

The Economics of Payment Finality

CHARLES M. KAHN AND WILLIAM ROBERDS

Kahn holds the Bailey Memorial Chair of Finance and Economics at the University of Illinois. Roberds is a vice president and economist in the Atlanta Fed's research department. They thank Richard Fraher, Karsten Jeske, and Larry Wall for comments on earlier drafts. Although the discussion below touches on certain legal issues, the authors are laypersons in the law and their remarks should not be construed as legal advice.

The U.S. economy runs on debt. Not counting the debt of financial institutions (which issue debt to some parties while holding others'), the total debt outstanding in the U.S. economy is almost \$19 trillion, or upwards of \$67,000 for each U.S. resident. But this debt is not static in any sense: each day about \$2.5 trillion worth of debt is incurred, paid off, or transferred through the operation of the nation's various payment systems.¹ This vast and incessant reshuffling of financial relationships—from \$10 credit card purchases to \$100 million bond trades—makes possible the complex patterns of exchange necessary for the economy to thrive and grow.

Payment systems based on “inside money,” or the transfer of debt (such as checks, credit cards, and wire transfers), are crucial to exchange because they allow one debt claim (for example, a utility bill, a mortgage payment, or a trading position in the financial markets) to be extinguished by the transfer of another (bank deposits). This type of payment increases economic efficiency because it induces a sort of economy-wide netting out of obligations, with two main benefits. The first benefit, well known to economists, is that the use of inside money can economize on the amount of cash—which pays no interest and hence bears an implicit tax—needed to carry out exchange.

A second benefit of inside-money systems, which this article focuses on, is that these systems, by allow-

ing some claims to effectively cancel one another, help ensure that outstanding debts that remain are those for which there is a strong incentive to repay. Key to this second benefit is the idea that certain transfers have the property of finality. Transfers of inside money, meaning nowadays the transfer of money held as bank deposits, are supported by a well-developed legal infrastructure consisting of laws, regulations, and private contracts. This infrastructure determines, among other things, the finality of a given transfer—the circumstances under which the transfer of an asset (in practice, almost always a deposit or line of credit with a bank) extinguishes a debt. Without finality, a transfer of bank funds would not necessarily constitute a payment and “money in the bank” would not function as money. When a person sends a check to pay a utility bill, for example, she expects that her debt to the utility company will be discharged so long as there are sufficient funds in her account to cover the check.²

While the concept of finality is well understood by lawyers and payment system practitioners, it has received relatively little attention from economists. This article partially rectifies this omission by providing a basic analysis of finality and its role in facilitating exchange. The discussion shows how finality can increase economic efficiency by suitably allocating the risks associated with decentralized exchange.³ In some cases, finality may be necessary for exchange to occur at all. It will also be argued

that finality is best understood as a general concept rather than a specific rule or type of risk allocation. Many things can go wrong in the process of decentralized exchange (default and fraud, for example), so that the notion of finality appropriate in one set of circumstances will not necessarily be the right one in other circumstances. Choosing the degree of finality for a given situation involves a trade-off between the benefits of finality versus the costs that can arise when an erroneous or fraudulent transfer occurs.

Payment by transfer of debt is desirable in the model economy because it can short-circuit credit chains and thereby lessen the chances for opportunistic default.

A Simple Model of Transferable Debt and Finality

One can better understand the role of payment finality in today's economy by understanding its role in a simpler context. To this end, one can analyze a model in which it is desirable for people to make payments by transferring debt. In the model economy, unlike present-day economies, there are no banks, but individuals can transfer debt that they issue on their own. While this model is not appropriate for most modern situations, it has some historical verisimilitude. In earlier times, the debt of private individuals and firms (often in the form of bills of exchange) commonly circulated as a type of money. This system—payment in negotiable instruments—originated in the Low Countries during the seventeenth century and was subsequently adopted in many other countries, including the United States.⁴

In the model economy, depicted in Chart 1, there are three people in a typical business setting: a supplier, A; a merchant, B; and a consumer, C. Supplier A can provide merchant B with certain goods, “supplies.” If B has access to supplies, then she can transform these into salable merchandise. Consumer C would like to purchase the merchandise from B right away, but is momentarily strapped for cash. However, C expects a cash inflow at some later date.

How should exchange be organized in this simple economy? In modern circumstances, both B and C could apply for bank loans. With the loan proceeds, B could buy supplies from A, and C could buy merchandise from B. Using the proceeds of her sale to C, B could then repay her loan.⁵ Finally, C would repay his loan as his income becomes available. This type of arrangement would have been problematic in earlier times, however, as there were relatively few banks, and often these could not directly transfer funds outside of a given city. So for long-distance trade other arrangements had to be developed.

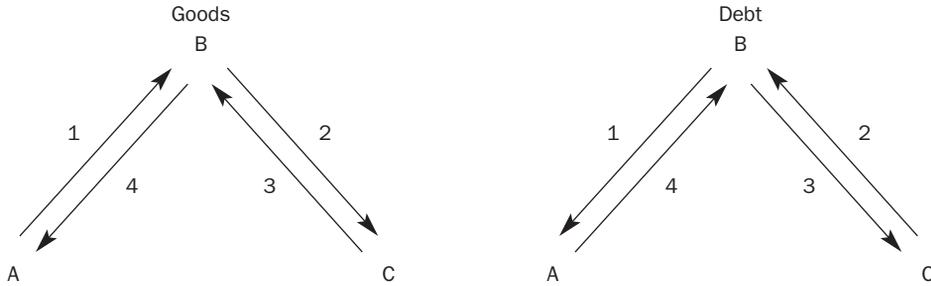
Returning to the model, trade could proceed if C were willing to wait until he has enough cash to purchase the merchandise. But if C values having the merchandise now rather than later, there may be better arrangements available. One alternative would be a credit chain (see Possibility 1 in Chart 1). B could issue debt to A in return for supplies, and C could issue debt to B in return for merchandise. Both forms of debt claims are payable later, meaning when C's cash arrives. When this happens, B presents C with the debt C issued earlier and A presents B with his. If B and C are reliable creditors, C can pay B, and B can use the proceeds from this payment to pay A. In an environment where creditors' rights are limited, however, there is a good chance that B may be unreliable. Because B no longer has possession of the merchandise when the time comes for her to repay, A's ability to coerce payment by using the merchandise as collateral is

1. Estimates of nonfinancial debt are available from the Federal Reserve's H.6 data release (available on-line at www.federalreserve.gov/releases/h6). Data on payment flows, not counting cash, are available from the Bank for International Settlements (2001).
2. One could attempt to pay such a bill by offering another good or service, that is, by barter. This possibility is acknowledged in the Uniform Commercial Code, but since a match between the needs of the buyer and seller is highly unlikely, as a practical matter virtually all payments are in money. See also the discussion in the Task Force on Stored-Value Cards (1997, 673–74).
3. Much of the discussion below is based on Kahn and Roberds (2001). Technical details of this article's arguments can be found in that paper.
4. See, for example, Kohn (1999) on the origins of negotiable instruments. Rogers (1995) explains how these instruments were incorporated into English and eventually American commercial law. Payment in bills of exchange enjoyed something of a revival in Russia during the 1990s (see Andrews 1997). There, short-term obligations of firms known as *veksels* often functioned as a makeshift form of money. Ickes (1998) describes how the use of *veksels* arose partly from the need to limit opportunities for default.
5. As in the arrangements discussed below, for this scheme to work, B must have an incentive to repay her loan. In modern settings, banks and other lenders have a variety of legal devices available to help guarantee repayment. For example, where the merchandise consists of automobiles, a lender may retain title to the automobiles until they are sold.

CHART 1

Possible Arrangements for Trade in a Simple Model Economy

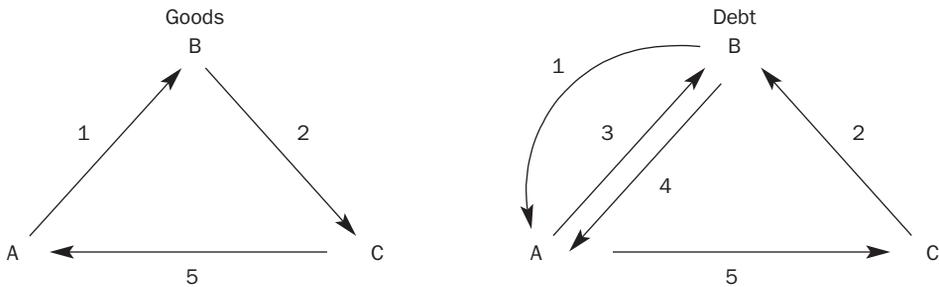
Possibility 1: Credit Chain



Sequence of Events

1. A sends supplies to B and receives B's debt in return.
2. B sends merchandise to C and receives C's debt in return.
3. B presents C's debt for redemption; C pays cash to B (or defaults).
4. A presents B's debt for redemption; B pays cash to A (or defaults).

Possibility 2: Credit Chain with Debt Transfer



Sequence of Events

1. A sends supplies to B and receives B's debt in return.
2. B sends merchandise to C and receives C's debt in return.
3. A presents B's debt for redemption.
4. B transfers C's debt to A (or defaults).
5. A presents C's debt for redemption; C pays cash to A (or defaults).

limited. In such cases the credit chain may fail: B may take C's money and run.

If transferable debt is introduced into the model (see Possibility 2 in Chart 1), another alternative becomes available: in lieu of cash, B can use C's debt to settle with A. That is, suppose that A grants credit to B but that both agree that the credit will be short-term and that B's debt to A can be discharged by the transfer of a third party's (such as C's) debt to A. B then sells merchandise to C in return for debt. In contrast to the earlier situation, however, B does not hold C's debt but instead passes (at least some portion of) C's debt to A in order to repay her short-term loan. Then, when C's cash becomes available, C pays A directly, avoiding a possible default by B.

Clearly, the desirability of this sort of scheme would depend on certain conditions being satisfied. First, C must be a stronger credit risk than B; for example, C could be wealthier or less indebted than B. Also, since C gains possession of merchandise in return for issuing debt, it could be that A could prevent C from making full use of the merchandise if B defaults on her debt. Second, B must have an incentive to hand over C's debt to A. Such an incentive will exist if, in case B defaults on her obligation to A, A may restrict B's ability to collect on C's obligation. The earlier widespread use of transferable debt suggests that these conditions were often satisfied in practice.

In summary, payment by transfer of debt is desirable in the model economy because it can

short-circuit credit chains and thereby lessen the chances for opportunistic default. But some failures in this process can occur in even the best-designed payment system. Rules governing payment finality are then needed to sort out who bears the losses. The discussion below will show how these rules can vary across different payment systems. Nonetheless, all of these arrangements share the fundamental feature, in the course of normal trade, of pulling the middle party out of the credit chain.

The introduction of negotiable debt greatly increased the efficiency of long-distance trade. In addition to providing strong incentives for repayment of debts, negotiability economized on the use of costly coin.

Finality and Negotiability

Kohn (1999) relates the failure of some early attempts to use transferable debt as a form of payment. One of the most critical problems was that of adverse selection; that is, when a debtor can settle a debt by transferring the debt of a third party to his creditor, there is always a temptation to pay with a lower-quality or even fraudulent instrument. In the context of the model, for example, the merchant B might try to settle a debt with supplier A by tendering the debt of a risky or nonexistent third party, C, or by forging the debt of a legitimate third party.

In response to such problems, societies had to develop a workable set of rules concerning liability in cases where there was a breakdown in the relationship among the three parties (buyer, seller, and third-party debtor). In other words, rules were needed to decide when a transfer (or a promised transfer) of debt discharged an obligation and when it did not. Since a transfer of coin would have unconditionally discharged the obligation, such rules effectively determined when debt would function as money.

The set of rules that resulted led to the establishment of the legal concept of negotiability, a concept which survives to some extent even today. Debt that was issued and transferred according to these rules was negotiable, or generally acceptable in exchange, because merchants understood in advance how losses would be allocated in case of

fraud or default. While a formal description of negotiability is somewhat involved, for the present purposes the two most essential aspects of negotiability are the ideas of “indorsement” and of a “holder in due course” (see Winn 1998 for a more complete discussion).⁶

Initially, a negotiable instrument could be transferred only by an indorsement or signature of the payee.⁷ The indorsement implied a conditional liability: if the original debtor could not pay, then the indorser’s obligation was not discharged and the indorser would become liable for the obligation.⁸ In other words, a payment in third-party debt was final only if the third party actually honored the debt. In the model, for example, B could pay A by indorsing over C’s debt to A, but such a transfer would be final only if C made the promised payment.⁹

Somewhat paradoxically, the indorsement requirement increased the acceptability of debt instruments in exchange by limiting the finality of their transfer. The implied contingent liability for a third party’s debt served as a powerful incentive for would-be transferors of such debt to screen the quality of any debt they might want to transfer. In his discussion of different forms of finality, Mengle (1990) classifies this type of loss-allocation rule as obeying a “least-cost-avoider” principle—liability is assigned to the party able to avoid the loss in question (here, fraud or default on a third party’s debt) at least cost. The indorsement rule, in effect, applied this principle through the presumption that someone attempting to pay with a third party’s debt would know more about the quality of that debt than someone receiving the debt in payment.

While indorsement amounted to a weakening of finality, paying a debt by indorsing someone else’s debt still represented an improvement over simple credit chains. As long as the original debtor (such as party C in Chart 1) honored his debt, intermediate parties (such as party B) were in effect removed from the credit chain. This removal reduced the need for payments in cash (coin or precious metal) and limited the scope for opportunistic default. Only when the original debtor could or would not pay did the chain of obligations come into play, and such defaults were relatively rare.

Indorsement served to limit the scope for buyer-side fraud, that is, fraud that would result from a buyer passing on a low-quality or forged debt instrument to a seller. The doctrine of a holder in due course played a similar role in limiting opportunities for fraud by the seller. That is, a buyer who paid for a purchase with negotiable debt could not repudiate that debt, when the debt was presented

by a good-faith third party, by claiming that the goods paid for with the debt were faulty or non-existent. For example, in the model this principle would mean that if C issued debt to B in order to pay for merchandise and B subsequently transferred that debt to A, C would have to honor the debt even if the merchandise were somehow defective (though C could still try to recover damages from B). As long as A had obtained C's debt in return for delivering goods to B in a legitimate transaction (not a sham transaction and without knowledge of fraud by B), then A would be a holder in due course of C's debt, with full rights to enforce the debt claim.

As with indorsement, this holder-in-due-course feature made negotiable debt more acceptable in exchange. In contrast to indorsement, however, this process was accomplished by extending the set of circumstances in which a debt transfer was final. The effect was to give buyers a strong incentive to screen the reliability of potential sellers and, consequently, an incentive to limit the incidence of seller-side fraud. Again, this allocation of losses can be defended as an application of the least-cost-avoider principle.

The introduction of negotiable debt greatly increased the efficiency of long-distance trade. In addition to providing strong incentives for repayment of debts, negotiability economized on the use of costly coin. Payment in negotiable debt thus allowed merchants to conduct a larger volume of trade than would have otherwise been possible. Many of the underlying ideas of these early payment arrangements, particularly rules concerning finality, remain relevant today. With appropriate modifications, these ideas have been incorporated into more modern and familiar systems.

Transferable Debt in Modern Payment Systems

History shows that payment systems based on the circulation of negotiable debt worked reasonably well, but their use was limited to wholesale transactions. A system based largely on the perceived value of a signature was bound to be

restricted to those individuals whose wealth and reputation were well known within the merchant community. By contrast, modern payment systems are widespread (if not quite universal) in their coverage: anyone with a bank account can pay anyone else with a bank account, without using cash. This arrangement is possible in part because banks have established an infrastructure that allows them to efficiently clear payments but also because there are well-understood rules that govern the finality of payments across banks.

To see how payment might work in a more modern context, consider the previous example of the supplier, the merchant, and the consumer. In a modern situation, each of the three parties would have his or her own bank account (see Chart 2), and trade would typically be financed by banks. That is, merchant B would obtain her supplies from supplier A by means of an inventory loan from her bank (that is, Bank II in Chart 2). Consumer C could also take out a loan from his bank (Bank III) to purchase merchandise from B. B and C would make their respective payments through traditional channels, say, by writing a check. B's check would effectively transfer an asset, her credit line with Bank II, to A's account at Bank I while C's check would transfer his credit with Bank III to B's account at Bank II (see Chart 2c). B could then use funds received from C to pay off her inventory loan from Bank II. With each check payment, the relevant check would be presented to the check writer's bank for payment.¹⁰ These obligations would typically be settled by transfers of "cash" (for example, by transfer of funds in each paying bank's account at the Federal Reserve).¹¹ Ultimately, C would repay his loan to Bank III, and A would draw down his funds at Bank I, completing the cycle of trade.

As in the earlier example, the efficiency of this system derives partly from the fact that the middle party B does not have to pay cash for her supplies so long as she can pay off one type of debt (her inventory loan from Bank II) with another (essentially, with C's deposit at Bank III). The system

6. In the legal literature, the spelling "indorsement" is preferred to "endorsement."

7. A person who received the debt in this manner (an indorsee) could transfer it to another party by indorsing it again. Instruments designed to be transferable without indorsement, such as bank notes, were not widely used until later.

8. In fact, the original debtor was most often the acceptor of a bill of exchange, that is, the party who agreed to be responsible for its payment.

9. In theory the promised payment was always due in coin, but in practice some netting of payments occurred.

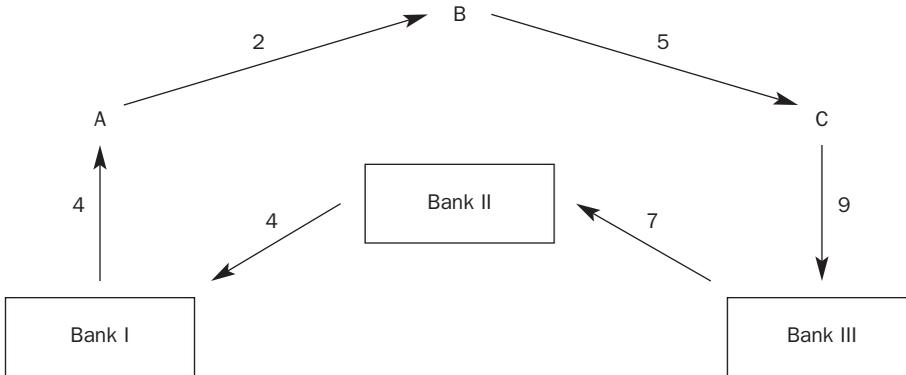
10. There are numerous ways checks can clear—by direct presentment (say, through a courier), through a private clearinghouse arrangement, or through the Federal Reserve System.

11. There are also different ways in which a check can settle, including settlement through a private clearinghouse or correspondent arrangement.

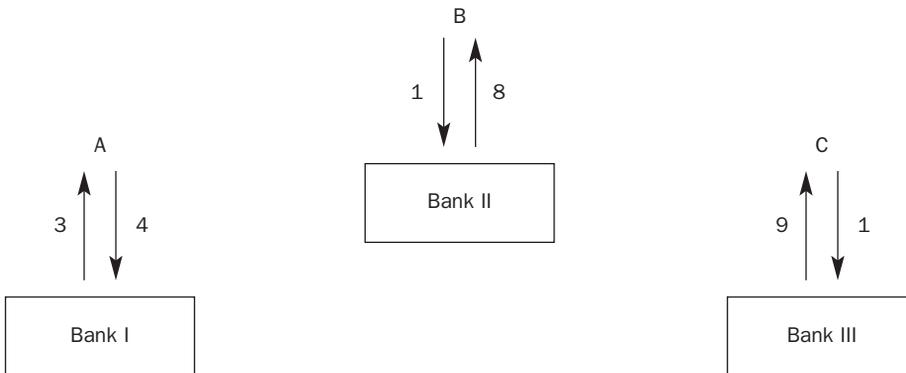
CHART 2

Trade with Checks

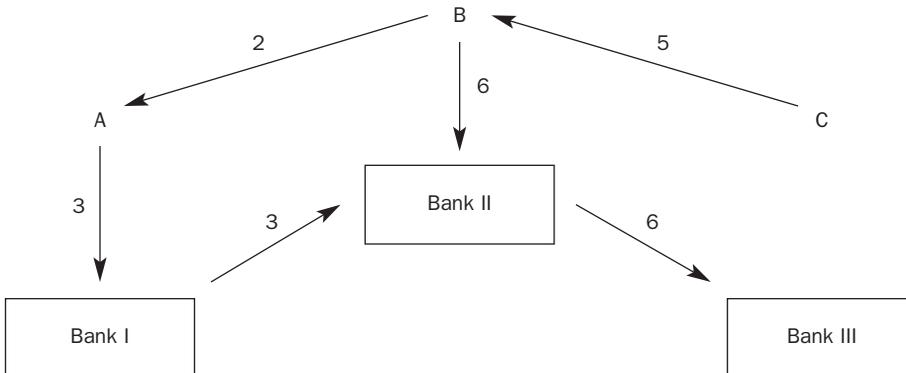
2a: Flow of Goods



2b: Issue of Debt



2c: Check Clearing



Sequence of Events

1. B issues debt to (obtains inventory loan from) Bank II; C obtains loan from Bank III.
2. A sends supplies to B in return for a check drawn on B's account.
3. B's check is deposited into Bank I; Bank I presents B's check for payment.
4. Bank II pays B's check and debits B's account; A can withdraw funds as needed.
5. B sends merchandise to C in return for a check drawn on C's account.
6. B deposits C's check into her account at Bank II; Bank II presents C's check for payment.
7. Bank III pays C's check.
8. Bank II debits B's account and B's loan is repaid.
9. C realizes cash inflow and repays loan to Bank III.

would not work nearly as well if each bank had to settle every payment one-for-one by transfer of reserves. But since on a normal business day a bank will have many incoming and outgoing funds transfers, on net the transfer of reserve balances necessary for interbank settlements is quite small relative to the gross volume of payments.¹²

Trade based on payment in bank funds is also desirable because it helps limit opportunities for default. Requiring middle parties (such as B in Chart 2) to pay in bank debt allows for reduced credit exposure to such parties as compared to a trading arrangement based on open credit chains. In other words, suppliers (such as A in Chart 2) need not wait for repayment until their customers' customers can afford to pay in cash but instead can count on immediate repayment or repayment after a short, fixed term. Banks end up bearing the residual credit risk in these arrangements (in the Chart 2 example, Bank III must bear the risk that C will not repay his loan), but banks are usually well equipped to bear such risks, having both good information about potential borrowers' creditworthiness and ample legal clout to enforce their creditors' claims in the courts.

The Role of Finality in Check Payment

Modern payment systems are useful because of their ubiquity, but the large volume and scope of payments means that there are many opportunities for things to go wrong.¹³ In the case of check payments, for example, checks are readily forged, altered, or written on insufficient funds. Who bears the losses in such cases?

In the case of checks, the rules for finality are at first glance not too different from those that governed the circulation of negotiable debt hundreds of years ago.¹⁴ Technically, checks are still regarded as negotiable instruments, and much of the law involving check payment is still based on classical concepts of negotiability. In particular, while the concept of

holder in due course is of limited applicability in modern situations (see for example, Mann 1999, 445–47) the notion of indorsement and the conditional liability that it implies are still relevant. In most cases, check payments are not final until they have been settled between the banks involved. For example, a bank may give a depositor conditional access to funds on a deposited check before payment has been finalized between banks. But by indorsing a check at the time of its deposit, a depositor recognizes his contingent liability in case the check is not good (drawn on insufficient funds or a nonexistent account, forged, and so on). In other words, in theory at least, the finality of a check payment made in 2002 is not terribly different than that of a bill-of-exchange payment made in 1602.

The implied risk allocations can be quite different in practice, however. For example, the check collection system no longer relies on examining indorsements and knowing the indorsers as a means of ensuring that a check is good. Anyone can indorse a check, and checks, unlike some earlier types of negotiable instruments, are payable on demand, or without delay. For many types of checks, banks may try to insulate themselves against fraud risk by delaying availability of a check deposit, but there are legal limits to this practice. The Expedited Funds Availability Act of 1987, which is implemented by Federal Reserve Regulation CC, requires banks to release deposited funds according to specific deadlines.¹⁵ It is sometimes difficult for banks to verify whether checks will be honored within the time of the required deadlines, and when checks are dishonored in such cases it is often the payee's bank (the depository bank) that bears the loss. These losses can be substantial: a 1995 survey of banks estimated banks' aggregate losses from check fraud to be more than \$600 million per year (see Board of Governors 1996).¹⁶ And in cases in which banks successfully avoid paying fraudulent checks, the loss is often borne by the party that took the

12. For the U.S. economy as a whole, this ratio (banks' reserves at the Fed to total daily noncash payments) is on the order of 0.5 percent.

13. For example, the Bank for International Settlements (2001) reports the following statistics on payments usage in the United States for the year 1999: 68 billion check payments, 26 billion credit or debit card payments, and 6 billion automated clearinghouse transfers.

14. Check payments within the United States fall under Federal Reserve Regulations J and CC and under a body of law known as the Uniform Commercial Code (UCC). Each of the states has incorporated the UCC into its own laws although there are some minor inconsistencies across states. A number of legal texts provide useful introductions to check payment law (see, for example, Mann 1999). For a discussion of issues regarding the finality of check payments, see Mengle (1990) or the Task Force on Stored-Value Cards (1997).

15. Regulation CC provides for some exceptions to these deadlines. For example, longer holds are allowed on funds deposited in new accounts.

16. A 1999 survey (quoted in Middlemiss 2001) puts banks' annual losses at \$679 million. The average loss on a bad check was \$1,518.

check as payment. Aggregate losses (to all parties involved) from check fraud are commonly said to exceed \$10 billion annually.¹⁷

Another way in which the modern-day check payment system differs from its historical antecedents is in the value of the check writer's ("drawer's" or "maker's") signature. When negotiable instruments were first developed, the signatures of the relevant parties were crucial in verifying the authenticity of the instrument involved. Only principals of merchant firms could write, accept, or indorse orders to pay; instruments were written out in longhand; multiple copies of each instrument were sent by different

There is nothing resembling the concept of negotiability for credit card payments. In other words, while just about anyone can pay anyone by check, credit card payments must flow through specific channels.

means; and many handwriting samples were kept (Kohn 1999). In modern circumstances, the sheer volume of checks makes it impractical for banks to employ these kinds of verification procedures. Nonetheless, a check's validity generally rests on the validity of the signatures on it. For example, if a bank pays a check on which the check writer's signature is forged, the paying bank usually ends up restoring the funds of the forged check to the depositor on whose account the check was drawn and bearing the loss from the forgery. Unfortunately, the depositor is not always completely protected by this right: until the forgery is discovered and the funds restored, the depositor does not enjoy access to the full amount of funds in his account and may be inadvertently bouncing checks.

Despite these and many other changes over the years, the check payment system has by and large adhered to the principle of assigning risk to the least-cost avoider. When a bank mistakenly pays a forged check and suffers a loss as a result, under the least-cost avoider principle this outcome is seen as appropriate because the payor bank was in a better position than any other party to detect the fraud. But these risks are not entirely borne by banks. A merchant who provides some good or service in return for a fraudulent check is also at risk (for the value of the good or service provided), as is a depositor who is grossly negligent in the management of

his bank account (for example, who never checks his bank records for unauthorized withdrawals).

In many instances this risk allocation rule has had the desirable effect of encouraging people to undertake measures to mitigate fraud and other types of risk. For example, businesses that receive large numbers of check payments are increasingly making use of check verification services. Banks are also using new technologies to limit fraud—for example, pattern recognition programs that look for unusual account activity. Many businesses employ measures designed to verify check disbursements on their accounts, such as "positive pay" and "reverse positive pay."¹⁸

Nonetheless the rather stark allocation of risk implied by the least-cost-avoider principle may not be appropriate in every payment situation. Consumers and small businesses, in particular, can end up bearing significant losses because of fraud or simply nonpayment of checks. In such cases, assignment of risk based on a least-cost-avoider principle is likely to be less effective as the parties bearing the risk may lack the expertise or wherewithal to effectively limit that risk. And since the losses involved may be large relative to the wealth of the individuals involved, many people will find it advantageous to either insure against the risks of loss or simply forgo the opportunity to transact.¹⁹ One reason for the popularity of alternative means of payment such as credit cards is that they provide some degree of insurance against various types of payment risk, as the discussion below will show.

Payment by Credit Card

After the check, the second most popular form of noncash payment in the United States is the credit card. Credit cards, like checks, are a means of transferring bank claims from payor to payee. A key difference between the two is that a payment by credit card does not directly draw funds from the cardholder's account but instead draws on a line of credit established in advance by the card issuer. In the context of the Chart 2 example, if the consumer C paid the merchant B by credit card, then C would not have to go through the formality of applying for a loan to make the payment but could instead simply offer the credit card to B.

Credit card payments are supported by a physical system for clearing payments, by the private contracts between various participants in the credit card networks (including cardholders, merchants, and card issuers as well as other parties), and by a body of law that governs (among other things) the finality of such payments.²⁰ However, these con-

structs differ substantially from their counterparts that support check payments. As a consequence, the implied allocation of risks for credit cards is also quite different from that for checks.

Perhaps the most basic difference is that there is nothing resembling the concept of negotiability (in the sense of general acceptability in exchange) for credit card payments. In other words, while just about anyone can pay anyone by check, credit card payments must flow through specific channels. To receive payments by credit card, one must have a merchant account with a credit card company. And before they can receive credit card payments, holders of merchant accounts generally have to follow certain fraud-abatement procedures, such as obtaining electronic authorization for the payment or obtaining the payor's signature in transactions when the payor is present. Procedures for the clearing of payments (which in the case of checks are largely left up to the bank where the check is deposited) proceed in a fairly uniform way according to the rules of the credit card company.²¹

Credit card payments are in this sense more convenient than check payments. Credit cards are also considerably more expensive to the payee since the credit card company deducts a merchant fee on all transactions made with its cards.²² Despite these disadvantages, there are many situations in which both buyer and seller will prefer payment by credit card over payment by check. For the seller, the greatest advantage of a credit card payment is its higher degree of finality: as the name credit card implies, the credit risk of the transaction (that is, the risk that the cardholder cannot pay) is absorbed by the card issuer. And while credit card companies do not unconditionally guarantee payment in cases of fraud, they are willing to absorb a significant percentage of losses from unauthorized transactions.²³ In short, payment by credit card offers sellers a greater degree of insurance against various types of risks than does payment by check.

Payment by credit card also offers advantages for buyers. Credit cards are useful simply because sellers are willing to accept them in circumstances in which checks would be unacceptable. Credit cards have other advantages as well. For example, in some instances cardholders have the right to withhold payment of disputed charges, and their liability in the case of a lost or stolen card is capped at \$50.

The rules for credit card payments thus offer a high degree of assurance to both sides of a given transaction. A seller who receives a payment by credit card enjoys a higher degree of payment finality relative to a check payment while the buyer enjoys widespread acceptance of her credit as well as additional assurance against theft and fraud. Providing this insurance to both buyer and seller is costly, but the costs are to a large extent shared by all customers of a given credit card issuer so that in most cases no individual buyer or seller faces an inordinate risk of loss.

By insulating people from the consequences of various types of payment risk, however, such risk spreading generates an undesirable side effect. As compared to a system based on the least-cost-avoider principle, a system that emphasizes risk spreading lessens its participants' incentives to limit risk and thus can significantly increase costs. Credit card companies have sought to contain these risks, particularly fraud risk, by strictly enforcing rules about how transactions can occur (for example, rules for merchant account holders concerning signatures and authorizations) and by aggressive use of antifraud measures such as pattern recognition software. Through such efforts the rate of fraud on credit card transactions has fallen sharply in recent years, from about 0.18 percent of all transactions in 1995 to 0.06 percent in 1999 (see Fickenscher 2000). Even so, there remains a trade-off of insurance against cost and convenience. In situations characterized by a relatively high degree of trust (for example, payment of utility bills or mortgages),

17. See, for example, the National Check Fraud Center at www.ckfraud.org.

18. Under positive pay, a bank will pay only checks whose amounts and other identifying information match the information previously provided by a depositor. Under reverse positive pay, the depositor himself performs the verification.

19. For example, Prescott and Tatar (1999) note that the risk of an overdrawn bank account is one reason that many low-income households choose not to have an account.

20. Principal among these laws is the Truth in Lending Act of 1968, which is implemented by Federal Reserve Regulation Z.

21. For a description of the processes for clearing checks and credit card payments, see U.S. General Accounting Office (1997).

22. A 1998 survey by the Food Marketing Institute, cited in Humphrey, Pulley, and Vesala (2000), puts the merchant's average cost of a credit card transaction in a grocery store at \$1.07 whereas the average cost of a verified check transaction is about \$0.45.

23. A recent study by the U.S. General Accounting Office (1997, 114) estimates the average share of credit card losses borne by credit card issuers to be 70 percent, with the remainder borne by merchants. The finality of a credit card transaction is considerably reduced when the cardholder is not physically present at the transaction and cannot sign a charge slip, as is the case with mail-order, telephone, or Internet transactions. See, for example, Winn (1999) for a discussion.

checks prevail whereas credit cards have come to dominate in many situations in which goods or services are exchanged in fairly anonymous situations with little recourse to the seller in case of fraud (for example, paying for restaurant meals).

Debit Cards and Check Conversions

Debit cards are the most rapidly growing form of noncash payment in the United States. From 1995 to 1999, transactions using debit cards grew from about 1.6 billion transactions (1.9 percent of all noncash transactions) to 7.5 billion transactions (about 7.5 percent) (Bank for International

Debit cards offer sellers many of the same advantages as credit cards do—principally, assurances of payment finality.

Settlements 2001). The finality rules for payments by debit card are interesting because in many respects they are intermediate between those for checks and credit cards.²⁴

While debit cards and credit cards are similar in appearance, their functions are quite different.²⁵ Rather than drawing on a pre-established line of credit, as is done in a credit transaction, a debit card draws on funds already in the cardholder's bank account. There are two types of debit cards, on-line (or PIN-based) and off-line (or "signature-based").²⁶ In an on-line debit card transaction, the cardholder authorizes the transaction by entering a personal identification number, or PIN, at the point of sale. In an off-line transaction, the cardholder authorizes the transaction by a signature, just as for a credit card transaction. Historically, the two types of debit card transactions have also been cleared in different ways, with on-line transactions being cleared through regional automatic teller machine (ATM) networks while off-line transactions are cleared through the same channels as credit card payments.

Debit cards offer sellers many of the same advantages as credit cards do—principally, assurances of payment finality. As discussed above, this finality is useful to buyers as well because it allows them to conduct transactions in situations in which check payments would be unacceptable to the seller. Buyers (consumers) are also afforded a

high degree of assurance against unauthorized transactions using their account. However, the degree of consumer protection guaranteed by law for debit cards is less than for credit cards. For example, in some cases, the maximum consumer liability for unauthorized debit card transactions can exceed the \$50 limit for credit card transactions although some debit card issuers have voluntarily chosen to enforce this limit for debit transactions. Also, consumers cannot withhold payment solely because of dissatisfaction with a purchase. And even in cases in which a bank restores a depositor's funds after an unauthorized transaction, the depositor faces the risk of a depleted bank account until the funds are restored.

By definition, debit cards do not involve an automatic extension of credit to the cardholder but instead draw directly on funds already available in the cardholder's bank account.²⁷ Since the debit card issuer absorbs little or no credit risk, debit card-based transactions should, at least in theory, be less costly than credit card-based transactions. This theoretical cost differential is reflected to different degrees in the marketplace. The merchant costs of on-line debit card transactions are well below those of other noncash forms of payment while costs of off-line transactions are more comparable to those for credit card transactions.²⁸

A related type of transaction is known as a point-of-sale (POS) electronic check conversion. In this type of transaction, a consumer writes a check at a retail location or similar point of sale. The check is then run through a terminal that reads the information necessary to convert the check into an electronic debit to the consumer's bank account and verifies the information against a database for fraud. After this "conversion" (sometimes called truncation), the check is returned to the consumer, and the transaction is then cleared through the automated clearinghouse network (the same network used for direct deposits of payrolls). Because each transaction is verified and can be settled the following day, payment by POS check conversion potentially offers merchants many of the same finality advantages as those for debit card payments. This type of transaction is relatively uncommon today, but several major retailers have announced plans to increase their use of POS check conversion technology, and many expect this payment method to find widespread usage within a few years.

Because a POS transaction originates as a check but is cleared as an electronic transaction, there has been some ambiguity about whether POS check

conversions should be treated legally as a check or as an electronic payment. Recently, the Fed issued an official interpretation (Board of Governors 2001) that such transactions do qualify as electronic funds transfers and hence are governed by the same legal framework as debit card transactions. Thus, a buyer whose check is converted at the point of sale enjoys essentially the same legal protections (under Regulation E) as would a buyer who made the same purchase with a debit card.²⁹

In short, payment by either debit card or POS check conversion offers a higher degree of payment finality than a check payment. Buyers who use these means of payment are not guaranteed the same degree of insurance against certain risks as those who pay by credit card. Nevertheless, these types of payments may be more efficient in certain situations, because their cost per transaction can be substantially lower than that of credit cards.

Summary and Conclusion

Payment finality is critical to the process of decentralized exchange. By specifying the circumstances in which the transfer of one type of claim extinguishes another, finality rules fulfill two important roles. The first is to minimize opportunities for default along the credit chains that arise in any developed economy. The second is to allocate other risks, such as fraud, that naturally arise in the course of exchange.

While all payment systems incorporate some notion of finality, different systems have different finality rules that imply different risk allocations among system participants. The finality of check payments is in general quite tentative, and the risks in the payment process are often concentrated on a single party. Credit and debit card payments, in contrast, are generally more final than check pay-

ments, and the liability for potential losses in the payment process tends to be shared among participants in these systems.

It should be emphasized that one set of finality rules is not necessarily better than another in all situations. The law governing check payments, for example, recognizes that such payments are by their very nature quite dispersed and cannot always be subjected to outside verification. This advantage of check payment makes it easy to pay virtually anyone by check but at the same time makes it very difficult to control risks. The solution to this problem, developed over centuries of experience, has been to encourage people to limit these risks by concentrating risk on the party or parties best able to avoid it.

Card payments work on a different principle. Because the laws governing such payments effectively insulate cardholders from many types of payment risk, issuers of credit and debit cards have no choice but to sharply restrict the circumstances under which such payments can be accepted and to intensively monitor payments for fraud. But, by imposing such limitations, card issuers are able to guarantee a higher degree of payment finality than for check payments.

Continuing progress in communications technology will offer the potential for improvements in these trade-offs. New technologies in on-line payments, for example, may allow for convenient, person-to-person payments between consumers with a higher degree of finality than is offered by check payments. And smart-card technology may enable card issuers to reduce the fraud risk associated with card payment. No matter what form these new technologies may take, however, the essential service they offer will be the same—the provision of payment finality.

24. Rules for debit card payments are governed in large part by the Electronic Funds Transfer Act of 1978 and Federal Reserve Regulation E.

25. In many cases the same physical card can be used as either a debit card or a credit card. For purposes of the discussion here, such cards will be treated as two distinct cards.

26. Again, a single debit card may function as either an on-line or off-line card.

27. Both on-line and off-line debit cards require the use of verification systems to insure that a cardholder can make a payment before the payment is authorized.

28. The survey quoted in Humphrey, Pulley, and Vesala (2000) puts the average merchant cost of an on-line debit transaction at \$0.29 versus \$0.80 for an off-line transaction. As of October 2001, the cost of a \$40 on-line transaction cleared over Visa's Interlink network was \$0.20 versus \$0.60 for an off-line transaction (ATM & Debit News 2001).

29. The Fed's interpretation (Board of Governors 2001) also allows for conversion of checks at payment processing or "lockbox" locations. Such conversions are sometimes subject to additional requirements beyond those for POS conversions. A plain-English summary of the rules for check conversion is available from the Fed (Board of Governors 2002).

REFERENCES

- Andrews, Edmund L. 1997. Want to make big rubles? Invest in a comrade's debt. *New York Times*, October 14.
- ATM & Debit News. 2001. Visa delays Interlink rate hike: Retailers delay dropping card. October 11.
- Bank for International Settlements. 2001. *Statistics on payment systems in the Group of Ten countries*. Basel: BIS. <www.bis.org/publ/cps44.pdf> (April 9, 2002).
- Board of Governors of the Federal Reserve System. 1996. *Report to the Congress on funds availability schedules and check fraud at depository institutions*. <www.federalreserve.gov/boarddocs/rptcongress/chkfraud.pdf> (April 9, 2002).
- . 2001. *Revisions to staff commentary on Regulation E*. <www.federalreserve.gov/boarddocs/press/boardacts/2001/20010313/default.htm> (April 9, 2002).
- . 2002. *When is your check not a check? Electronic check conversion*. <www.federalreserve.gov/pubs/checkconv/default.htm> (April 9, 2002).
- Fickenscher, Lisa. 2000. Visa shores up Web position, ends fees on theft of card numbers. *American Banker*, February 22.
- Humphrey, David, Lawrence Pulley, and Jukka Vesala. 2000. The check's in the mail: Why the United States lags in the adoption of cost-saving electronic payments. *Journal of Financial Services Research* 17, no.1:17–39.
- Ickes, Barry W. 1998. Notes on vekselns. Pennsylvania State University. Unpublished manuscript. <econ.la.psu.edu/~bickes/ickres.htm> (April 9, 2002).
- Kahn, Charles M., and William Roberds. 2001. Transferability, finality, and debt settlement. Federal Reserve Bank of Atlanta Working Paper 2001-18a, October. <www.frbatlanta.org> (April 15, 2002).
- Kohn, Meir. 1999. Bills of exchange and the money market to 1600. Dartmouth College Department of Economics Working Paper 99-04. <www.dartmouth.edu/~mkohn/99-04.pdf> (April 9, 2002).
- Mann, Ronald J. 1999. *Payment systems and other financial transactions*. New York: Aspen Law and Business.
- Mengle, David L. 1990. Legal and regulatory reform in electronic payments: An evaluation of payment finality rules. In *The U.S. payment system—efficiency, risk and the role of the Federal Reserve: Proceedings of a symposium on the U.S. payment system sponsored by the Federal Reserve Bank of Richmond*, edited by David B. Humphrey, 145–80. Boston: Kluwer Academic Publishers.
- Middlemiss, Jim. 2001. Banks jumping aboard ECP wagon, proponents say. *Bank Systems & Technology*, October 1.
- Prescott, Edward S., and Daniel D. Tatar. 1999. Means of payment, the unbanked, and EFT '99. Federal Reserve Bank of Richmond *Economic Quarterly* 85 (Fall): 49–70. <www.rich.frb.org/pubs/eq/pdfs/fall1999/prescott.pdf> (April 9, 2002).
- Rogers, James Steven. 1995. *The early history of bills and notes: A study of the origins of Anglo-American commercial law*. Cambridge: Cambridge University Press.
- Task Force on Stored-Value Cards. 1997. A commercial lawyer's take on the electronic purse: An analysis of commercial law issues associated with stored-value cards and electronic money. *The Business Lawyer* 52 (February): 653–727.
- U.S. General Accounting Office. 1997. *Payments, clearance, and settlement: A guide to the systems, risk, and issues*. Washington, D.C.: GAO. <www.gao.gov> (April 9, 2002).
- Winn, Jane Kaufman. 1998. Couriers without luggage: Negotiable instruments and digital signatures. *South Carolina Law Review* 49. <faculty.smu.edu/jwinn/> (April 9, 2002).
- . 1999. Clash of the titans: Regulating the competition between established and emerging electronic payment systems. *Berkeley Technology Law Journal* 14. <faculty.smu.edu/jwinn/> (April 9, 2002).