ELSEVIER

Editorial

Frontiers in payment and settlement systems: Introduction

This special issue of the Journal of Banking and Finance focuses on papers at the frontiers of research into payments and settlement systems. The range and importance of this topic can be seen by the variety and diversity of the papers published in this volume. The papers published here examine payments and settlement systems from many different angles, ranging from very micro-level analysis to very macro-level discussions. Furthermore, the large diversity in the types of payment and settlement systems can be seen by those being discussed in this volume, which include consumer payment systems, securities settlement systems and settlement flows between countries.

There are a variety of research themes that emerge from the papers in this volume. The first major theme to emerge, concerns developments in payment systems in consumer financial markets. There are four papers in this volume that address this issue. Milne, examines why banks have incentives to under invest in consumer payments networks at the expense of their own networks. Humphrey et al. examine the impact on bank costs of using ATMs rather than branches, and electronic rather than paper processing of payments. Calem Gordy and Mester examine the role of information asymmetries and the possible impact of new screening technology in the market for credit cards. Finally Barnes and Lopez examines how much the US banks should pay the Fed for the provision of payment services as required under the Monetary Control Act.

A second important theme to emerge from the papers published here is the role of technology in payment systems. The paper by Bauer and Hein poses the question of why consumers are sometimes reluctant to adopt new payment technologies. Roth and Richter address the topic of technology and payment systems from another angle, by proposing a new methodology which increases the security of ATMs without imposing undue burdens on consumers.

A third important theme is the macro and international context of payment and settlement systems. Rosati and Secola use a gravity model to examine European payment flows. The gravity model considers distance and economic magnitude to be the key determinants of payment flows between countries. Schmiedel et al. compare securities settlement systems in different countries in order to determine if economies of scale exist in these systems.

The fourth theme to emerge from these papers concerns research into securities settlement systems (as opposed to consumer payment systems described above). The paper by Schmiedel et al. (described above) is one such paper, as is the paper by Devriese and Mitchell, which examines what happens when there is a major default in a securities settlement system.

In the sections below we describe in more detail the main contributions of the different papers.

1. Payment systems in consumer markets

1.1. Milne

Milne's paper "What's in it for us? Network effects and bank payment innovation" addresses the important puzzle of why the technological adoption and diffusion of bank payment systems seem to lag behind the technological diffusion of so many other technologies (e.g. cell phones etc). Why, in an environment of almost instant cell phone and internet connections, does it still take three days or more to clear a check in the UK, and two or more days in the US? This question is particularly important, given that research indicates that the costs of payment services can amount to 3% of GDP.

In order to answer this puzzle, Milne provides a theoretical model which focuses on the different incentives faced by banks when investing in payment services. The key observation is that the cost of a payment services network is shared by all banks in the system. Thus all (or many banks) in the system have to cooperate in order to increase investment in the payment system network. On the other hand, banks face a significant incentive to invest in their own within-bank payment networks in order to exploit account externalities. The argument that Milne makes is that when a bank faces a choice about whether to invest in the combined payment services network, it will ask the question "What's in it for us?" with the answer "Not Much".

A key contribution of the theoretical model is that it tries to explain why there is such a variation across countries in payment system networks (e.g. the UK is so much slower than the Scandinavian countries). This is particularly puzzling because standard microtheory would predict that large countries with competitive banking systems would have faster rates of technological innovation compared to smaller countries with concentrated systems (such as the Scandinavian countries). One explanation to this puzzle, proposed by Milne, is that one incentive for investing in the combined system (often ignored by banks) is that they will benefit from the increased macro-economic growth that will flow from greater investment in the payment system. The argument made by Milne is that these social benefits would be more likely to be internalized by banks in a small country where banking system is dominated by a few banks, which would explain why banks in Scandinavia have invested so much in their payment system.

1.2. Humphrey, Willesson, Bergendahl and Lindblom

In most of the existing literature on the impact of technical change on bank costs, technical change has usually been estimated using a time dummy variable. The main contribution of the Humphrey et al. paper is to use new data to examine the specific impact on bank costs of two different elements of technical change (i) the movement from paper

to electronic payment systems and (ii) the movement from branch offices to ATMs. By specifically examining the impacts of these technical changes on overall costs, the authors can provide an estimate of the total savings achieved by the banks from these changes. The authors test their model using data from 12 European countries.

The authors find that the economic magnitudes of the impacts of these technological changes are clearly substantial, with savings in the order of $32 billion, or 0.38% of GDP. The authors are also able to extrapolate that if banks moved to a situation of all payments being processed electronically rather than using paper, then the cost savings could be as high as 1% of GDP.

The paper differs from many other studies in that it specifically relates banking costs to two measures of physical services delivery (ATMs, branches) and four measures of payment processing alternatives (check, paper giro, electronic giro and debit/credit card), rather than traditional cost function measures such as bank loans and bank deposits in the balance sheet. The authors find that a 10% increase in all six of these output characteristics would result in a 9% increase in operating expenses – implying that average operating costs would fall. The authors are also able to conclude that (as expected) the costs of paper based payments is increasing over time, while the costs of electronic based payment systems is declining over time. Furthermore, scale economies in ATMs are also reducing their costs to the banks.

1.3. Calem, Gordy and Mester

This paper addresses an old puzzle in the economics of credit cards – why do credit card interest rates appear to be sticky downward, even though the market resembles a competitive market with many thousands of competitors? A variety of different explanations for this puzzle have been proposed in the literature, including the existence of search costs and switching costs. The focus on this paper is on empirically reexamining the switching costs hypothesis first proposed by Calem and Mester (1995). This argument states that if a high balance customer intends to switch card providers, the provider will not be able to distinguish between whether the applicant merely intends to switch cards, or alternatively, whether the intention is to accumulate more (unsustainable) debt. For this reason card providers may be reluctant to provide credit cards to high balance switchers – thus resulting in reduced competition for these customers. In testing this hypothesis Calem and Mester (1995) used the 1989 survey of consumer finances and found that households with larger balances were more likely to be rejected, which is evidence that is consistent with the switching costs hypothesis.

The aim of the current paper is to retest the hypothesis in the light of significant institutional changes in the credit card market over the last decade. The advent of sophisticated credit scoring technology, which has automated the prescreening of applicants, seems to imply that that banks now have improved ability to judge the creditworthiness of applicants – thus reducing the importance of asymmetric information issues related to the switching cost hypothesis. On the other hand the last decade has also seen the rapid increase in delinquency which could imply that information asymmetries are still of importance.

The main finding of the current paper is that using the 1998 as well the 2001 survey of consumer finance data, the authors find that high balance households are more likely to be denied credit or receive a lower than requested credit limit when applying for credit. This

result thus replicates the results from the earlier research using the 1989 survey of consumer finance data. The important conclusion that can be drawn from this result is that even after all of the institutional changes in the credit card market in the 1990s involving the use of credit screening technologies, information asymmetries still appear to play an important role in the allocation of credit cards.

1.4. Barnes and Lopez

This paper address the question of how much US banks should pay the Federal Reserve for its provision of payment services such as check clearing and electronic fund transfers. The Monetary Control Act requires that the price the banks have to pay for these payment services reflects the costs that a private sector provider would incur. In order to impute these costs, the Fed uses a methodology called the private sector adjustment factor, to determine these costs.

A key element of these costs that the Fed has to impute is the average cost of equity capital which would be appropriate for these services. What the Fed has been doing is to select a peer group of bank holding companies and determining their cost of equity capital using methods such as the CAPM model. The problem with this approach however, is that, as argued by Fama and French (1997) and many others, using the CAPM is an extremely poor measure of cost of equity capital. On the other hand, this approach is still the most commonly used in practice despite its known flaws.

The aim of this paper is to examine in detail the various choices that can be made when using the CAPM in order to determine which would be the most appropriate approach for the Fed to use. These choices include; which publicly traded firms should be included in the peer group, which additional factors should be added to the CAPM model, which econometric method should be used and should firm specific factors be included. Based on these different decisions, the authors generate more than 200 different cost of equity estimates for the Fed's payment business, and using a variety of different criteria, attempt to determine which are the most appropriate for use in practice.

The authors conclude that the most appropriate approach for the Fed to use is to estimate the cost of equity capital using a single factor CAPM model, using individual bank holding company regressions for a relatively large bank holding company peer group. The authors find that even though a variety of other modeling decision could be made ex ante, these choices do not greatly alter the results from the simpler methods proposed. An important advantage of using a simpler method is that it allows for the public replicability of the Fed's estimates, which is an important element in terms of the transparency of the methodology used.

2. The impact of technology on payment systems

2.1. Bauer and Hein

This paper examines the question of why different bank customers adopt new technologies to access banking services at different rates – a topic of considerable importance to many banks. For example, it is important for banks to understand when they introduce a new technology such as phone banking or internet banking, how many of their consumers will adopt these new technologies. In order to examine this issue they build a theoretical

model where the probability of adoption of a new technology will depend on the consumers perceived added utility from adopting the new technology, the cost of the technology as well as the risk premium.

In order to test this model the authors use data from the 1998 survey of consumer finance (SCF). This survey provides data on which households used internet banking and which used phone banking. It also includes a variety of other types of data such as the number of bank accounts held as well as demographic variables. Of most interest in the context of this paper is that the survey has data on each consumer's degree of risk aversion, based on their answers to a list of subjective questions regarding their risk taking behavior.

Their empirical results are in line with the theoretical predictions of the model. They find that the probability of a consumer using an internet account is significantly related to the number of checking accounts a consumer has, the younger the age of the consumer, the higher the consumer's income, and the higher the familiarity the consumer has with the internet. Of particular interest is the finding that the higher the consumers self assessment of risk aversion the lower the likelihood of adopting a new technology such as internet banking.

The authors extend this analysis by interacting the risk term with the age term. They find that older consumers are more hesitant about adopting new technologies, no matter what their self described level of risk aversion. For younger consumers on the other hand, the degree of risk aversion plays a key role in determining the adoption of new technologies. The implications of this finding for banks are that they should focus their efforts on educating younger consumers on the advantages of adopting these new technologies.

2.2. Roth and Richter

The paper by Roth and Richter address a very important problem for the providers and users of ATM machines – the threat of fraudulent acquisition and use of personal identification numbers (PIN Numbers) as well as the "skimming" of information on the magnetic strips on the back of cards. A particular security threat is "shoulder surfing" which means that perpetrators of the fraud access an individual's ATM PIN number by means of hidden cameras etc.

The aim of this paper is to propose a solution to this problem. An important constraint in any solution is to ensure that the consumer who owns the ATM card is not required to remember more than the four numbers of the PIN number. On the other hand, the aim is to ensure that for somebody simply observing the consumer typing in the PIN number they would not be able to recall the required information contained in the PIN number. Roth and Richter introduce a mechanism called "cognitive trapdoor games" which they claim make it very difficult for a criminal to obtain the information in a PIN number, even though they could fully observe all the information that the consumer types into the key pad.

The main mechanics of their proposed procedure is that ATM users are shown screens with different numbers highlighted in either black or white. For the first number of the PIN the user only responds to the ATM whether that number is displayed in either black or white. This color identification is repeated 4 times for each of the 4 numbers in the PIN code. The implication of this mechanism is that all the consumer has to remember is the 4 numbers in the individual PIN. On the other hand, a criminal watching this transaction

will only observe the consumer responding white or black. In this way it will be impossible for a fraudulent observer to obtain the PIN number – even though they can watch all the numbers being entered by the consumer.

The authors have extensively tested this proposal for both security as well as usability, and have found that it performs well. Although they found that their method takes a short while longer to key in, the additional benefits it provides in terms of additional security are significant.

3. Payment and settlement in the macro and international context

3.1. Rosati and Secola

While gravity models have been used extensively in the international trade literature to measure the impact of country size and distance on trade flows, Rosati and Secola use such a model to explain the large payment flows between European countries. They examine data from two different European payment systems, the TARGET system which is operated by the European Central Banks, and the EURO1 system which is operated as a privately owned system by the European Banking Association.

The aim of their analysis is to understand the payment flows between the different European countries in order to examine the impact of different financial centers, and also the possible dependence of one country's banking system on the liquidity provision of other countries. The authors find that liquidity within the Euro area flows in what they describe as three different circuits. The first , and largest, circuit, is comprised of the core or largest countries in the EU – Germany, France and the UK. The second circuit consists of these major countries providing liquidity to the small or medium sized countries. The final circuit is the provision of liquidity between smaller countries which have geographic or cultural connections (e.g. a Scandinavian Circuit and an Iberian Circuit).

Besides the gravity model, which examines how payments flow across space, the authors also undertake a study of the determinants of payments flows over time using daily data from 1999 to 2002. Their aim is to distinguish between economic determinants of these flows and technical determinants. They find that the key determinant of these flows is trading volumes in the euro money market, and in particular in the overnight unsecured segment of the market. In terms of the technical determinants, they find that bank holidays in the US, the UK or Germany will reduce the amount of cross border flows in the whole system. One interesting finding is that banks within Europe seem to begin redistributing liquidity on the day of the Eurosystem's main refinancing operations, which implies that the banks are anticipating these refinancing operations.

3.2. Schmiedel, Malkamaki and Tarkka

The aim of this paper is to examine the extent of scale economies in the settlement of securities such as equities, interest-bearing instruments and derivatives. Clearance and settlement costs are a key element of the transactions costs faced by investors involved in trades. The paper addresses the issue of whether economies of scale will result in the emergence of one or a few securities settlement systems around the world. The authors also focus on the fragmented nature of the European settlement systems, and examine whether integration across Europe will result in increased economies of scale. A key argument in

the paper lies in the need to examine both the supply as well as the demand side of the industrial structure of market places. The demand side has been evaluated in many market micro-structure studies, but this paper examines the supply side in terms of the provision of settlement services.

This is the first study to examine economies of scale across a wide variety of different settlement systems in Europe, Asia and North America. The paper uses a panel-based analysis to examine economies of scale all of the major settlement institutions from 1993 to 2000. The authors follow the methodology used to examine economies of scale in stock exchanges and apply it to settlement systems.

The main result of the study is the finding of substantial economies of scale for both depository as well as settlement services. As expected, the highly centralized US system is found to be the most cost effective settlement system. On the other hand, for most of the smaller systems examined in the study, there would be a high percentage of unit cost savings if the systems were to increase in size. For example, the authors report that a doubling of the size of the smaller settlement systems would only increase costs by two-thirds. Furthermore, the study finds that the current complexity of the European systems has significantly increased the costs of international settlement relative to purely domestic systems.

The main policy conclusion of the paper is that there should be increased mergers and acquisitions of settlement systems so that investors can take advantage of the increased economies of scale. Furthermore the removal of institutional impediments currently in place in Europe could also provide significant additional benefits.

4. Securities settlement systems

4.1. Devriese and Mitchell

The aim of this paper is to examine the impact on securities settlement systems of a major default – in particular the default of its largest participant. As was evident in the week after September 11, failure of a securities settlement system can have very severe consequences, thus it is important to understand the dynamics of such a crash. A key point emphasized by the authors is that securities settlement systems are different from payment systems in that they involve both a cash leg as well as a securities leg. Thus the default of a major participant could have an impact on other members of the system in terms of both their cash holdings as well as their securities holdings. This is an important difference to payment systems where only cash is involved.

Another important difference between payment systems and securities systems is that the settlement in securities systems takes place two days after the trade. This increases the risk to the system, because asset prices may have changed thus making it more expensive to settle the trade. Furthermore, because of this settlement lag, the impact of a default can last longer than the two-day settlement period. This is because agents have to make predictions about the impact of the default on their future trades. If however settlement failure turns out to be higher than agents' predictions, then new settlement failure could occur because of these inaccurate predictions.

Using simulation based analysis the paper is able to reach some important conclusions concerning the dynamic impacts on settlement following the default of the largest participant in the system. The results of the simulation indicate that the crisis situation initially

worsens before it improves, and that the crisis lasts longer than the two-day settlement lag. These results follow from the impact of unfulfilled expectations and contagion. Another important result concerns the position of the large institution, whose default initiates the crisis. If the defaulting institution is in a large net buy position, then this will have a significantly greater impact compared to when the defaulter is in a large net sell position. However this can be ameliorated to some extent if the system itself (or the central bank) provides additional liquidity during the crisis. However the crisis will still be severe because liquidity providers can have an impact on the cash leg of a trade but not the securities leg.

Anthony Saunders
Leonard Stern School of Business
New York University
Salomon Brothers Center
44 West 4th Street
New York, NY 10012-1126
United States
Tel.: +1 212 998 0711; fax: +1 212 995 4232
E-mail address: asaunder@stern.nyu.edu

Barry Scholnick
School of Business
University of Alberta
Edmonton
Alta., Canada T6G 2R6
Tel.: +1 780 492 5669; fax: +1 780 492 3325
E-mail address: Barry.Scholnick@ualberta.ca

Available online 26 October 2005