Bruce Chapman*

TRUST, ECONOMIC RATIONALITY, AND THE CORPORATE FIDUCIARY OBLIGATION†

I *Introduction*

The restructuring of corporations affects all of us. Whether it takes the form of a friendly merger, a hostile takeover, a sell-off, a plant closing, or a leveraged buy-out, it is difficult to think that we are not affected somehow, either directly in our role as employee or corporate client who experiences the loss or gain of a job, or indirectly as a member of the particular community that is made richer or poorer by the same transaction. Yet few of us will have very much to say about the running of those corporations that affect us so much. At most we will be able to address our concerns to the political regulators of corporations, and there it is clear that we will often have to stand in line behind those who have reached the corridors of power before us.

Of course, it is not at all clear that we *should* have very much to say about fundamental corporate changes. These changes, it is said, are the stuff of economic competition and the efficient reallocation of resources. When we are benefited, we should be happy but not surprised; that is only what economically efficient contracting has always promised to provide. And when we are burdened, we should recognize that that too is only a part of the process of efficient reallocation of resources to higher-valued uses, the process through which, it is said, we can all reasonably expect to benefit in the long run.

However, at times the burdens seem to be especially severe and concentrated. The employee who loses a job, the creditor left with unpaid debt, the supplier without a client to buy custom-made goods, will all feel the exit of a corporation from the community more seriously than those of us who can turn more easily to other opportunities. Moreover, these corporate stakeholders will feel a sense of betrayal that the rest of us will not. They will have relied on the restructured corporation in a way that

* Faculty of Law, University of Toronto

† I wish to thank Ian Ayres, Paul Halpern, Roberta Romano, and William Simon for helpful comments on an earlier draft. I am also grateful to the Yale Law School for providing me with the opportunity to work on this paper by way of a 1991–92 visiting faculty fellowship funded by the John M. Olin Foundation.

distinguishes them, indeed, in the way that probably prevents them from so easily turning to the other opportunities that are now available to us.

It will be argued that this sense of betrayal is misplaced. These stakeholders have only themselves to blame; if they had special concerns about the corporation carrying on as before, or special reliance interests that needed protecting, then the argument will be that they should have contracted for such protection in advance. Then the benefits of a seemingly efficient corporate restructuring would have been properly tested against these reliance interests and, if the benefits failed to match this competition, the reliance interests would be protected. Indeed, the proponents of this argument will suggest that it is quite possible that the significance of these reliance interests has already been tested in market contracting and found wanting. That is why they are unprotected now. Or, perhaps, the contract has already been signed that compensated the stakeholder for this risk ex ante, in which case we are confronted with someone who simply wants to have her cake and eat it too.

This response is symptomatic of the view that sees the corporation, and all it does, as contractual, as a 'nexus of contracts,' to use the current phrase.[1] All dealings with, and within, the corporation are to be interpreted according to contractual principles. Corporate problems are 'contract failure' problems, amenable to contractual solutions to the extent that rational, fully informed parties can see their way around transactions costs. Never mind the fact that this contract, unlike so many others, requires registration with the state and a declaration of corporate purpose. Never mind that the corporate statutes governing this contract are, to a large extent, mandatory and not merely enabling. And do not be distracted by all the discussion of a separate entity to which corporate directors are said to owe a fiduciary obligation. This is all 'transcendental nonsense'[2] through which contractual analysis will safely guide us and, if more is needed, help us to reform. The corporate fiduciary obligation, for example, is properly owed to shareholders, not because there is anything special about that relationship, and certainly not because there is anything non-contractual in it, but only because all efficient corporate contractors would recognize that shareholders, as residual claimants, put the highest value on control of the corporation.

1 The phrase seems to have originated with Jensen and Meckling 'Theory of the Firm: Managerial Behaviour, Agency Costs and Ownership Structure' (1976) 3 *J. of Fin. Econ.* 305.

2 See F. Cohen 'Transcendental Nonsense and the Functional Approach' (1935) 35 *Columbia LR* 809 at 820.

This paper provides the beginnings of a challenge to this contractual view of the corporation. It is premised on the belief that competitive corporate contracting cannot achieve all that this view promises unless it is aided by the very value that a contractual understanding of the fiduciary obligation denies, namely, the duty of loyalty and trust. Trust plays an essential role in all modern economies, and without it, or without the coordination that is provided by institutional loyalty, even efficient wealth-maximizing corporate contracting can make us all worse off. That too, unfortunately, is the stuff of competition. Nor can the concept of trust be very easily accommodated into the contractual model of the corporation. Properly interpreted, the concepts of trust and loyalty present a deep challenge not only to that contractual model, but also to the very conventions of instrumental rationality upon which the model is based.

Section II of the paper shows how the concept of trust is important to the achievement of Pareto efficiency in even the most simple of economies involving the consumption of purely private goods (that is, where there are no externalities). Section III introduces the possibility of (pecuniary) externalities through the use of Kaldor-Hicks-efficient (or wealth-maximizing) reallocations and shows that these too can produce Pareto-inferior results. Thus, the mere 'mimicking' of trustworthy exchange will not do. Section IV shows that the problem discussed in section III is merely the contractual analogue to the well-known majority voting cycle that has already attracted the attention of some economic theorists of the corporation, most notably Jeffrey Gordon and Henry Hansmann. The section argues that the solutions proposed by Gordon and Hansmann for dealing with the voting cycle, namely restrictions on political 'voice' in the corporation, have their analogues in the need for restrictions on the 'exit' of capital from the corporation. This insight provides us with some understanding of recent initiatives to broaden the scope of the corporate fiduciary obligation to include some consideration of non-shareholder constituencies during a corporate control transaction. The contractual cycle is exemplified in section IV with an analysis of the breach of trust scenario that has recently been advanced by Andrei Shleifer and Lawrence Summers as a partial explanation of corporate takeovers, and some criticisms of that scenario, most notably those advanced by Bengt Holmstrom and Roberta Romano, are addressed. The paper concludes with a discussion in section VI of the deep challenge, and promise, that a proper understanding of institutional loyalty will provide for the conventions of economic rationality in general, and as applied in the theory of the corporation in particular.

II *Trust, money, and the Pareto principle*

While economists recognize that underdeveloped economies are often characterized by, among other things, a lack of trust between persons,[3] neither the concept of trust nor its close relation, the concept of loyalty, receive much explicit attention in their theories of efficient production and exchange.[4] Yet the most simple of examples shows that, without institutions of trust, efficient exchange may not occur, and the result may be that some social state that everyone prefers to the status quo, that is, a Pareto-superior social state, may not be achieved.

To see this, suppose that there is an economy of three individuals, 1, 2, and 3, each of whom begins with a particular piece of fruit. Specifically, individual 1 has an apple, individual 2 has a pear, and individual 3 has an orange. Assume that the three individuals, while they each like all three kinds of fruit somewhat, have preferences over the three sorts of fruit in the way described by table 1. That is, reading from top to bottom, individual 1, for example, prefers oranges to apples, and apples to pears. The initial endowment of each individual with each piece of fruit is indicated in table 1 by an asterisk. Thus, the three initial endowments described above indicate that we are beginning in social state y.

It is easy to see that there is a feasible reallocation of the three pieces of fruit that will make everyone better off, namely, the reallocation described in social state x. However, it is also easy to see that, in a pure barter economy of bilateral exchange, where each individual is assumed to be rational in the sense that she will not willingly enter into an exchange that makes her worse off, this Pareto-superior reallocation will not be achieved. For example, individual 1, who has the apple that individual 2 wants, is not interested in the pear that individual 2 can offer in exchange; individual 1 wants an orange instead. But individual 3, who has the orange, would not want to exchange that for individual 1's apple. Thus, there is no exchange possible between individual 1 and either individual 2 or individual 3. The same sorts of problems exist for any exchange that might be proposed for individual 2 and individual 3.

3 See, e.g., G. Akerlof 'The Market for "Lemons": Qualitative Uncertainty and the Market Mechanism' (1970) 84 *Q. J. of Econ.* 488; K. Arrow *The Limits of Organization* (New York: Norton 1974) 26; and P. Dasgupta 'Trust as a Commodity' in D. Gambetta (ed.) *Trust* (Oxford: Basil Blackwell 1988) 49.

4 The authors listed in the preceding note are worthy exceptions to the rule. See, in addition, G. Akerlof 'Loyalty Filters' (1983) 73 *Am. Econ. Rev.* 54. The most notable contribution from economics to the analysis of loyalty is, of course, A.O. Hirschman *Exit, Voice and Loyalty* (Cambridge: Harvard University Press 1970).

TABLE 1

	Individual 1	Individual 2	Individual 3
state *x*:	orange	apple	pear
state *y*:	*apple*	*pear*	*orange*
state *z*:	pear	orange	apple

However, the introduction of money into this economy seems to make all the difference. Suppose that one dollar is given to individual 1, and that each of the three individuals is willing to pay at least one dollar for her most preferred piece of fruit, but not as much as one dollar for the other fruits.[5] Thus, '$1' appears as the second-best outcome for each individual in table 2, which sets out this new situation in the same sort of schematic way in which the original situation was set out in table 1. The fact that individual 1 has now been endowed initially with one dollar as well as an apple is also indicated with an asterisk.

It is now easy to see that individual 1 and individual 3 can trade. Individual 1 prefers having an orange to having a dollar, and individual 3 prefers the reverse. Once this first trade is complete, of course, then individual 3 can trade the dollar for individual 2's pear, which dollar individual 2 will trade for individual 1's apple. The end result after three such trades is that the dollar is returned to individual 1, and that each of the three individuals has ended up with her most preferred piece of fruit. Pareto would surely be pleased.

It is important to appreciate, however, that money can act here as the lubricant for what would otherwise be recalcitrant bilateral exchange only if money is, and continues to be, a trusted currency, that is, a currency that continues to have value. For suppose that this is not so. Suppose in particular, for example, that money will have no value at the end of this series of trades. Then, on the same individual rationality assumption made above, namely, that no individual will willingly enter into a trade making her worse off, it is hard to see why the last trade in the sequence would be made at all. Why would individual 1 give up a perfectly edible apple for a now useless dollar? Suppose further that the irrationality of individual 1's doing so is also perfectly predictable by the other equally rational individuals 2 and 3.[6] Then it is hard to see how the earlier two

5 Since money has only exchange value in this example, any amount of money would do, and that amount would always appear as second-best in any individual's preference ordering.

6 In the theory of games, this amounts to assuming what is termed the 'common knowledge' of rationality among all players in a game. This has turned out to have

TABLE 2

Individual 1	Individual 2	Individual 3
orange	apple	pear
$1	$1	$1
apple	*pear*	*orange*
pear	orange	apple

trades involving money would take place either. Why, for example, should individual 2 accept money from individual 3 if the money will not in the end be accepted by individual 1? It seems that, without the continued value of money, the whole sequence of deals is in danger of unravelling from back to front and, therefore, of not getting off the ground at all.[7]

There is, of course, another way. Rather than assume that money is to be trusted, that is, that it will continue to have value after all the trades are complete, one could assume that individual 1 is to be trusted to accept money on the last trade with individual 2 despite the fact that money will not have any value for her at that point. Indeed, it is easy to believe that individual 1 may well have promised as much to get the first exchange off the ground. But this solution has its own problems, not least of which is why, in the first exchange between individuals 1 and 3, individual 3 should take much comfort (as she takes the dollar) from the fact that individual 1 has promised to take the dollar back from individual 2 in the third trade, a promise that is crucial to individual 2 (who is not even present at the first exchange) if she is going to agree to take the dollar from individual 3 in the second trade. At a minimum, there is a coordination or third-party beneficiary problem here; more seriously, one might think that it is time to hook up a 'conference call' and to

serious implications for determining the equilibrium of a game. On this, see K. Binmore and P. Dasgupta 'Game Theory: A Survey' in K. Binmore and P. Dasgupta (eds) *Economic Organizations as Games* (Oxford: Basil Blackwell 1986) 1–45.

7 This is an example of what is now known in game theory as the 'backward induction paradox.' Reinhard Selten seems to have been the first to provide a formal account, although he called it 'the chain store paradox' because of the particular market structure example he used to illustrate the problem; see his 'The Chain Store Paradox' (1978) 9 *Theory and Decision* 127. For excellent general discussion of the paradox, see Binmore and Dasgupta, supra note 6, 19–21; and R. Sugden 'Rational Choice: A Survey' (1991) 101 *Econ. J.* 751, 771–4. For critical analysis of the paradox and a suggested solution, see P. Pettit and R. Sugden 'The Backward Induction Paradox' (1989) 86 *J. of Phil.* 169; and for some discussion of the paradox in the context of the theory of the firm, see D. Kreps 'Corporate Culture and Economic Theory' in J. Alt and K. Shepsle (eds) *Perspectives on Positive Political Economy* (New York: Cambridge University Press 1990) 104.

abandon the idea of purely bilateral exchange altogether. But that, as we shall see, is to get ahead of our story.[8]

The other serious problem with the idea of trusting individual 1's promise to take the money in exchange for the apple on the last trade of the sequence is that, within the economic theory of rational choice, a promise to make such a last-period exchange is simply not credible.[9] Individuals choose what they prefer on the economic model, and individual 1, once she has used the money to get her orange, would clearly prefer to keep her apple to getting back the money, which is now quite useless. To be trustworthy, therefore, individual 1, at least on the economic model, must appear to be credibly irrational.

Nor is it an answer to say that there is more to the economic model of preference maximization than the consumption of fruits; some, presumably most, individuals like to get on well with others, and individual 1's threatened defection from the trustful sequence of monetary exchanges would, in all likelihood, go against any such preference. However, while this objection might suggest that this particular example is fast wearing out its welcome, the point is that it *is* just an example. For the argument to go through, the goods being made available for exchange could have been anything at all so long as the individual preferences and endowments over those goods were as described. But initial endowments and individual preferences are usually taken by the economist as given, that is, something that is prior to and beyond the reach of the usual economic analysis, and something upon which the efficient processes of voluntary exchange are said merely to operate. Thus, the general point of the argument, which the example was only meant to illustrate, comes so far to this: For some configurations of individual preferences, and some initial endowments, efficient exchange will only effect a Pareto-optimal outcome if either money can be trusted to retain its value, or some individual can be trusted to make what, on the economic model at least, is an irrational choice not to maximize her preferences.[10]

Nevertheless, it may be better that, having now provided some motivation for the more general analysis, we take the hint being offered

8 See infra, section VI.

9 The analysis of exactly when interpersonal assurances (and threats) can be made credibly is a major issue in game theory. See discussion in Sugden 'Rational Choice' supra note 6, 778–2; and Binmore and Dasgupta, supra note 6, 20.

10 That some kind of 'irrationality' (at least as an ex ante possibility) might be needed to avoid the 'backward induction paradox' has now been generally recognized in game theory. See, e.g., Kreps, supra note 7, 104.

in the last paragraph, and leave behind all discussion of individual preferences for fruit. By first abstracting to a more general set of individual preferences over a range of unspecified goods, such as A, B, and C, we will be in a better position to concentrate on the pure *structure* of what I believe is a very general problem plaguing both market and political processes alike. Then, in sections IV and V, we can fill in these abstract scenarios with some details from the sorts of corporate control transactions that are the more immediate concern of this paper. In the final analysis, we can reasonably hope, I think, that all of this will allow us both to understand better the general sources of obligation that exist between different corporate constituencies, and to make some quite specific points about how to structure better both market 'exit' from, and political 'voice' within, the corporation.[11]

III *Kaldor-Hicks-efficient exchange and the Pareto principle*

So, in aid of this, consider table 3, which essentially reproduces table 2 except that all references to fruit have been replaced with references to three generic goods, A, B, and C. The three individuals have the corresponding preferences over these goods that they had in table 2, and there is also for each a willingness to pay some amount of money, denoted by the symbol $, for her most preferred good, which exceeds the amount she is willing to pay for the other two goods, whatever they might be in each case.

Now consider the possibility of a series of Kaldor-Hicks 'trades.'[12] Kaldor-Hicks trades, of course, are not really trades at all. Rather, under the Kaldor-Hicks efficiency criteria, goods are to be allocated to those whose willingness-to-pay for the goods is greatest. This could involve allocating a good away from one individual to another. However, and this is the crucial difference between a Kaldor-Hicks-superior and a Pareto-superior reallocation, there is no need to pay compensation to those against whom the reallocation is made to ensure that they are not made any the worse off by it. In this sense Kaldor-Hicks exchanges are involuntary exchanges.

The importance of the Kaldor-Hicks efficiency test is usually thought to be that it is more permissive than Pareto efficiency. It is probably

11 The terms voice and exit are, of course, borrowed from Hirschman, supra note 4.

12 For an excellent summary and analysis of the notions of efficiency that are used in economic analysis, including Kaldor-Hicks efficiency as used in the economic analysis of law, see J. Coleman *Markets, Morals and the Law* (New York: Cambridge University Press 1988) 67–132.

TABLE 3

Individual 1	Individual 2	Individual 3
A	B	C
$	$	$
B	*C*	*A*
C	A	B

inappropriate as an ethical matter to insist that all reallocations be Pareto-superior (some losses are deserved after all, and should not be compensated as a matter of right[13]), and it is certainly far too demanding as a practical matter that all reallocations secure the voluntary acquiescence of all possible losers by paying them actual compensation in every case. The best that can be hoped for, and the regime that rational individuals would agree to in some suitably defined ex ante position, it is often argued,[14] is the one that promises to provide a sufficient number of Kaldor-Hicks-efficient reallocations over time that everyone can reasonably expect ex ante to come out ahead.

This is the general argument that justifies the imposition of competitive harms, or pecuniary externalities,[15] for example. Displaced competitors are simply those individuals who were unwilling, or unable, to pay as much as some other party in the relevant auction that took place for some scarce resource. To be sure, in these cases the extra willingness-to-pay takes the form of an additional willingness to reduce prices, or wages, to secure contracts, but that is a difference of no real moment in the market model. Moreover, the fact that there are utility losses in such cases that lie beyond the reach and measurement of willingness-to-pay also makes no difference; that is true of all Kaldor-Hicks-efficient reallocations. Again, the promise is not to secure individuals against losses, but rather to increase total wealth and, therefore, everyone's ex ante prospects for a welfare gain after a sequence of such Kaldor-Hicks reallocations have been completed.

13 Ibid. 92–3

14 See, e.g., R. Posner *The Economics of Justice* (Cambridge: Harvard University Press 1981) chapter 4.

15 The term pecuniary is used to convey the idea that the externality occurs through the market mechanism rather than outside it; for discussion, see W. Heller and D. Starrett 'On the Nature of Externalities' in S.A. Lin (ed.) *Theory and Measurement of Economic Externalities* (New York: Academic Press 1976) 16. For the efficiency argument in favour of not compensating such competitive harms, see H. Demsetz 'The Exchange and Enforcement of Property Rights' (1964) 7 *J. of Law & Econ.* 11, 25-6.

That there is some promise in this approach is easily illustrated by way of the example laid out in table 3. Unlike the situation discussed in table 2, where actual voluntary exchange was to take the parties by a sequence of apparently Pareto-superior transactions to a final genuinely Pareto-superior social state, under Kaldor-Hicks 'trades' it is enough to observe (somehow[16]) that, for example, individual 1's willingness-to-pay for good A exceeds individual 3's, and to reallocate the good from 3 to 1 for that reason. Similarly, because individuals 2 and 3 have a greater willingness-to-pay for goods B and C respectively, the reallocation to them of those goods is also Kaldor-Hicks-efficient. Of course, each reallocation, taken in isolation, produces a loser. But that is exactly what Kaldor-Hicks efficiency, unlike Pareto efficiency, permits. Moreover, what has been accomplished after a sequence of such involuntary trades is precisely what is endangered by an insistence upon only voluntary exchanges, namely, the final achievement of the Pareto-superior social state described by the first row of the preference matrix in table 3.

This is the good news about Kaldor-Hicks efficiency and the process of competitive contracting for which it provides some normative support. The bad news is that, without more, and in particular without some attention to the very same sorts of individual preferences that, taken as given, were so problematic for the model of monetary exchange discussed above, we cannot be sure that a sequence of Kaldor-Hicks-efficient exchanges will not, just as easily, bring about a result that is Pareto-*inferior* to the social state from which the sequence started.

To see this, consider what it is that we actually know if all we can observe is the three Kaldor-Hicks-efficient reallocations of the three goods A, B, and C described above. Refer to the status quo social state from which the sequence of trades begins as social state y. Then, the first Kaldor-Hicks reallocation, call it social state z, which on willingness-to-pay criteria reallocates good A from individual 3 to individual 1, clearly makes individual 3 worse off and individual 1 better off. Alternatively, we could say (borrowing from the conventions of social choice theory) that individual 3 prefers y to z, denoting this symbolically as yP_3z, and that individual 1 prefers z to y, denoting this symbolically as zP_1y. ('P' should be read 'is preferred by __ to,' with the blank filled in by the individual

16 How these observations would be made, of course, is not always obvious. That is why the economist is predisposed, so far as is possible, towards the use of actual rather than merely hypothetical bargains; the former pose fewer empirical problems. However, as the argument in this paper shows, the advantages of actual over hypothetical exchange are more than merely empirical. They go to the very possibility of Pareto-inferiority itself. See infra, text at note 17.

named by the subscript.) Individual 2, whose initial endowment of goods is unaffected by this first Kaldor-Hicks reallocation, is, therefore, indifferent between z and y. This we can denote symbolically as zI_2y (where 'I' is read 'is considered indifferent by __ to,' again with the blank referring to a subscripted individual). Now consider the second Kaldor-Hicks reallocation, call it social state w, which transfers good C from individual 2 to individual 3 without compensation on the basis of individual 3's greater willingness-to-pay. Using the symbolic notation just introduced, it seems fair to say that wP_3z and zP_2w, because of the transfer, and that zI_1w, because this Kaldor-Hicks reallocation leaves individual 1 unaffected. Finally, consider the third Kaldor-Hicks reallocation, social state x, which transfers good B from individual 1 to individual 2 without compensation. Clearly, wP_1x and xP_2w, because of the transfer, and wI_3x, because individual 3 is unaffected.

Table 4 summarizes, therefore, what we know of these individuals' preferences after the sequence of three Kaldor-Hicks reallocations that brought us from the status quo ante, social state y, to our end point, social state x. The important thing to note about these preferences is that for each and every individual the preference ranking between social state x and social state y is so far undetermined. Could it be that the sequence of Kaldor-Hicks-efficient trades has only resulted in a Pareto-inferior social state, that is, that for all individuals i, yP_ix?

To see that such a result is not inconsistent with the preferences which are implied by the three Kaldor-Hicks transfers, it is enough to introduce this possibility (that is, of yP_ix for each individual i) explicitly into each of the three rows in table 4. This allows for a full and consistent determination of each individual's preference ordering over all four of the social states, w, x, y, and z, as shown in table 5 (with social states ordered by preference for each individual from top to bottom in each column, and indifferent social states for each individual appearing in the same row). Given these orderings over the social states, the three Kaldor-Hicks-efficient transfers from y to z, from z to w, and from w to x, add up to a Pareto-inferior outcome, or an outcome in which every individual is worse off than when she started.

This cannot happen if compensation is actually paid to those individuals who lose in Kaldor-Hicks reallocations. For if compensation is actually paid, then each individual in table 3 actually occupies one of the three top slots in her preference ordering as the sequence of (now) Pareto-efficient exchanges unfolds. But it is impossible for individual 2, for example, to prefer actually having the '$' to her initial endowment with good C (a preference that was necessary for her contract with individual 3 to go through), and also to prefer actually having good B to

TABLE 4

	Trade 1	Trade 2	Trade 3	By transitivity
Individual 1	zP_1y	wI_1z	wP_1x	wP_1y , zP_1x
Individual 2	zI_2y	zP_2w	xP_2w	yP_2w
Individual 3	yP_3z	wP_3z	xI_3w	xP_3z

TABLE 5

Individual 1	Individual 2	Individual 3
w, z	z, y	y
y	x	x, w
x	w	z

Note: Assume for all i, yP_ix

the '$' (a preference essential to the last trade with individual 1), and yet also prefer having good C to good B, as would be implied if all individuals preferred social state y to social state x. Such a preference ordering over the goods (including the money) would be intransitive. So, not surprisingly, a sequence of Pareto-efficient exchanges, with actual compensation being paid, cannot bring about a Pareto-inferior result.

What may be surprising to some, however, is the fact that the Kaldor-Hicks exchanges that are designed to 'mimic'[17] these very same Pareto-efficient exchanges can do precisely that. Thus, the difference between Kaldor-Hicks-efficient and Pareto-efficient exchanges over time may not simply be a matter of the distribution, or who gets what share (positive or negative), of the aggregate social gains, as is so often implied. Rather, the difference might be that the latter sorts of exchange ensure there are genuine social gains in a way that the former do not. The former, as illustrated here, can leave us all worse off.

Before going on to consider in the next section whether the individual preferences giving rise to these sorts of difficulties are likely to be confronted in the real world, it is worthwhile trying to summarize the general thrust of the argument so far. We began in section II by showing that certain individual preferences could present difficulties for the achievement of Pareto-superior social states by way of bilateral exchange unless there was a certain amount of trust, or a trusted currency, in the economy. However, we also suggested that the very idea of trust, which seemed so essential to achieving Pareto efficiency, might be problematic

17 Richard Posner is the best-known proponent of 'mimicking' the market; see generally his *Economic Analysis of Law* 3d ed. (Boston: Little, Brown 1986).

within the usual economic understandings of individual rationality. Trust means that some individual, somewhere in the economy, must be making a voluntary choice against the alternative that would maximize her preferences or, at least, must be able to get others to believe that she will do just that. In either case, this would involve a relaxation of the usual rationality assumptions of economic theory.

We then considered in this section the possibility of Kaldor-Hicks-efficient exchanges. Part of the motivation for doing so was the fact that Kaldor-Hicks efficiency held out the promise of 'mimicking' the exchanges of a purely voluntary market without paying actual monetary compensation. Thus, no one individual's reluctance to take money on the last trade of a voluntary sequence of monetary exchanges could possibly hold everything up, thereby frustrating the achievement of Pareto efficiency. So far, so good. The problem was that a sequence of Kaldor-Hicks-efficient exchanges could give rise to its own problems of Pareto inferiority. When one recognizes that competitive contracting, and the harms that follow from market competition, are typically justified because they are Kaldor-Hicks-efficient, then this last argument would seem to have some rather far-reaching implications.

The only ways to ensure that a Pareto-inferior social state is not the consequence of a sequence of Kaldor-Hicks competitive contracts are:

(i) By providing some compensation for competitive harms, thus effectively converting Kaldor-Hicks-efficient exchanges into something more like Pareto-efficient exchanges. In the context of corporate control transactions, this has already been proposed by John Coffee, although for political reasons that are quite different from those being offered here;[18]

(ii) By restricting the domain of preferences that are admissible into Kaldor-Hicks competitive contracting so that the particular profile of individual preferences that gives rise to the difficulties of Pareto inferiority is not present. Since there is always something problematic about 'laundering preferences'[19] in advance, as a practical matter this control technique usually reduces either to expressing the hope or belief

18 See J. Coffee 'Unstable Coalitions: Corporate Governance as a Multi-Player Game' (1990) 78 *Georgetown LJ* 1495, 1548. Coffee's argument is that it may be essential to compensate the losers from Kaldor-Hicks efficient takeovers to preserve the political viability of these largely efficient corporate control transactions. Otherwise, the losers from takeovers, who are well placed and concentrated within the relevant jurisdictions to organize effectively on the political front, will have legislation passed to restrict the number of takeovers that will be permitted.

19 But see R. Goodin 'Laundering Preferences' in J. Elster and A. Hylland (eds) *Foundations of Social Choice Theory* (Cambridge: Cambridge University Press 1986) for discussion sympathetic to this idea.

that such preferences, while problematic in theory, are unimportant in real life, or to adopting control technique (iii) (see below). In section IV we shall see that this particular problematic preference profile can be expected to occur quite frequently in real life and, indeed, has been recognized as such in some recent analyses of possible voting problems, or problems associated with political 'voice,' in corporate governance.[20] One of the goals of the present analysis is to show that the same profile of individual preferences presents analogous difficulties for competitive contracting, or market 'exit,' in this context; and

(iii) By accepting that the problematic preference profile is present, and then adjusting either initial endowments or rights of contract accordingly so that the preference profile cannot enter decisively into a sequence of Kaldor-Hicks transactions leading to a Pareto-inferior result. This control technique might show up in the context of corporate control transactions as a proposal to give certain stakeholders new property rights,[21] but it is more likely to surface as a claim that corporate capital should have a more restricted contractual right to 'exit' from the corporation. As sections V and VI will show, this is one plausible way to interpret recent legislation that imposes on corporate management a fiduciary obligation to non-shareholder constituencies in the corporation.

In the next section of the paper, we begin with (ii), or the possibility that the problematic preference profile discussed above is not likely to be present in any decisively relevant way in a real-world economy. We show that this is likely to be a vain hope in general and, indeed, that the problematic preference profile has already surfaced in some economic analyses of corporate governance. Most of these analyses focus on political governance, concentrating on the problems that arise for stability of decision-making if there is unrestricted voting or 'voice' in the corporation. But others, to be examined in section V, are more traditionally contractual in nature and show the importance of trust and loyalty, or limitations on market 'exit,' for efficient contracting. It is the purpose of the next section, building on the analysis so far, to show that both sorts of analysis, the political as well as the contractual, are essentially dealing with the same problematic preference profile. We will then go on to

20 See J. Gordon 'Shareholder Initiative and Delegation: A Social Choice and Game Theoretic Approach to Corporate Law' (1991) 60 *U. Cin. LR* 347; H. Hansmann 'When Does Worker Ownership Work? ESOPs, Law Firms, Codetermination, and Economic Democracy' (1991) 99 *Yale LJ* 1749; and L. Benham and P. Keefer (1991) 'Voting in Firms: The Role of Agenda Control, Size and Voter Homogeneity' 29 *Econ. Inquiry* 706.

21 See, e.g., J. Singer 'The Reliance Interest in Property' (1988) 40 *Stan. LR* 611.

provide interpretations of some of the institutional and public policy responses that have been proposed in this context to see what light our analysis might be able shed on them.

Section VI of the paper will then return to the theme of trust and loyalty introduced in section II, and will attempt to sketch some of the larger implications for the theory of corporate law obligations that might be at stake in this analysis. As at least one commentator has already suggested elsewhere, what might be at stake is a proper understanding of the important role that must be played by 'corporate culture' in coordinating the economy of corporate control transactions.[22]

IV *Voting cycles and contractual cycles in corporate law*

In several recent papers, different authors have pointed to the difficulties that might arise for the corporation from the so-called problem of cyclical majorities.[23] The problem of cyclical majorities arises when each and every option selected by a majority vote can in turn be defeated by another majority coalition. Since majority rule seems doomed under such a scenario to settle (if settle it can) on a minority-preferred option, the problem is also referred to as the paradox of majority voting.

This paradox is just one of the many real-world manifestations of Arrow's general impossibility theorem,[24] but it is clearly the best known. Since Arrow himself first got interested in the more general problem of social choice by way of certain speculations about voting in the corporation,[25] it should not be surprising, perhaps, that recent analysts have returned to the corporate voting context to find that the general pathologies of social choice are manifested there most clearly. However, it is the burden of sections IV and V of this paper to demonstrate that it is a mistake to think that the general problem of social choice is restricted

22 See Kreps, supra note 7; see also R. Daniels 'The Law Firm as an Efficient Community' (1992) 37 *McGill LJ* 801. Daniels uses the idea of law firm culture to explain the different forms in which law firms might choose to grow (e.g., lateral recruitment, merger, branch office, etc.). I am grateful to Ron Daniels for first suggesting to me the possible importance for economic efficiency that might inhere in corporate culture.

23 See, e.g., Gordon, Hansmann, and Benham and Keefer, supra note 20. Also see F. Easterbrook and D. Fischel 'Voting in Corporate Law' (1983) 26 *J. of Law & Econ.* 395, 405 and 409–10, for an argument that the additional representation that is given to different groups by cumulative voting for executive board members will lead to voting cycles.

24 See K. Arrow *Social Choice and Individual Values* rev. ed. (New Haven: Yale University Press 1963).

25 See the interview with Arrow published in *Social Choice and Welfare*.

to voting. The same problematic profile of individual preferences that plagues majority voting in the corporation also shows up in the context of competitive corporate contracting, and to the same effect. Thus, just as 'restricted voice' has been trotted out by the above-mentioned authors as a method of securing a 'structure induced' equilibrium against political cycling,[26] so 'restricted exit' needs to be used to ensure that there is no contractual cycling, particularly the sort of contractual cycling that takes the contractors through Pareto-inferior social states.

Jeffrey Gordon is one of those authors who has argued very effectively that the corporation might confront a majority voting cycle among its own shareholders if those shareholders were given unlimited access to the control of the day-to-day business of the corporation.[27] Gordon suggests that 'the absolute delegation doctrine,' which effectively restricts shareholder voting to overall control contests in which a whole new board of directors must be elected to replace an old one, and which strictly controls more focused shareholder business initiatives, is a device designed to restrict shareholder voice and to avoid problems of majority cycling.

Henry Hansmann has made a similar point in the context of worker ownership.[28] Rejecting the usual arguments that have been offered to explain the existing pattern of worker ownership and control across industries, Hansmann argues that a better explanation can be found once one recognizes the sorts of problems that can arise for worker ownership and control if worker interests in the corporation are very diverse. Hansmann's view is that worker ownership and control is much less likely to be found, whatever else the advantages, in situations where worker interests are heterogeneous. While in no way limiting his analysis to the problems of majority cycling that might be the result of such worker-as-voter diversity, it is clear that this is a large part of what Hansmann has in mind.[29] Where there does appear to be worker ownership combined

26 A 'structure induced' equilibrium is one in which the equilibrium is achieved because of some feature of institutional design rather than because there is no problematic preference profile existing among the voters of the relevant jurisdiction. An equilibrium achieved because of the latter rather happy set of circumstances is said to be 'preference induced.' See K. Shepsle 'Institutional Arrangements and Equilibrium in Multidimensional Models' (1979) 23 *Am. Pol. Sci. Rev.* 27. Richard McKelvey has shown that it is unlikely that equilibrium will be preference induced since the sorts of preferences that give rise to the majority voting paradox would be the rule rather than the exception in cases where goods are valued on more than one dimension. See his 'Intransitivities in Multidimensional Voting Models and Some Implications for Agenda Control' (1976) 12 *J. of Econ. Theory* 472.

27 Gordon, supra note 20

28 Hansmann, supra note 20

29 Ibid. 1781

with a great deal of worker heterogeneity, as, for example, in many of the large employee stock ownership plans (ESOPs) in the United States, Hansmann argues convincingly that this is usually a case of worker ownership without worker voting power or control.[30] On this last point he is supported by some recent work by Lee Benham and Philip Keefer, which focuses almost exclusively on the problem of majority cycling.[31] Thus, just as for Gordon's stockholders, so for Hansmann's worker-owners, the solution to the problem of majority cycling takes the form of restrictions being imposed on voting, or voice, in corporate governance.

For all of these authors, the view seems to be that the problem of cycling is an essentially political problem, or a problem that only plagues the corporation when it chooses to make its decisions by a method of unrestricted majority voting. There is never any suggestion that the same kind of difficulty might show up for the corporation in market contracting. This apparent advantage of the market over political mechanisms of choice is, perhaps, emphasized most explicitly by Hansmann in the conclusion to his paper:

> It is paradoxical that the aspect of worker ownership often extolled as its principal virtue – participation in governance of the firm through democratic institutions – appears in fact to be its greatest liability. In business firms as in other settings, political mechanisms operate crudely. In comparison to market contracting, they evidently aggregate and represent the interests of a heterogeneous group relatively inefficiently in all but the most severe situations of market failure. Even then, fiduciary mechanisms may generally be a more effective substitute for the market than are political institutions: Firms managed *for* their workers, it appears, often perform better than firms managed *by* their workers.[32]

The mention of fiduciary mechanisms in this passage is a reference to, among other things, the fact that in most ESOPs, for example, voting power is effectively exercised, not by the workers themselves, but by a trustee acting on behalf of the workers in the plan.[33] Thus, any movement away from market contracting in firms adopting an ESOP should not be viewed too simply, according to Hansmann, as the substitution of political devices to protect worker interests in the firm for the more usual devices of labour market contracting. Rather, the fiduciary obligation that effectively operates within such plans, Hansmann suggests, is better

30 Ibid. 1797–1800. Hansmann also argues that the prevalent use of ESOPs can be explained by the particular tax advantages that such plans enjoy.

31 Benham and Keefer, supra note 20

32 Hansmann, supra note 20, 1816

33 Ibid. 1797–1800

interpreted as some kind of third alternative to both market *and* political methods of making social choices.[34]

This last point is an interesting one, but it is limited by its suggestion that the fiduciary mechanism operates as a solution to prevent inefficient cycling only within an ESOP, and only on the political side of the social choice continuum between market contracting and political voting. In the discussion to follow in section V, I shall argue that a cycling problem, exactly analogous to the majority voting cycle, can be seen to be operating within the context of corporate control *market* transactions. Moreover, I shall argue that the same sort of fiduciary obligation as that proposed by Hansmann, only this time one limiting the exit of capital into the market rather than the voice of labour in voting, may well be what is required to prevent this contractual cycle's worst inefficiencies.

However, to draw the parallel between the two sorts of cycle, we have to consider first the sorts of individual preferences that must exist within a voting population for a majority voting cycle to occur.[35] Suppose, in particular, that a majority of voters have indicated a preference for social state w over social state z, for social state z over social state y, and, consistent with a cycle, for social state y over social state w. It is easy to prove that, for such a cycle to occur, the profile of individual preferences shown in table 6 must be present somewhere in the voting population.[36] To see that this is so, consider, for example, that group of individuals who formed the decisive majority for w over z. Not everyone in this

34 Hansmann characterizes the ESOP more as a non-profit institution, like a voluntary association, than an institution allowing for significant workplace democracy and control; see Hansmann, supra note 20, 1799. This characterization also suggests that ESOPs are better thought of as part of 'the third sector' than as part of the more conventional political or market-based sectors.

35 This will give us a more precise sense of what Hansmann might mean when he argues that heterogeneity of interests among workers is problematic for worker control and political participation.

36 The following proof is based on R. Sugden *The Political Economy of Public Choice* (Oxford: Martin Robertson 1981) 157. It should be noted how this proof requires the use of the neutrality condition, or the condition that requires social choice to be 'invariant' as a given profile of individual preferences is permeated across a different set of alternatives for choice. Thus, if a given majority of individuals is decisive within a population of voters for a particular alternative, then that same majority should continue to be decisive for another alternative if the preferences of all the individuals in this decisive majority correspond. The neutrality condition was first introduced into the social choice literature by K.O. May as one of the defining conditions of simple majority rule. For a proof and formalization, see May 'A Set of Independent, Necessary and Sufficient Conditions for Simple Majority Decision' (1952) 20 *Econometrica* 680. For a general discussion of the important role which various 'invariance' conditions play in social choice theory, see A. Sen *Choice, Welfare and Measurement* (Cambridge: MIT Press 1982) 228–9.

TABLE 6
Preference orderings

Type-I	Type-II	Type-III
w	z	y
z	y	w
y	w	z

group of individuals can also prefer y to z since that would imply that the same majority would be decisive for y over z, a contradiction. Therefore, someone in this group who prefers w to z must also prefer z to y. But this means that someone must have the type-I preference ordering in table 6. Similarly, one can show that someone in the group that was decisive for z over y must prefer y to w, or have the type-II preference ordering in table 6, and that someone in the decisive majority for y over w must prefer w to z, or have preference ordering type-III in the table. Thus, all three preference orderings in table 6 must be present somewhere in the voting population for this particular majority voting cycle to occur.[37]

There are two points worth emphasizing about this derivation of a preference profile that is problematic for majority voting. First, while the preference profile shown in table 6 is *necessary* for the majority voting cycle, w over z, z over y, and y over w, to occur, it is not *sufficient* for it. Such a preference profile could be present in the population, for example, yet not give rise to the cycle if, say, a simple majority of all the individuals had preference ordering type-I. Then, that same group would be decisive for all majority votes, and there would be no voting cycle despite the possible presence in the voting population of the other two types of preference orderings shown in table 6.

Second, and this is a point of more immediate concern to this paper, it is not necessary that *all* the social states that might form part of a majority voting cycle be involved in a preference profile like that shown in table 6; it is enough if one proper subset of *three* of the alternative social states is so involved. For example, consider the individual preference profile shown in table 7. Table 7 reproduces the individual preference orderings of table 6 over the three alternative social states w, z, and y, but adds social state x, which is a Pareto-inferior social state to social state y. Obviously, a Pareto-inferior social state cannot be the best

37 For closely related but more general discussion of the conditions that, if imposed on individual preferences, will be sufficient for avoiding the majority voting paradox, see P. Pattanaik and A. Sen 'Necessary and Sufficient Conditions for Rational Choice Under Majority Decision' (1969) 1 *J. of Econ. Theory*, reprinted in Sen, supra note 36, 135–57.

TABLE 7
Preference orderings

Type-I	Type-II	Type-III
w	*z*	*y*
z	*y*	*x*
y	*x*	*w*
x	*w*	*z*

available alternative for anyone, and so it cannot be part of an individual preference profile analogous to that shown in table 6 (where *x* would be substituted for one of the other three alternatives shown there), since that profile shows every alternative as occupying top spot for at least one individual. Nevertheless, a majority might cycle through the Pareto-inferior social state *x* (and, possibly, end up choosing it), by choosing *z* over *y*, *w* over *z*, and *x* over *w*. Thus, table 7 shows a way for the method of majority voting to choose, or cycle through, not only minority-preferred alternative social states, but Pareto-inferior ones as well.

By now it should be apparent that we are coming back to the discussion we temporarily left behind in section III where we first introduced the possibility that a sequence of Kaldor-Hicks-efficient (or market-competitive) trades might lead to a Pareto-inferior social state. Compare the problematic individual preference profile in table 5, which gave rise to that result, with the profile of individual preferences in table 7, which, suitably distributed across a population of voters, could give rise to a majority voting cycle, and, therefore, the consequent choice by a majority of a Pareto-inferior social state. The similarities are striking. Where, in table 7, a majority of voters might begin by choosing *z* over *y*, then *w* over *z*, and finally *x* over *w*, thus ending up with a Pareto-inferior social choice (unless, of course, the cycle is made complete and all is returned to the status quo ante), in table 5 the same moves are accomplished, not by majority voting, but by a sequence of Kaldor-Hicks-efficient trades.

For example, social state *z* is a Kaldor-Hicks-efficient choice over social state *y*, a choice that favours individual 1, prejudices individual 3, and leaves individual 2 indifferent. Thus, what is accomplished *politically* by the type-I voters in table 7, when they have sufficient political support from type-II voters to form a majority coalition against the type-III voters, is accomplished as a matter of Kaldor-Hicks-efficient exchange, or (wealth-maximizing) *contractual right*, by those individuals in table 5 who have the type-I ordering, for example, individual 1, at the expense of those individuals with the type-III ordering, for example, individual 3.

Indeed, we can draw the parallel even tighter if we allow the individuals who are unaffected by a given Kaldor-Hicks exchange, and who, therefore, might be expected to be indifferent about it, to actually show a slight preference for the Kaldor-Hicks transfer. Such an assumption, while not important for the Kaldor-Hicks recommendation in any way,[38] would convert the orderings of table 5 into exactly those shown for the majority voting cycle in table 7. Moreover, the same can be said of each of the other two moves in the sequence leading to the Pareto-inferior social state x. Thus, given the profiles of individual preferences revealed in tables 7 and 5 respectively, each of the moves towards Pareto inferiority can be accomplished either politically, by the method of majority voting, or contractually, in a series of Kaldor-Hicks-efficient trades.

It is a mistake, therefore, to think that the very same sorts of heterogeneous interests that give rise to the majority voting cycle, and that, quite rightly, have been identified as problematic for majority voting in the corporation, cannot show up, at least as a theoretical matter, in fully analogous form in competitive market contracting in general, and in competitive corporate contracting in particular. However, it remains to be seen whether the sorts of solutions referred to by Gordon and Hansmann in their respective discussions of the problems posed by majority voting in the corporation, namely those involving restrictions on political voice and the use of a fiduciary mechanism, have an analogous role to play in the avoidance of some kind of Kaldor-Hicks contractual cycle involving the corporation. In the next section of this paper we shall see that indeed they do. However, to demonstrate this convincingly, we must first say something about how corporate control transactions might affect different constituencies, or stakeholders, in the corporation. This will provide the context necessary for showing both that there is a Kaldor-Hicks contractual cycle at work in the corporate world, and, further, that there is a need for restrictions on the exit of capital from the

38 It seems plausible to think that the Kaldor-Hicks decision rule satisfies 'non-negative responsiveness.' This is the condition, sometimes used in social choice theory, which requires that a social decision rule not be perverse in its response to a change in individual preferences. In particular, if the social decision was for some alternative t over another alternative s before the change in individual preferences, and then some of the individuals who previously preferred s to t came to prefer t to s (with no other changes in individual preferences being involved), then the social decision rule should not, perversely, see this as an opportunity now to indicate a choice for s over t. For a formal definition and discussion of non-negative responsiveness, see A. Sen *Collective Choice and Social Welfare* (San Francisco: Holden-Day 1970) 67–77. Simple majority voting has the analogous but stronger condition, 'positive responsiveness,' as one of its defining conditions; see Sen, ibid. and May, supra note 36.

corporation that are fully analogous to those restrictions on political voice that have been proposed by both Gordon and Hansmann for the purpose of avoiding the majority voting cycle.

V *Contractual cycles and shareholder breach of trust*

A. AN INTERPRETATION OF THE BREACH OF TRUST SCENARIO

The previous section showed the close relationship that exists between the Kaldor-Hicks contractual cycle, which was first introduced in section II, and the majority voting cycle, which has received so much attention from the corporate theorists discussed in section IV. The argument was that the same problematic profile of individual preferences can explain both sorts of cycle, and both sorts of cycle can lead to the same choice of a Pareto-inferior social state. Since the majority voting cycle has already received some attention from corporate theorists, the suggestion was that it was now time to give the same attention to the Kaldor-Hicks contractual cycle. It is quite possible that the solutions proposed for the voting cycle, in particular those restrictions that have been imposed on the exercise of political voice and the need that has been suggested for a fiduciary to act as an intermediary between corporate decision-making and certain corporate beneficiaries, will have exact analogues in the contractual context.

It might be objected, however, that while the discussion to this point has demonstrated the *theoretical* possibility of a contractual cycle that is fully analogous to the more frequently analysed voting cycle, there is little in the argument so far to show that the contractual cycle is a problem with *practical* significance for the running of the corporation. However, this is mere wishful thinking. As it happens, there is nothing particularly 'odd' about the problematic preference profile that appears in table 7, and some recent analyses of the effects certain corporate control transactions have on different non-shareholder constituencies in the corporation, particularly employees, very readily lend themselves to interpretation in its terms.

To see this, consider the argument originally provided by Andrei Shleifer and Lawrence H. Summers,[39] and more recently adopted in large part by John Coffee,[40] which puts forth the idea that a source of at least some of the large shareholder gains from hostile takeovers and other corporate restructurings is to be found in the breach of certain implicit

39 A. Shleifer and L. Summers 'Breach of Trust in Hostile Takeovers' in A. Auerbach (ed.) *Corporate Takeovers: Causes and Consequences* (Chicago: University of Chicago Press 1988) 33–56

40 Coffee, supra note 18

contracts understood to exist as a matter of trust between the incumbent management and employees of the target firm. Breach of these implicit contracts, it is said, takes the form of the opportunistic appropriation by new corporate acquirers of certain up front (ex ante) wage concessions or, more generally, firm-specific investments, which the employees might have provided to the corporation on the understanding that higher wages would later be forthcoming from the corporation as part of a longer-run wage deal or (ex post) 'settling up.'

This breach of trust scenario can be interpreted in terms of the problematic preference profile that gives rise to the Kaldor-Hicks contractual cycle, the only difference being that the Pareto-inferior social state is achieved under this scenario as a way of avoiding the cycle rather than as a consequence of going through it. To see this, suppose that a firm is in a social state in which there is much protracted bargaining between its employees and management. Every detail is worked out in advance, but the costs of transacting this way are high, both in terms of the direct costs of negotiation and drafting, and in terms of the adversarially charged atmosphere that seems to be the result of the whole process. Call this status quo social state x.

Now suppose that someone proposes that the employees and management meet on their own without the lawyers, that they proceed less 'legalistically,' trusting that they will be able to work problems out as they meet them according to some much more general and implicit 'understanding' of appropriate behaviour. Call this social state y, and suppose that it is better than x for employees, management, and shareholders alike. This seems plausible, at least if one thinks of social state x as a state of affairs in which large amounts of the firm's resources are being consumed by a process of costly rent-seeking.

However, the employees are worried that a large part of the implicit understandings in social state y will involve the corporation making deferred payments to employees for past efforts. In particular, they are worried that these deferred payments will represent a growing wage fund that might attract a corporate acquirer, a person who would not feel bound to any prior understandings to which it was not a party. The threatened loss of this wage fund, call it social state z, would make the employees even worse off than they were in social state x, although the shareholders in the corporation, together with management (if they can engineer a share of the deal through, say, some 'golden parachute' arrangement), would probably prefer z, or selling into such an acquisition, to carrying on in either social state x or social state y.

The employees have a further option, however. Confronted with the threat of a takeover by a corporate acquirer that is not bound by any prior implicit understandings, they might be in a position to propose to

management that they take a large equity stake in the company in the form of an ESOP. Call this social state *w*. This has proven to be an effective defensive tactic in the past, and one which the courts have been more reluctant to set aside.[41] It also gives the employees some voice in the running of the corporation even if it is only through a trustee. While social state *w* would involve the employees in even more non-diversified risk than they have had up until now as workers with all their human capital locked into one particular firm,[42] and, therefore, is inferior for the employees to both social state *x* and social state *y*, it is clearly to be preferred by them to social state *z* where their wage fund is threatened with opportunistic appropriation by an acquirer. Management prefers social state *w* most of all, of course; in *w* they would have a legally enforceable defensive tactic against all takeovers and, therefore, the sort of job security they have not hitherto enjoyed. Shareholders would consider social state *w* the worst of all possible outcomes for exactly the same reasons; such job security would undermine the incentive management currently has for the efficient running of the firm by isolating it from the disciplinary effects of the market for corporate control.

Thus, we have the preference profile shown in table 8 for these three different corporate constituencies over these four possible social states. These preferences correspond exactly to those in table 7, which are those of the majority voting cycle and (with only the slight variations on table 5 already mentioned) those of the Kaldor-Hicks contractual cycle.

Thus, the breach of trust scenario exemplifies the Kaldor-Hicks contractual cycle as follows: The status quo social state *x*, in which labour negotiations are costly and adversarial, is considered Pareto-inferior to social state *y* in which more of the contracting is left to implicit understandings. Thus, for all individuals i, $yP_i x$. However, once in social state *y*, management and shareholders would want to form a coalition to arrange the sale of the corporation to a new acquirer since for both management and shareholders, $zP_{ms}y$, where now the subscripts refer to the Kaldor-Hicks decisive contractual coalition of management (m) and shareholders (s) who both share a preference for *z* over *y*. Confronted with the

41 Ibid. 1527, note 119. Coffee cites the case of *Shamrock Holdings Inc.* v. *Polaroid Corp.* 559 A.2d 257, 275–6 (Del. Ch. 1989) in which the holding appears to be that a 'shareholder neutral' ESOP (that is, one funded through employee wage concessions rather than share dilution) is intrinsically fair and, therefore, not subject to the same degree of judicial supervision as other defensive actions undertaken by the board in the course of a takeover bid.

42 See J. Macey 'Externalities, Firm-Specific Capital Investments, and the Legal Treatment of Fundamental Corporate Changes' (1989) 92 *Duke LJ* 173, 194.

TABLE 8

Management	Shareholders	Employees
w	z	y
z	y	x
y	x	w
x	w	z

Note: x = the status quo explicit bargaining; y = trust and implicit bargaining; z = threat of acquisition and appropriation of deferred wages; and w = adoption of ESOP as defensive tactic.

worst of all possible social states z, however, the employees (e) would want to contract for management's cooperation in securing social state w, or the ESOP alternative, perhaps by providing some kind of wage concession to management in exchange for an equity stake. This is an offer management (m) would be keen to accept. Thus, we have the Kaldor-Hicks decisive contractual coalition $w\mathrm{P}_{me}z$, threatening to take us from social state z to social state w.

However, social state w is inferior as a final outcome to the status quo ante, or social state x, for both shareholders and employees alike (that is, $x\mathrm{P}_{se}w$), and it is difficult to see why employees and shareholders, anticipating the possibility of this cycle through to social state w, would be prepared to set off on this cyclical journey in the first place. Thus, the corporation might remain in social state x, a social state that is Pareto-inferior to social state y, which was also attainable. This, in a stylized form which is reminiscent of the more general cycling problems referred to above, is essentially the breach of trust scenario for ex ante inefficiency that has been proposed by Shleifer and Summers.[43]

B. SOME CRITICISMS AND IMPLICATIONS OF THE BREACH OF TRUST SCENARIO

Now, to be convincing, this breach of trust scenario must have four essential elements. First, there must be some reason why employees

43 It should be noted that this particular interpretation of one part of the Shleifer and Summers argument does not carry with it any kind of endorsement of the argument's other aspects. For example, it is surely difficult to argue, in the face of empirical work showing otherwise, that a large part of takeover gains come from the appropriation of deferred wages. See, e.g., the empirical studies cited in R. Romano 'A Guide to Takeovers: Theory, Evidence, and Regulation' (1992) 9 *Yale J. of Reg.* 119, 139, which show that, at most, employee losses are 10 to 20 per cent of the overall shareholder gains. Admitting this much about the ex post efficiency of takeovers, however, does not preclude Shleifer and Summers from claiming that ex ante inefficiencies will result from the breach of trust involved in the appropriation of even these smaller employee losses.

would make firm-specific investments and accept deferred compensation rather than insist on being paid for the value of the marginal product of their labour for each and every period of their contract. Without deferred compensation, there is no wage fund for the new acquirer to appropriate given a change in control of the corporation.

Second, it must be that the appropriation of the deferred compensation is not legally actionable. This is the role played by implicit as opposed to explicit contracting, and it typically requires that the contract have terms that, while they might be observable to the contracting parties themselves, are not verifiable to third parties such as courts.[44] It is this lack of verifiability that explains the lack of enforceability of the implicit contract and, therefore, that holds open the possibility of an opportunistic breach of trust. Trust might even play a more essential role in the implicit contract than this if the terms of the contract are not only not verifiable to third parties, but also are not observable to one of the parties themselves; in that situation the trusting party cannot even be sure that a breach of the contract has really occurred.

Third, there must be some reason for thinking that this form of implicit contracting can occur with some managers, in particular incumbent managers, but is threatened by any replacement of those incumbents with a new management group, introduced, for example, by some new acquirer of the corporation. Without such an asymmetry in the proclivities of old and new management for breach of trust, there is a sense in which enforcement of the implicit contract is always an empty dream. Finally, there must be some kind of social or corporate advantage to implicit contracting, an advantage that will be lost if all contracts must either be made explicit or not made at all.

With respect to the last element, it is not difficult to see that there will be some advantages to implicit contracting in some contexts. Because future contingencies are difficult to predict and specify fully in advance, complete contingent claims contracting will always be costly. In such situations, it will be beneficial to leave some of the contractual terms open, or implicit, with one of the parties to the contract being trusted to deliver on it according to some incompletely specified 'understanding.' The alternative of haggling over every possible contingency involves either the high costs of going through that bargaining process itself, or incurring the costs of forgoing certain contractual benefits altogether. In

44 For discussion of implicit contracting in these terms, see Kreps, supra note 7, 110. Also see A. Schwartz 'Relational Contracts in the Courts: An Analysis of Incomplete Agreements and Judicial Strategies' (1992) 21 *J. of Legal Studies* 271, 279–80, and Roberta Romano, supra note 43, 139.

either case, it seems plausible to argue that there are social costs if the process of implicit contracting is threatened by a breach of trust.

The real challenges to the breach of trust scenario typically focus on its other three elements. Roberta Romano, for example, has criticized the Shleifer and Summers version of this argument for its failure to provide a convincing account of the exact nature of the firm-specific investments that the employees whom Shleifer and Summers talk about (usually unionized airline employees) are supposed to have made in the corporation that comes to be targeted for an opportunistic acquisition.[45] If the focus is on human capital, Romano argues, there does not seem to be much that is firm-specific about the worker skills involved. Yet, without firm-specificity, there is little opportunity for the new corporate acquirer to appropriate the benefits of the productive investment in human capital while still paying the target-firm employees only their opportunity cost wage; they will simply take their new skills elsewhere. On the other hand, Romano notes that the other sorts of firm-specific investments that are mentioned as possibilities by Shleifer and Summers, such as increases in the value of the employee pension fund benefits, for example, seem to be both sufficiently observable and verifiable as to be the subject of explicit and, therefore, legally enforceable contracting. In effect, therefore, Romano is challenging the first and second elements of the breach of trust scenario.

However, while they may be appropriately targeted at the specifics of the Shleifer and Summers argument, Romano's objections turn on too narrow an understanding of firm-specific employee investments to provide a generally applicable response to the breach of trust argument. *Any* deferred compensation package is a firm-specific investment by an employee in the targeted firm. The appropriate question to ask is why an employee and an employer might agree to such a deferred form of compensation for the employee's services in the first place.[46]

45 Romano, supra note 43, 140

46 Actually, since the more limited purpose of this paper is only to show the opportunity for a corporate acquirer to appropriate deferred wages, it is enough that such deferred compensation plans *exist*, and that they condition on non-observable or non-verifiable events. The *explanation* for them could reasonably be left for another occasion; hence, the brevity of the explanation that follows in the text. For empirical evidence on the existence of such compensation schemes, see J. Medoff and K. Abraham 'Experience, Performance, and Earnings' (1980) 95 *Q. J. of Econ.* 703 (salaries tend to grow more rapidly than 'productivity' as measured by supervisor's performance ratings). See also E. Lazear and R. Moore 'Incentives, Productivity, and Labour Contracts' (1984) 99 *Q. J. of Econ.* 275.

Problems of asymmetric information would seem to provide a number of possible explanations. When an employee is first hired, her potential productivity may not be observable to the employer. She may either be unskilled, or inclined to low levels of effort, or prone to quit. For any one of these reasons, an employer may offer the employee an initial wage that is lower than the value of her expected marginal product, with the understanding that if she proves herself skilled and productive, and stays on at the job, she will later receive a wage sufficient to make up for the difference between this initial wage package and her opportunity cost. This is the sort of firm-specific investment that an employee who is confident of her own skills, industrious, and not prone to quit would be willing to make.[47] The problem, however, and this is where the implicit contracting would come in, is that the employer would have private information on the relative productivity of the employee's efforts for the employer's profits ex post. This would be a case of the non-observability of a key contractual term. Alternatively, if the contract conditioned on effort rather than productivity, the effort, while observable to both parties, would not be verifiable to others and, therefore, not enforceable in the courts. In such a situation, therefore, the contract will turn on whether the employee, having made her labour commitment in advance, can trust the employer to settle up later as implicitly agreed. A new corporate acquirer may be in a position to breach this trust and, therefore, appropriate the deferred compensation as a gain from making the corporate acquisition.

Romano and others[48] have gone on to object at this point to the third element of the breach of trust scenario, arguing that it provides no good explanation for why the management team of a new acquirer would have any greater tendency to breach the trust than would the old management team that was already in place before the takeover. Presumably, on any forward-looking, purely instrumentalist account of shareholder rationality, what is a good or bad investment for the new acquirer and its management team is also a good or bad investment for the incumbent shareholders and management team. Why, therefore, should there be

47 The employer cannot offer an 'average wage' to productive and non-productive workers (or to 'quitters' and 'non-quitters') alike because of the adverse selection problem; that is, such an average wage would tend to attract the relatively unproductive worker. On this problem, see Akerlof 'Lemons,' supra note 3.

48 Romano, supra note 43, 140; see also B. Holmstrom 'Comment' in Auerbach, supra note 39, 56 at 57–9. Coffee, supra note 18, 1526, recognizes the importance of this point as well, but seems to accept the Shleifer and Summers explanation.

any special threat of a breach of trust around the time of a takeover or other such change in corporate control?

This is an extremely important objection, which goes to the heart of what it might mean for an incumbent management to be loyal to the corporation itself as distinct from its shareholders and, therefore, of what it might mean for managers to be able, credibly, to commit themselves against any breach of the trust that might have been placed in them by a non-shareholder constituency such as corporate employees. Therefore, in section VI of this paper, it will be important to analyse this objection in some greater detail. However, for the moment, it is worth considering again Jeffrey Gordon's account of the delegation doctrine in corporate law. This account goes some of the way towards explaining why incumbent managers, as a consequence of law rather than loyalty, may breach trust less frequently than those managers caught in the middle of a corporate control transaction.

It will be recalled that Gordon's argument was that the absolute delegation by shareholders of the day-to-day running of the corporation to management could be explained as a device to restrict shareholder voice in such matters and, therefore, to avoid the perils of the majority voting cycle. He recognized, however, that such an absolute delegation of power did not preclude shareholders from having a voice on more fundamental corporate changes such as changes in corporate control. Thus, on Gordon's view, shareholders would be in a better position to bring management's discretionary powers into line with their own interests when there is the possibility of a control contest involving new management than they would be on a day-to-day basis with incumbent management. This is the beginning of a possible answer to Romano's question as to why there might be more breaches of trust against employees in corporate control contests than one sees in the day-to-day running of the firm. Incumbent managers can be trusted to hold out against shareholder interests (and the possibilities of a voting cycle) in the latter situation in a way that they cannot when corporate control is at stake.[49]

However, this way of putting the point, when it is combined with the insight that Gordon's majority voting cycle can be reproduced as a

49 There is some evidence that deferred compensation plans (that is, those for which trust would be important) are less common where there is concentrated shareholder ownership, that is, a situation in which managers would be more subject to shareholder control. On this, see G. Garvey and N. Gaston 'Delegation, the Role of Managerial Discretion as a Bonding Device, and the Enforcement of Implicit Contracts' (1991) 9 *Advances in Econometrics* 87, 90.

Kaldor-Hicks contractual cycle in the context of corporate control transactions, suggests that management's independence from shareholders, which Gordon has so effectively argued for on a day-to-day basis under the delegation doctrine, should also somehow be provided for in the corporate control context. In terms of our stylized example, this would mean destabilizing the Kaldor-Hicks decisive coalition between management and shareholders that would otherwise take the corporation from social state y to social state z. This is, of course, a restriction on shareholders being able to 'exit' with their capital, to use the Hirschman terminology, but it is fully analogous to, and grounded in the same reasons as, the restrictions on labour 'voice' that both Gordon and Hansmann have proposed in other contexts. My point has only been to suggest that this insight, as its origins in social choice theory might suggest, is more generally relevant than these analysts seem to realize.

As a more practical legal matter, the destabilization of the management-shareholder contractual coalition, which would otherwise choose social state z over social state y, will mean that management should not always be held to a narrow fiduciary obligation to maximize share values in a corporate control contest. Such a narrow obligation is effectively what is being manifested in the $zP_{m}y$ Kaldor-Hicks decisive coalition, and it is what is usually mandated in much of contemporary corporate law.[50] Permitting a broader fiduciary obligation would have the effect of breaking up this close and exclusive association between management and shareholders, and would allow management to consider the interests of non-shareholder constituencies in a corporate control contest as well. In our example, such a broad fiduciary obligation would likely prevent any breach of trust given by employees to incumbent management, and would hold the corporation to social state y, the alternative social state which is Pareto-superior to the status quo.[51]

50 The recent wave of stakeholder statutes is, of course, the exception; for a discussion of these statutes, see M. O'Connor 'Restructuring the Corporation's Nexus of Contracts: Recognizing a Fiduciary Duty to Protect Displaced Workers' (1991) 69 *N. Carolina LR* 1189, 1229–35. In the case law, *Tech Corp Ltd.* v. *Millar* 23 DLR (3d) 288 (BCSC) is also well known for its support of a broader fiduciary obligation.

51 Typically, the stakeholder statutes alluded to in the preceding note are only permissive; they do not mandate that directors address the needs of non-shareholder constituencies. See O'Connor, supra note 50, 1230. In the last section of this paper, we shall inquire into the possibility that managers might naturally be predisposed to consider the interests of all the corporate constituencies, not just those of the shareholders, and even when their own self-interest, narrowly construed, might be thought to tempt them otherwise. In the absence of this more general loyalty to the corporation, it might be necessary, if the contractual cycle is to be avoided, to make the broad fiduciary obligation mandatory and not just

To be sure, it is still open for Romano and others to argue that management's day-to-day concessions to employees are not the stuff of implicit contracting and efficiency-enhancing trust at all, but are more easily explained with the agency cost theory that management is seeking a peaceful wage settlement with labour at shareholders' expense.[52] A new corporate acquirer would have an advantage over the dispersed shareholders it replaces of being more able to coordinate shareholder interests against these slack managers and, therefore, of being more able to recapture the excess wages that the managers have been paying out to labour to secure an easy way of life for themselves. Such a theory might explain wage concessions as a source of gains in corporate control transactions, but the efficiency implications would obviously be very different from those under the breach of trust scenario.[53]

However, the point of this discussion has not been so much to settle finally, as an empirical matter, whether the wage concessions being granted in these different contexts are to be explained better by accounts based on breach of trust than accounts based on agency cost theory. Rather, by embedding it in the larger theoretical context provided by social choice analysis, the point has been to support the analytical respectability, within economics, of the former sort of explanation against what has fast become a 'prevailing orthodoxy' among economic theorists of corporate law.[54]

This orthodoxy views the corporation as a web of merely contractual relationships, or a 'nexus of contracts,' in which all the different interest groups – shareholders, employees, management, creditors, suppliers, and so forth – compete in the usual market fashion, seeking to maximize their

permissive. This would amount, in some sense, to a restriction on management's freedom of association. However, for a general argument that many of the difficulties of social choice might be solved if certain individuals were not allowed to freely associate or, equivalently, if freedom of association were not permitted in certain contexts, see F. Frohock *Rational Association* (Syracuse: Syracuse University Press 1987), e.g., at 25.

52 This is the thrust of both Romano, supra note 43, and Holmstrom, supra note 48, and it is also Oliver Williamson's preferred view of the particular takeovers involving airlines, which are so much discussed by Shleifer and Summers. See Williamson 'Comment' in Auerbach, supra note 39, 61–7.

53 However, this agency cost account would not require that the higher wages that are being (re-)appropriated from labour in this way be *structured* as deferred wages. Yet, that is the evidence provided by Garvey and Gaston, supra note 49. This would be consistent with what is assumed in the breach of trust theory.

54 This term, and much of the next paragraph's characterization, is borrowed from Coffee, supra note 18, 1495.

own interests through the corporate form.[55] Shareholders are given the ultimate right to control the corporation under this series of bargains, because the various participants can see that, as residual claimants, shareholders have the appropriate incentives at the margin to maximize the overall profitability of the firm, something that can rebound to the advantage of all.[56] Thus, the defining feature of shareholders is their control, not ownership, of the corporation; and all talk of ownership, as if there were 'something else out there' independent of the contractual understandings between the different owners of the various inputs that go into production is viewed as 'legal metaphysics' and a source of great intellectual confusion.[57] Moreover, although shareholders have ultimate control of the corporation, as principals they delegate the day-to-day running of the corporation to management, their agents. Thus, on the orthodox view, all problems of corporate governance boil down to problems that arise in this principal-agent relationship, and most corporate law, including most corporate fiduciary law, comes down to how these attendant 'agency costs' can be most effectively reduced.

Against this orthodox contractual view of the corporation, the analysis in this paper has claimed so far that trust and institutional loyalty play a larger role in economic arrangements generally, and the corporation specifically, than the usual economic analysis allows. That this is so is evidenced by the central role that fiduciary relationships play in corporate law. Moreover, fiduciary relations are not usually thought to be identical to the principal-agent relations generated by contract.[58]

The important role of the fiduciary as someone who, while attending to the beneficiary's interests, does not accede totally to the beneficiary as decision-maker has been explicitly recognized by some economic theorists of the corporation, by Hansmann for example, when they analyse problems of voting in the corporation, but not when they attend to the

55 See, e.g., Jensen and Meckling, supra note 1; E. Fama 'Agency Problems and the Theory of the Firm' (1980) 88 *J. of Pol. Econ.* 288. For a systematic and impressive application of this approach to corporate law, see F. Easterbrook and D. Fischel *The Economic Structure of Corporate Law* (Cambridge: Harvard University Press 1991).

56 See Easterbrook and Fischel 'Voting' supra note 23, 403–4.

57 See, e.g., Fama, supra note 55, and Macey, supra note 42, 179.

58 See, e.g., O'Connor, supra note 50, 1247–51; D. Demott 'Beyond Metaphor: An Analysis of the Fiduciary Obligation' (1988) *Duke LJ* 879; and A. Stinchcombe 'Reason and Rationality' in K.S. Cook and M. Levi (eds) *The Limits of Rationality* (Chicago: University of Chicago Press 1990) 285, 304–5. Stinchcombe goes further than merely characterizing the fiduciary obligation as non-contractual; he finds the fiduciary's proper reasons for acting to be 'institutional.' This is very close to the argument for institutional loyalty that is presented in this paper. See infra, section VI.

corporation as a nexus of contracts with its attendant problems of agency costs.[59] However, this partial view of the importance that fiduciary obligations have in corporate law is only a manifestation of the deeper partiality which the economist has for the market over political institutions, a partiality further reflected in the commonly held view that the general problem of social choice, so long ago identified by Arrow, is only a problem for voting, the solution for which is to be found in institutions that restrict political voice.[60]

The analysis in this paper has attempted to remedy this last misunderstanding by showing that the majority voting cycle is reproduced in the market context as a Kaldor-Hicks contractual cycle, and that the problematic preference profile that is necessary for the former also shows up as an explanation for the latter. Indeed, as illustrated here, the breach of trust scenario that recently has surfaced as an explanation for stakeholder dissatisfaction with corporate control transactions can easily be interpreted in these terms. Thus, the 'restricted voice' solution to majority voting cycles in the corporation has an analogue solution in the context of corporate contractual cycles in the 'restricted exit' solution that must sometimes be imposed on target shareholders tempted by the gains of a takeover. Predictably, given Hansmann's insight on the nature of a fiduciary in the political context, restricted exit takes the form of management having a broader fiduciary obligation in the context of corporate control transactions than would be implied if management's obligations were those of an agent *contractually* bound to its principal and not really a *fiduciary* at all.

A further problem remains, however. To this point the analysis in this paper has spoken of trust and the concept of a fiduciary in purely instrumental terms. Specifically, the idea has been that the introduction of trust and the fiduciary into the contractual model of the corporation might be useful for the achievement of Pareto-optimal social states, that is, social states in which every one is better off than they would be in some other social state, *given* their preferences. While that is undeniably the motivation for the analysis, and while one can surely *model* the proposed fiduciary solution in this forward-looking way, a way that is consistent with the conventional economic accounts of a purely instru-

59 See supra, text at note 32. An important exception is Robert Clarke. See his 'Agency Costs and Fiduciary Duties' in J.W. Pratt and R. Zeckhauser (eds) *Principals and Agents: The Structure of Business* (Boston: Harvard Business School Press 1985).

60 Hirschman, supra note 4, 15 and 17, notes the propensity economists have for the virtues of 'exit,' or the market, over the vices of 'voice,' or politics. Hirschman's own analysis, by contrast, is wonderfully balanced.

mental rationality, there is a real danger that if the concepts of loyalty and trust are actually *experienced* this way by individuals, they will be self-defeating because non-authentic.[61]

It is the burden of the next section of the paper to show why that is so. Moreover, an appreciation of this fact will go a long way towards explaining why the law views the corporation as an entity separate from its different and changing constituencies, an entity that, in principle at least, is capable of commanding their loyalties and disciplining their preferences, particularly those problematic preferences that underlie the cycles of social choice described above. Only in this way can a world in which we are all better off, something that the economic theory of instrumental rationality can promise but not provide, actually be achieved.

VI *Corporate loyalty and the limits of a purely instrumental rationality*

Let us begin this concluding section of the paper by reconsidering the example with which we introduced the concept of trust in section II. Recall that there were three individuals, each in possession of a particular good that one of the other three individuals preferred to have. However, using bilateral exchange to bring about the preferred reallocation of goods was problematic because no individual in the example who preferred what another particular individual had could offer a good that the second individual would want to take in return.

The introduction of money into this simple economy promised to provide an effective solution to the problem until it was recognized that there would be a predictable problem with the last of the three monetary exchanges. Specifically, individual 1 would be reluctant to give up her particular good in exchange for the money on the last trade with individual 2 unless the money continued to have some kind of further exchange value. But suppose it did not. Then her predictable defection

61 See Sugden 'Rational Choice' supra note 7, 781. By non-authentic, I mean that the obligations of loyalty will be experienced as irrational in the same way that it is 'irrational' not to breach the trust of another on the last play of a prisoners' dilemma game. To be authentically loyal, therefore, one must either be irrationally non-instrumental in the satisfaction of one's preferences, or have changed one's preferences as a result of having played the game. While the latter allows for the rational satisfaction of preferences in their changed form, and so appears quite conventional in its continued instrumental approach to *individual* choice, it is, nevertheless, still a quite unconventional and non-instrumental approach to *institutions*, or games, in that now these are seen not only as devices through which individuals seek to satisfy *given* preferences, but also as mechanisms through which their given preferences are changed.

from the sequence of cooperative monetary exchanges would undermine the possibility of any earlier exchanges, and the preferred reallocation of goods would continue to be frustrated.

What if individual 1 simply *promised* to take the money on the last exchange? Surely, if anything is predictable, then that is. After all, it is as much in individual 1's interest as anyone else's to get the sequence of exchange going. The question, however, is whether such a promise can be a credible one in the eyes of the other individuals. The difficulty is that a promise made for only an instrumental or forward-looking reason, the sort of reason that a rational maximizer of her own preferences would provide, is a promise that should, rationally, be broken when it pays to do so. But that is exactly the sort of reason which individual 1 will face on the last exchange. Having promised, for purely instrumental reasons, to take the money on the last exchange, individual 1 will find that these *same instrumental reasons* tell her to break the promise that she has made.[62] Furthermore, all this instrumentally rational behaviour can be predicted perfectly by the other individuals; after all, they too are rational in this way. That is why the Pareto-superior reallocation of the goods cannot be achieved.[63]

62 See ibid. The following remarks by Arrow, supra note 3, 23, also suggest the problems involved in the rational choice approach to trust: 'Trust is an important lubricant of a social system. It is extremely efficient; it saves a lot of trouble to have a fair degree of reliance on other people's word. Unfortunately this is not a commodity which can be bought very easily. If you have to buy it, you already have some doubts about what you've bought.'

63 It will be objected that this is precisely why we make such promises (when they are supported by at least token consideration) legally enforceable. However, this objection simply serves to shift the focus of concern from individual 1 as a contracting party to the enforcer of the contracts (for example, the judge). Suppose that this enforcer was also an instrumentally rational maximizer of her preferences. Then she would have little incentive to enforce the contractual terms as written, especially if individual 1 was offering the enforcer some small part of her goods in return for 'looking the other way' (that is, a bribe). (This will only seem far-fetched to those who have not contemplated the prospect of hiring mercenaries to defend their national borders.) The point is a perfectly general one, as Kenneth Arrow has recognized in his 'Gifts and Exchanges' in E.S. Phelps (ed.) *Altruism, Morality and Economic Theory* (New York: Russell Sage Foundation Press 1975) 24: '[T]he price system, in order to work at all, must involve the concept of property ... Property systems are in general not completely self-enforcing. They depend for their definition upon a constellation of legal procedures, both civil and criminal. The course of the law itself cannot be regarded as subject to the price system. The judges and the police may indeed be paid, but the system itself would disappear if on each occasion they were to sell their services and decisions. *Thus, the definition of property rights based on the price system depends precisely on the lack of universality of private property and of the price system ... The price system is not, and perhaps in some basic sense cannot be, universal.* To the extent that it is incomplete, it must be supplemented by an implicit or explicit social contract. Thus one might loosely say

We can now appreciate why, in their argument for the breach of trust account of hostile takeovers, Shleifer and Summers insisted that the incumbent management team, which employees were to trust, be thought of as the sorts of people who simply *were* trustworthy. It would not be enough that they merely *chose* to be trustworthy for purely instrumental reasons. As in our example of monetary exchange, no one can really trust the latter sort of person to carry through on her promises. Thus, Shleifer and Summers suggested that the trustworthy manager needed for their argument would likely be the sort of person who has spent a long time with the corporation and worked her way up through its organization with the help of other employees.[64] In such a situation, the manager would feel committed to these other employees, and would even come to *identify* herself as 'a corporation X sort of person' with the kinds of preferences that these people 'properly' have. She would naturally feel that her loyalties were owed first to those who were like her in this respect, namely those with the same long-run stakes in the corporation. By contrast, her obligations to shareholders, a largely anonymous group whose interest in the corporation is only as transient as rate of return and portfolio diversification considerations will allow, would be secondary. Thus, it is the corporate attribute or *character* of these managers, something that is part of what they are, not what they rationally choose to be, which makes them trustworthy in the way the Shleifer and Summers argument requires.[65] It is also another reason why, whatever strategic disadvantages there might also be for a new management team's ability to resist shareholder interests after a corporate control contest,[66] incumbent managers, naturally predisposed in their loyalty to other

that the categorical imperative and the price system are essential complements' (emphasis added).

64 See Shleifer and Summers, supra note 39, 40. 'From the *ex ante* viewpoint, such dedication to stakeholders might be a value-maximizing managerial *attribute* (not choice!) ... [M]anagers pass through a 'loyalty filter' ... before reaching the top. Having done so, they find stakeholder welfare has now entered their preferences ...' For an extended argument that emotional attachments rather than rational choice may be essential to the achievement of improved outcomes, see R. Frank *Passions Within Reason: The Strategic Role of the Emotions* (New York: Norton 1988).

65 Thus, the requirement here is for something stronger than the 'detached role' postulated by Meir Dan-Cohen as the usual sense of attachment one has to an economic 'organization.' See his 'Law, Community, and Communication' (1989) *Duke LJ* 1654, 1655. It may be that what is required is even as much as what is implied by his conception of the 'non-detached' role, something which, he thinks, forms the constituent element of much closer communities (for example, the family). (But see the qualification at 1658 where he admits that economic organizations might be characterized, in their 'upper echelon management,' for example, by non-detached role play.

66 See supra, text following note 48.

corporate stakeholders, would be less inclined in any event to accede to shareholders' demands in the day-to-day running of the firm than new managers might be. Thus, in response to one of Romano's earlier-mentioned concerns about the breach of trust scenario, it would seem that a hostile takeover, accompanied by the displacement of this naturally loyal incumbent management team, might be essential for the appropriation of any deferred wages that have been given to this incumbent management in trust.

It may still be objected that the analysis to this point has paid too little attention to the importance of reputation in controlling for breaches of trust. The single-play character of the fruit reallocation example introduced in section II, it will be said, has been misleading from the start. If only individual 1 had a reputation to worry about, a reputation that would decisively affect her ability to secure the benefits of ever being trusted again by other individuals in other transactions, then she could, on a purely instrumental account of her reasons, credibly commit to being trusted now.[67] This objection would naturally go on to make the further point that this reputational account of trust requires some sort of ongoing entity to which the reputational label can attach, for example, an individual or a firm. Moreover, in the case of a firm, where the individual members are subject to constant change, it will be important that a position in the entity to which this reputational label attaches be transferable at a price. This purchase price will act as a bond that is forfeitable against the owner should she choose to renege on the entity's reputation for trustworthiness.[68] Thus, we can begin to make sense of the corporation as a separate entity capable of ownership in the way that the law seems to require. And all this we can do without abandoning an instrumentally rational account of the corporation.

This is a compelling objection, but it is important to appreciate exactly what it can accomplish as a reply to the breach of trust argument as developed, for example, by Shleifer and Summers. First, if there is any predictable 'last play,'[69] any possibility for trust will unravel from the back in the usual inductive way. Second, and more important, it is absolutely crucial that each owner of the corporate reputation feel that her breach of trust would be decisive (negatively) for the continued reputation of the firm and, therefore, that her bond in the firm would thereby be forfeited.

67 Kreps, supra note 7, 106
68 Ibid. 109–10
69 Ibid. 108

However, this last point need not hold true. If the corporate acquirer believes that its own reputation for trust is quite unaffected by the breach of trust by the incumbent shareholder-management coalition (after all, it has not committed any breach of trust, since it did not make any promise), then it may be prepared to offer an acquisition price for the target shares that does not impose any forfeiture costs on the incumbents and that, therefore, secures for all shareholders some share of the appropriated wage concessions. Admittedly, this assumes that the acquirer's reputation can continue sufficiently unsullied so that, even after an acquisition in which it subsequently appropriates the deferred wages that the target corporation's employees had entrusted to the former management-shareholder team, it can recapture all the benefits of implicitly recontracting with those same employees again. If this were not so, then the only real source of gain in the acquisition would be in the fact that the gains from breach, G, exceed the losses, L, from losing completely one's reputation for trust. But then, as several commentators have pointed out,[70] the original team of shareholders and incumbent management, who would be looking at the same magnitudes of G and L, would already have breached the trust themselves, and no takeover would occur, at least for that reason. So one is forced back to accepting the view that the acquirer can breach another's trust without any loss of its own reputation for trustworthiness, that is, that its L is zero or at least very small, hardly a plausible assumption.

It is small wonder, therefore, that Bengt Holmstrom, commenting on the Shleifer and Summers scenario, but still working from a reputational model, felt obliged to work out a different asymmetry between the incentives for breach of the acquirer and the incumbent management team by postulating a difference, not in the losses L, but in the gains G. Holmstrom conjectured that the incumbent managers may not be able to generate the same gains from appropriating the wage concessions as a new acquirer because the incumbents would already have capitulated to wage demands in the past and would, therefore, have a 'softer' reputation than would the new acquirer, which could come to the labour negotiations unburdened in this respect.[71] However, Romano, commenting in turn on Holmstrom, has found this unpersuasive.[72] She suggests that employees might just as easily capitulate more to the incumbent managers they know, and whom they know more particularly to have

70 Holmstrom, supra note 48, 57; Romano, supra note 43, 140
71 Holmstrom, supra note 48, 58–9
72 Romano, supra note 43, 140

provided them with concessions in the past, as wait around for a less accommodating new acquirer. In the final analysis, Romano concludes that the Shleifer and Summers thesis, even under the aspect of Holmstrom's reinterpretation, is 'clever but unconvincing.'[73]

But neither Holmstrom nor Romano really address the fundamentally different account of loyalty and trust that the Shleifer and Summers scenario provides. This is because both are still committed, as the reputational model is, to understanding these concepts from a point of view that is purely instrumentalist and forward-looking from a given set of individual preferences. But the incumbent managers in the Shleifer and Summers argument are not loyal or trustworthy for instrumental reasons at all. They simply *are* loyal in virtue of the actual lives they have lived, and the promises they have made. Moreover, the corporation, which has been so large a part of their lives, is something that has *shaped* their preferences, and is not something they approach anew every day as if they had no such past, and through which they seek the satisfaction of given preferences as if these were always determined elsewhere. Until this possibility of a non-instrumental, preference-forming account of institutions is taken seriously in its own terms, the arguments provided by theorists like Shleifer and Summers will continue to appear very mysterious indeed.

Moreover, the returns to having a non-instrumental understanding of social institutions like the corporation, as well as a non-instrumental understanding of the motivations and loyalties of the individuals who form their constituent groups, should take us even further. Often, whether the choices we make as individuals turn out well or badly for us when they involve interaction with others depends less on the particular choice we have made than it does on whether the choice in question coordinates well with the choices these other individuals have made. One has only to imagine the case of choosing which side of the road to drive on as one approaches a blind corner. Presumably, it does not matter at all whether one drives on the left or right so long as the party coming from the other direction chooses the *same* way. There is no conflict of interest in this example (it is a pure coordination problem), but there is a possible result that is very clearly Pareto-inefficient. The question is how best to avoid it. An economist might say that there is no real puzzle here and that the evidence of how we solve this trite problem is all around us. We have avoided the inefficiency of crashing at every corner by simply agreeing, in North America at least, to drive on the right. Once

73 Ibid. 139

so agreed, moreover, there is no real conflict of interest, and no real need for legal enforcement. As the economists would say, we are blissfully in Nash equilibrium.[74] Surely, this must be the easiest possible case for the economic efficiency theorist.

But the analysis has already gone too fast. A forward-looking efficiency theorist is not someone who really takes past agreements, or promises, very seriously. Promises and agreements are not meant to be upheld unless there is a good forward-looking reason for doing so, for example, that the agreement will continue to pay off better than the alternative of breaching it (as in the reputational model), or that defection from the agreement will be legally sanctioned (as in the legal deterrence model). But, by assumption, there is no more reason to abide by the particular agreement arrived at in the blind corner case as opposed to the other one that might have been signed calling for driving on the left. Moreover, all like-minded efficiency theorists know that this is true and, further, they all know that they all know that this is true. This is already enough for there to be real problems, in theory at least, when two efficiency theorists drive towards each other at the blind corner, *at least if they think about it too much.*

However, that is exactly the point. They will not, if they are genuinely social beings, think about it at all. Rather, they will, in a perfectly non-instrumental yet rational way, easily circumnavigate the blind corner, just as agreed. Each will successfully play her part in the social practice that is involved here, that is, each will follow the rule, as Wittgenstein says, 'blindly' and 'without reasons.'[75] It is only if they exercise too much instrumental rationality, that is, if they think too much (again) about what they should do and why they should do it and, further, if they think that others are thinking likewise, and thinking likewise about them, that, by

74 A Nash equilibrium is an outcome reflecting the strategy choices of all the players in a game such that, given the strategy choices of all the other players, no player can improve her pay-off by unilaterally changing her choice. See Binmore and Dasgupta, supra note 6, 4–5.

75 See L. Wittgenstein *Philosophical Investigations* Anscombe (trans.) (Oxford: Basil Blackwell 1958) sections 211 and 219. Margaret Gilbert relates this aspect of Wittgenstein's work to the economic theory of games in her 'Rationality and Salience' (1989) 57 *Phil. Studies* 61 at 75. For a closely related, if not identical, invocation of Wittgenstein into game theory, see M. Hollis, 'Moves and Motives in the Games We Play' (1990) 50 *Analysis* 49 at 59: 'The Wittgensteinian objection is that there is no neutral currency of satisfactions for assessing actions which belong to different practices. Underlying it is a suggestion that to be rational is to follow an appropriate rule appropriately, rather than weighing options as if in the omni-competent pans of a pair of scales.' Also see Hollis's book *The Cunning of Reason* (Cambridge: Cambridge University Press 1987), especially at 130–45, where he discusses a related distinction between 'regulative rules' and 'constitutive rules.'

infinite regress, they will open up so many new possibilities for action that all chance of coordinating on the specific option made salient by their prior agreement becomes impossible. The result may well be inefficiency, even disaster.

It is the great insight of the common law that it recognizes that non-instrumentally rational coordination, at least as much as instrumentally rational contract and exchange, is a source of efficiency gains for society.[76] It is often thought that it is a great barrier to socially beneficial reform to have a myriad of cases that have been already decided in the past, and that burden us now with their precedential weight. But this is to assess the law exclusively as a forward-looking instrumentalist would, and it forgets that it might be just as important to get on with *some* convention for the coordination of human action as it is always to seek out some uniquely best one. Indeed, as the blind corner example suggests, the lack of coordination that plagues those who ignore their past and keep only a constant eye on a rational approach to their future, may result in even greater efficiency losses being experienced by these rationalists when their future arrives.

This is also the insight that Shleifer and Summers are encouraging us to bring to our understanding of the corporation. There is an advantage, they suggest, to having managers (and, presumably, other corporate role players with whom management itself might be attempting to coordinate) who simply *are* loyal to the corporation, and who, therefore, play their part in it 'blindly' and 'without reasons,' at least without reasons of the forward-looking kind that inform the constant flux (or cycling) of efficient (re-)contracting and exchange, the stuff of 'contractual nexus' analysis. The advantage is that those individuals who are attempting to coordinate with the corporation can do this better if they can rely on some general understandings of what is 'done' and 'not done' (and, more specifically, what counts as a real breach of trust[77]) by a 'corporation X sort of person.'

76 For an eloquent development of this argument, see R. Hardin 'The Social Evolution of Cooperation' in Cook and Levi, supra note 58, 358 at 368, 374–5, 376. Also, for a more detailed exposition of this argument, with implications for the usual rationality axioms of economics, see B. Chapman 'The Rational and the Reasonable: A Comparison of Social Choice Theory and Legal Adjudication' (1993) *U. Chi. LR* (forthcoming).

77 This is the crux of the argument provided by Kreps, supra note 7, for corporate culture. Without some prior understanding of which of the many Nash equilibria (all of which are possible in the reputational model) should count as the baseline from which breaches of trust are to be measured, any action is in danger of being misunderstood as a breach and, therefore, of threatening efficient cooperation. The analysis here shows that this corporate culture, to be genuinely coordinating, needs to be experienced, but it cannot be rationally chosen.

There is an irony, of course, perhaps even a paradox, in the fact that *more* efficient exchange will actually be achieved this way than if we sought to achieve it self-consciously in the way that the instrumentally rational contractual nexus analysis of the corporation suggests we should.[78] But some consequences, even those consequences that constitute efficient social states, are what Jon Elster has termed 'intrinsically incidental.' That is, these social states 'can never ... be brought about intelligently or intentionally, because the very attempt to do so precludes the state one is trying to bring about.'[79]

Perhaps the final irony for economic theorists of the corporation, therefore, is that there is something of Adam Smith's invisible hand working here after all. Smith, it will be recalled, also argued that certain transactions that were prompted by self-love would achieve a greater social good than if each transaction actually aimed at achieving that social good self-consciously. The difference is that this paper has shown that the private and highly localized virtues of loyalty and trust within an institution like the corporation, at least as much as the unconstrained pursuit of self-interest through contracts, can also add up to the unconscious attainment of a greater good for all.

78 Albert Hirschman has also commented on the somewhat paradoxical fact that 'loyalty' is its most functional (for preserving the possibility of beneficial change from within) for an organization when it appears to be least rational, that is, when loyalty means a strong attachment to an organization that does not seem to warrant such attachment because it is so much like another one that is also available. This is very much like the problem of rationalists failing to coordinate across outcomes because no one outcome has any *instrumentally rational* salience for them. See Hirschman, supra note 4, 81. This is also an appropriate point to introduce a qualification into the analysis that has so far not been mentioned. Hirschman distinguishes in his analysis between 'loyalty' and 'faith.' Loyalty raises the costs of exit for committed members so that they will not exit at the first opportunity promising a higher rate of return; faith is a die-hard refusal to exit regardless of costs. Hirschman considers the former advantageous, the latter disadvantageous for society. The mention in this paper of following rules 'blindly' may suggest that something more like 'blind faith' than loyalty is at stake. That would be to misinterpret the thrust of the word 'blind.' When one is loyal, one defers, or is blind to, the calculus of a purely instrumental rationality, but only up to a point.

79 See J. Elster *Sour Grapes* (Cambridge: Cambridge University Press 1983) 43.

University of Toronto LAW JOURNAL

volume XLIII, number 3, summer 1993