Measuring the Costs of Retail Payment Methods

By Fumiko Hayashi and William R. Keeton

Information on the costs of retail payment methods is more critical than ever to central banks concerned with the efficiency of their nations' payments systems. The last two decades have witnessed a dramatic shift in retail payments—all those other than large-dollar payments by wire—from paper-based methods to electronic-based payment methods. However, paper-based payment methods are still widely used in many countries, including the United States. Accurate and up-to-date information about the relative costs of paper-based and electronic payment methods would help central banks decide how hard to push for complete transition to electronic payment methods. Better cost information would also help central banks decide which electronic methods to promote, whether in the banks' roles as providers of payments services or their roles as regulators or catalysts for change.

To obtain such information, a number of central banks have recently conducted comprehensive studies of the costs of retail payment methods. These studies have reached some common conclusions, such as that debit cards are less costly than credit cards. However, the studies have reached different conclusions about the relative costs of other

Fumiko Hayashi is a senior economist and William R. Keeton was an assistant vice president and economist at the Federal Reserve Bank of Kansas City. *Research associates Elizabeth Antonious, Jacob Schak, and Thad Sieracki helped prepare the article. This article is on the bank's website at* ***www.KansasCityFed.org.***

payment methods, suggesting that cost rankings can depend on the specific characteristics of a country's payments system and the scale at which a payment method is used in the country. Estimates of the aggregate costs of making retail payments have also varied across the studies, from 0.5 percent to 0.9 percent of GDP.

These differences suggest a need for each central bank to conduct its own cost study. The danger of relying on other countries' cost studies is particularly apparent for the United States, where checks and credit cards are used on a larger scale and more parties are involved in the payments process. Nevertheless, the cost studies of other central banks can serve as useful models for a central bank undertaking its own study. For example, the earlier studies can provide ideas on whether to distinguish fixed and variable costs, whether to scale costs by the number or value of transactions, and how to allocate shared costs among payment methods.

This article reviews the methodology and results of previous cost studies with an eye to helping central banks decide whether and how to conduct their own cost studies. The first section discusses the benefits and limitations of cost studies to central banks in meeting their payments-related policy goals. The second section explains the key cost concepts involved in comparing the costs to society of different payment instruments and the key decisions that must be made in gathering cost data. The third section reviews four recent cost studies by central banks, comparing their key features and findings. The fourth section discusses the lessons to be learned from these studies by a central bank contemplating its own study of retail payment costs.

I. BENEFITS AND LIMITATIONS OF COST STUDIES TO CENTRAL BANKS

Information on the costs of different retail payment methods can be very useful to a central bank in assessing the efficiency of the nation's payments system. However, information on costs is not all that is needed for such an assessment. Information on the benefits of different payment methods is also important. This section first describes how central banks can use cost information to help meet their policy goals and then explains why cost information alone is not sufficient to meet those goals.

How cost information can help central banks meet their policy goals

In most developed countries, the efficiency and safety of retail payments are viewed as appropriate policy objectives of the central bank (Bank for International Settlements 2003). Retail payments are all payments other than large-dollar payments, which consist of wire transfers among businesses and financial institutions that can run into hundreds of millions of dollars. Retail payments used to receive less attention from central bankers in developed countries than large-dollar payments because disruptions to the latter posed more risk to financial stability. However, a safe and efficient retail payments system is now viewed by most central bankers as a key requirement of a smoothly functioning economy.

For a nation's retail payment system to be efficient, the payment methods used by consumers, businesses, and government entities should be those that impose the lowest cost on society for a given level of benefits. To determine if this condition is being met, a central bank needs to know how much of society's resources are absorbed by each retail payment method in current use. Suppose the central bank determined that two payment methods could be used for the same type of payment and yielded the same total benefit to end users, but one method used less of society's resources. The central bank could then legitimately conclude that the existing payment system was inefficient—the same benefits could be achieved at lower cost to society by substituting the lower-cost payment method for the higher-cost method.

While the policy goal of a safe and efficient payments system justifies central bank interest in retail payment costs, experts differ as to how central banks should act on such information. Central banks are traditionally viewed as playing three possible roles in the retail payments system: as operator, overseer, or catalyst (Bank for International Settlements 2003; Weiner). How a central bank uses information on retail payments costs to promote a more efficient payments system depends on which of these three roles it plays.

Operator role. Central banks serve as operators by providing retail payments services. In some countries, these services are confined to settlement, which is the transfer of reserves among banks to discharge the obligation of the payer to the payee. In other countries, the services include clearing, which refers to the exchange of payment information

among banks and the calculation of net amounts owed or due in preparation for settlement.

In its role as operator, a central bank could use information about the costs of retail payments to improve the efficiency of its own payment services. For example, in the 1990s, the Federal Reserve adjusted the prices of its automated clearinghouse (ACH) services and took steps to make those services more attractive to businesses. A major goal of this effort was to encourage businesses to shift from paper checks to ACH payments, which were viewed as less costly to society because they were direct electronic transfers between bank accounts (Connolly). More recently, the Federal Reserve has offered discounts to banks that agree to accept electronic presentment of checks, in an effort to reduce the cost to society of payments that continue to be made by check rather than ACH (Bauer and Gerdes).

Overseer role. In the role of overseer, central banks regulate retail payment services provided by the private sector. Developed countries vary significantly in the authority granted to the central bank for such oversight. In Australia, the central bank has explicit authority to enact regulation to ensure efficiency and safety of retail payments. In the United States, by contrast, the Federal Reserve's regulatory authority tends to be limited to implementation of specific laws. Examples include the Electronic Funds Transfer Act, which authorizes the Fed to enact regulations on the rights and responsibilities of consumers using electronic payments; the Expedited Funds Availability Act, which allows the Fed to regulate certain aspects of check collection by private-sector entities; and the Durbin Amendment to the Dodd Frank Act, which requires the Federal Reserve to regulate fees paid by merchants to banks on debit card transactions.

In those countries in which the central bank has broad regulatory authority over retail payments, cost information can be used to improve the efficiency of payments services provided by the private sector. In the early 2000s, for example, the Reserve Bank of Australia (RBA) used its regulatory authority to cap fees paid by merchants to banks on credit card transactions. The RBA argued that these fees encouraged banks to promote the use of credit cards at the expense of alternative payment methods such as PIN debit cards and ACH payments. According to the RBA, slowing growth in credit card use by capping the fees would benefit society because the costs of credit card payments exceeded the costs of the alternative payments methods (Lowe).

Catalyst role. Central banks in almost all developed countries act as catalysts or facilitators by working with the private sector and government to promote a more efficient retail payments system. Many central banks have research staffs that study retail payments instruments and markets. In addition, central banks often have cooperative relationships with bank regulatory agencies and competition authorities. Finally, central banks tend to have close contacts with banks and other financial institutions as a result of their supervisory and monetary policy responsibilities. All these factors make central banks well suited to lead multiparty efforts to develop payments standards, improve payments infrastructure, or pass payments legislation.

As catalysts, central banks can use cost information to identify efficiency-improving changes in retail payments services and work with other parties to bring about those changes. In the United States, analysis of internal data on check-processing costs convinced the Federal Reserve that the costs to society of check payments could be greatly reduced by converting paper checks to electronic images after deposit. However, a large-scale shift toward check electronification could not occur unless all banks in the collection process agreed to accept images. To overcome this obstacle, the Federal Reserve led the effort for enactment of the Check Clearing for the 21st Century Act (Check 21). This law, passed in 2003, assured banks they could process checks electronically as long as they were prepared to provide a new paper instrument called a "substitute check" to any bank that refused to accept electronic check images.

Why cost information may not be sufficient to meet policy goals

To determine whether one retail payment method is more efficient than another, a central bank would need to know not only the costs to society of using each payment method, but also the benefits. In some comparisons, such as a check cleared in paper form versus an electronic image, a central bank may have good reason to believe the benefits of the two payment methods are comparable. In such cases, efficiency can be judged on the basis of cost alone. But in other comparisons, the benefits of the payment methods to end users may differ significantly. In these cases, the relative efficiency of two payment methods cannot be determined without knowing both their benefits and costs.

One reason a retail payment method could yield higher benefits to end users than another is that the method is more widely accepted. Widespread acceptance is one of the most important attributes of a payment method. Consumers derive little benefit from a payment method that few merchants accept, and merchants find little value in accepting a payment method that few consumers use. A payment method that uses up few resources but lacks widespread acceptance is not necessarily more efficient than a higher-cost method that is available to all consumers and accepted by all merchants.

Another factor that could cause one payment method to yield higher benefits than another is convenience. Paying with a check can be less convenient than paying with a debit or credit card because more time is needed at checkout to complete the transaction. In other cases, a payment method may be more burdensome to set up and learn to use than another method. As before, a payment method that absorbs few resources but is slow or difficult to learn need not be more efficient than one that has higher cost but also greater convenience.

Finally, some new payment methods may have benefits that were available to only a limited degree with traditional payment methods. For example, mobile payment methods may allow consumers to achieve much greater control over their finances and spending by allowing them to check their account balances prior to making a payment. Mobile payment methods may also allow merchants to tailor ads and promotions to the needs of individual consumers. To the extent that mobile payments yield higher benefits to end users than traditional debit and credit cards, comparing the payment methods in terms of cost alone would be misleading.

II. CONDUCTING A COST STUDY

Policymakers need to consider various cost concepts when comparing the costs to society of different payment instruments. This section explains these concepts and discusses the key decisions that the managers of a cost study must make in gathering the cost data.

Key cost concepts

Policymakers need to understand four key cost concepts to compare the costs of payment instruments. Those concepts are social costs versus private costs, total costs versus variable social costs, costs per

transaction versus costs per unit of payment value, and constant versus increasing returns to scale.

Social costs versus private costs. In evaluating the relative efficiency of different payment methods, what matters is the social cost of each payment instrument. The *social cost* of a payment instrument is the sum of the resource costs incurred by all parties in transactions using that instrument. This cost is distinguished from the *private cost* of a payment instrument to an individual party. The latter cost includes not only the resource costs incurred by an individual party to a payment transaction, but also the fees paid by that party to other parties as part of the transaction. These fees are excluded from social costs because from society's point of view, the fees paid by one party to a transaction are offset by the fees received by another party.[1]

To see how social costs are calculated, consider the example of a cash purchase at a retail store. The first step in computing the social cost is to identify the parties involved in the transaction. At the point of sale, the two parties involved are the consumer and the merchant, who interact to initiate the transaction. In most cases, however, the consumer will have obtained the cash beforehand by visiting an ATM or a branch of a depository institution (a commercial bank, thrift, or credit union). At the end of the day, the merchant will also need to deposit cash receipts in its depository institution. Finally, in most countries, the central bank is responsible for distributing cash to depository institutions and storing any excess not needed by the banks to meet customers' demands. Thus, for this type of transaction, four parties play a role—consumers, merchants, depository institutions, and the central bank. Transactions that use different payment instruments may involve different sets of parties.

The social costs of a transaction equals the sum of the resource costs incurred by the different parties, whether for labor, capital, or raw materials. In the cash transaction above, the most obvious resource costs are those of the consumer and merchant. The consumer must spend time at an ATM or bank branch to acquire the cash and time at the checkout counter to make the purchase. Although the consumer may not incur a monetary cost, the time spent carrying out the transaction is still treated as a labor resource cost, because in theory, the time could have been spent in other productive ways. The merchant making

the sale must pay the wages and benefits of the cashier at the checkout counter and of the back-office employees who prepare the cash for deposit. The merchant must also pay the rental or depreciation costs of equipment used in the transaction (cash registers and computer systems for record keeping) plus the costs of any raw materials used (electricity for operating the cash registers and paper used for printing receipts).

To conduct a cash transaction, additional resource costs must be incurred by the consumer's and merchant's depository institutions and by the central banks. The depository institutions that provide cash to consumers and receive cash from merchants must incur the cost of operating the ATMs and branches that provide these services. They must also incur resource costs to record cash withdrawals and deposits and calculate customers' resulting account balances. Finally, the central bank must incur resource costs to maintain an adequate supply of currency and coin and to distribute cash to depository institutions as needed.

Adding up the resource costs incurred by all parties to a payment transaction yields an appropriate measure of the cost of the transaction to society as a whole. In contrast, adding up the private costs of the transaction to the parties would overstate social cost, because it would double count resource costs that payment participants recover from other parties through fees.

Total versus variable social costs. Social costs can reflect either the variable cost of using the payment method or the total cost, including fixed costs in addition to variable costs. Fixed costs are those incurred regardless of the number or value of payments, while variable costs are those that increase with the number or value of payments. Fixed costs are generally associated with capital equipment, while variable costs are associated with labor and raw materials.

For questions involving the long-run efficiency of the payments system, the appropriate measure of social costs is total cost, including both fixed and variable costs. For example, electronic payment methods tend to require large investments in infrastructure, such as card-reading terminals in stores and computer equipment in depository institutions and card networks. However, once that investment is in place, the additional costs of processing transactions tend to be low. Thus, electronic payments may have low variable social costs but high total social costs because of their high fixed costs. Cash payments, on the other hand, require

relatively small investment in infrastructure but large inputs of labor and raw materials to print, process, and safeguard cash. As a result, these payments may have high variable social costs but low total social costs. Focusing on the variable social costs of the two methods rather than the total social costs could bias policy makers in favor of a purely electronic payment system, by causing them to ignore the large capital investment required to switch to such a system. The Appendix provides an example in which looking at variable costs could mislead policy makers into preferring a more capital-intensive payments technology, even if that technology was not more efficient over the long run.

While information on total social costs is needed for long-run policy questions, information on variable social costs may be useful for some short-term policy issues. For example, if changes in the payment infrastructure are likely to take a long time, policymakers may want to know if the existing infrastructure could be used more efficiently by shifting more payments from cash to electronic methods. To answer this type of question, variable social cost is the relevant cost measure.[2]

Cost per transaction versus cost per unit of value. The aggregate cost of using any payment instrument depends at least partly on the total volume of payment made with that instrument. Thus, to compare the cost efficiency of alternative payment instruments, the aggregate social cost of each instrument must be scaled by some measure of total payments volume. Two approaches are commonly used in payment cost studies—dividing aggregate costs by the total *number* of transactions made with the instrument, and dividing aggregate costs by the total *value* of transactions with the instrument.

Neither approach is perfect because the costs of using a payment instrument can depend on both the number and the value of transactions. Some costs depend mostly on the number of transactions rather than the value. Examples of such "transaction-related" costs include the cost to a bank of approving a purchase by one of its debit card holders and the cost to a cashier of processing a card payment at the checkout counter. The infrastructure costs of electronic payment methods also seem likely to depend on the number of transactions rather than the value—the same amount of computer capacity is likely to be required to process a million $10 card payments as a million $100 payments.[3] However, other costs of using a payment method may depend

not only on the number of transactions but also on their average value. For cash transactions, for example, the costs of counting bills and coins, transporting cash to and from banks, and providing protection against theft are likely to increase with the total amount of cash handled, and thus with the total value of cash transactions. Similarly, in the case of electronic payment methods, fraud losses and fraud prevention costs are likely to be greater for high-value payments than low-value payments.[4]

In the general case in which social costs depend on both the number and value of transactions, each of the two methods of scaling costs may give a distorted view of the relative costs of payment methods. As illustrated in the Appendix, focusing on cost per transaction will tend to favor payment methods with low average transaction sizes, such as cash. Conversely, focusing on cost per unit of value will generally favor payment methods with high average transactions sizes, such as credit cards and checks. As a result, a good case can be made that both ways of scaling social costs should be used in comparing payment methods—dividing by total transactions and dividing by total value. If one payment method has both lower cost per transaction and lower cost per unit of value than another, it can be safely assumed that the first method is more cost-efficient. If the two measures point in opposite directions, however, the possibility must be considered that one method is more cost-efficient for small transactions, while the other method is more cost-efficient for large transactions.[5] Thus, in designing a cost study, policymakers are well-advised to collect information on both the total number and the total value of transactions carried out with each payment instrument.

Constant versus increasing returns to scale. The last key cost concept is the role of scale. Comparing the social cost per transaction or social cost per unit of value of different payment methods can help assess whether one method is more cost-efficient than another. In making such comparisons, however, it is important to remember that the per-unit costs of a payment method may depend on the scale at which the payment method is used. Ranking payment methods by observed per-unit costs makes sense if each method is subject to constant returns to scale, in the sense that per-unit costs remain unchanged as volume rises and infrastructure is adjusted in the optimal way. Focusing on per-unit costs may be misleading, however, if some payment methods are subject to

increasing returns to scale. For example, a relatively new payment method may have high unit costs only because it requires a large minimum investment in infrastructure and is still being used on a small scale. As illustrated in the Appendix, such a payment method could appear to have a higher unit cost than traditional payment methods even if it could be operated at substantially lower cost than those methods if used on a large enough scale. Thus, while comparing the per-unit social costs of different payment methods can be informative, policymakers should not rely too much on such comparisons and should be alert for evidence that methods with high unit costs are being operated at suboptimal scale.

Key decisions in conducting a cost study

The above discussion suggests that a payments cost study concerned with efficiency should focus on social rather than private costs, cover fixed as well as variable costs, and collect data on the value as well as the number of payments. But other key decisions must also be made before conducting a cost study. These include which payment instruments and types of retail payment to cover, which payments participants to survey, and how to allocate shared costs among payment instruments.

Which payment uses and instruments should be included? Costs of making payments depend both on the type of payment being made and the particular instrument used for making the payment. Table 1 lists the major types of retail payment, sometimes known as "use cases," while Table 2 lists the major payment instruments.

For use cases, the main distinction is between payments made by consumers in stores in return for goods and services, known as "point-of-sale" (POS) transactions, and payments made to and from consumers outside of stores. The latter include telephone and online purchases by consumers, bill payments by consumers, person-to-person (P2P) payments between individuals, and payments to individuals from businesses and governments, such as wages and benefits. The final use case is business-to-business payments other than large dollar payments, indicated in the last row.

Among payments instruments, three main categories can be distinguished. The oldest payment methods are paper-based. In the United States, these methods include cash and the traditional check, which

Table 1
MAJOR TYPES OF RETAIL PAYMENTS (USE CASES)

Payments by consumers for in-store purchases (point-of-sale)
All other payments to or from consumers (non-point-of-sale) • Purchases by consumers outside of stores (mail, telephone, or Internet) • Bill payments by consumers (one-time and recurring) • Person-to-person payments (P2P) • Payments to consumers from businesses or government agencies, such as wages and benefits
Payments from businesses to businesses (other than large-dollar)

remains in paper form throughout processing. Another paper-based payment method that has been used in some European countries is the paper giro. When payment is by paper check, the payer gives the payee a piece of paper authorizing the payer's bank to transfer funds to the payee. When payment is by paper giro, the payer gives his own bank a piece of paper instructing it to transfer funds to the payee's bank account.

The next category of payment instruments are purely electronic instruments, in which the payment order neither starts nor ends in paper form and is processed entirely by electronic means. This category includes both payments cards (prepaid, debit, credit, and charge cards) and direct transfers between bank accounts. In a direct debit, the payer electronically authorizes the payee to have the payee's bank pull funds from (i.e., debit) the payer's bank account. In a direct credit, the payer electronically authorizes the payer's bank to transfer funds to (i.e., credit) the payee's bank account. In the United States, ACH debit transactions are an example of direct debits, while ACH credits and wire transfers are examples of direct credits.

The final category consists of hybrid instruments that start in paper form but are processed largely by electronic means. In the United States, for example, enactment of Check 21 made it possible for paper checks to be converted to electronic images by the bank of deposit and presented in that form to the paying bank. Merchants can also

Table 2

MAJOR TYPES OF CONSUMER PAYMENT METHODS

Type of payment instrument	Explanation	U.S. examples
Paper-based methods		
Cash	Currency and coin issued by central bank or government	
Traditional checks	Payer gives payee a paper authorization for payer's bank to transfer funds to payee. Check is presented by payee's bank to payer's bank for collection.	
Paper giros	Payer writes a paper order instructing payer's bank to transfer funds from payer's bank account to payee's bank account.	
Electronic methods		
Payment cards	Payment is initiated with a card and processed by a card network.	
Prepaid card	Payment is from funds contributed by cardholder in advance and held at a bank for use only with the card.	Green Dot, NetSpend, Wal-Mart MoneyCard, H&R Block Emerald Card
Debit card	Payment is from funds held by cardholder in a transactions account at bank issuing the card.	
• PIN debit	Personal identification number (PIN) is used to authenticate the cardholder. Cardholder's bank account is debited at time of purchase in the U.S.	Interlink (Visa), STAR, Pulse, NYCE
• Signature debit	Signature is used to authenticate the cardholder. Cardholder's bank account is not debited until a day or more after purchase in the U.S.	Visa, MasterCard
Credit card	Payment is from funds advanced to cardholder by card issuer, and balance can be carried over to next billing cycle.	Visa, MasterCard, American Express, Discover
Charge card	Payment is from funds advanced by card issuer, and balance must be paid at end of current billing cycle.	American Express

Table 2 Continued

Direct transfers	Payment is initiated without a card and is processed electronically.	
Direct debit	Funds are transferred from payer's bank account to payee's bank account through an order authorized by payer and originated by payee's bank.	ACH debit authorized on telephone or merchant website
Direct credit	Funds are transferred from payer's bank account to payee's account of payee through an order authorized by payer and originated by payer's bank.	ACH credit initiated by payer on his bank's website
Electronic bill payment service	Bills are paid by direct debit or direct credit and are processed by a non-card payment network	CheckFree (ACH debits), EBIDS (ACH credits)
Person-to-person (P2P) payment through a nonbank intermediary	Funds are transferred from payer to intermediary by direct debit and from intermediary to payee by direct credit.	PayPal
Hybrid methods		
Paper checks converted to electronic form for collection	Payer gives payee a paper check, which payee or payee's bank converts to electronic form for processing	Check truncation in accord with Check 21, ACH check conversion

take paper checks received as payment and convert them to ACH debit transactions if consumers have been properly notified.

Ideally, a cost study would calculate the social cost of each payment instrument in each use case for which the instrument is being used. As a practical matter, however, the designers of a cost study must make difficult choices about which instruments and which use cases to cover. To ensure that a cost study is not prohibitively expensive and obtains accurate results, it may be necessary to exclude some important payment instruments. And for those instruments that are included, it may only be feasible to estimate the cost of the primary use case for that instrument (for example, POS transactions for debit cards), or the average cost of the instrument for a mix of use cases (for example, the average cost of using a debit card for POS, telephone, and online purchases). Another difficult issue is how to treat the hybrid instruments. For example, should truncated checks be lumped with paper checks in computing social costs, or should they be treated as a separate payment instrument? Similarly, should checks converted to ACH be included with other checks, included with other ACH transactions, or treated separately?

How should cost information be collected? It is generally not practical to collect data on resource costs for all major participants in retail payments transactions. The participants included most often are banks, merchants, consumers, card networks, the central bank, and private interbank clearing organizations.[6] But other parties may also play key roles in payments transactions, either as processors or intermediaries between payers and payees. If data on these parties' resource costs are unavailable, the only alternative may be to use the fees paid to them by banks, merchants, and consumers as a rough estimate of their costs.[7]

Another key question is whether to collect information on the resource costs of major participants through surveys or independent sources. The advantage of using surveys is that the questions can be tailored to the information needs of the cost study. However, because surveys can be expensive and time-consuming, some studies have used information from independent sources to estimate resource costs for key payments participants. For example, rather than ask consumers directly about their costs of using cash for purchases, some studies have estimated these costs using independent data on cash purchases and

plausible assumptions about the time required by consumers to withdraw or deposit cash and the opportunity cost of their time.[8]

How should shared costs be allocated among payment methods? A final issue—a common thorn in the side of cost accountants—is how to allocate shared costs across payment instruments. Some payments costs are associated with specific payment instruments. For example, the cost to central banks of distributing cash to banks can be viewed as a cost of using cash for payments, and the cost to Visa or MasterCard of processing a credit card transaction can be viewed as a cost of using a credit card for payments. But other costs are associated with multiple payment instruments. An example is the cost of a bank's customer services. The same customer service personnel who handle customer inquiries related to checks may handle customer inquiries related to debit cards and credit cards. Thus, allocating all the costs of customer service to checks would overstate the social costs of checks and understate the social costs of debit cards and credit cards. One solution would be to allocate the total cost of customer service personnel according to the share of customer inquiries related to each payment method. Another approach would be to divide the total cost of customer service personnel according to the share of transactions using each method.

In addition to allocating shared costs of payment instruments, a cost study must also allocate costs shared by payment and non-payment activities. For example, securely storing customer data supports both a bank's payment operations and its funding and lending operations. Some studies have allowed participants in the cost study to decide how to allocate such costs. However, to obtain consistent data it may be preferable to provide participants at least general guidance on how to allocate the costs to the payment operation and to each payment instrument. In the case of data storage, for example, participants could be instructed to divide costs between payment and non-payment activities according to the volume of stored data associated with each activity.

III. PREVIOUS COST STUDIES

To fulfill their oversight responsibility for the payments system, a number of central banks have conducted studies to estimate the social costs of retail payment methods. This section reviews studies conducted in four countries during the last decade—Australia, Belgium,

the Netherlands, and Norway.[9] The section compares the key features of the studies and summarizes what they say about the social costs of different payment methods.

Key features of the four studies

The four studies can be compared in three dimensions—how payment costs were measured, what types of payments were covered, and what groups of payment participants were included (Table 3).

How costs were measured. The main goal of each study was to estimate the social cost of different retail payment methods. The Australian and Norwegian studies went further, however, by also estimating the private costs to payment participants of providing various payment services.[10] While social costs are appropriate for assessing payments efficiency, private costs can be useful for other policy questions. For example, data on private costs can help determine why consumers prefer a particular payment method or whether consumers are paying higher fees for a payment service than it costs the provider to produce it.

Each study estimated the total social costs of each payment method, including both the variable resource costs associated with inputs of labor and raw materials and the fixed resource costs associated with payments infrastructure. As explained earlier, including the fixed cost of infrastructure is essential in determining whether one payment method would be more cost-efficient than another over the long run. However, all the studies except the Norwegian one also reported separate estimates for variable social costs, allowing policymakers to consider whether costs could be reduced in the short run by shifting payments from one method to another.

Finally, each study collected sufficient information to compute both the social cost per transaction and the social cost per unit of transaction value.[11] As noted in the previous section, looking at both measures can help control for the influence of transaction size on cost. Specifically, it can rule out the possibility that a method has low cost per transaction only because the method tends to be used for small-value transactions, or that a method has low cost per unit of value only because it tends to be used for large-value transactions.

What types of payments were covered. The studies covered somewhat different sets of payment methods. All four studies included cash, PIN

Table 3

KEY FEATURES OF RECENT PAYMENT COST STUDIES

	Australia	Belgium	Netherlands	Norway
Year of study	2006	2003	2002	2007
How were costs measured?				
Social versus private costs	Social and private costs	Social costs only	Social costs only	Social and private costs
Total versus variable costs	Total and variable costs	Total and variable costs	Total and variable costs	Total costs only
Scaling of aggregate social costs	By number and value of transactions	By number and value of transactions	By number and value of transactions	By number and value of transactions
What payments were covered?				
Payment methods	Cash, credit card, PIN debit card, direct credit, direct debit, check, electronic bill payment	Cash, credit card, PIN debit card, prepaid card	Cash, credit card, PIN debit card, prepaid card	Cash, credit card, PIN debit card, prepaid card, direct credit, direct debit, paper giros
Payment context	POS, non-POS	POS only	POS only	POS, non-POS
Which participants were included?				
Groups who were surveyed directly	Merchants, banks	Merchants, banks, card networks	Merchants, banks, clearing organization[1]	Merchants, banks, consumers
Groups whose resource costs were reported	Merchants, currency issuers, banks, consumers	Merchants, currency issuers, banks and card networks combined	Merchants, currency issuers, banks and clearing organization combined	Merchants, currency issuers, banks, subcontractors[2], consumers

[1] Interpay, which clears ACH payments and PIN debit transactions for member banks.

[2] Include BBS, a central clearing agency for payments; EDB, an e-billing company; and NOKAS and Loomis, cash handling companies.

debit cards, and credit cards. The Belgian and Dutch studies also included prepaid cards, known as "e-purse" in both cases. The Australian and Norwegian studies were the only studies to consider direct debit and direct credit, and the Australian study was the only one to consider checks. The Australian study also included BPAY, a popular bill payment service similar to direct credit.

As noted earlier, the cost of using a payment method can depend on the type of transaction, and specifically, whether it is a POS transaction. The Belgian and Dutch studies focused exclusively on the payment costs of POS transactions, while the Australian and Norwegian studies also considered the costs of non-POS payments. For payment methods used in both POS and non-POS transactions, such as checks, the Australian study estimated the costs separately. The Norwegian study did not try to estimate costs separately for POS and non-POS transactions. Instead, the study classified some payment methods as primarily POS (debit and credit cards) and other methods as primarily non-POS (direct debit and direct credit).

What groups of payments participants were included. Information on costs to participants was collected through surveys and independent sources. Each study surveyed banks and merchants. In addition, the Belgian study sent surveys to card networks, the Dutch study surveyed the central clearing organization for banks, and the Norwegian study surveyed consumers. The Australian study did not survey consumers but instead estimated the cost of their time to make payments. These estimates were based on information collected from merchants on the time required for in-store checkout, other studies on time spent travelling to ATMs to withdraw cash, and assumptions about the value of consumers' time. In all the studies, data on the costs of producing cash were obtained from the public currency-issuing entities.

The groups of participants whose resource costs were reported varied across the studies. Each study reported resource costs of merchants and public currency-issuing entities. However, the studies differed in how they treated banks and clearing organizations such as card networks. The Australian study reported resource costs for banks alone but included the fees paid by banks to "subcontractors," among which were clearing organizations. The Belgian and Dutch studies reported resource costs for banks and clearing organizations combined,

using the information collected from surveys of the two groups. The Norwegian study followed still another approach, reporting resource costs separately for banks and subcontractors but estimating subcontractor costs by subtracting an estimate of profits from fees. A final difference among the studies was the treatment of consumers. The Belgian and Dutch studies excluded consumer resource costs from social costs, while the Australian and Norwegian studies included them.

Main findings on social costs

Each study estimated the aggregate social cost of retail payments relative to GDP. The estimates were 0.49 percent in Norway, 0.65 percent in the Netherlands, 0.74 percent in Belgium, and 0.94 percent in Australia.[12] The estimates are not strictly comparable because the Belgian and Dutch studies included POS payments only, while the Australian and Norwegian studies added non-POS payments (Table 3). Nevertheless, the estimates provide a rough indication of the potential gains to these nations from increasing the cost efficiency of their retail payments systems.

Each study also estimated the social costs of individual payment methods. Such information is especially valuable because it can help policymakers determine if efficiency could be increased by changing the relative use of different payment methods. For long-run policy decisions, the relevant costs are the *total* social costs of different payment methods, including both fixed and variable costs. This subsection points out similarities and differences in the studies' finding on total social costs, focusing on the two main ways of measuring these costs—cost per transaction and cost per unit of transaction value. For short-run policy decisions on how to make the best use of the nation's existing infrastructure, variable short run costs are the relevant measure. The subsection explains how the findings of two studies on these costs can be used to estimate the short-run cost savings from changing the composition of payments.

In assessing the studies' findings, it is important to keep two factors in mind. First, differences in methodology and coverage could cause the estimated social cost of all payment methods in a country to appear higher in one country than another, even if payment methods were equally efficient in the two countries. Thus, this subsection looks

at what the studies say about the *relative* social costs of payment methods within each country (for example, the social costs of cash relative to debit cards), rather than the *absolute* social costs of those methods. Second, with increasing returns to scale, a payment method could have a high estimated cost in a particular country only because the method was being used on a small scale. To identify such cases, it is important to look not only at the relative cost of each method but also the volume of transactions conducted with the method.

Total social cost per transaction. Among electronic payment methods used at point of sale, total social cost per transaction was significantly higher for credit cards than for PIN debit cards in all four countries. As shown in Table 4, the social cost of a credit card transaction was about twice that of a PIN debit card transaction in Australia and more than four times in Belgium, the Netherlands, and Norway. In the two cases in which prepaid cards were included—the Belgian and Dutch studies—the cost of a credit card transaction was also significantly higher than the cost a prepaid card transaction.

The high cost of credit card payments relative to PIN debit card and prepaid card payments in the European countries could be due partly to low use of credit cards and economies of scale in card processing (Table 5). Credit cards accounted for only 1 percent or less of POS payments in Belgium and the Netherlands. Credit cards were used more often in Norway but still much less than PIN debit cards, which accounted for two-thirds of POS payments. To the extent credit cards were used below their optimal scale, their unit costs would tend to be higher than for debit cards. However, the Australian case suggests that the high relative cost of credit cards in the four countries may not have been entirely due to a failure to exploit economies of scale in credit-card processing. Unit costs in Australia were also noticeably higher for credit cards than PIN debit cards, even though credit card usage was only slightly less than debit card usage there.

The studies' findings on cash versus electronic payment methods were not as consistent but generally indicated that social cost per transaction was at least as low for cash as for PIN debit cards. Specifically, cash was estimated to be less costly than PIN debit cards in Australia and the Netherlands and about the same cost as PIN debit cards in Belgium. The glaring exception was Norway, where the cost per

Table 4

TOTAL SOCIAL COSTS PER TRANSACTION (U.S. DOLLARS)

Payment Method	Australia	Belgium	Netherlands	Norway
POS Payments				
Cash	.60	.70	.39	2.05
PIN Debit Card	1.15	.72	.64	.69
Credit Card[1]	2.02	3.44	4.71	3.49
Prepaid Card	—	.71	1.22	—
Check	8.11	—	—	—
Non-POS Payments				
Credit Card[1]	2.09	—	—	—
Direct Debit	.99	—	—	.44
Direct Credit	—	—	—	.78
Electronic Bill Payment	1.44	—	—	—
Check	5.63	—	—	—

[1] Excludes cost of credit collections and write-offs.
Note: Cost in local currency was converted to U.S. dollars using the exchange rate for the year of the study: .9974 for Australia, 1.3133 for Belgium and Netherlands, .1671 for Norway.
Sources: Schwartz and others (Tables 11 and 12); National Bank of Belgium (Table 1); Brits and Winder (Table 4.1); Gresvik and Haare (Table 7).

Table 5

TRANSACTON SHARES FOR MAJOR PAYMENT METHODS (PERCENT)

POS Payments

Payment Method	Australia	Belgium	Netherlands	Norway
Cash	79.1	81.3	85.5	24.0
PIN Debit Card	10.9	14.8	12.9	67.8
Credit Card	10.0	1.0	0.6	8.2
Prepaid Card	—	2.9	1.1	—
TOTAL	100.0	100.0	100.0	100.0

Note: For each method, the share is the number of transactions conduced with that method relative to the number of transactions conducted with all the methods listed. Shares may not add to 100.0 due to rounding.
Sources: Schwartz and others (Table 13); National Bank of Belgium (Table 2); Brits and Winder (Table 4.1); Gresvik and Haare (Table 7).

transaction was estimated to be almost three times as high for cash as PIN debit cards. One possible explanation for the high relative cost of paying with cash in Norway is the low scale at which cash is used in that country. As shown in Table 5, cash transactions represented only a quarter of POS payments in Norway, versus more than three-quarters of payments in Australia, Belgium, and the Netherlands. The production, distribution, and storage of cash may be subject to increasing returns to scale, in the sense that even a low volume of cash payments requires a substantial investment in equipment such as vaults, ATMs, and armored vehicles. If so, countries with low rates of cash usage should tend to have high social costs per cash transaction, as observed in Norway.[13]

Two other findings on social cost per transaction are of some interest, though less conclusive because they are based on only one or two studies. First, checks had by far the highest social cost per transaction in the only study in which they were included—the Australian study. In that country, the cost of paying by check was four times the cost of paying by credit card in POS transactions (upper half of Table 4), and two-and-a-half times in non-POS transactions (lower half of Table 4).[14] Second, in the two studies that considered non-card electronic payment methods, direct debit had a low cost per transaction relative to other payment methods—lowest of all methods in Norway, and second only to cash payments in Australia. In Norway, the cost of direct credit was also low although higher than that of direct debit. In Australia, electronic bill payment ranked closer to the middle among payment methods, reflecting the layering of services on top of direct credit to facilitate payment of bills.

Total social cost per unit of value. The studies' findings on total social cost per unit of value serve as a useful check on the findings on total social cost per transaction. Using cost per transaction as the metric, the studies found that debit cards were less costly than credit cards and (except for Norway) that cash was less costly or no more costly than debit cards. But as noted in Section II, the costs of both cash transactions and credit card transactions tend to increase with the size of the transaction—in the case of cash, because costs depend on the volume of currency handled, and in the case of credit cards, because fraud losses depend on the size of the payment. Furthermore, Table 6 shows that average transaction value was generally higher for credit cards than

Table 6

AVERAGE TRANSACTION SIZE FOR MAJOR PAYMENT METHODS (U.S. DOLLARS)

POS Payments

Payment Method	Australia[1]	Belgium	Netherlands	Norway
Cash	18.95	23.07	12.32	36.41
PIN Debit Card	72.81	65.42	57.96	61.86
Credit Card	67.82	130.04	151.32	104.98
Prepaid Card	—	6.76	3.56	—
Check	356.07	—	—	—

[1] As reported by merchants.
Sources: Schwartz and others (Table 11); National Bank of Belgium (Table 2); Brits and Winder (Table 4.1); Gresvik and Haare (Table 7).

debit cards and higher for debit cards than cash. These facts suggest that the differences in cost per transaction among the payment methods may have been due to differences in average transaction size rather than differences in the underlying efficiency of the payment methods. As Section II noted, one way of ruling out this possibility is to see if a payment method that had higher cost per transaction also had higher cost per unit of transaction value.

All four studies found that total social cost per unit of value was significantly higher for credit cards than debit cards (Table 7). This finding suggests that the higher cost per transaction of credit cards was likely due to the lower underlying efficiency of credit cards rather than the higher average transaction size.

While the relative ranking of credit cards and debit cards was unchanged when cost per unit of value was used as the metric, the relative ranking of cash and debit cards was reversed in three of the countries. With social cost per transaction used as the metric, cash was significantly less costly than debit cards in Australia and the Netherlands and slightly less costly in Belgium. However, when social cost per unit of value was the standard of comparison, cash was much *more* costly than debit cards in all three countries. These findings suggest that cash was efficient in all three countries for small transactions but not necessarily for large transactions.

Table 7

TOTAL SOCIAL COSTS PER UNIT OF TRANSACTION VALUE (PERCENT)

POS Payments

Payment Method	Australia	Belgium	Netherlands	Norway
Cash	3.21	3.03	3.20	5.63
PIN Debit Card	1.58	1.10	1.10	1.12
Credit Card	2.99	2.65	3.11	3.33
Prepaid Card	—	10.49	34.32	—
Check	2.29	—	—	—

Sources: Schwartz and others (Tables 8 and 11); National Bank of Belgium (Tables 1 and 2); Brits and Winder (Table 4.1); Gresvik and Haare (Table 7).

Variable social costs. The Belgian and Dutch studies also estimated the variable social costs of each payment method. Cost per transaction was assumed to be independent of the number of transactions but potentially varying in the size of the transaction.[15] The Dutch study found that prepaid cards had the lowest cost at all transactions sizes, and that cash had a lower cost than debit cards for small and medium-size transactions (Table 8). The study also found that credit cards had a higher cost than debit cards and prepaid cards at all transaction sizes and a higher cost than cash for all but the largest transactions. The findings of the Belgian study were similar except that prepaid cards were more costly than debit cards for large transaction sizes.

From these estimates of variable social costs, the studies were able to calculate the short-run cost savings from hypothetical shifts in the composition of payments. For example, the Belgian study calculated that shifting 25 percent of small and medium-size cash payments to a mix of debit cards and prepaid cards would reduce variable social costs by 6 percent. Similarly, the Dutch study estimated that shifting 21 percent of cash payments to debit cards and prepaid cards would reduce variable social costs by 7 percent.[16] These simulations indicate that relying less on cash and more on debit cards and prepaid cards to make small and medium-size payments could make better use of existing infrastructure. Because the simulations are based on variable costs, however, they do not prove that the shift from cash to debit cards

Table 8
VARIABLE SOCIAL COSTS PER TRANSACTION (U.S. DOLLARS)
By Transaction Size

Netherlands

Payment Method	Transaction Size		
	$5	$15	$100
Prepaid Card	0.04	0.04	0.04
Cash	0.18	0.25	0.84
PIN Debit Card	0.25	0.25	0.26
Credit Card	1.06	1.09	1.30

Belgium

Payment Method	Transaction Size		
	$5	$15	$100
Prepaid Card	0.12	0.15	0.35
Cash	0.21	0.29	0.96
PIN Debit Card	0.28	0.28	0.28
Credit Card	0.74	0.75	0.82

Note: Cost in local currency was converted to U.S. dollars using the exchange rate for the year of the studies, 1.3133.
Sources: Brits and Winder (Table 4.3) and National Bank of Belgium (Table 3)

and prepaid cards would also reduce social costs in the long run, when infrastructure can be adjusted.

IV. IMPLICATIONS FOR COST STUDIES BY OTHER CENTRAL BANKS

Two conclusions follow from the discussion of key cost concepts and the review of cost studies in Australia, Belgium, the Netherlands, and Norway. First, a central bank is well advised to conduct its own study of retail payment costs rather than rely on the findings of costs studies in other countries. Second, a central bank should give careful attention to the choice of cost measures and methodology for the study.

Need for a central bank to conduct its own cost study

Central banks need to conduct their own cost studies because both the technology and relative use of different payment methods can differ significantly across countries. Some countries may use a more cost-efficient technology for a payment method, causing the method to have a lower social cost. In addition, some countries may make greater

use of a payment method, taking greater advantage of economies of scale. Only by conducting its own cost study can a central bank be confident that it is using accurate estimates of the social costs of different payment methods to assess the efficiency of its payments system.

The need to conduct a separate cost study is especially evident for the United States, given the unique characteristics of its retail payments system. One important feature of the U.S. payments system is the continued heavy reliance on checks. Checks accounted for 22 percent of the number of noncash payments in 2009 (Federal Reserve System). Because checks are so important, considerable effort has been devoted to converting them to electronic image to reduce costs. As a result, the vast majority of checks in the United States are now processed electronically. Because of such electronification, the social costs of checks are likely to be lower in the United States than in countries where checks are rarely used and thus largely processed manually.[17]

Another unique characteristic of the U.S. payments system is the larger volume of card transactions. Although the number of card transactions per capita is smaller in the United States than in some European countries, the total number of card transactions is about seven times greater in the United States than in the U.K., the country with the largest number of card transactions in Europe (BIS 2011; European Central Bank). If economies of scale depend more on the total number of transactions than on the number of transactions per capita, the social costs per card transaction could be lower in the United States.

A further distinctive characteristic is the large number of participants in the U.S. retail payments industry. In most countries, debit cards are processed by only one or two card networks. In the United States, by contrast, debit cards are processed by more than 10 networks. The number of banks providing payment services is also larger in the United States, a legacy of unusually tight restrictions on the geographic expansion of banking in this country. Nonbank payment service providers are more numerous in the United States as well, due partly to the favorable climate for innovation (Bradford and others). In the case of debit cards, the large number of networks could prevent economies of scale in debit card processing from being fully realized, raising the social costs of such payments. However, the greater degree of competition among bank and nonbank payments providers in the United States could enhance

productivity, reducing social costs (Holmes and Schmitz). Conducting a separate cost study for the United States would help ensure that these potentially important effects were captured.

Need to choose appropriate cost measures and methodology

The second implication of the discussion of key cost concepts and the review of previous cost studies is the importance of choosing the appropriate cost measures and interpreting them carefully. For questions of short-run efficiency, measures of variable social costs may be sufficient because infrastructure can be taken as given. To assess long-run efficiency, however, information on the total social cost of each payment method is needed, including both fixed and variable costs. In interpreting findings, it is also important to remember that total social costs may be high because the payment method is being used at too low a scale to exploit economies of scale rather than because the method is inherently inefficient. This factor could partly explain the high social costs of credit card payments in the four studies reviewed in Section III.

Central banks are also well advised to collect information on both the number and value of transactions, so that costs can be scaled by both measures of payments volume. Comparing payment methods in terms of both methods can help a central bank determine whether cost per transaction is high because a payment method is inherently inefficient or because costs depend partly on transaction size and average transaction size is high. For example, the cost studies reviewed in Section III found that credit cards had both higher social cost per transaction than debit cards and higher social cost per unit of value. That finding suggested that the higher cost per transaction of credit cards was not due solely to their higher average transaction size.

In designing a cost study, a central bank should also pay careful attention to methodology. It was noted in Section III that the treatment of consumer costs varied across the four studies reviewed. The Belgian and Dutch studies excluded consumer costs of using a payment method from social costs. The Australian and Norwegian studies included consumer costs but measured them differently. If consumer costs are more important for some payment methods than others, such differences in treatment could influence countries' cost rankings of the payment methods.

One way to avoid such biases is for central banks to follow a common methodology in their cost studies. An example of this approach is a cost study being conducted by the European Central Bank (ECB) in cooperation with the European System Central Banks. To collect consistent and comprehensive information on the social costs of different payment methods in European countries, researchers have developed a common methodology (Ruttenberg). Other central banks could follow the same methodology as the ECB, making it easier to compare their cost rankings with those of European countries.

V. CONCLUSIONS

To meet the policy goal of an efficient retail payments system, central banks need accurate and comprehensive information about the costs of making retail payments. The potential gains from reducing retail payment costs may be considerable, with such costs estimated to absorb 0.5 percent to 0.9 percent of annual economic output in a number of countries. As this article has noted, reductions in retail payment costs will represent a net gain to society only if the benefits of the payments system to merchants and consumers are maintained or increased. Thus, to assess the overall efficiency of the payments system, central banks need information on the benefits of each payment method in addition to the costs. Nevertheless, collecting cost information can represent a key step in promoting payments efficiency, whether central banks are acting in their roles as operators, overseers, or catalysts for change.

In an effort to acquire such cost information, central banks in developed countries have recently conducted comprehensive studies of retail payments costs. These studies have reached similar conclusions on some questions, such as the relative cost efficiency of credit cards and debit cards but different conclusions on other questions, such as the relative cost efficiency of cash and debit cards. The divergence in some findings and the significant variation across countries in payments technology and rates of use of payment methods suggest that central banks conduct their own studies of retail payments costs. This implication is especially important for countries such as the United States, whose retail payments system has a number of unique characteristics. Central banks wishing to conduct such cost studies need not start from scratch,

however. This article has shown that they can learn many lessons from previous cost studies about how to measure and interpret costs and what kind of information to collect.

APPENDIX

This Appendix uses a simple example to illustrate the key cost concepts discussed in Section II.[18] Let Q be the quantity of payment services produced with a particular payments technology. Assume that providing such services requires two inputs—capital K, which is fixed in the short run but variable in the long run, and labor L, which is variable in both the short run and long run. Assume that the maximum quantity of payments services that can be produced given K and L is $Q = AK^{\alpha}L^{\beta}$, where $0 < \alpha < 1$ and $0 < \beta < 1$. Finally, let r be the annual cost of a unit of capital, and let w be the annual cost of a unit of labor.

For any given capital input K_o, the *short-run variable cost* of producing total payment services Q is the cost of the labor needed to produce Q. Under the assumptions above, this cost can be expressed as

$$\text{(A1) } SRVC(Q, K_o) = wL = w(Q/(AK_o^{\alpha}))^{1/\beta}$$

The *short-run total cost* of producing Q is the variable cost plus the cost of the fixed capital input:

$$\text{(A2) } SRTC(Q, K_o) = wL + rK_o = w(Q/(AK_o^{\alpha}))^{1/\beta} + rK_o$$

Finally, the *long-run total cost* of producing any quantity of payment services Q is the minimum possible cost of producing Q when both capital and labor are treated as variable:

$$\text{(A3) } LRTC(Q) = wL + rK = (\gamma Q)^{1/(\alpha+\beta)},$$
$$\text{where } \gamma = r^{\alpha}w^{\beta}(\alpha/\beta + \beta/\alpha)/A$$

As noted in Section II, the costs of different payment technologies can be compared only if total costs are scaled by some measure of payments volume. Accordingly, each of the expressions above is divided by the quantity of payment services to obtain corresponding measures of the *average* cost of producing payments services:

$$\text{(A1')} \; SRAVC(Q, K_o) = SRVC(Q, K_o)/Q = (w(Q/(AK_o^{\alpha}))^{1/\beta})/Q$$
$$\text{(A2')} \; SRATC(Q, K_o) = SRTC(Q, K_o)/Q = (w(Q/(AK_o^{\alpha}))^{1/\beta} + rK_o)/Q$$

(A3') $LRATC(Q) = LRTC(Q)/Q = \gamma^{1/(\alpha+\beta)}Q^{(1-\alpha-\beta)/(\alpha+\beta)}$

1. Total versus variable costs

Using the example above, Chart A1 illustrates why it is appropriate to focus on variable costs when considering short-run payments efficiency but total costs when assessing long-run efficiency. Of the two payment methods shown, method A has a more capital-intensive technology than method B. As a result, method A has only half as great an average variable cost as method B given the existing capital stocks and quantities of payments services. In particular, method A is used to produce quantity 17 at an average variable cost of 4, while method B is used to produce quantity 12 at an average variable cost of 6. Given the existing capital stocks, total payment costs could be reduced by shifting some payment services from method B to method A, using less labor in method B and more in method A. Over the long run, however, method A has no cost advantage over method B. As capital depreciates, it will have to be replaced, and these replacement costs will be higher for method A because it uses a more capital-intensive technology. In fact, in the example, the two payment methods have an identical long-run average cost of 12, indicating that neither method is superior.

2. Cost per transaction versus cost per unit of value

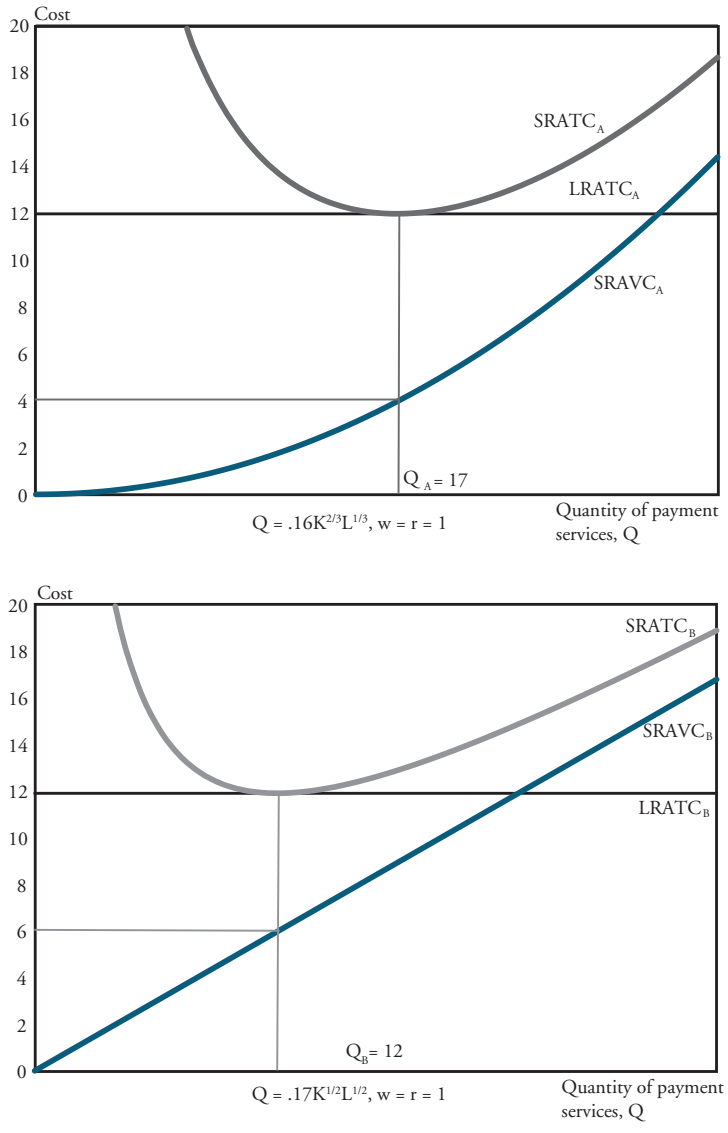
A simple extension of the above example shows why it is important to look at both cost per transaction and cost per unit value in comparing different payment methods. Suppose that there are constant returns to scale in the production of payments services—that is, $\alpha+\beta=1$. Suppose further that the quantity of payment services that must be provided is a linear function of the number of transactions, n, and the average transaction size, s. Specifically, $Q = n(d+es)$, where d and e are constants. Under these assumptions, it is easily shown that the long-run total cost of carrying out n payments of size s is

(A4) $LRTC(n,s) = n(a+ bs)$,
where $a = \gamma d$, $b = \gamma e$, and γ is given by (A3).

Dividing by the number and value of transactions yields the two cost measures for comparing the efficiency of payment methods:

Chart A1

TOTAL VERSUS VARIABLE COST OF PRODUCING PAYMENT SERVICES



Sources: Authors' calculations

(A5) Cost per transaction = LRTC/n = a+bs
Cost per unit of transaction value = LRTC/(ns) = a/s + b

Now consider two payment methods A and B that use the same technology but are operated with different transactions sizes. As shown in the top panel of Chart A2, method A with the smaller transactions size will have lower cost per transaction but higher cost per unit of transaction value. Thus, even though the two methods use the same technology, A will appear superior to B when the methods are ranked by cost per transaction, and inferior to A when the methods are ranked by cost per unit of value.

If a cost study could collect enough information to estimate the function LRTC(s,n) for each payment method, it would be a simple matter to determine which method, if any, was more cost efficient at each possible transaction size. Typically, however, a cost study will only determine the total cost, total number of transactions, and total value of transactions at the time the study was conducted. In terms of the bottom panel of Chart A2, the study will only be able to identify one point for each payment method—say point A for method A.

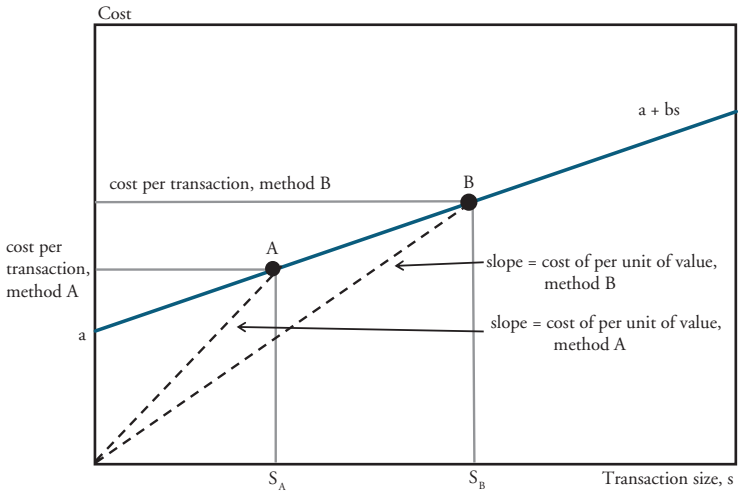
Even with such limited information, however, it may still be possible to gain insight into the relative costs of two payment methods. If a second payment method B happens to fall in the hatched region in the diagram, with a higher cost per transaction but lower cost per unit of transaction value, there will be no way to know which method is more cost efficient at each transaction size. Suppose, however, that method B falls in the shaded region in the diagram, with both a higher cost per transaction and a higher cost per unit of transaction value. Then as long as the cost function takes the linear form assumed in (A4), method A must be more cost efficient than method B. Specifically, method A must have lower cost than method B at transaction size s_B, and method B must have higher cost than method A at transaction size s_A.[19]

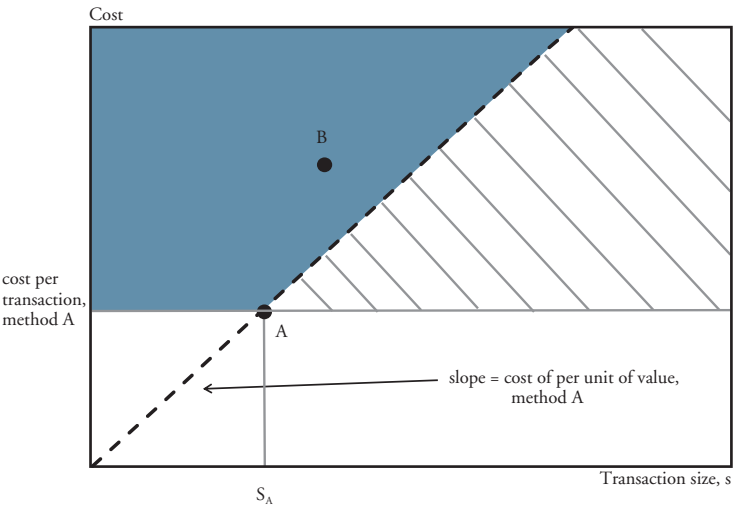
3. Constant versus increasing returns to scale

In the examples above it was assumed that the payments technology is subject to constant returns to scale, in the sense that doubling capital and labor inputs doubles the quantity of payment services that can be produced ($\alpha+\beta= 1$). Chart A3 shows that in the presence of increasing returns to scale, the long-run average total cost of producing payments

Chart A2

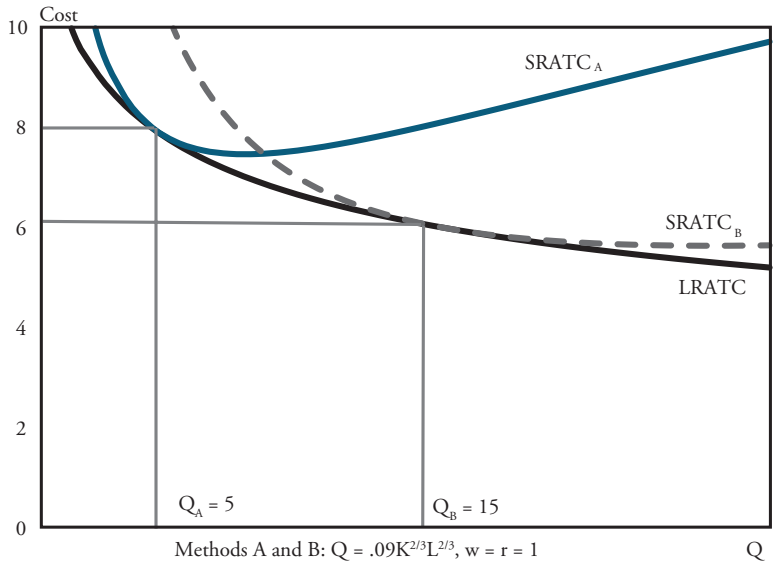
COST PER TRANSACTION VERSUS COST PER UNIT OF VALUE





Sources: Authors' calculations

Chart A3
INCREASING RETURNS TO SCALE



Sources: Authors' calculations

services can give a misleading indication of the relative cost efficiency of two payment methods. In the example shown, payment methods A and B have the same increasing-returns-to-scale technology ($\alpha+\beta > 1$) but are used to produce different quantities of payments services. Method A is used to produce quantity 5 at a long-run average cost of 8, while method B is used to produce quantity 15 at a long-run average cost of 6. Thus, method B has lower long-run average cost than method A, but only because the two methods are operated at different scale and the technology of each method is subject to increasing returns to scale.

ENDNOTES

[1]Another example of a private cost is float cost—the interest that a payee forgoes where there is a delay between the initiation of payment and the transfer of funds from payer to payee. Float costs represent a transfer from the payee to the payer, who earns additional interest as a result of the delay.

[2]While the focus of this article is on social costs, there may also be good reasons for a study of private costs to collect information on the variable component of those costs. In some situations, policymakers may conclude that the price that one party is charged by another party for a payment service should depend on the incremental cost to the other party of providing the service. For example, in the Durbin Amendment to the Dodd-Frank Act, Congress took the view that the "interchange fee" a merchant has to pay when a consumer uses a bank-issued debit card should be proportional to the incremental cost to the issuing bank of processing the payment (Board of Governors). Implementing a standard of this type requires information on private variable costs.

[3]Consistent with this view, econometric studies of economies of scale in electronic payment methods generally use the number of transactions as the measure of output, rather than the value of transactions (Bauer and Ferrier; Hancock and others).

[4]For this reason, the Board of Governors allowed an ad valorem component in debit interchange fees in its final rule implementing the Durbin Amendment to the Dodd-Frank Act (Board of Governors).

[5]In some countries, official cost studies have estimated variable cost functions of the form $c = a + bs$, where c is the cost of carrying out a transaction of size s, a is the component of the cost that is independent of size, and b is the cost per unit of value. These cost functions can be used to compare the variable costs of two payment methods at each possible transaction, taking the existing infrastructure as given. However, because the cost functions are for variable costs only, they cannot be used to determine which payment method is less costly at each transaction size when the existing infrastructure is *not* taken as given.

[6]In the United States, card networks include not only Visa and MasterCard, which process all major types of card transactions, but also smaller networks such as STAR and NYCE, which process PIN debit transactions only. Examples of private interbank clearing organizations in the United States include CHIPS for wire payments and the Electronic Payments Network (EPN) for ACH transactions.

[7]Fees paid to outside parties may overestimate social costs of a payment instrument if those fees include above-normal profit margins.

[8]Some earlier studies have attempted to estimate the social costs of retail payment methods relying entirely on information from independent sources (Humphrey and Berger; Garcia-Swartz and others; and Simes and others). As noted by Shampine (2012), the estimates of social costs in such studies can be highly sensitive to the underlying assumptions and the reliability of the sources.

[9]Other studies of retail payment costs have been conducted by central banks in Canada, Finland, and Portugal (Arango and Taylor; Takala and Viren; Banco de Portugal). These studies are not reviewed here because they focused on private costs of merchants or financial institutions and did not attempt to estimate social costs. Another study by the central bank of Sweden estimated social costs but did not include the fixed costs of opening and maintaining bank accounts (Bergman and others).

[10]See, for example, Table 8 in Schwartz and others and Table 17 in Gresvik and Haare.

[11]The Australian study reported only the social cost per transaction. However, for each method, social cost per unit of value can be computed by dividing social cost per transaction by average transaction size, which is reported separately.

[12]In each case, aggregate social cost is the sum of total social costs for the various payment methods included in the study. The Norwegian study notes that mainland GDP may be a better measure of economic activity, because it excludes offshore activities such as oil extraction and shipping that generate wholesale rather than retail payments. When mainland GDP is used, the ratio for Norway is 0.65 percent (Gresvik and Haare, p. 13).

[13]Two other factors may have contributed to the higher relative cost of cash in the Norwegian study. First, the Norwegian study included consumer resource costs while the Belgian and Dutch studies excluded them. Consumer resource costs tend to be greater for cash payments than debit card payments, for example, because consumers relying on cash must make regular trips to ATMs to withdraw funds. As a result, including consumer resource costs raises the estimated social cost of cash more than that of debit cards. Second, while both the Norwegian and Australian studies included consumer resource costs, the Norwegian study used a higher estimate of time spent in cash withdrawals and assigned a higher value to consumers' time.

[14]The social cost of checks relative to other payments methods appears even greater in Australia than in other countries such as the United States, where checks are used in payments. One reason may be that the low volume of checks in Australia makes it impossible to exploit economies of scale in check processing. The Australian study reported that checks accounted for less than 1 percent of consumer payments in 2006 (Schwartz and others, Table 13). In the United States, by contrast, checks were estimated to account for 14 percent of consumer payments in 2008 (Nilson Report).

[15]See note 5.

[16]Such simulations require assumptions about both the number and the average size of the transactions that are shifted. The Belgian example assumes that two-thirds of the cash payments have an average value of $5 and are shifted to prepaid cards, while one-third have an average value of $20 and are switched to debit cards.

[17]While the social cost of checks has fallen in the United States, it has tended to exceed the social cost of other payment instruments. As a result, estimates of the total social cost of retail payments have historically been higher for the United States than other developed countries—as much as 2 percent of GDP (Humphrey).

[18]For further information on the cost functions used in this Appendix, see a microeconomics textbook such as Varian.

[19]From the diagram, it can be seen that if the cost functions were convex rather than linear (that is, if cost increased with transaction size at an increasing rather than constant rate), method B could fall in the shaded region and still have lower cost than method A at transaction size s_A.

REFERENCES

Arango, Carlos, and Varya Taylor. 2008. "Merchants' Costs of Accepting Means of Payment: Is Cash the Least Costly?" *Bank of Canada Review*, Winter 2008-2009: 15-23.

Banco de Portugal. 2007. Retail Payment Instruments in Portugal: Costs and Benefits.

Bank for International Settlements, Committee on Payment and Settlement Systems. 2003. "Policy Issues for Central Banks in Retail Payments," May.

_____. 2011. "Statistics on Payment, Clearing, and Settlement Systems in the CPSS Countries," September.

Bauer, Paul W., and Gary D. Ferrier. 1996. "Scale Economies, Cost Efficiencies, and Technological Change in Federal Reserve Payments Processing," *Journal of Money Credit and Banking*, 28(4), November.

Bauer, Paul W., and Geoffrey R. Gerdes. 2009. "The Check is Dead! Long Live the Check! A Check 21 Update," Federal Reserve Bank of Cleveland, *Economic Commentary*, June.

Bergman Mats, Gabriela Guibourg, and Bjorn Segendorf. 2007. "The Costs of Paying – Private and Social Costs of Cash and Card Payments," Sveriges Riksbank Working Paper No 212.

Bradford, Terri, Fumiko Hayashi, Christian Hung, Simonetta Rosati, Richard J. Sullivan, Zhu Wang, and Stuart E. Weiner. 2009. "Nonbanks and Risk in Retail Payments: EU and U.S." in Managing Information Risk and the Economics of Security, edited by M. Erick Johnson. Hanover, N.H., USA.

Brits, Hans, and Carlo Winder. 2005. "Payments are No Free Lunch," De Nederlandsche Bank Occasional Studies, 3(2).

Connolly, Paul M. 1996. "The Future Retail Payments System and the Role of the ACH," Federal Reserve Bank of Boston, speech to NACHA 1996 Payments Conference, April 16.

European Central Bank. 2011. "Payments Statistics," September.

Federal Reserve System. 2011. "The 2010 Federal Reserve Payments Study." April 5.

Garcia-Swartz, Daniel, Robert Hahn, and Anne Layne-Farrar. 2006a. "The Move Toward a Cashless Society: A Closer Look at Payment Instrument Economics," *Review of Network Economics*, 5(2): 175-197.

Garcia-Swartz, Daniel, Robert Hahn, and Anne Layne-Farrar. 2006b. "The Move Toward a Cashless Society: Calculating the Costs and Benefits," *Review of Network Economics*, 5(2): 198-228.

Gresvik, Olaf, and Harald Haare. 2009. "Costs in the Norwegian Payment System," Staff Memo, No. 4, Norges Bank.

Guibourg, Gabriela, and Bjorn Segendorf. 2007. "A Note on the Price- and Cost Structure of Retail Payment Services in the Swedish Banking Sector 2002," *Journal of Banking and Finance*.

Hancock, Diana, David B. Humphrey, and James A. Wilcox. 1999. "Cost Reductions in Electronic Payments: The Roles of Consolidation, Economies of Scale, and Technological Change, *Journal of Banking and Finance*, 23, 391-421.

Holmes, Thomas, and James Schmitz. 2010. "Competition and Productivity: A Review of Evidence," *Annual Review Economics*, 2: 619-642.

Humphrey, David B. 2010. "Retail Payments: New Contributions, Empirical Results, and Unanswered Questions," *Journal of Banking and Finance*, 34: 1729-1737.

Humphrey, David B., and Alan N. Berger. 1990. "Market Failure and Resource Use: Economic Incentives to Use Different Payment Instruments," in Humphery, David B. (ed.), *The U.S. Payment System: Efficiency, Risk, and the Role of the Federal Reserve*, Boston: Kluwer.

Lowe, Philip. 2005. "Payments System Reform: The Australian Experience," Interchange Fees in Credit and Debit Card Industry: What Role for Public Authorities? Proceeding a Conference at the Federal Reserve Bank of Kansas City.

National Bank of Belgium. 2006. "Costs, Advantages and Drawbacks of the Various Means of Payment," *Economic Review*, June: 41-47

Ruttenberg, Weibe. 2011. "Status of the ECB Study: The Social Cost of Payment Instruments," presentation, October.

Schwartz, Carl, Justin Fabo, Owen Bailey, and Louise Carter. 2007. Payment Costs in Australia: A Study of the Costs of Payments Methods. Reserve Bank of Australia.

Shampine, Allan. 2007. "Another Look at Payment Instrument Economics," *Review of Network Economics*, 6(4): 495-508.

Shampine, Allan. 2012. "An Evaluation of the Social Costs of Payment Methods Literature," Jan.12, at *http://ssrn.com/abstract=1984248*.

Simes, Ric, Annette Lacy, and Ian Harper. 2006. "Costs and Benefits of Alternative Payments Instruments in Australia," Melbourne Business School Working Paper No 8.

Takala, Kari, and Matti Viren. 2008 "Efficiency and Costs of Payments: Some New Evidence from Finland," Bank of Finland Discussion Papers 11.

Varian, Hal. 2009. *Intermediate Microeconomics*. 8th edition, New York: W.W. Norton.

Weiner, Stuart E. 2008. "The Federal Reserve's Role in Retail Payments: Adapting to a New Environment," Federal Reserve Bank of Kansas City, *Economic Review*, Fourth Quarter.