a form such as will place human society onto any particular path, rational or otherwise, and will not (unless it completely suppresses freedom and creativity) generate conditions that will allow for the identification and realisation of a rational mode of life.[3]

Rationality and ideology

In the context of this essay, a "rational human existence" might mean either of two things. On the one hand, it might refer to a specifically desired form of collective social life. The quality of rationality here does not hinge upon the pursuit of reason itself, but upon the organised pursuit of some anterior goal which is itself taken to be rationally desirable (welfare, say, or peace, or the "caring" or the "just" society, perhaps even wealth). Where, as is often the case, a substantive aim is taken to be rational precisely insofar as it is not merely individually preferable, the goal around which social institutions are interpreted and organised is naturally seen as, in some sense, a collective one. Traditionally, therefore, the category of "the rational" (in this first sense) was taken to exist on a level separate from and above that of individuals; to be an object of the good life rather than a subject of personal preferences. Within the dominant patterns of Protestant political thought that have shaped the modern Western political consciousness, by contrast, this first sense of "the rational" takes shape not as some independently determined object of contemplation, but as that upon which human preferences converge: disagreements being intelligible as such only by virtue of arising within a shared framework where things have an approximately agreed sense and value. Whichever path is taken, the implication is clear that the existence of a rational form of life in this sense presupposes the presence of an impersonal aggregate, capable of identification in abstraction from the concrete human personalities which are its varied expressions. Accordingly, we may refine our understanding of the ideal form of life by clarifying our beliefs concerning the substance of this shared humanity.

The other way in which the idea of a "rational human existence" might be understood is by reference to a less substantial conception of the impersonal aggregate. This is best explained in contrast to the more full-bodied idea outlined above. Typically, the eschatological understanding of the human being implied in the suggestion of an impersonal aggregate is subsumed within the sphere of politics by way of an ideology. The uniting characteristics of "the human being", considered apart from the particular lineaments of individual personality, can be understood to refer to those dimensions in which all human beings possess a basic equality. Since politics (at its best) concerns the

2 See L Kolakowski, *Modernity on Endless Trial* (Chicago: University of Chicago Press 1990), ch. 22.

3 There are strong grounds for supposing that the complete suppression of freedom, even if taken as a deliberate aim of law, cannot be realised in fact: see N E Simmonds, "Straightforwardly false: the collapse of Kramer's positivism" (2004) 63 *Cambridge Law Journal* 98.

advancement of the human lot through the organisation and use of power, such basic equalities are naturally suggestive of an ethic in the light of which our conception of the proper end of politics is refined, and its practical exercise guided.[4] Because the nature of such equalities, and hence also the content of the ethic, are capable of endless interpretation, we may reasonably refer to the variant understandings of the guiding ethic as "ideologies". Politics then consists in the rhetorical and practical exercise of encouraging convergence upon preferred ideological principles; and in most cases it will be the law that is the primary instrument through which the dominant ideology will be advanced.

By contrast, the second, more austere understanding of rational existence takes the same impersonal substrate as demonstrating nothing more than the independence of each person (as an instance of it) as a centre of thoughts and impulses from all others. The basic equalities expressed by the substrate form the means by which each such centre is distinguished, as an *ens completum*, from the rest. This "necessary" substrate is to be contrasted with those accidental features of the concrete personality which arise adventitiously or contingently, and which give rise in turn to the pursuit, not of one common or unique set of values the attainment of which is regarded as the proper end for human beings, but rather a variety of different values and projects which may conflict with and impede one another. Thus, it is believed, no single ideology is implied in the form of the substrate, law and politics being an unfortunate if necessary body of rules and standards for the maintenance of a world in which conflicting ideologies must find some way to coexist.

Here also, nevertheless, there is an ideology at work in the transition from eschatology to politics. For, on the one hand, the features of the person that are taken to constitute the necessary substrate rather than the variable particulars of the personality are determined, at whatever level of abstraction, from a historically conditioned understanding of humans as actual social beings. On the other hand, the belief that the purpose of politics and the sphere of law lie in the protection rather than the suppression of individuality, the maintenance of many means and ends rather than the organised means to a common end, may itself display an ideological concern with the values of freedom and autonomy as central to political thought. It is important to stress here that "concern" does not necessarily connote support: the eschatological recognition of the human person as an "individual" (that is, an *ens completum*) necessitates the restructuring of political thought around ideologies specifically relating to that character, even if the resulting ideological standpoint is one that is hostile to individuality. Hobbes's philosophy perhaps provides an example of a politics of the individual in which the need to set strict limits to the expression of individuality is keenly felt. Hobbes does not celebrate diversity but seeks to contain it: it is not individuality (still less toleration or democracy) that he strives to uphold, but rather peace and stability in a world where the unrestrained expression of individuality would lead to the war of each against all.[5]

This second notion of a "rational human existence" might be abridged in the following way. We may see it as the effort, not to bring about convergence upon some definable ideology, but to reduce randomness in the simultaneous pursuit by distinct individuals of their own ideologically driven projects and preferences. It is therefore concerned with the suppression or management of conditions of uncertainty in social interaction (such as the possibility of violence, or lack of coordination in expectations). The lines quoted from

4 The long history of such efforts is sufficiently well known to excuse illustrative treatment of these themes in the writings of Grotius, Pufendorf, Locke, Wolff (to name but a few). Interested readers may wish to refer to excellent discussions in J B Schneewind, *The Invention of Autonomy* (Cambridge: Cambridge UP 1997), and R Tuck, *Philosophy and Government* 1572–651 (Cambridge: Cambridge UP 1993), for example.

5 See T Hobbes, *Leviathan*, R Tuck (ed.) (Cambridge: Cambridge UP 1996), ch. 15.

Fuller at the beginning of this essay are more suggestive of this second conception of rational existence than of the first outlined above; but insofar as both conceptions can be seen to animate the dominant approaches to modern politics, the remainder of this essay shall be concerned with both.[6]

Ideology and history

Though not exhaustive of the ways in which a "rational existence" might be understood, the above suggestions are sufficient to show that the notion of rationality is inherently an eschatological and, therefore (when it becomes the object of political contemplation), ideological, one. We might thus refer to such conceptions as "ideologies of rationality". Fuller's claim, that the purpose of law is to create and secure the conditions necessary for a rational human existence, is equivalent to the proposition that the purpose of law embodies an ideology of rationality. The burden of this essay then concerns an argument to the effect that the phenomenon of law implies no such coherent ideology.

In order to understand the complex relationship between law and ideology it is necessary to begin by examining the common law's association with what is sometimes referred to as a "historical" conception of justice. This expression can be taken to describe a form of adjudication in which remedies are determined on the basis, not of static principles of distribution or notions of fairness (for example), but of a detailed consideration of the way in which the present relationship between the litigants came about. But the common law method is historical also in a deeper sense, for the practice of determining the relevant law by reference to the interpretation of previous decisions itself embodies a distinctive historical ethos. This ethos, which predates the conscious articulation of the doctrine of *stare decisis*, is most clearly visible in the writings of Coke. It is animated by the notion that "reason" as such is too wide a basis for legal decision, as the treatment of each situation as a unique particular demanding consideration *de novo* invites the introduction of potentially endless sequences of argument and counter-argument in which there is no limit to the number of factors and perspectives which can be brought to bear. The avoidance of moral paralysis thus demands that the unique particularity of each case be recognised only within limits, by a comparison of the extent to which its various features resemble or differ from previous cases which bear some recognisable similarity to one another.

Because any two cases may be judged alike or unalike in countless different respects, such comparisons necessarily involve a determination to follow certain criteria of judgment to the exclusion of possible others. It is in this sense that jurists of Coke's era refer to common law adjudication as "the artificial perfection of reason": the artificiality resides in the submission of judgment to the continual reinterpretation of previous decisions as a means of channelling moral thought along a manageable number of lines. The relevant image is of a body of rules (that is, reasons for decision) being constantly refined and improved as they are reapplied in successive cases. Law in this sense represents a society's accumulated wisdom about the way in which the perennial problems afflicting it ought to be addressed.[7]

6 Having dealt with these understandings of politics elsewhere, only a brief precis will be given here. For more extensive discussion, see my *From Positivism to Idealism* (Aldershot: Ashgate 2007).

7 See E Coke, *Institutes*, vol. I (1628), s. 128. Coke uses the image of a blacksmith gradually hammering a piece of metal into shape. I discuss this conception of common law reasoning in more detail in my "Two concepts of legal analysis", in S Coyle and G Pavlakos (eds), *Jurisprudence or Legal Science?* (Oxford: Hart Publishing 2005), p. 15. For an extended discussion, see G Postema, "The roots of our conception of precedent" in L Goldstein (ed.), *Precedent in Law* (Oxford: Clarendon Press 1987), p. 9.

This historical conception of justice is itself expressive of an ideology of rationality. For it embodies a commitment to the image of reason as the product of reflective immersion within a body of experience, rather than (say) the deduction of a system of general principles of morality which derive their validity in abstraction from that experience. The ideological dimension to this distinction is clearly articulated in the first chapter of Blackstone's *Commentaries*, where he contrasts the shallow and regrettable attempts of the younger generation of MPs to legislate general solutions to social problems, with the deep, patient moral sensitivity exhibited by those who have devoted themselves to long study of common law precedents.[8] This ideological elevation of the common law, though never embraced unequivocally, was nevertheless visible in the writings of prominent jurists throughout the 17th and 18th centuries as a means of indicating certain limitations to the powers of Parliament and the monarchy. The fact that such a view of the common law's authority was but one argumentative standpoint in the search for a constitutional settlement underlines the ideological character of the variant understandings of moral rationality.

The ideology of rationality which lies at the heart of common law adjudication is not easily accommodated within the currents of modern philosophical thought, for which the idea of immanence is deeply problematic. Informed by a changed understanding of the nature of philosophical analysis, modern philosophy is animated by the belief that the possibility of moral judgment depends upon the independence of moral "validity" from the practices to which the relevant norms stand in a critical relationship. An obvious source of this conception lies in the need to preserve the necessity of the critical standards vis-a-vis the variable contingencies to which they apply: the critical function of morality being lost precisely where the standards of judgment are as contingent as the "facts" being judged. But the inherent oddness of this conception is never far from the surface: it can be seen, for example, in the sceptic's question of how the truth of standards which are *independent* of human practices can be demonstrated to apply to those practices?

The distinction between the two ideologies is captured quite clearly in the reflection that moral judgment depends upon an appreciation, not simply of facts, but of the *meaning* of the facts. Modern philosophy has by-and-large assumed that if history has any meaning, it must be in virtue of human interpretation. Spurred on by rapid developments in logic and the philosophy of language at the beginning of the 20th century, it was thought that meaning is not simply present within the facts, awaiting discovery; rather, meaning or significance are imputed to the facts on the basis of values that derive from elsewhere. Two distinctive lines of philosophical thinking served to reinforce this tendency. First, the search for a new and logically satisfactory theory of the *a priori* and the apparent success of philosophical semantics in explaining certain problems of mathematics led to a general belief in the dispensability of metaphysical concepts and a strong focus upon the conventional nature of language.[9] The second was an increasing tendency towards moral voluntarism, in which "values" were regarded as emanating from a source in human will and reason, rather than entering the consciousness via the intellect. Though he was not the originator of either of these developments, both received important (and in some ways definitive) expression in the writings of Kant.[10]

The form and substance of common law judgment stand in opposition to these developments and in this we find, within the historical conception of justice, an echo of an

8 W Blackstone, *Commentaries of the Laws on England*, vol. I (1769), "Introduction: The study of laws".

9 For an informative discussion, see J A Coffa, *The Semantic Tradition from Kant to Carnap: To the Vienna Station* (Cambridge: Cambridge UP 1991).

10 On Kant's involvement with the second of these intellectual developments, see e.g. Schneewind, *Invention*, n. 4 above.

earlier tradition of thought about morality and reason. Here, the aim of law lies not in the production of a set of ideal arrangements, nor the elucidation of a quintessential body of rules to secure their attainment, but rather in the creation of a means of fair and openly intelligible judgment in relation to occasions of social conflict. Law in this sense does not directly embody an ideology of the human good (that is, of reason-in-the-world), but instead seeks to give precise expression to aspects of the background of shared understandings upon which social interaction of all kinds is based. Thus, law, more perhaps than any other social institution, articulates the meaning of historic forms of human association; and in so doing it both reflects and refines those understandings which underpin social existence.[11]

It is not difficult to see how this understanding of common law judgment can blossom into an ideology of rationality in either of the two senses outlined above. For, in giving expression to the conditions of mutuality which constitute a civilised existence, we also clarify (as it would seem) a body of ideal arrangements in which the high watermark of that form of life can be identified. Thus, the progressive refinement of present rules of social engagement naturally appears as a journey from the imperfect society of the here-and-now towards the future Utopia. There is no doubt that a significant body of political and juristic writings, both of the early modern period and of the present day, reinforce this congratulatory myth; but the *artificiality* of the idea of reason to which the older jurists were (on the whole) much more sensitive ought to highlight the limitations of this view. This might be explained as follows.

In embodying a historical conception of justice, the juristic method of the common law displays an awareness far greater than that of the majority of ethical theories of the present day of the fact that a moral consciousness is, in the end, a form of historical consciousness. Amongst modern philosophers it is perhaps Gadamer who gives clearest expression to this idea: "Within the concrete conditions of his own historical existence – not from some position suspended above things," he writes, "[the historian] sets himself the task of being fair."[12] What is true of the historian here is also true of the philosopher and, more particularly, the jurist: for it is within the historic forms of language, culture and engagement that both the reality and the knowledge of our essential similarity and involvement in conditions of mutuality are manifested. Thus,

> [j]ust as understanding connects the individual ego with the moral commonalities to which it belongs, so also these moral commonalities themselves – family, people, state, and religion – can be understood as expressions.[13]

Law, more than any other aspect of the institutional fabric of society, gives direct expression to the meaning of these moral commonalities. But what does it mean to express such meanings, which are inevitably historical meanings? The human consciousness is essentially a finite source of understanding, and it embodies an intellect which is itself limited in terms of its powers of arrangement and perception: it is not something before which everything enjoys a simultaneous and equal presence. The arguments of the early natural lawyers, including Grotius, Locke and others, articulated this limitation as an inescapable frailty or imperfection of human reason, cause and consequence of the turbulent conditions of Fallen man. Yet it is these same limitations which make intelligence possible: memory, perception and the submersion of the individual power of judgment

11 I do not intend to provide extensive argument for that proposition here, as it seems to me fairly self-evident. However, a deeper argument for it can be found in my "Positivism, idealism and the rule of law" (2006) 26 *Oxford Journal of Legal Studies* 257.

12 H-G Gadamer, *Truth and Method* (London: Continuum Books 1975), p. 211.

13 Ibid., at p. 213.

within the context of historical effect are precisely functions of understanding, for they form the basis of priority and distinction. And it is the power of distinction which allows an escape from an existence of mental paralysis in which everything is equally present and co-significant. The historic consciousness, therefore, is inevitably a partial one, for it can perceive meaning only in the part and not the whole of history; and it is this partiality that constitutes the power of judgment. It is in this sense that (I believe) the artificiality of common law reason must be understood.

That being so, the narrow limits within which the law can be said to pursue the goal of a "rational existence" begin to take shape. The notion of a rational existence in the first of the senses explored above (as a specifically desired form of social life) depends upon the idea that history may be read in such a way as to reveal a general arc or trajectory belonging to human existence, terminating in some ultimate condition or state of perfection. Historical progress is thus viewed as a step-by-step transition from a primitive condition of existence to one of enlightened civility. The emergence of concepts or political phenomena, such as rights, can then be conceived as identifying the attainment of stages of enlightenment in which some previously hidden potentiality is actualised. (Grotius's elucidation of the *suum*, for example, may in this sense be thought to give expression to a necessary dimension of human personality hitherto undisclosed. If we believe this, then we believe such intellectual discoveries to involve the recognition of historical inevitabilities.)

The general suppression of metaphysical perspectives in modern philosophy is ironic in this regard, for the idea of human history as a journey from a more primitive condition towards a standpoint of enlightenment is one for which historic inevitability is substitutable for teleology. This teleological dimension to political philosophy is disguised only because, by an intellectual sleight of hand, the notion of a prized final state becomes detached from its historical context and becomes an abstract model of "the good society", history itself being appropriately reduced to a mere arena of application for moral insights derived from elsewhere. Thus, "historical necessity" becomes "logical necessity" or "moral necessity", comfortably shorn of metaphysical associations.

In the following section of this essay I shall attempt to reveal the shortcomings of this notion of a rational existence. It will then become clear that the second understanding of "rational existence" follows directly on from the failure of the first. The subsequent discussion will then focus upon the difficulties with this second concept of the rational life, and its applicability to notions of legal order only in a reduced and heavily attenuated sense.

Law and the pursuit of a rational existence (1)

It is important to be clear from the outset on the proposition that is to undergo criticism. The proposition is that the purpose of law is to create the conditions essential to a rational human existence, in the sense of a specific set of desired arrangements that are either absent or (more likely) only partially and imperfectly embodied in the society of the present. Law, in this sense, might be said to embody an ideology of rationality in that it is believed to give shape to a particular vision of what a "rational" existence is.

As we have already seen, the ideal of a rational existence is an eschatological one insofar as it implies a direction to human life, terminating in some finally desirable condition of being. In less evasive language (that is, the language of the layman rather than the philosopher), such an end-point might be termed "the meaning of life"; and to express the proposition in these terms is to claim that law exists to bring about the institution of an ideal set of arrangements which give clear and direct expression to this meaning. Expressed in these terms, the central problem becomes immediately clear: for an agreed understanding of the meaning of life is precisely the happy condition which is absent from human

existence. Politics, then, cannot concern in any straightforward way an advance toward the realisation of that meaning through law; rather, the question of the meaning of human existence is just that which defies agreement in all facets of human social engagement.

It is not, however, this fact which is most damaging in political thought (though the creation of blinkered and evangelical ideologies has certainly played a significant part in determining modern political realities). For the error of political and jurisprudential ideologies lies not in mistaking specific dimensions of a contested truth for the agreed core; rather it concerns the very pursuit of a rational existence, in those terms, at all. Upon reflection, it becomes apparent that the attempt to institute an enhanced mode of existence carries the suggestion that the full or true meaning of life is something postponed to a future time, and hence that such meaning cannot completely emerge in the conditions of the present. This is, of course, a direct if implicit reduction of the meaning associated with the life of the here-and-now. Such a reduction betrays a basic misunderstanding of human existence: for to view present arrangements as unsatisfactory is one thing; but to view them as a mere stage of development on the road to an imagined higher state is to contrast the actual imperfection of man with a vision of human perfectability which is alien to its essential nature.

It is possible to challenge this vision of perfectability in a number of ways, but one such explanation is as follows. Historians have frequently pointed out the tendency, in some degree inevitable, of understanding historical events in the light of present knowledge. This is not true simply of history but of all fields of human enquiry: despite stubbornly embedded philosophical intuitions to the contrary, there are not "facts" and "interpretations" of the facts, of which only the latter are knowable. Rather, the "facts" are always different ways of looking at the facts. (This is inevitable in any context involving discussion of facts: the notion of some objective bedrock lying behind our characterisations is something which must itself lie beyond characterisation, and is hence useless and irrelevant to human thought.) If this is true of our relationship to present events, then the status of our understandings of historical events as interpretations does not affect their validity, simply because they differ in numerous respects from those of contemporaries. The very possibility of history as a subject depends on this. But there is an important difference between the proposition that the past is always understood as a projection of the present, and the proposition that the significance of the present and future are determined on the basis of the meaning attributed to this constructed past. The latter, but not the former, implies the existence of a reliable method of historical explanation.

As Kolakowski amusingly recounts, there is a simple way of demonstrating the absence of any such method, which is at the same time a demonstration that the status of all explanations depends upon a decision to treat them as explanations:

> My late friend Lucien Goldmann displayed admirable ingenuity in linking up the smallest details of Pascal's *Pensées* with the plight of the French *noblesse de robe* after the Fronde. One would think that he could really write the *Pensées* without reading them, solely on the basis of the historical evidence concerning the class conflicts of the time. And it is here that the crucial point lies. For if there were a reliable method for a historical explanation of culture, we would also be able to use it as a tool for prediction. To be able to explain what has happened is also to be able to predict what has not yet happened, otherwise the word *explain* would not have the meaning normally attributed to it.[14]

14 Kolakowski, *Modernity*, n. 2 above, at p. 244.

Thus, he concludes,

> Whoever claims to be able to explain particular phenomena in the history of music, or of the novel, can prove this claim only by writing a novel or a piece of music which does not yet exist but which will be created tomorrow by someone else.[15]

Similar thoughts hold true of the attempt to characterise an imagined state of human perfection. To claim to have discovered a general trajectory in human existence, from an initial state of barbarism to one of urban civility, and to project that into the future as a vision of the ideal society (or, at least, the ideal set of principles for running such a society), is to express belief in a form of historical explanation that does not exist. It is here that the historical conception of justice implicit in the common law differs from an ideology of rationality of this kind. For the body of rules and ideas inherited from the past are not taken as together revealing "the meaning of life"; they are simply a convenient set of mundane arrangements which have been tried and tested in a vast array of previous contexts, and which thus form a useful point of departure for the judgment of new contexts which we decide to treat as relevantly similar to those that have gone before. The appropriate image is of a set of standards for managing the tensions endemic to a social existence, rather than a body of principles which aim, through their own gradual self-transformation, to resolve all tensions and bring about a heaven-on-earth.

The goal of a historical conception of justice is not the production of a general understanding of how things ought to be done, or of how things must be refined further and further to some ideal point; it is instead a form of self-knowledge, that is, a society's attempt to give concrete expression to its own form of life. The ability to furnish self-knowledge is limited in just the same way as human understanding lacks a general method of historical explanation. For history is not simply a collection of causal processes and outward appearances, but a living tradition "riddled with countless breaking-off points, and each creative act, each creative individual, is such a point".[16] That being so, there is no single direction in which human existence moves, nor a collective destiny by which human existence is shaped: "historical reality is not merely a heavy, opaque medium, mindless matter, rigid necessity against which the spirit beats in vain and in whose bonds it suffocates".[17] The view that legal order gives expression to a form of life by *determining* that form of life (that is, by uncovering its eschatological meaning) is thus a false picture.

The idea of a "form of life" is roughly that of the various relationships which connect the individual person to the moral commonalities to which it belongs. As previously noted, the meaning of these commonalities are inevitably historical meanings. Truly to understand such meanings we must relinquish the false view of historical progression. History (as well as law) is shaped by human choices; but these choices are not made by reference to abstract understandings of historical time or direction, nor of their significance for the era in which they come about.[18] Rather they emerge as direct responses to current events. In the same way, the judgment of single instances, which forms the main business of the common law, is not exhausted by the explanation of such instances as particular examples of a general rule or concept, for each case is identified precisely by reference to that which makes it unique.

15 Kolakowski, *Modernity*, n. 2 above, p. 245.

16 Ibid., p. 246.

17 Gadamer, *Truth*, n. 12 above, p. 199.

18 This comes across most strongly in the context of those actions performed with a conscious regard for posterity: for the meaning of the actions of the French Revolutionaries (for example) are finally explicable only as responses to particular events which provoked them. Robespierre is as much an unrepeatable product of his time as any other historical figure. The uniqueness of the individual demands recognition of this point.

Legal judgment, too, is therefore inherently particular: just as there are no "facts" but only different ways of looking at the facts, there are no "values" which structure judgment in the manner of abstract rules of morality that exist independent of the world; there are only judgments about the nature of particular "facts". A morality, as Oakeshott remarked, "cannot exist in a book or a vague ideal or anywhere except in an active sensibility".[19]

It follows from this that the attempt to give expression to a form of life must necessarily remain incomplete. For the significance of each judgment is completed neither by its contemplation of the uniqueness of the situation in reference to which it exists, nor in the relation of the particular case to the general concepts or understandings in relation to which it is identified. In direct consequence, no given set of principles or ideals which might be abstracted from practice as those which guide legal decision can ever amount to a coherent and complete statement of a "form of life", nor to an unambiguous governing ethic capable of producing a higher state of social existence or a rational mode of life. Lacking full comprehension of the meaning of the present, we also lack any reliable means of predicting the future significance of human choices.

Law and the pursuit of a rational existence (2)

The failure of the ideology of rationality lately described gives rise to idealism of a different kind. This can be explained as follows, but it is worth mentioning that the failure of the first ideology does not imply its irrelevance to modern society, for it is still experienced as a significant pull in legal and political thought. Indeed, the relationship between the two ideologies (the first in which a common end is the object of organised pursuit, the second seeking to maintain a balance between various competing means and ends) is much closer than might be supposed, as the argument below serves to demonstrate.

At some level, law and government consist in the production of an average condition of being. The pursuit of a particular mode of rational existence, as a single end in social life, serves to place that average condition at the heart of political and legal aspirations. As noted above, such an average condition of being is but an incomplete understanding of the complex relationships that exist between the individual and the common moral structures which that persona inhabits. This is, indeed, the reason why legal rules are capable of endless refinement and adaptation. An ideology of rationality in the second sense mooted above perceives perhaps more clearly than the first the truth that wherever organised government exists, people are governed as abstractions. This immediately discloses an inherent tension in political and legal order: for the duties, rights and privileges that constitute the person under law will never fully or completely correspond to the actual range of interests, desires, beliefs and dispositions which belong to the concrete person. No matter how sophisticated it may be, a mere concept inevitably implies a *contrast* with real life, not its description. (It is, as Gadamer points out, essential to an experience that it cannot be exhausted by what can be said about it or grasped as its meaning.)[20]

This tension is proper to law, for it is in effect part of the function of law to suppress human creativity. At first distasteful, the truth of this proposition becomes clear when we reflect on the fact that unrestrained human creativity embodies anarchy (precisely the lack of regulatory boundaries to human endeavour). Thus, a society, in the proper sense of the term, demands the presence of an organised bureaucracy and the rule of law in order to contain just those forces of human creativity which mark the breaking-off points that

19 M Oakeshott, *Religion, Politics and the Moral Life* (New Haven: Yale UP 1993), p. 45.

20 Gadamer, *Truth*, n. 12 above, p. 58.

preclude the emergence of historical explanation.[21] It might then be supposed that central to the notion of a "rational" existence is the need to preserve large and concurrent areas of freedom in which to explore the self in all its particularity. Indeed, it may be thought that a rational existence involves a search for the minimal bureaucratic and legal arrangements needed to secure this personal freedom from the will of others.

Such a vision of a rational existence is just as chimerical as that which seeks the implementation of a higher state of social being based on a view of human perfection. One could mention, of course, that the possibility of agreement upon such a set of minimal arrangements is on a par with that of a general consensus on the meaning of life; or that a view of human perfection is every bit as present in this vision as in the first (for how else are we to determine which forms of behaviour it is necessary to suppress within these minimal arrangements?). But I wish here to explore a different aspect of the problem. For what is common to both visions of "rational human existence" is the assumption that the conditions required to bring that state of perfection about are external rather than internal to the self.

The elevation of social philosophies and theories of justice to the status of secular religions is to be despised precisely insofar as they present the meaning of life as a set of external conditions as opposed to an inner state of reflective awareness. It has been the hallmark of modern political thought to seek for an optimal pattern of distribution of rights and liberties by reference to which all persons are treated as equals. This has led to a damaging tendency for present arrangements to harden into ideologies. But the aspect of this transformation that is most damaging is not the ongoing series of military attempts to liberate "backward" or "unenlightened" regimes for which such ideological developments are otherwise unattainable; it is rather the spiritual impoverishment that results from a view that a form of life is to be comprehended in terms of its outward arrangements. To have discovered the meaning of life or sociality in a particular set of arrangements or mode of governance is to assume that life has less significance, or less value, where such external features are absent. But this is manifestly absurd: both the inner life of the spirit and the outer life of action and performance remain undiminished even in contexts wherein fundamental rights, say, do not exist or are unimaginable, or where magic has replaced ordinary causality. Passing familiarity with Tennyson's poem "The Lady of Shalott" or with Siegfried's life as presented in the *Nibelungenlied* are enough to confirm this.

If, instead, that which gives meaning to life is not to be found in the external conditions within which the individual moves, but rather in a person's inner spiritual life, it becomes apparent that that meaning is articulated in the endlessly variable ways in which the individual is able reflectively to transcend the present state of affairs and approach the eternal. Here, society is not irrelevant to the spiritual quality of life, but it is not the goal. It rather forms an essential context in which spiritual reflection is able to take place. This becomes clearer when we reflect upon the fact that the "individual" self only has meaning *within* society, as part of the whole world of ideas in which it is understood and given meaning. The nature of humanity as it is made transparent to us is derived from this body of ideas, and thus the ability to reflect upon the human situation is itself determined by involvement in the shared world of ideas. As Michael Oakeshott observed, a society is not finally understood as a collection of bodies in proximity, but of minds in relation.[22]

21 I use the word "creativity" here to refer equally to the production of evil or destructive forces as to those of "progress" or the good.

22 Oakeshott, *Religion*, n. 19 above, p. 50.

It follows from this that, if the production of an average condition of being falls far short of a genuine understanding of whole and part within the social world, then so too does the attempt to give maximal expression to "the individual". For an individual's "experience" of life, which is given meaning through concepts and language (phenomena that are essentially public rather than private), forms the basis for deliberation, understanding and decision. Given that our very understanding of an individual (whether as an "agent", or locus of choice, or passive victim of circumstance) is conceptual, the meanings of individuality and of individual experience remain fused with the totality of social life in which they move, and thus the meaning of an individual's creative expressions, thoughts, desires and actions are constantly accompanied by the meaning of that social whole. Here, therefore, it is necessary to think of the part *neither* as determined by the whole in which it operates, *nor* as independent of it (even relative to certain aspects or dimensions only), but rather of whole and part as organically connected at every point.

So long as juridical and political understandings focus upon the differentiation between individuals and the external conditions which affect them as forming the basis for discussion and analysis, such a notion of organic connection will remain elusive. For it will continue to seem that the philosophical meanings that are sought in jurisprudence and politics lie neither in present institutions and arrangements nor in the social commitments and personal freedoms they generate, but instead in some abstract world of concepts beyond them, and in the light of which the present social forms are imperfect realisations of a governing ideal. Life (specifically the prized life of liberal society under the rule of law) is misrepresented as a journey towards this ideal, or towards a fuller expression of the free self, rather than as a process of understanding in which the goal is to "know thyself" (in the words of the Delphic injunction). The precise distinction between the person, as an individual, and the broader social whole is irrelevant to this enterprise except insofar as it forms a context in which reflection on that question takes place. Self-understanding is finally inseparable from the conceptual nexus supplied by language and culture, for it is these dimensions of life which shape the identity of all individuals subject to them. It is because such social forms and arrangements are themselves historically undeterminable and evolving outcomes of human creativity that the shaping effect of these forces is not to be *contrasted* with the existence of personal freedom, however much they may limit immediate possibilities. One recognises oneself not in contradistinction to these forces, but rather through them:

> In language, customs, and legal forms the individual has always already risen above his particularity. The great shared moral world in which he lives represents a fixed point through which he can understand himself in the face of the fluid contingency of his subjective emotions.[23]

It is this organic connectedness which undermines the two alternative ideologies of rationality explored above. For within the creative world underpinned by common structures, there is no common object of pursuit or independence of individual direction in which to anchor the notion of a "rational existence". Kolakowski points to the

> simple fact that all of us, both in politics and in private life, pursue various independent objectives, irreducible to each other, inexpressible in homogeneous units, and unattainable jointly; the means we employ to achieve one objective usually limit, sometimes even destroy, the hopes of achieving another. Since we may not evaluate the objectives on a hierarchy of preferences in terms of

23 Gadamer, *Truth*, n. 12 above, p. 229. Gadamer is here commenting upon Dilthey.

> rationality, we are often helpless in assessing the rationality of actions if they imply a choice between incompatible or mutually limiting aims.[24]

This is especially true when seeking to assess the rationality or otherwise of a society's life, or that of a single individual, over time. For even supposing a hierarchy amongst available choices and values were possible (something that is itself difficult to ground in "rationality"), the defence or criticism of particular choices or historical trends in hindsight would require the ability to construct accurate counterfactual chains of causality in which an optimal solution is reached. Again, for this we would require a truly scientific method of historical explanation and prediction concerning the overall effects of particular decisions, which the very presence of creativity denies.

More than this, the options around which individual and collective choices and desires are structured do not come readily demarcated into discrete units; nor are such options generally functionally independent of one another. This is most clear at the political level, where party politics ensure that all but the most committed ideologues must settle for indicating support for a range of policies and objectives of which only a percentage are actually desired. Here, the mechanics of constantly shifting majorities belie the attempt to analyse political progression in terms of the rational pursuit of consciously desired aims. But this is also true at the level of individual choice and desire, for each exercise of an option serves to eliminate or restrict others, or to invite unavoidable and unwanted consequences of its own. In the light of this, "rationality" for the individual is probably limited to the deliberate avoidance, where possible, of foreseeably counterproductive actions. The complexity of all mechanisms of social choice, and the randomising effect of interaction between millions of everyday individual decisions, means that even this restriction is insufficient for capturing a workable notion of rationality at the social level. Thus, in neither sense can the purpose of law be said to consist in the pursuit of a rational human existence.

An alternative suggestion

The structure of a human society of any size or complexity is such that the connection between visible means and clearly defined ends lies beyond the scope of criteria of "rationality" or "irrationality". It is perhaps a necessary feature of such societies that they seek the joint realisation of objectives that are known to be mutually limiting or in conflict. Thus, the desire for minimum standards of food safety is apt to undermine the desire for low prices; the welfare state comes at the price of unwieldy bureaucracy, duplication and waste; popular elections bring about the coarsening and dilution of political debate; and all measures aimed at security (of property, of the person, of the market etc.) are bought at the expense of freedom. That these familiar facts have not dimmed the attempt to characterise legal and political arrangements as elements in the creation of a "rational" society is due to the fashionable philosophical belief that the goal of juridical and political theory concerns the achievement of a satisfactory or correct balance between the competing aims. But the notion of "balance" between goals which pull in mutually exclusive directions is entirely spurious: the tensions that exist between the various goals are not binary or opposite, but complex. No direct comparison of such tensions with physical forces in the context of which "balance" has an established meaning is available; hence the key ingredient of a successful analogue, that the meaning of the conceptual transition is direct and clear, is missing.

24 Kolakowski, *Modernity*, n. 2 above, p. 193.

It would have been possible, if overly dismissive, to object to Fuller's statement about the purpose of law at the beginning of this essay, by recalling the simple fact that only persons, and not systems, have purposes. Systems may be created or adapted for a purpose, but they do not mechanically give expression to that purpose in their operation. In order to understand the significance of law, therefore, it is necessary to begin with an idea of the person.

We might begin with the observation that the person has being within a social world to some extent composed of competing causes and random effects. Being random, such forces will only ever be imperfectly and incompletely understood. Moreover, the external forces (fashions, manners, prices, availability, political majorities etc.) affecting the person are not fixed or static but constantly changing and in motion. Given that such forces are a permanent fact of social existence, Fuller's statement that law exists to "rescue man from the blind play of chance and to put him safely on the road to purposeful and creative activity" must be read with care. An uncareful reading of this statement might suggest that law exists as a means of organising the activities of persons so that they serve to converge upon specifically agreed goals; or that law operates to create domains of equal freedom from the will of others in which each person can coherently pursue medium to long term projects of their own. Both suggestions are indeed useful as convenient shorthands; but if taken for deep truths about the nature of law they are apt to mislead. For when are important social goals (those significant enough to become the focus for collective pursuit) ever agreed upon, rather than converged on more-or-less obliquely, and with greater or lesser passion or interest, often incidentally on the road to other goals? And when can the person be said to be free of the will of others, when we operate in a world of interpersonal cause and effect, of market interaction and multilayered influence: in short, a world of shared ideas?[25]

Law does not exist to *rescue* humankind from the blind play of chance, but rather (in forming a body of rules and decisions) to offer greater stability to expectations, that is, to make the decisions of others more predictable, in a world where chance cannot be eliminated. Thus, the decisions of others (deliberate or otherwise), including those of judges, may be made more transparent and predictable without ever becoming fully predictable or clear as to their total effect. We know this; but the ideologies of rationality that lie at the heart of modern jurisprudential and political thought operate as if phenomena such as "freedom" can be given clear boundaries within which they possess an absolute, or at least rigid, existence. This is no more than the observation that concepts are always but incomplete descriptions of realities which constantly outrun them. But, moving within a world of ideas, we have lost touch with the feeling that society is not simply a collection of externally fixed rules and boundaries and interstices of freedom in which to pursue subjective desires. It is, rather, a malleable and constantly shifting product of those same ideas; and our understanding of it is shaped by exactly the same linguistic and conceptual forms that give expression to our desires.

Such an existence might well seem to deserve description as "rational"; yet if it is thus described, this really is to say no more than that we enjoy a "human existence" or a "social existence". The adjective "rational" is in that case redundant. But the implicit connection of

25 This is sadly true even of philosophy, where a market in ideas certainly exists and operates to define the orthodoxy and the accepted centre ground. For example, the starting points of this essay are determined by what is acceptable and readily recognisable to others. It is still more true of political argument: perhaps the thought that the perceived value of ideas depends insidiously upon their popularity might do something to dent the belief that academic commentary, in virtue of being more refined, is somehow more "rational" than lay opinion?

that term with measurements of the quality or validity of arguments and chains of reasoning has the damaging effect of suggesting that the nature of human society is other than it actually is: it suggests, falsely, that human societies can be and are directed and underpinned by coherent ideologies that are capable of abstract expression and "rational" analysis.

I do not intend to offer any general conclusions regarding the ramifications of accepting this thesis. Rather, I shall briefly consider one issue that I believe to be central to jurisprudential and political thought: that is the altered conception of the individual's existence in a modern society.

By and large, the political and jurisprudential philosophy of the modern day has concerned the relationship between the individual and the state, and the nature of the relationship between citizens within a state. The approach I intend to suggest is different in that it focuses upon the nature of the individual person's own existence within a society composed of ideas (or, in Oakeshott's words, of minds in relation). In seeing society as something *external* to the individual, something to which the individual is *related*, philosophers and social theorists have committed a fundamental error. For what is the character of such a society? Scholarly and popular opinion often refer to the concept of the "secular state", for example. But what does this mean except to reflect a substantial body of opinion that religious belief, and perhaps religiously inspired morality, are matters for private decision rather than public goods? Given the manifest failure of efforts to remove altogether religious discussion from public debates of all kinds, we must understand "the secular state" as an abstraction, as part of a world of ideas which is both broader and deeper and which constitutes human society.

Labels such as these are undeniably useful in describing vast currents of experience, but become limiting towards philosophical thought when mistaken for the whole. Once it is understood that human society is in fact a constantly shifting mixture of ideas (in which large and noticeable shifts occur but slowly), we can appreciate the error involved in analyses of the relationship between the individual and society. For although we may artificially isolate dimensions of this supposed relationship, there is in fact no settled object which is in the relevant sense "externally" related. Consequently, the goal of philosophical reflection might be said to be the partially autobiographical one of understanding the way in which an individual exists and develops within a world of ideas of which they are themselves an element. And this is simply to know thyself, to understand the meaning of one's life as it takes place within the possibilities that are open to it.

If there is an overall moral to be drawn from this, it might be expressed along the following lines. Philosophies of politics and of law should not concern the reduction or elimination of tensions and conflicts within society according to some overarching political ideology or theory of justice. Tensions are endemic to human society, and in that respect their particular form and substance at a given point in history is peripheral to analysis of the kind I am suggesting. But if tensions and conflicts are endemic and permanent features of the human condition, our responses to them are not. The meaning of "free will" is disclosed in the fact that our reaction to adversity of all kinds is open and unpredictable, which is to say that each thought and decision is a breaking-off point or creative act, not something fixed by the mechanical operation of grace or its opposite. Thus, the goal of self-understanding, of the good or virtuous life, is just the recognition and banishment of hatred and its replacement with the institution of the Christian ethic of neighbourly love: the avoidance of depersonalising forces and alienating tendencies in human thought and action in favour of a compassionate understanding and fellow-sympathy. In knowing oneself, therefore, one comes to understand others. The difficulty then lies, not in understanding what the good life consists in, but in actually living that life.

This is an approach to the understanding of human societies with which modern philosophical approaches are fundamentally out of sympathy. For they seek a mode of analysis which aims to emulate in key respects the scientific method of clarifying general laws. As G K Chesterton's Father Brown observed:

> Science is a grand thing when you can get it; in its real sense one of the grandest words in the world. But what do these men mean, nine times out of ten, when they use it nowadays? . . . They mean getting *outside* a man and studying him as if he were a gigantic insect: in what they would call a dry impartial light, in what I should call a dead and dehumanised light. They mean getting a long way off from him, as if he were a prehistoric monster . . . When the scientist talks about a type, he never means himself, but always his neighbour; probably his poorer neighbour. I don't deny that the dry light may sometimes do good; though in one sense it's the very reverse of science. So far from being knowledge, it's actually the suppression of what we know. It's treating a friend as a stranger, and pretending that something familiar is really remote and mysterious.[26]

To return to my initial question, law does not "create" the conditions for a rational human existence if this is intended to refer to deliberately constructed conditions which emerge at some posterior time. Rather, law forms an essential and organic part of a civilised human existence, and cannot be analysed in separation from it. In what sense, then, can law be said to play a role in determining human progress? I would like to close this essay with the following brief observations. First, the notion of "human progress", much like that of "rational human existence" is essentially an ideological abstraction which finds no foothold in history. For how is such progress to be measured? Is it a matter simply of greater average standards of living? This suggestion is problematic in that greater material wealth and comfort might be said to lead to decadence, complacency, waste and spiritual impoverishment. Nor can progress be measured simply by technological advancement, for every such advance can be used to increase evil in the world as well as good. At every point, therefore, the general notion of "human progress" demands the arrangement of values that are irreducible to one other into a hierarchy: values which, being mutually independent sources of good (or evil), do not admit of hierarchical arrangement on the basis of rationality.

Rather than relating law to an ideology (of progress, of rationality), I suggest that law must be seen as an integral part of the fabric of social existence. It does not play any part in creating or sustaining something which is abstract and historically impossible, but is rather an element of something concrete and familiar but historically indescribable. Thus, finally, understanding of the role of law in human existence is inseparable from an understanding of the human person or of human nature in general.[27]

26 G K Chesterton, "The secret of Father Brown", in *The Complete Father Brown* (Harmondsworth: Penguin 1981), p. 465.

27 A previous version of this essay was presented as a public lecture at the LSE. I am most grateful for the helpful comments I received on that occasion, and to my colleague Fiona Smith for her insights and suggestions.