Abstract

When contracts are incomplete, the law must rely on default rules to resolve any issues that have not been explicitly addressed by the parties. Some default rules (called ‘majoritarian’ or ‘market-mimicking’) are designed to be left in place by most parties, and thus are chosen to reflect an efficient allocation of rights and duties. Others (called ‘information-forcing’ or ‘penalty’ default rules) are designed not to be left in place, but rather to encourage the parties themselves to explicitly provide some other resolution; these rules thus aim to encourage an efficient contracting process. This chapter describes the issues raised by such rules, including their application to heterogeneous markets and to separating and pooling equilibria; it also briefly discusses some non-economic theories of default rules. Finally, this chapter also discusses economic and non-economic theories about the general question of why contracts should be enforced at all.

JEL classification: K12
Keywords: Contracts, Incomplete Contracts, Default Rules

1. Introduction

This chapter describes research bearing on the general aspects of contract law. Most research in law and economics does not explicitly address these general aspects, but instead proceeds directly to analyze particular rules of contract law, such as the remedies for breach. That body of research is described below in Chapters 4100 through 4800.

There is, however, some scholarship on the general nature of contract law’s ‘default rules’, or the rules that define the parties’ obligations in the absence of any explicit agreement to the contrary. The phrase, ‘complete contingent contract’ is sometimes used to describe an (imaginary) contract that would spell out in complete detail the exact legal rights and duties of each party under every possible state of affairs. While no real contract ever achieves this level of completeness, the concept is still useful to define one endpoint of a spectrum of completeness. If any contract ever succeeded in reaching this endpoint, the law’s default rules would then be irrelevant, as no issue would ever arise that could not be settled by the terms of the contract itself. Indeed, the same is true
of the interpretation of most other legal documents, such as wills (see Chapter 5830) or legislative enactments (see Chapter 9200).

As long as legal documents fall short of this level of completeness, the law must have some set of presumptions or default rules, in order to resolve disputes that are not settled by the terms of the document itself. The law could, of course, simply refuse to enforce any contract (or any will, or any statute) that fell short of absolute completeness. But such a rule would itself be a ‘default rule’: it would be a legal rule defining the obligations (or lack of obligations) that result when a contract does not itself specify what rules should govern. As long as actual contracts fall short of full completeness, then, the existence of default rules is not so much a choice as a logical necessity. The only question is what the content of those default rules ought to be.

In choosing the content of default rules for contractual relationships, it is often useful to distinguish default rules chosen to increase efficiency if they are allowed to remain in force (as discussed in Sections 2-4) from those chosen to increase efficiency if many parties contract around the default rule (discussed in Sections 5-9). In some cases, the steps required to contract around a default rule could themselves increase efficiency, by inducing one party to reveal private information (as discussed in Sections 6-8). In other cases, it may be too hard for the law to identify a single rule that would increase efficiency if allowed to remain in force, so the default rule may instead be selected purely for its simplicity or ease of administration (as discussed in Section 9). The choice between these approaches depends in part on the ability of courts or lawmakers to identify efficient default rules; this issue is addressed in Section 10. The possible effect of default rules on the parties’ own preferences is discussed in Section 11, while Section 12 discusses some non-economic theories of default rules.

Finally, there is also a good deal of scholarship on the general question of when contracts ought to be enforceable, either through informal social sanctions or more by formal legal mechanisms. Section 16 introduces these issues, and discusses the choice of enforcement mechanisms. Sections 18-20 focus on the question of whether promises should be enforceable at all (whatever the mechanism), with Section 18 discussing noneconomic theories of enforcement and Sections 19-20 discussing the economic theories.

A. ‘Majoritarian’ or ‘Market-Mimicking’ Default Rules

2. Introduction to Majoritarian Default Rules

One way to select a default rule is to identify the rule that would be most efficient if that rule were allowed to remain in place (for example, if the parties
did not specify some other rule in their contract). For example, if the expectation measure of damages were determined to be the most efficient remedy for breach of contract, this approach to selecting a default rule would argue for making expectation damages the default remedy.

Much of the economically-oriented research on specific topics of contract law falls into this category. For example, many analyses of the remedies for breach aim to identify the remedy that would produce the most efficient result if that remedy were allowed to govern the parties’ relationship (for example, if the parties’ contract did not stipulate that some other remedy would govern). The same is true of many analyses of implied excuses such as impracticability and mistake, or of other terms such as implied warranties. This work is described at more length in Chapters 4500 (unforeseen contingencies), 4600 (remedies), and 4700 (warranties). This section addresses only those issues common to all default rules of this sort.

3. Majoritarian Rules and Hypothetical Consent

Default rules selected on this basis - that is, on the grounds of their efficiency if allowed to remain in force - are sometimes described as the rules that the parties themselves would have chosen, if they had taken the time to agree explicitly on a rule to govern their relationship. In most cases, parties to a contract have an interest in maximizing the efficiency of their relationship, so the rule the parties themselves would have chosen will be the same as the rule that would be most efficient if allowed to remain in place. Thus, default rules chosen on this basis have also been labeled ‘market mimicking default rules’, or default rules based on the principle of ‘hypothetical consent’. To be sure, if there are third-party effects, or if the parties to the contract are imperfectly informed or are subject to any other market failures, the rule that would be chosen by the actual parties might no longer coincide with the rule that would in fact be most efficient. For general discussions of the relationship between efficient default rules and hypothetical consent, see Posner and Rosenfield (1977, p. 89), Ayres and Gertner (1989, pp. 89-93), Coleman, Heckathorn and Maser (1989), and Craswell (1992). Section 4 discusses some additional complexities that arise in heterogeneous markets, in which the same rule would not be chosen by every pair of contracting parties.

This basis for selecting default rules has been supported by two economic arguments, both of which involve transaction costs. (Some noneconomic arguments will be discussed in Section 12 below.) The first economic argument applies when it would be prohibitively expensive for the parties to make their contract more complete by specifying the rule they want to govern a particular contingency. This is especially likely for extremely low-probability
contingencies, where the expected benefits of specifying a rule in advance are likely to be low. In such a case, any rule the law selects as a default rule will inevitably be left in place by the parties, so the only way to get the benefits of whatever rule is most efficient is to make that rule the default rule. The second economic argument applies if it would be costly, but not prohibitively expensive, for the parties to make their contract more complete by specifying their own rule. In such a case, selecting a default rule that matches whatever rule the parties prefer may save some parties from having to incur those transaction costs, thus producing all the benefits of the most efficient rule with lower total transaction costs.

These arguments become more complicated, however, if a ‘market-mimicking’ default rule is being applied to a market that is imperfect in some respect. For example, if a remedy of expectation damages would in fact be most efficient, but if imperfect information has led the parties to believe that some other remedy would be more efficient, should the law adopt as its default remedy the one that is in fact most efficient, or should it adopt the less efficient remedy that the (imperfectly informed) parties would choose if left to their own devices? The transaction-cost argument discussed in the preceding paragraph could suggest that, if the cost of contracting around the default rule is sufficiently low that the parties are likely to do so, then the law might as well adopt the less efficient remedy as the default rule, because that is the remedy the parties will contract for anyway. Alternatively, if the costs of transacting around the default rule are so high that the parties are likely to leave the default rule in place, perhaps the law should adopt the remedy that in fact is most efficient, as that will give the parties the benefit of the more efficient rule. But this latter approach is complicated by the fact that the identity of the rule that is most efficient (if allowed to remain in force) may itself change if the parties are imperfectly informed, or if there are other market imperfections. Moreover, in some cases the presence of information asymmetries or other market imperfections may call for an entirely different approach to the selection of default rules, in which the default rules are intentionally chosen not to be left in force by many parties. This approach is discussed below in Sections 5-9.

4. Majoritarian Rules in Heterogeneous Markets

Market-mimicking or ‘majoritarian’ default rules raise additional issues in markets characterized by heterogeneity, where different rules would be efficient for different contracting pairs. If a single default rule must be chosen to govern all contracts in such a market, the rule selected will determine which contracting pairs can do best by leaving the default rule in place and which pairs could potentially do better by incurring the transaction costs needed to
specify some other rule. It is sometimes assumed that the most efficient single default rule would be that which was most efficient for most of the contracting pairs (hence the label ‘majoritarian’ default rule, coined by Ayres and Gertner, 1989). But if transaction costs differ across heterogeneous parties, this conclusion actually depends on the exact levels of transaction costs that each member of each contracting pair would have to incur to contract around whatever default rule the law adopts (see Ayres and Gertner, 1989, pp. 112-115).

If different rules would be efficient for different contracting pairs, the law must also to decide the extent to which its default rules should be ‘tailored’, or customized to match the rule that would be most efficient for each individual contracting pair (see Ayres, 1993; Goetz and Scott, 1985; Baird and Weisberg, 1982). If the law adopts a single, untailored default rule, that will be most efficient for only some of the contracting pairs: all other pairs will have to incur either (1) the transaction costs of contracting around the default rule, or (2) the efficiency loss from leaving in place a default rule that is less than efficient for their contract. A set of default rules tailored to each individual contracting pair can in theory eliminate or reduce these costs, but such ‘tailoring’ will usually introduce other costs. For instance, individually tailored rules will usually be more complex, thus increasing the drafting costs that must be incurred by the legislature or other lawmaking body. Moreover, it is often impossible to spell out a complete set of individually tailored rules in advance, so the law will instead have to rely on vague standards to be applied by courts on a case-by-case basis (for example, an implied excuse in cases where performance is no longer commercially ‘reasonable’). Vague standards such as these usually entail higher litigation costs; they also introduce the possibility of case-by-case error in the application of the standard, and may make it hard for the parties to predict what rule will be applied to their relationship. In short, the question of how finely to tailor a default rule - and, indeed, whether to tailor it at all - raises most of the same issues that are raised whenever the law faces a choice between specific rules and vague standards (see Chapter 9000).

B. ‘Information-Forcing’ or ‘Penalty’ Default Rules

5. Introduction to Information-Forcing Default Rules

In many transactions, the two parties begin with differing amounts of information. In some cases, they may be differently informed about the relevant legal rules, or about the risks involved in the transaction. They may also be differently informed about the characteristics of the other party - for example, the seller may not know how much the buyer will lose if the seller’s product turns out to be defective. Differences such as these are often referred to as
‘information asymmetries’. While information asymmetries are often addressed directly through disclosure regulations and the like (see Chapter 5110), they can also have implications for the law’s choice of default rules. Section 6 below discusses cases where the two parties are differentially informed about the governing legal rule. Sections 7-8 then discuss cases where the two parties are differentially informed about some other aspect of the transaction.

6. Forcing Information about the Legal Rule

In some cases, the choice of default rule may help correct one party’s information about the legal rule itself. For example, suppose that the default remedy for breach provides for only a small monetary payment. Suppose also that the seller (the potential breacher) knows this, but the buyer (the potential nonbreacher) does not. Suppose that the buyer instead thinks, incorrectly, that the law’s default rule provides for a very large payment in the event of breach. If left uncorrected, this information asymmetry could lead to either of two problems. First, if a larger remedy would be more efficient for this contracting pair, they may not alter their contract to provide for the larger and more efficient remedy, because the buyer will think that he or she already has the benefit of a larger remedy. Second, even if the smaller remedy provided by the default rule is in fact most efficient for this contracting pair, the buyer’s ignorance about the actual rule may keep him or her from optimally adapting his or her behavior to that rule. For example, the buyer may purchase insufficient insurance, or take insufficient precautions to reduce the losses that would be caused by breach. (The effect of legal remedies on the incentives governing decisions such as these is discussed at more length in Chapter 4600.)

By contrast, suppose now that the default rule were changed to provide for a very large payment in the event of breach. If this remedy is left in place, that is, if the parties do not contract around it), then both parties will optimally adapt their behavior: the buyer will adapt because this rule matches what he or she thinks the rule is; while the seller will adapt because, by hypothesis, he or she is correctly informed about whatever the default rule is. And if this default remedy is not left in place (that is, if the parties contract around it by stipulating a smaller remedy), the act of stipulating some other remedy should serve to inform the buyer about what remedy will apply in the event of breach. Indeed, if the default rule is one that the parties will be certain to contract around - for example, if it is chosen to be highly unfavorable to whichever party is better informed about the law - that will virtually ensure that the other party will also become correctly informed about the resulting rule, by the very process of stipulating to some other rule in their contract. Hence, these default rules are sometimes described as ‘information forcing’ rules (Scott, 1990, p. 609), or as
‘penalty defaults’ slanted against the better-informed party to induce him or her to contract around the default rule (Ayres and Gertner, 1989, p. 97). A number of analyses of particular topics in contract law have recommended default rules chosen on this basis, or have suggested that actual legal rules could be understood as serving this function. See, for example, Ayres and Gertner (1989, pp. 98-99, 104-106), Goldberg (1984, pp. 295-296), Muris (1983, p. 390), Verkerke (1995, pp. 885-890), Isaacharoff (1996, pp. 1793-1795).

One difficulty raised by default rules slanted against the better-informed party is that it may not be clear which party is better informed. This raises another ‘tailoring’ question of the sort discussed above in Section 4: should the party against whom the default rule is slanted be determined in advance for a broad category of cases, or individually on a case-by-case basis?

Another difficult issue, even in cases where it is clear which party is least well-informed, concerns the comparison between the benefits and the costs of correcting the information asymmetry. The benefit of correcting the asymmetry will vary from case to case, depending on the information involved, so it is difficult to generalize. The costs of correcting the asymmetry clearly include the direct costs of contracting around the original default rule (which will depend in turn on the procedures that are required to contract around a default rule, as discussed in Section 14 below). The costs of correcting the asymmetry also depend on the risk that the parties will simply forget to contract around, thereby unintentionally leaving in force a default rule that was chosen not to be efficient for these contracting parties.

7. Forcing Information about the Other Contracting Party

In many markets there is heterogeneity among the potential parties on a single side of a proposed transaction. For example, sellers may differ in the reliability of their products, or buyers may differ in the losses that they would suffer if the product they purchase fails to perform. At the outset of the transaction, this information is often known to one party but not to the other - for example, each buyer may already know how much loss he or she would suffer, but sellers may have no way of knowing how much any individual buyer has at stake. This is another example of information asymmetry.

When this asymmetry is present, contract law’s default rules can again create incentives for one party to take actions that will reveal information to the other party. For example, if the default remedy for breach of contract provides for only a small monetary payment, buyers who would lose a larger amount in the event of breach may have an incentive to try to contract around that rule, by negotiating for a clause stipulating a larger payment in the event of breach. By contrast, if the default remedy is already set at a larger payment, buyers who
would lose less may be the ones with an incentive to contract around that rule, by negotiating for a smaller stipulated damage to get a more favorable price). In other words, the choice of default rule will determine which set of buyers - buyers with little at stake, or buyers with lots at stake - have an incentive to try to contract around the default rule. As a result, sellers may be able to infer something about the amount that any given buyer has at stake by observing whether that buyer does or does not try to contract around the default rule. And if buyers also differ in the gains they would get from transacting around the default rule, or in the costs they would face in doing so, the resulting equilibrium may depend critically on which default rule the law adopts. (The welfare consequences of these differences are discussed below in Section 8.)

These differences are often analyzed using signaling models from game theory (see generally Chapter 0550). If sellers have no information about how much any given buyer would lose in the event of breach, the result may be a pooling equilibrium in which all buyers are charged an identical price. But if this asymmetry can somehow be overcome, the result may be a separating equilibrium in which buyers who face large losses in the event of breach will have a right to collect those losses, but will pay a higher price (to compensate the seller for its greater potential liability), while buyers who face lower losses will pay a lower price. One way to achieve this separating equilibrium is to select a default rule that induces one of these classes of buyers to signal the amount they have at stake, by their actions in trying to contract around the default rule. For formal models of this effect see, for example, Ayres and Gertner (1989, 1992), Johnston (1990), Allen and Gale (1992), Hviid (1996).

8. Information-Forcing and Economic Welfare

Unfortunately, it is often difficult to assess the welfare implications of default rules that produce separating equilibria. Part of the difficulty is that, while there will usually be efficiency gains from eliminating this information asymmetry, that will not always be the case. Gains could arise if information about each individual buyer’s potential losses may let some sellers make customized adjustments to the reliability of their products. Gains could also arise if this information let sellers calculate a price that better reflected the true risk of selling to that particular buyer, thus sending buyers the correct signals about whether to purchase that product (see Ayres and Gertner, 1989; Quillen, 1988). In some cases, though, information about the buyer’s potential losses may be of no relevance to sellers - as, for instance, when sellers operate mass businesses that cannot practicably adjust their actions or prices for individual buyers (Danzig, 1975; Eisenberg, 1992, pp. 591-596). In such markets, there will be no efficiency gains at all from curing the information asymmetry. In
still other markets, there may be positive gains from curing the asymmetry but those gains may not outweigh the costs of communicating the necessary information (Bebchuk and Shavell, 1991).

Moreover, even when net efficiency gains are possible, the parties’ private incentives may lead them to act in a way that fails to achieve those gains. For example, if buyers face a price-discriminating monopolist who can charge higher prices to buyers who have more at stake in a particular transaction, buyers may be reluctant to do anything to reveal the amount they have at stake for fear of having to pay a higher price (Wolcher, 1989; Johnston, 1990; Ayres and Gertner, 1992). In other cases, some buyers may get private benefits (such as a more favorable price) by distinguishing themselves from other buyers, thus leading to socially excessive signaling (Rea, 1984; Aghion and Hermalin, 1990). At present, it is difficult to generalize about when the signaling or separating effects of a default rule will improve overall efficiency (see Hviid, 1996).

9. ‘Formalities’ as Default Rules

Still another approach to selecting default rules aims entirely at ease of administration. That is, rather than trying to identify the default rule that would (1) be most efficient if left in place, or would (2) induce an efficient disclosure of information, this approach aims for default rules that are easily administered by courts, and easily learned and understood by contracting parties. In an influential early article, Fuller (1941) coined the term ‘formality’ to refer to rules designed to induce parties to specify their intentions in a way that could easily be recognized by courts. In this respect, Fuller’s rules were designed to force the parties to disclose information to the court, rather than (or in addition to) forcing one party to disclose information to the other contracting party.

For example, if the parties to a contract fail to specify the quantity of goods they intend to convey, many jurisdictions refuse to enforce any obligation whatsoever, thus implicitly filling the gap with a quantity of ‘zero’ (see Ayres and Gertner, 1989, pp. 95-96). This default rule clearly is not designed to match what any contracting pair would have intended anyway. Instead, its purpose is to induce each pair to contract around the default rule by specifying the quantity they prefer, thus sparing the court from having to guess about the parties’ preferred quantity. Thus, unlike majoritarian or market-mimicking default rules, formalities are not necessarily intended to be left in place by the parties.

Unlike other ‘penalty’ or ‘information-forcing’ rules, however (see Sections 5-8), formalities need not be slanted against any particular party. For example, the default rule for contracts that fail to specify any quantity of goods could just as easily be set at any specific number, not necessarily zero. True, a default rule of ‘zero’ has certain administrative advantages over any other number: there
is no way to breach an obligation to deliver a quantity of zero, and hence no need to ever measure damages. Still, the essential feature of a formality is merely that some number be fixed in advance and be easily learnable by the parties, so that they can determine whether they will have to specify some other number in their contract. The formality could even be set at the number that most parties would prefer (if that number were known), thus giving it one feature in common with a market mimicking or majoritarian default rule. The only key is that, since formalities are chosen for their simplicity and ease of administration, they will necessarily be ‘untailored’ in the sense defined earlier in Section 4. As a result, no matter what specific rule is adopted as the default formality, many (perhaps most) contracting pairs will find it in their interest to specify some other rule in their contract.

Because formalities are designed to be easily administered, they will at least have the virtue of reducing litigation costs in all cases when the parties fail to specify some other rule in their contract, so the default rule remains in effect. This same ease of administration may also minimize another component of transaction costs: the cost to the parties of becoming informed about the legal rule, and of predicting how that rule might be applied. Also, if the formality is not chosen to be deliberately unfavorable to either party, it may reduce the costs imposed when parties forget to specify some other rule in their contract: in this way, formalities may reduce the cost of remaining uninformed about the law. (These two effects have offsetting influences on parties’ incentives to become informed about the law.)

C. Other Issues in Designing Default Rules

10. Default Rules and Institutional Capabilities

The preceding sections have discussed three bases on which default rules might be selected: market-mimicking or majoritarian defaults (Sections 2-4), information-forcing or penalty defaults (Sections 5-8), and pure formalities (Section 9). The choice among these various approaches depends in part on the competence and capabilities of legal institutions. For example, the first two approaches - market mimicking or majoritarian defaults, and information-forcing or penalty defaults - place obvious demands on those institutions, either to identify the rule that would in fact be most efficient (if left in place) for most contracting pairs, or to identify the default rule that would induce the most efficient disclosure of information. These demands are only increased if either of these approaches is to be adopted on a relatively ‘tailored’ basis, requiring the legal decision makers to assess the effect on individual markets or even individual contracting pairs. Thus, it is obviously an oversimplification to
analyze how either of these approaches would work if it were to be administered by an omniscient legal decision maker. The real question is how well either of these approaches will work when administered by actual courts and legislatures.

Two views of this question have developed in the contracts literature. The first begins by positing that certain rules - often, the rules that would have been efficient in a first-best world with perfect legal institutions - are simply beyond the capability of real legal institutions. This assumption is often made when the rule in question depends on information that, although observable by one or even both of the parties, cannot be demonstrated or verified publicly in court. For example, if the default remedy for breach were based on the net profits the breacher made, this would require a court to be able to measure the breacher’s revenues and costs, which might in some cases require more accounting expertise than real courts possess. Scholars often posit that such a rule would be unworkable in order to focus their analysis on alternative rules: rules that, while they might be less efficient in a first-best world, place fewer demands on the court or other legal decisionmaker (for example, Hermann and Katz, 1993). For a general discussion and defense of this approach, see Schwartz (1992, pp. 279-280, 1993, pp. 403-406).

By simply positing that certain rules are ‘unworkable’, though, these scholars implicitly assume an unfavorable balance between (1) the sum of all costs that would be imposed if courts tried to apply the unworkable rule, with the large number of errors that would entail; and (2) the sum of all costs associated with whatever rule is second-best. Other scholars have attempted a more finely-grained analysis by explicitly modeling the costs associated with judicial implementation of an unworkable rule. However, those costs depend in part on the nature and the probability of various errors courts might make in applying such a rule, and there is no consensus (and, regrettably, no empirical data) on how best to model such an error function. For two applications of this approach to contract issues, see Hadfield (1994), Hermann and Katz (1991) and Allen and Gale (1992). The general discussion of legal errors and uncertain rules in other bodies of law (see Chapter 0790) is also relevant here.

11. Default Rules and Preference Formation

In some situations, the law’s choice of default rules could alter contracting parties’ beliefs or preferences, thereby changing the value of the costs and benefits associated with different rules. For example, if the law adopts an implied warranty making sellers liable for all defects in their products (unless explicitly disclaimed), this could conceivably lead buyers to increase the value they place on such a warranty. At the same time, the adoption of the opposite default rule (no implied warranty) might lead buyers to the opposite view, if it
caused them to reduce the value they place on such a warranty. In that event, it would be possible for either default rule (an implied warranty, or the absence of an implied warranty) to be left in place by the parties, if the law’s adoption of each default rule changed the parties’ preferences sufficiently. It would also be possible for both default rules to be perfectly efficient, at least when judged by the preferences the parties would have once the law adopted either default rule.

The possibility that default rules might influence parties’ preferences is only just beginning to be explored. Experimental tests of this possibility can be found in Schwab (1988) and Korobkin (1998); the latter also discusses many of the potential policy implications. (Other experimental work regarding the effect on preferences of legal rules generally is discussed in Chapter 0570). Brief discussions of the implications for default rules in particular can also be found in Charny (1991, pp. 1835-1840); and in Schwartz (1993, pp. 413-415), who refers to default rules with this effect as ‘transformative default rules’.

12. Non-Economic Theories of Default Rules

While default rules have received less attention in scholarship outside of law and economics, there are several non-economic theories that deserve mention. For general discussions in the legal literature, see Charny (1991), Barnett (1992), and Burton (1993).

The first, and least well-developed (at least as a general theory), depends on the identification of certain rules as morally superior, as a sort of ‘merit good’. That is, just as it is sometimes argued that all citizens ought to have certain goods (education, health care, and so on), it is sometimes said that all citizens ought to have certain contract rights, such as the right to complete compensation in the event of a breach; or that the law should especially promote certain kinds of relationships, such as those based on long-term cooperation. Arguments of this sort could be particularly compelling if the law’s choice of default rule influenced citizens’ own beliefs about the value of certain relationships, as discussed above in Section 11.

A second non-economic theory rests on the idea of hypothetical consent. There is an entire family of non-economic theories that traces the binding force of contracts to individual autonomy, and to the fact that the contractual obligation was freely chosen by the contracting party (see Section 17 below). To be sure, these theories might seem to have little to say about the content of default rules, which (by definition) fill in the content of obligations that were not explicitly chosen by the parties (see Craswell 1989a). However, some have argued that, even in the absence of explicit consent, a party should still be bound to the rule that he or she would have consented to if he or she had explicitly negotiated an agreement on that point. Interestingly, this argument
converges with the market-mimicking or majoritarian approach discussed earlier in Sections 2-4. As a result, some economically-oriented scholars have sought to rest their recommendations on this philosophic position, as well as on more conventional economic grounds (for example, Schwartz, 1988, pp. 357-360). In the philosophical literature, however, it is disputed whether this invocation of hypothetical consent carries any of the moral force of ‘real’ consent, or whether it adds anything to standard efficiency arguments. For discussions in the legal literature of this issue, see Barnett (1992), Charny (1991), Coleman, Heckathorn and Maser (1989), Craswell (1992), and the exchanges between Coleman (1980a, 1980b) and Posner (1980, 1981). In the philosophical literature, useful discussions include Scanlon (1982), Gauthier (1986), and Brudney (1991).

A third non-economic approach suggests that default rules should be designed to mimic the norms or customs that are already observed, either in the community at large or in the parties’ prior relationship. This approach is often employed in the ‘relational contract’ theory usually identified with Ian Macneil (1978, 1980, 1981); see also Brown and Feinman (1991), Feinman (1993), and Craswell (1993b). In addition, some scholars working from philosophical theories of individual autonomy and consent have seen the prevailing norms and customs as acceptable sources of default rules, at least in the absence of any explicit agreement to the contrary (Barnett, 1992; Burton, 1993). In particular cases, this approach could of course converge with any of the economic theories discussed here, depending on whether the norms or customs happened to be efficient (see the general discussion in Chapter 0800). Among the issues raised by this approach are questions about the manner in which the norms or customs are to be identified (see Bernstein, 1996, pp. 1787-1795; Craswell, 1998), and about the ability of courts or other legal decisionmakers to properly carry out this identification (see Section 10 above).

D. Contracting Around Default Rules

13. Displacing Default Rules by Agreement

A default rule, by definition, leaves parties free to specify some other rule to govern their relationship if they so choose. There is, however, very little scholarly analysis of the steps the parties must take if they wish to specify some other rule. Most scholars implicitly assume that specifying some other rule would require a valid provision to that effect in a valid contract. This could suggest that the rules governing those steps can be left to the law of contract formation and contract interpretation, two bodies of law that are addressed at more length in Chapters 4300 and 4400.
Unfortunately, those bodies of law are themselves among the least analyzed portions of contract law. Moreover, the choices made by a legal system in selecting the rules of contract formation and interpretation can both be influenced by, and exert an influence upon, the choice of the original default rule. Sections 14 and 15 briefly discuss some of the interactions between these sets of rules.

14. Procedural Requirements for a Valid Agreement

Any attempt to displace a default rule will normally have to satisfy (at a minimum) all the requirements of an enforceable contract. For example, depending on the legal regime, the agreement may (or may not) have to be in writing, and it may (or may not) have to be supported by consideration. These requirements, and the other rules governing contract formation, are discussed at more length in Chapter 4300.

Should there be extra procedural requirements for parties who wish to contract around a default rule? If the default rule is presumptively valid, and especially if it is presumptively valid for moral or other non-economic reasons, it might be argued that parties who wish to choose a presumptively less valid rule should have to take extra steps, if only to ensure (and to demonstrate to a court) that this is what the parties really intend. For example, it might be argued that one party should not be able to displace a default rule by means of a single clause buried in a 30-page document that the other party will not normally bother to read. Arguments of this sort are often made in connection with standard form contracts and the legal doctrine of unconscionability (see Chapter 4100). Some of the links with standard default rule analysis are noted briefly in Craswell (1993a, pp. 12-14).

Indeed, arguments of this sort may fit best in connection with information-forcing or penalty default rules. As discussed earlier (see Section 6), such rules may be deliberately designed to induce the better-informed party to contract around the default rule, in the hope that the process of contracting around the rule will itself give the other party more information about the rights that he or she now has. The efficacy of this approach, however, depends on whether the process of contracting around the rule really does inform the other party - and this, in turn, depends on the procedures that are required to contract around the rule. Again: if a default rule can be altered merely by inserting a clause in a 30-page contract that nobody ever reads, the process of contracting around the default will not inform the other party at all.

On the other hand, extra procedural requirements will usually increase the transaction costs required to contract around the default rule. For example, if altering a default rule requires the clause in question to be pointed out and
explained to the other party, the process of contracting around the default might indeed increase the other party’s information. However, this requirement may also be costly to satisfy, especially if a single contract alters a large number of default rules, and if the other party has no desire to sit through a lengthy set of explanations. In that event, the cost of altering the default rules via this procedure might turn out to be effectively prohibitive, so many parties would end up leaving the original default rules still in force. If so, this would require reevaluation of the original decision to select a default rule that was designed not to be left in force (that is, a default rule that was chosen to be inefficient precisely in order to induce the parties to contract around it). In this way, the rules governing the process of contracting around a default rule can themselves affect the principles on which the original default rule should be chosen.

15. Interpreting Contractual Terms that Alter a Default Rule

Contracting parties often use language which appears to address an issue that would otherwise be governed by a default rule, but which is vague or ambiguous, and thus requires interpretation. For example, imagine a contract providing that the remedy to be paid in the event of a breach should be an amount that would ‘reasonably compensate’ the nonbreacher. Should the courts treat this vague language as leaving a gap in the contract, to be filled by the normal default remedy? Or should they instead apply some other set of principles?

In conventional legal terminology, default rules are said to apply only when there is an actual gap in the contract, while questions of vague contractual language are usually described as questions of interpretation. There is, however, no consensus (and very little theory) on what should count as a ‘gap’ (see Ayres and Gertner, 1989, pp. 119-120). Fortunately, this distinction often will not matter, because the approaches used to interpret contracts (see Chapter 4400) have much in common with the approaches used to select default rules. In many cases, for example, vague or ambiguous language is interpreted so as to fit whatever the parties probably would have agreed to if they had discussed the matter, thus producing the same result as the majoritarian or market-mimicking default rules discussed in Sections 2-4 (see also Goetz and Scott, 1985). In other cases, vague or ambiguous language is interpreted against the party who drafted it, just as in the case of a penalty or information-forcing default rule designed to induce more careful and explicit communication (see Sections 5-9).

Notice, though, that some decision must still be made to determine when the parties have taken enough steps to make their language sufficiently clear to avoid the rule construing ambiguities against the drafter, or when their language is sufficiently clear to avoid being interpreted in accordance with the
court’s idea of what most parties would have wanted. This is the question of precisely how far the parties must go to contract around an otherwise applicable default rule (or, in this case, around an otherwise applicable rule of interpretation). As discussed in Section 14, whatever requirements the law adopts here will have to be considered in deciding what default rule (or what interpretive presumption) to adopt in the first place.

E. The Enforceability of Contracts Generally

16. Introduction to Theories of Enforceability

The analysis of default rules takes it as given that contracts are enforceable, and concerns itself with determining the content of the enforceable obligation. This Section addresses the opposite issue: given a contract of (let us assume) definite content, when should that contract be legally enforceable?

One branch of this question asks whether contracts should be enforced by a legal system, or whether non-legal enforcement mechanisms might be superior (see, for example, Bernstein, 1992, 1996; Charny, 1991). Non-legal enforcement mechanisms could include arbitration panels or trade association boards that function much like a court, except for not being backed by official state sanctions. They could also involve much less formal mechanisms, such as the threat of withholding business from anyone who had broken a promise in the past. In both of these respects, they are similar to the non-legal mechanisms that might be used to enforce any other obligation (see Chapters 0780 and 0800), with which this literature has much in common.

By contrast, most of the literature on whether contracts ought to be enforced typically abstracts from the choice of the enforcement mechanism, and focuses instead on the question of whether (and why) contracts ought to be enforced by any mechanism whatsoever. Non-economic theories are discussed briefly in Section 17 below, followed by a discussion of the law and economics literature in Sections 18-19.

17. Non-Economic Theories of Enforceability

Moral philosophers have written extensively about the question of whether (or why) a promise should be morally binding. Some of their theories rest on utilitarian accounts that have much in common with the economic theories; these will be discussed in Sections 18-19. This section focuses instead on the non-utilitarian or non-instrumental accounts. In the legal literature, general
surveys of these accounts can be found in Atiyah (1981), Barnett (1986), Craswell (1989a, pp. 491-503), and Gordley (1991).

One family of theories derives the obligation to keep one’s promise from considerations of individual freedom and moral autonomy. For example, Fried (1981) argues that if the law did not allow individuals to bind themselves with respect to their future conduct, it would thereby fail to respect the individuals’ status as autonomous moral agents. (For a variant account of when individual autonomy requires that promises be enforced, see Barnett, 1986.) These accounts do not imply that promises ought to be kept in every circumstance: this theory, like all of the others discussed in this section, allows for the possibility of a set of implied conditions or implied excuses. Instead, the goal of these theories is to explain why promises are prima facie binding, and this family of theories rests that binding force on the promise’s status as the voluntary commitment of a autonomous moral agent.

Another, different set of theories rests the obligation to keep a promise on the substance of what has been promised. Perhaps the best-known theories in this set focus on harm to others: if one’s promise has led others to change their behavior in reliance on the promise, so that they would now be injured if the promisor failed to perform, that promise is binding because it is wrong to harm other agents. Making a promise and then failing to keep it might also be seen as a form of lying, for such a promisor has misled others about the future. Other theories in this set focus on reciprocity or restitution: if one has received a benefit and promised to pay for it, that promise is binding because it is only right to pay for the benefits one has received (Atiyah, 1981, pp. 34-36). Under some theories, the fairness of the terms of the promise is also relevant to whether the promise is binding (for example, Gordley, 1995).

Under any of these substance-based theories, the promise’s status as the voluntary commitment of a free moral agent is less significant, for the agent is (in some sense) promising to do what he or she ought to be doing anyway. Indeed, under many of these theories, if there is no independent reason for the agent to perform the action - for example, if the agent has received no benefits, and if no one has yet relied on the promise - the obligation to keep the promise is seen as weaker or non-existent. These theories thus have difficulty explaining why the obligation should be stronger in the case of an individual who has made a promise than it is in the case of an individual who has acted identically but has explicitly disclaimed any promise (for example, ‘I think I’m going to do x, but I warn you that I’m not promising to do it, so you should rely only at your own risk’, or ‘you can confer those benefits on me if you want, but I warn you now that I have no intention of paying you anything for them’). Atiyah (1981, pp. 184-202) has argued that an explicit promise could serve as evidence, perhaps even conclusive evidence, that the underlying obligation really is a just one that ought to be enforced. If one asks why an individual’s
promise ought to be given such evidentiary weight, however, it is difficult to avoid falling back either on autonomy-based explanations of the sort discussed earlier in this section, or on utilitarian theories of the sort to be discussed in Sections 18 and 19.

18. Economic Theories of Free Exchange

Economic theories of why promises should be enforced can be divided into two categories. One set of theories, to be discussed in this section, focuses on the utility created by the eventual performance of the promise. The other set, to be discussed in Section 19, focuses on the utility created by the ex ante incentives that the promise sets up.

The first economic argument for enforcing promises rests on the utility that will be produced when the promise is eventually performed. Since voluntary transactions generally increase the welfare of all parties to the transaction, whenever a promise is voluntary it could be argued that welfare will usually be increased if the promise is carried out. Viewed in these terms, the economic argument for enforcing promises is very similar to the economic argument for free exchange and free markets generally (see Chapter 5000).

This argument, however, could imply that a promise should not be binding if conditions have changed sufficiently that the promised transaction would not increase both parties’ welfare. For example, if a promisor’s costs have increased to the point where it is no longer efficient for him or her to perform the promise, this theory could suggest that the promisor ought to be released with no obligation to pay any damages at all, because this particular efficiency rationale for enforcing promises no longer applies. Under this theory, the only way to save the prima facie obligation to keep one’s promises (even after conditions have changed) is to argue for the virtues of a general rule over a case-by-case inquiry into the efficiency of any given transaction - in philosophical terms, by shifting from act-utilitarianism to rule-utilitarianism. That argument, in turn, must rest on empirical claims about the ability of courts to determine whether and when performance of the promise was still efficient.

More generally, there is an important difference between permitting free exchange and permitting (or enforcing) binding promises. An exchange can take place instantaneously, but a promise necessarily involves a commitment to act in a certain way at some time in the future. Once this temporal element is recognized, it can be seen that enforcing promises does not simply transfer existing goods from one owner to another. Instead, the enforcement of promises creates a new good: it allows people to exchange ‘wheat to arrive on September 1’ (for example), where without a binding promise they could only exchange actual bushels of wheat. Viewed in these terms, the economic case for enforcement rests on the proposition that this new good (that is, the good
represented by a future commitment) is a socially useful good which people frequently will want to exchange. To explain why such a commitment is useful and valuable, economists have focused more on the ex ante effects created by such a commitment, as discussed in Section 19 below.

19. Economic Theories of Advance Commitments

Most economic justifications for enforcing promises have focused on the ex ante effects that would be created by a general rule of enforceability. In essence, these economic theories analyze promises as effectuating a present transfer not of goods and services, but of rights and duties. (For a similar view expressed in non-economic terms, see Barnett, 1986, pp. 291-300.) Such a transfer of rights and duties may itself produce efficiency gains, independently of whether it is efficient to actually carry out the promise.

The efficiency gain that is analyzed most often stems from the effect on the promisee’s incentive to rely on the promised behavior (see Goetz and Scott, 1980). If the law denied promisees any compensation whenever a promisor could show that performance of the promise had become inefficient, this would shift more of the risk of a change in conditions to the promisee, and thus would reduce the promisee’s incentive to rely. To be sure, if the law instead guarantees that promisees will be compensated whenever the promisor fails to perform, this could create incentives for excessive reliance on the part of the promisee (see the discussion of reliance incentives in connection with remedies for breach in Chapter 4600). If the incentives for excessive reliance can be constrained by other legal doctrines, however, the net effect on the promisee’s reliance incentives could still be positive. The argument that enforceability is needed to induce one party to take the risky step of beginning his or her own performance (when the other party’s return performance is not due until later) is really a special case of this argument, since beginning performance is one of the many ways in which parties may rely on a contract.

More generally, there are many other ex ante effects that could be produced by a rule requiring compensation even when conditions have changed to make performance unprofitable. If the promisor and promisee differ in their attitudes toward risk, a rule requiring compensation may achieve a better allocation of risk between the two parties (see, for example, Polinsky, 1983). A rule requiring compensation may also increase the promisor’s incentive to tell the truth about the conditions that will affect his or her performance, thus increasing the promisee’s information about those conditions (see Shavell, 1991; Craswell, 1989b; Katz, 1996, pp. 1289-1291). A compensation requirement may also give the promisor an incentive to affect those conditions directly, if the probability of performance rests on factors that are within the
promisor’s control. Generally speaking, all of the economic analysis of efficient remedies for breach (see Chapter 4600) is relevant here, since to make a promise enforceable is to set a non-zero remedy for the promise’s breach. The less-developed economic analysis of the conditions under which formation of a binding contract should be inferred (see Chapter 4300) is also relevant here.

Acknowledgments

Richard Craswell is a professor of law at the Stanford Law School. Jack L. Goldsmith, Eric A. Posner, Cass R. Sunstein, and two anonymous referees provided helpful comments. Financial support was provided by the Lynde and Harry Bradley Foundation and the Sarah Scaife Foundation at the University of Chicago Law School.

Bibliography on Contract Law: General Theories (4000)


Marini, Giovanni (1990), *Promessa ed Affidamento nel Diritto dei Contratti (Promise and Reliance in Contract Law)*, Napoli, Jovene.


Other References


