The economics of credit cards, debit cards and ATMs: A survey and some new evidence

Barry Scholnick a,*, Nadia Massoud b, Anthony Saunders c,*, Santiago Carbo-Valverde d, Francisco Rodríguez-Fernández d

a School of Business, University of Alberta, Edmonton, Canada T6G 2R6
b Schulich School of Business, York University, Toronto, Canada
c Stern School of Business, New York University, New York, NY 10012-1126, United States
d University of Granada, Spain

Available online 6 May 2007

Abstract

This paper provides a critical survey of the large and diffuse literature on credit cards, debit cards and ATMs. We argue that because there are still many outstanding issues and questions about the pricing, use and substitutability of these payment mechanisms, that there are significant further opportunities for research in these areas. A large number of questions are examined in this survey, including the pricing of credit cards, the impact of networks on the provision and pricing of ATMs, as well as the tradeoffs that consumers make between different types of payment mechanism, including debit cards, credit cards and ATMs. Importantly, this paper is also amongst the first to provide new evidence on this latter question from bank level data (from Spain). We conclude that point of sale (debit card) and ATM transactions are substitutes, and that ATM surcharges impacts point of sale volume significantly.

© 2007 Elsevier B.V. All rights reserved.

JEL classification: G21

Keywords: Credit cards; Debit cards; ATM

1. Introduction

Thirty years ago most consumers accessed banking services, such as making deposits or determining balances, by visiting their bank in person. In the intervening thirty years, however, there has been enormous technological change in the provision of banking services to consumers. Today many consumers go for many months, or even years, without having to physically visit their bank, or a bank teller. The most important technological developments that have caused this change have been the emergence of credit cards, debit cards and automated teller machines (ATMs). Given the importance of these technologies, to both banks as well as consumers, it is not surprising that these technological developments have lead to a large and growing body of academic research.

This paper provides a survey of the literature on the economics credit cards, debit cards and ATMs. As we argue below, this literature is not only large and growing, it is also remarkably diffuse. Researchers have examined these products in the context of a large variety of different sub-disciplines including, financial economics, banking, monetary economics, macroeconomics, industrial organization, regulatory economics, consumer behavior, and network economics among others. The aim of this paper is to provide a survey of some of the major questions, puzzles and issues in this literature. We also provide some new evidence on the relationship between ATMs and debit cards. It should be emphasized that this survey focuses on academic...
Section 5 provides our new results taken from a very large hypotheses concerning credit cards, debit cards and ATMs. the data necessary to empirically test a large number of large panels of individual banks etc. which can provide seen the recent emergency of large and richly detailed data- ATMs can provide large rewards for economists from a studying the economics of credit cards, debit cards and cerning consumer financial decisions, it seems evident that tion. Given the complexities inherent in these products and technologies we examine here are highly complex products and services, which bundle together a large variety of attributes. For example, credit cards provide a means of payment services and line of credit etc. ATMs involve the customers’ bank as well as other banks in networks etc. Furthermore, all of these products involve consumer behavior, which could entail less than rational decision making. The complexities and distortions in these products and technologies have provided the starting points for many of the studies examined in this survey. Most of these studies attempt to explain how these products and technologies work, and the impacts that they have on consumers, banks and markets.

The main conclusion of this survey, is that while a large amount of research has already been undertaken on credit cards, debit cards and ATMs, we believe that there are still a great many issues and puzzles that remain to be resolved. For example, the issue of credit card interest rates has generated a large number of alternative hypotheses, with very littler convergence towards a commonly accepted explanation. Given the complexities inherent in these products and services and the possible irrationality of consumers concerning consumer financial decisions, it seems evident that studying the economics of credit cards, debit cards and ATMs can provide large rewards for economists from a variety of different sub-disciplines. Furthermore, we have seen the recent emergency of large and richly detailed data- sets involving large panels of individual consumers, or large panels of individual banks etc. which can provide the data necessary to empirically test a large number of hypotheses concerning credit cards, debit cards and ATMs.

This survey has four main parts. Sections 2-4 discuss in turn the literatures on credit cards, ATMs and debit cards. Section 5 provides our new results taken from a very large database on ATM and debit card (point of sale) usage in Spain. Finally Section 6 concludes.

2. Credit cards

Credit cards are highly complex financial instruments. Their usage reflects a large number of different characteristics and motivations (transactions, debt, consumer benefits etc.), involve a large number of prices (interest rates, teaser rates, grace periods, penalty fees, annual fees etc.) and quantity constraints (credit limits, minimum payments). These characteristics and their associated services are supplied by a large variety of different card providers (banks, non-banks etc.). Furthermore, because credit card markets involve decisions by consumers (rather than corporations or markets) issues of consumer behavior and consumer rationality play a far more significant role in this market relative to other financial markets.

In the last 15 years or so Credit cards have essentially become a “test case” for examining how far standard finance theory (risk and return, competitive markets etc.) can be extended in an environment where some of the usual characteristics of financial markets (e.g. competitive buyers and sellers of assets) are not present. The aim of this section is to provide a literature survey to draw together the various strands of research on credit cards as a financial instrument and payment mechanism.

2.1. Credit card pricing puzzles

One of the main strands of research on credit cards has been an examination of pricing issues, such as whether credit card interest rates are “too high”. As we argue in this section, a number of alternative hypotheses have been proposed in the literature to explain credit card pricing with very little consensus emerging. The clear implication from the array alternative explanations for credit card pricing is that significant further empirical research is required in order to provide a more robust understanding of how credit cards are priced. This section examines these various hypotheses regarding card pricing and the empirical evidence to date.

2.1.1. Adverse selection, search costs and switching costs

The article by Ausubel (1991) was one of the first to examine in detail the significant pricing distortions inherent in the market for credit cards, and to propose possible explanations for these distortions. Ausubel (1991) documents a significant puzzle in that even though the market for credit cards resembles a competitive market (with many thousands of suppliers and few barriers to entry), the actual interest rates charged by credit card providers are both high as well as sticky downward. Ausubel (1991) provides a variety of evidence in support of his view concerning the profitability of credit card providers, including premia in the resale of credit card accounts as well as the comparison of credit card profitability in the banking industry with other banking products. The puzzle Ausubel (1991) poses is why credit card interest rates are so sticky and why profits are so high, even though the market structure resembles that of a competitive market.

Ausubel (1991) poses a variety of different explanations for this puzzle, largely based on the issues of asymmetric information and consumer behavior. He highlights two possible costs to consumers in seeking to change credit card products. These are search costs (the costs of finding information on alternative providers) and switching costs, (the costs of switching to an alternative provider). He also proposes an adverse selection theory whereby banks are reluctant to cut interest rates because this action will attract high risk borrowers (based on the assumption that these
high risk borrowers are more likely to search for low interest rate cards because they assume that they are very likely to be using that debt in the future). Ausubel’s argument is that the credit card pricing puzzle can be explained by a combination of search costs, switch costs, adverse selection as well as some element of consumer irrationality.

The paper by Calem and Mester (1995) is a follow up paper to that of Ausubel (1991). These authors propose and test a wider variety of possible explanations for the credit card pricing puzzle than that of Ausubel (1991), all of which are based on issues of switch costs, search costs and adverse selection. In terms of search costs for example, Calem et al. (2005) argue that consumers who have higher credit card balances also face higher search costs because they have higher disutility from searching. For this reason, a bank that unilaterally lowers its credit card interest rate will tend to attract those (less profitable) consumers who have lower balances. An alternative explanation is that consumers who have large outstanding balances may have more difficulty in switching to a new card compared to other consumers. This is because such consumers would face a high likelihood of being rejected by an alternative bank. The implication of this is that if a bank unilaterally lowered its interest rate it would attract those (less profitable) borrowers who maintain lower balances. Once again the argument of Calem and Mester (1995) is that banks face an incentive not to unilaterally lower their interest rates – which could possibly explain Ausubel’s observation that credit card interest rates are sticky downward. Calem and Mester (1995) test their hypotheses on the 1989 Survey of Consumer Finance. They conclude that each of the three factors of (i) search costs, (ii) switching costs and (iii) adverse selection have contributed to explaining Ausubel’s credit card pricing puzzle.

Stango (2002) provides a new empirical test of the switching cost explanation for credit card pricing originally proposed by Ausubel (1991) and Calem and Mester (1995). In particular, he argues that the outstanding balances of individual credit card borrowers serve as a switching cost. Borrowers with higher credit card balances will have less of an ability to switch to alternative credit cards. This is because higher card balances are correlated with consumer characteristics that make consumers reluctant to switch cards (as also argued by Calem and Mester (1995)). Furthermore, lenders are less likely to provide new credit cards to potential borrowers if they have high balances, which will also increase the switching costs of such borrowers who would like to switch. He also argues that switching costs will be positively related to the levels of indebtedness at competitor banks – because this will also reduce the ability of borrowers to switch banks. In order to test this hypothesis, Stango (2002) regresses credit card interest rates of different banks against the outstanding balances of card consumers at those banks. His results show that the higher the outstanding balances of a provider, as well as the higher outstanding balances of competitor banks, the higher the interest rate charged. He interprets these findings to conclude that switching costs can explain as much as a quarter of within-firm variation of credit card interest rates of different banks.

Berlin and Mester (2004) examine in more detail the hypothesis that search costs are an element in explaining the high levels of credit card interest rates. The key empirical tool they use is to examine the distribution of different credit card interest rates across a large sample rather than any one bank specific interest rate. They base their empirical test on a model of search where buyers are divided into searchers and non-searchers. Non-searchers select a firm at random, while searchers will examine the prices from a sub-sample of firms and then select the firm with the lowest price from that sample. The three possible outcomes that are consistent with the searching hypothesis are either a (1) competitive outcome where all firms charge the competitive price, or alternatively (2) a mass point equilibrium (where a fraction of firms charge the competitive price and the rest charge a higher price) and finally (3) a no mass point equilibrium in which the distribution of prices is continuous between the monopoly price and a price above the competitive price.

The intuition from this model is that a competitive equilibrium will occur when there are many searchers, a mass point equilibrium will occur when there is a smaller fraction of searchers, while non-mass point equilibrium will occur when there are only very few searchers who sample a small sample of firms. The main empirical result of Berlin and Mester (2004) is that their data are not consistent with any of these possible searcher/non-searcher equilibria. They find for example, that higher priced firms do not have lower sales, thus it does not appear that consumers are searching for lower priced firms. Furthermore, the search models imply that the there should be a positive relationship between the lowest rate charged in the distribution (which reflects the competitive rate) and the marginal cost of funds. Berlin and Mester on the other hand find a significantly negative relationship between the lowest interest rate in the distribution (as measured in a variety of ways) and the cost of funds.

2.1.2. Rational or irrational consumers

While the above papers all focus on issues of asymmetric information, another strand of the literature has examined the credit card pricing issue built around issues of consumer rationality and irrationality.

Brito and Hartley (1995) propose an alternative explanation to that of Ausubel (1991) and Calem and Mester (1995) and others as to why card interest rates are sticky downward, which they claim is consistent with both competitive equilibrium and based on completely rational individuals. These consumers will continue to use credit card debt even though it is at a significantly higher interest rate than other forms of bank debt. The basis for credit card use in their model is the attempt by consumers to smooth their income and consumption streams over time, and also the very high transactions costs involved in accessing non-
credit card bank credit for small periods of time. Brito and Hartley (1995) argue that even small transactions costs of acquiring non-credit card bank credit (even at low interest rates) would make it rational for consumers to use credit cards for transactions purposes.

Brito and Hartley (1995) also provide an explanation for the significant spread between credit card rates and bank loan rates. They argue that credit cards allow borrowers to borrow if and when consumption exceeds income, while many bank loans are costly to set up, in particular when the periods for which credit is required are relatively short or are unpredictable. Given these advantages in credit card debt over bank loan debt, they argue that a significant spread between credit card and bank loan interest rates is not inconsistent with an equilibrium model. The implication of their argument is that a borrower may not switch from a credit card to a bank loan simply because the bank loan is cheaper than the credit card, because of the advantages of credit card debt over some bank loan debt. However, if the spread between credit card and alternative forms of debt widens enough, the benefits of using credit card debt will decline and more low risk consumers who have access to alternative sources of credit will stop using credit card debt.

Della Vigna and Malmendier (2004) examine the same issue as much of the credit card literature – credit card pricing – but again provide a very different explanation of its determinants. The basic premise of their argument is that firms are rational, but consumers are to some extent biased and irrational. Thus their basic assumption on consumer rationality is very different from that of Brito and Hartley (1995). Specifically, they attempt to model how rational firms price products such as credit cards if they assume that consumers have time inconsistent preferences (e.g. they have higher discount rates between the current and first future period, compared to periods in the future). They consider both agents who are aware of this time inconsistency (sophisticated) as well as agents who are naïve about it.

The main theoretical predictions of Della Vigna and Malmendier (2004) are that for products like credit cards where there are immediate benefits (from consumption) and delayed costs (when debt becomes due), the time inconsistency of borrowers implies that firms should charge below marginal costs in the short run (e.g. credit card teaser rates) but above marginal costs in the long run. This is in order to exploit the fact that consumers have lower discount rates into the future, which implies that naïve consumers will underestimate how much of their credit lines they will use in the future. Della Vigna and Malmendier (2004) argue that their prediction of longer term pricing above marginal cost applies to both naïve borrowers (who underestimate future credit card usage) as well as sophisticated borrowers (who are aware that they have time inconsistency). For these sophisticated borrowers, Della Vigna and Malmendier (2004) argue that the higher marginal costs are essentially “commitment devices” which act to limit the usage of credit card debt by sophisticated borrowers. These authors back up this conclusion by noting the empirical regularity that a large number of credit card holders do not make use of any interest bearing credit card debt.

Della Vigna and Malmendier (2004) also extend their model to incorporate a key element of credit card pricing – that of switching costs. Using a three period model, they show that firms have an incentive to impose switching costs on consumers because of the time inconsistency of consumers. The key intuition is that if naïve consumers underestimate the probability of renewal of debt contracts (i.e. continuing to utilize debt into the future) then firms have an incentive to impose switching costs in order to capture such borrowers. Della Vigna and Malmendier (2004) argue that a variety of institutional facts about the credit card industry are consistent with their theory. While many credit card providers introduce “teaser rates” in order to attract naïve borrowers, the renewal of these accounts after the switch to higher than marginal cost interest rates is usually automatic. This backloaded pricing structure is consistent with naïve borrowers underestimating the possibility that they will renew their debt contracts after the expiry of the initial period.

2.1.3. Fixed and variable interest rates

Stango (2000) provides yet another possible explanation for the puzzle of why credit card interest rates in much of the 1990s where flat and sticky downward. His explanation is based on the fact that there are two different kinds of interest rate usually charged by credit card providers – fixed rates which remain unchanged for long periods and variable rates which move simultaneously with market interest rates. The main question asked by Stango (2000) is whether this pricing structure affects the competitive structure of the market. Using a simple game theory model, he shows that the size of firms will have a strong impact on the pricing structure chosen (i.e. fixed or variable rates). He shows that smaller firms will price more aggressively to capture market share, while larger firms will price less aggressively in order to exploit their existing market share. He provides support for this conclusion with empirical results which show that market share is a significant variable in explaining the interest rate margin (card rate minus market rate) of card providers.

2.1.4. Tacit collusion

Knittel and Stango (2003) provide yet another possible alternative for explaining credit card interest rates based on possible tacit collusion between banks. In the 1980s many States in the US imposed State level price ceilings based on Usury Laws. In some States however, there was no binding ceilings. Using this institutional detail, Knittel and Stango (2003) examine if credit card providers resort to tacit collusion by using the interest rate level of the binding ceiling in some states as the focal point for interest rates in those states that did not have binding ceilings. They base
their discussion of focal point theory on the game theory results of Thomas Shelling and others. Their empirical approach is to develop testable implications regarding the pattern of pricing that would occur with or without focal points being used by firms to tacitly collude. They predict that if tacit collusion was occurring then they would observe greater clustering around the focal point than would otherwise be expected. The focal point is defined as a price ceiling that is not binding in that particular state. They present results which are consistent with their hypothesis that the reason for the clustering and stickiness of credit card interest rates is because of tacit collusion between card providers. It is important to note however that their empirical conclusions are based only on data up to 1989.

2.1.5. Option values

Park (2004) also provides an explanation for the high level of credit card interest rates, but his explanation is based on a discussion of the option value of credit lines. His main argument is that there is an option value inherent in the open credit lines of credit card borrowers, because they can keep on borrowing on their existing credit cards even as their degree of risk changes. Adverse selection also occurs because the cardholders will borrow more when they become riskier. Park (2004) derives the open value of these credit card credit lines in situations where the card borrower is getting riskier over time. He argues that it is for this reason that credit card interest rates are so much higher than the zero profit interest rate.

Park (2004) provides a model where the borrower is better informed about whether they will become riskier in the future than the card provider, and the expected gain of such a borrower obtaining a credit card rises with the increasing probability of becoming riskier. Park (2004) examines various pricing strategies available to card providers to offset this option value. He argues that if a bank unilaterally lowers its credit card rate it will attract those risky borrowers who are likely to become riskier. Furthermore, an upfront fee charged by the bank at the initiation of the credit contract will also result in increased adverse selection issues from those borrowers who become riskier at a later date. According to the Park (2004) model, when card issuers are well informed about current but not future risk levels, the use of “teaser” rates (which start off low and then increase) will best serve to mitigate the adverse selection problem. Thus, he argues that the option value of open credit lines explanation he offers is also an explanation for the increased popularity of teaser rates in the credit card market.

2.1.6. Credit card fees and risk

While much of the focus on the credit card pricing in the literature has been on the determinants of interest rates, Massoud et al. (2007) examine, theoretically and empirically, the determinants of credit card penalty fees (i.e. late fees and overlimit fees). As yet, this is the only paper in the literature to examine this topic, even though the issue of credit card penalty fees has become very prominent in public policy debates. For example, as part of his 2004 Presidential campaign, Senator John Kerry called for these fees to be regulated because they were “abusive”. The American Bankers Association (ABA) responded that such penalty fees are charged by banks to compensate for consumer default risk.

In order to examine the impact of risk on credit card penalty fees, Massoud et al. (2007) examine (1) bank level risk of credit card default as measured by the chargeoff ratio from each bank’s balance sheet, (from the FDIC), and (2) an exogenous measure of default risk as measured by bankruptcies per capita in the specific states where each card is marketed. Overall, they find that card penalty fees do reflect the risk of consumer default and that these penalty fees are negatively related to card interest rates. However, they also find evidence that banks with a large credit card market share charge higher penalty fees than those with a lower market share.

2.2. Examining consumer behavior using credit card data

In the credit card literature reviewed in Section 2.1, the issue has been to explain issues of credit card pricing – such as credit card interest rates and penalty fees etc. A different strand of the credit card literature has been to use evidence from credit card markets to specifically test issues of consumer behavior, such as personal bankruptcy decisions and the permanent income hypothesis etc. These issues are discussed in this section.

2.2.1. Personal bankruptcy

Gross and Souleles (2002a) use a very large dataset of individual credit card accounts to examine the issues of credit card delinquency and personal bankruptcy. Gross and Souleles (2002a) try to explain why personal bankruptcy filings in the United States rose by about 75% between 1994 and 1997. There are two possible explanations for this. The first is that the risk of borrowers may have deteriorated, because, for example, of the increased availability of credit cards. Gross and Souleles (2002a) call this the “risk effect”. The second, alternative, explanation concerns the cost of default, including social, information and legal costs. This second argument is based on the social stigma associated with declaring bankruptcy which can be both non-monetary (e.g. disgrace) as well as monetary (e.g. the future costs of a bad reputation). The hypothesis being examined is that if these costs are declining then this can explain why bankruptcy filings are increasing. The declining costs of declaring bankruptcy can also be explained by the increased supply of bankruptcy lawyers as well as the increased availability of “how to” bankruptcy books. Gross and Souleles (2002a) label this the “demand effect”, because it examines the demand for bankruptcy after controlling for risk characteristics.
Gross and Souleles (2002a) use a new data set of individual credit card accounts in order to disentangle the risk and demand effects of increased bankruptcy and delinquency. A key advantage of their data is that it includes detailed data on the credit risk of each individual consumer, including both credit risk scores as well as information on other assets and liabilities in the consumer’s portfolio. The authors estimate duration models to evaluate the relative importance of alternative variables in explaining delinquency and default. Their key result concerns how the estimated relationship between risk (as measured by consumer level variables) and bankruptcy has changed significantly over time. Even after they control for consumer level risk factors, they find that the propensity to default increased significantly in a relatively short time. While Gross and Souleles (2002a) are able to claim that consumer risk is not the key explanation for the significant increase in bankruptcy filings, they are not able to specifically show which if any of the various “demand effects” (e.g. lower social stigma of bankruptcy etc.) are significant because of lack of data.

2.2.2. Credit cards and the permanent income hypothesis

While the above paper examines why consumers choose to default or become bankrupt, a companion paper by Gross and Souleles (2002b) examines the impact that increases in credit supply (as measured by increases in credit limits or credit lines) impacts consumer borrowing behavior. This issue is of considerable importance in light of predictions of the permanent income hypothesis, which serves as a cornerstone of much of macroeconomics. The permanent income hypothesis predicts that if consumers receive an increase in the supply of credit – i.e. an increase in their credit card line of credit, then they will not increase their demand for credit – i.e. their debt as measured by their credit card balance. If however, consumers increase their credit card debt after they receive an increase in their credit card credit limits, then this implies that they have been liquidity constrained, and that the predictions of the permanent income hypothesis are violated.

In their results, Gross and Souleles (2002b) find that after one year, each extra $1000 of liquidity (higher credit limit) results in approximately $130 increase in debt – a result that is highly significant. These results are robust to a variety of tests to ensure that causality runs from credit supply to credit demand, and not vice versa. Furthermore, Gross and Souleles (2002b) also find that the interest elasticity of debt is highly significant. A 1% increase in the interest rate results in a $110 deduction in debt (as measured by the credit card balance) after nine months.

2.3. The pricing of the network interchange fee

While much of the credit card pricing literature has focused on examining credit card prices and consumer behavior, another key element of credit card pricing is the determination of the network interchange fee. In any credit card transaction, the bank that issued the card to the consumer is known as the issuer, while the bank that process the transaction on behalf of the merchant is called the acquirer. When these two banks are different, the acquirer pays the issuer an interchange fee. The fee paid by the merchant to the acquiring bank to process the transaction is called the merchant discount.

An important public policy issue is that the level of the interchange fee is determined collectively by all the banks that own the network. Because these banks can be assumed to be acting as profit maximizers when setting the interchange fee, it has been suggested that the collective setting of the interchange should be treated as collusion in a cartel like setting, and thus regulated.

Schmalensee (2002) analyses the issue of whether the credit card interchange fee should be considered as anti-competitive price fixing. His model is applicable to credit card networks as well as debit card and ATM networks described below. The key element of his model is that the behavior of issuer banks affects the behavior of acquirer banks (and vice versa) in determining the value of the payment system to both sets of banks. He argues that this network externality can be determined at the level of the whole system by using the interchange fee. The main argument of his model is that the interchange fee does not act to exploit the market power of all the banks in the system, but rather to shift the costs from the bank who issued the credit card, to the bank who acquired the credit card transaction. By doing this the interchange fee acts to increase the value of the network of all the banks who own the network.

Schmalensee (2002) argues that a remarkable aspect of his model is that under imperfect competition, the privately optimal interchange fee set by the banks also maximizes social welfare (as conventionally defined). The main argument in his model is that the interchange fee is not used to increase profit by reducing output – which is the essential element in most instances of price fixing. Rather, his model predicts that under most assumptions, the interchange fee actually maximizes output in order to maximize the systems benefit to the banks (its owners).

Another theoretical examination of the interchange fee is provided by Rochet and Tirole (2002). In their model both banks and merchants have market power. The main finding of their model is that an increase in the interchange fee increases the consumers use of credit cards, a long as the interchange fee does not rise above a certain threshold. This threshold is where merchants no longer have an incentive to accept credit cards because the costs to the merchant are equal to or greater than the benefit to the consumer.

An important element of the Rochet and Tirole (2002) model is the concept of merchant resistance to paying the interchange fee. Rochet and Tirole (2002) argue that the key reason why merchants agree to pay the interchange fee is a desire to “obtain a competitive edge” over other merchants. Because of this inter-merchant competition to obtain consumers, merchants will accept the interchange fee involved with credit cards, even though the cost exceeds
the technological and payment guarantee benefits that are associated with accepting credit cards. In other words, merchant resistance to accepting credit cards and paying the interchange fee will be lowered if consumers no longer purchase from the consumer if credit cards are not accepted.

Wright (2003) extends the models of Schmalensee (2002) and Rochet and Tirole (2002), by examining the optimal interchange fee under different sets of assumptions. He concludes that the socially optimal interchange fee occurs when the interchange fee equals the average transactions benefits obtained by the merchants that accept the credit cards. While this fee structure favors cardholders over merchants, this optimal fee does not distort retail prices.

3. Automated teller machines (ATMs)

Automatic teller machines (ATMs) began to be introduced in the 1970s with the aim of lowering bank costs. Their introduction has generated a large and growing research literature examining for example the role of technology on banking, the impact of network effects as well as the impacts of ATMs on bank competition.

3.1. ATMs in a network

A key element of ATMs is their role as part of a network. Saloner and Shepard (1995) use data from the introductions of ATMs to test various predictions from the networks economics literature. An important prediction of the network literature is that the value of a network to its users increases, as the number of locations in the network increases which is called the “network effect”. A related prediction is that the value of a network will increase as the number of users increase (called the “production scale effect”).

Saloner and Shepard (1995) argue that the greater the geographic dispersion of ATMs the greater the benefits to bank card holders, who are able to access ATMs for their banks in a wide variety of locations (i.e. the network effect). In other words, a bank can increase the value of its network by increasing the size of its ATM network.

An implication of the network effect hypothesis is that a bank that expects to have a larger ATM network in equilibrium (in order to benefit from network effects) will begin to implement the new ATM technology sooner. This is the format of the network hypothesis tested by Saloner and Shepard (1995) using data on the initial adoption of ATMs by banks in the 1970s, when the technology was just beginning to be introduced. As a proxy for the unobservable expected ATM network size in equilibrium, the authors use the current number of branches owned by banks, based on the argument that many ATMs are initially located inside branches. In other words, they argue that the network effect predicts that the greater the number of branches a bank has, the sooner it will have initially introduced the ATM technology.

Saloner and Shepard (1995) also test the production scale hypothesis (which states that the value of a network increases with more users). They argue that the greater the number of bank depositors (i.e. users) the greater the value of having an ATM network, and by implication the earlier they are likely to introduce the ATM technology.

The key finding of Saloner and Shepard (1995) is that the greater the number of branches a bank has, the earlier will have been its introduction of the new ATM technology in the 1970s. This they argue is consistent with the network effect hypothesis, which states that the greater the number of users of a network the greater will be the value of a network. In terms of the magnitudes involved, they find that a single additional branch that a bank owns would increase the probability of adopting ATMs in the first nine years of their availability by between 5.4% and 10.2%. Similarly, if a bank increased the number of its depositors by the average number of depositors at a branch, would increase the probability of ATM adoption by 4.7%. This evidence, they argue, strongly supports the network effect hypothesis.

While, one of the main empirical assumptions of Saloner and Shepard (1995) is to ignore the fact that banks can share ATM networks, a theoretical paper by Matutes and Padilla (1994) focuses specifically on the implications of such network sharing. Indeed the grouping of individual banks into larger networks of ATMs providers is one of the key empirical regularities in describing the way ATMs are organized.

The model developed by Matutes and Padilla (1994) examines the tradeoff that banks face between cooperating with their competitors in providing larger ATM networks, while at the same time competing with these same ATM partners for deposits etc. In the model they argue that depositors will be willing to accept lower deposit rates in exchange for having access to larger and thus more convenient ATM networks (network effects). In this environment though, consumers may be able to benefit from higher deposit rates paid by other banks in the ATM network, while still accessing the whole (large) ATM network. Matutes and Padilla (1994) label as the substitution effect, the fact that banks as part of the same network can be substituted for each other by consumers. These two effects are contradictory because while the network effect provides an incentive for banks to join large ATM networks, the substitution effect creates a rivalry between banks based on price (interest rates on deposits), which will provide incentives against joining ATM networks.

The main result of their model is that either a subset of all banks will share an ATM network, or no banks will join a network. The reason why there model does not predict the emergence of full compatibility (i.e. all banks in a single ATM network) is that with full compatibility no individual bank would be able to extract greater rewards from network externalities compared to its competitors. Furthermore, if all banks joined a single network then all would become substitutes for each other, resulting in fiercer price competition and lower profits for all banks.
3.2. Pricing of ATM services

While the paper by Matutes and Padilla (1994) examines theoretically the determinants of the quantities of ATMs within networks, Massoud and Bernhardt (2002) provide a theoretical model to explain the observed pricing of ATM services. There is much evidence showing that while banks set very high prices (ATM fees) for the use of ATMs by foreign consumers (i.e. consumers belonging to non-member banks), banks usually allow their own members to use ATMs for free. On the other hand banks charge their own members very high what are called non-discretionary charges (i.e. where customers paying for services only there own bank can provide, e.g. account fees).

In order to explain these stylized facts, Massoud and Bernhardt (2002) develop a special model, where both the pricing of ATM and other services by banks as well as the choices of banks by consumers are endogenized. In equilibrium banks price discriminate between their members and non-members for ATM services. They charge members the marginal cost of ATM services, but extract all their profits from non-discretionary services paid by members and ATM services paid by non-members.

Massoud and Bernhardt (2002) introduce the argument that high ATM fees charged to foreigners can have both direct as well as indirect effects on bank revenues. The direct effect is simply the increased revenue the bank receives from fees paid by foreign borrowers. However there is also an important strategic element in the banks behavior. A customer may have an incentive to switch to a bank because that bank has a higher ATM fee. This is because the only customers who pay the ATM fee are those who are not members of a bank. Thus by joining a bank a customer will have free access to that banks ATM network. This becomes particularly valuable to a customer when the bank has a large number of ATMs. For this reason, Massoud and Bernhardt (2002) argue that the equilibrium level of ATM surcharges charged to non-members exceed the price that would maximize expected profits that flow from the direct effect (i.e. directly from ATM fees paid). This implies that the indirect effect (i.e. revenues flowing from customers switching banks) is also an important revenue generator for banks when determining their ATM surcharges.

Two papers have attempted to empirically test some of the theoretical predictions of Massoud and Bernhardt (2002). These are Hannan et al. (2003) and Massoud and Bernhardt (2006). The paper by Hannan et al. (2003) attempts to provide empirical evidence on which banks impose ATM surcharges and which do not, thus they use a probit type methodology. Their data is taken from 1997, when a large number of banks had not yet imposed surcharges. An important dependent variable in their regressions is the individual banks share of ATMs in the local banking market. They argue that the market share of ATMs should impact a bank’s choice about whether to impose ATM surcharges through both the direct as well as the indirect effects. In terms of the direct effect, they argue that a bank with a large market share of ATMs will be able to extract higher ATM fees from non-members simply because their ATMs will be located at a greater number of locations, which implies greater convenience for non-members who require ATM services. The argument that larger market share of ATMs results in higher ATM surcharges also holds for the indirect effect. Banks with more ATMs are more likely to induce customers to switch to them as ATM surcharges rise, because those switching customers will have the benefit of free access to a larger ATM network, that is increasingly expensive to non-members.

The main findings of Hannan et al. (2003) are that firms with higher market shares are more likely to charge ATM surcharges – as predicted by both the direct as well as the indirect effects. In terms of their specific tests of the indirect effects they find that their measures of in-migration are indeed significant. However they find no evidence in support of the hypothesis that large banks are able to extract rent by raising ATM surcharges relative to smaller banks.

Massoud et al. (2006) also develop testable hypotheses from the existing theoretical literature on ATMs, in particular the models of Massoud and Bernhardt (2002) and Massoud and Bernhardt (2006). These papers consider the possibility that ATM surcharges can impact banks profitability, both directly as well as indirectly through a so-called customer switching effect. The direct effect on profitability stems from the direct revenue generated from “foreigners” (i.e. non-bank customers) paying higher ATM surcharges. The indirect effect results from customers at smaller banks with relatively few ATMs switching their deposit accounts to larger banks with a larger number of ATMs in order to avoid paying ATM surcharges. The higher the surcharge at larger banks the greater incentive customers have to switch to that bank in order to avoid paying the higher surcharges. If switching occurs from smaller to larger banks, then higher ATM surcharges should enhance the market share of bank products (e.g. deposits) of larger banks and reduce the market share of deposits of smaller banks.

The aim of Massoud et al. (2006) is to empirically examine whether the ability to impose ATM surcharges has indeed benefited larger banks and hindered smaller banks. They use a panel database, with bank level data on ATM surcharges, the size of ATM networks and the percentage of foreign (i.e. non-bank) users of ATMs who have to pay the surcharge. Using this database they examine the impact that ATM surcharges and ATM network size have.
on the market share of deposits of larger and smaller banks.

Massoud et al. (2006) find that the level of ATM surcharges is positively related to the market share of deposits of larger banks in the following year. On the other hand they find that for smaller banks the level of the surcharge is negatively related to their market share in the following year. This is consistent with bank customers switching accounts from smaller to larger banks to avoid paying surcharges at larger banks. However they also find that smaller banks can impact their market share of deposits in the following year by adopting a larger ATM network. Their results thus indicate that larger banks (who already have large ATM networks) are able to generate a larger market share of deposits by setting ATM surcharges at a higher level, while smaller banks are only able to influence their market share of deposits by establishing larger ATM networks.

Prager (2001) also investigates the issue of whether ATM surcharges negatively impact smaller banks. She examines State level data from 1987 to 1995, when only some states allowed ATM surcharges and some did not. She compares the market share of small banks in those states that allowed surcharging with those states that did no allow surcharging. Her main finding is that in this period there was no statistical difference between the market share of small banks operating in the States where surcharging was allowed, compared to small banks operating in states were ATM surcharges was prohibited.

There are two possible reasons why the findings of Massoud et al. (2006) are different from those of Prager (2001). The first is that Massoud et al. (2006) examine data from 1996 to 2001, while Prager examines data from 1987 to 1995. The second is that Massoud et al. (2006) examine bank specific data, while Prager (2001) examines data for markets or regions (MSAs etc.).

### 3.3. ATM networks and antitrust

When Schmalensee (2002), (described above), examined whether the collective setting by banks of the credit card interchange fee amounted to price fixing or collusion, he concluded that market structure had no impact on that particular element of the credit card network. McAndrews and Rob (1996) also examine the impact of market power on networks, but in this case examine ATM networks and ATM pricing.

They examine the situation where the ATM “network switch” is jointly owned by a large number of members of the network. The puzzle that they analyze is why banks that compete at the downstream (retail) level, cooperate at the upstream (wholesale) level of the ATM switches (i.e. the physical infrastructure that makes up an ATM network). A related issue that they raise is that empirically, they observe that there is a correlation between the amount of joint ownership (i.e. a collection of banks own an ATM network) and monopoly power in ATM networks. They argue that in many areas of the US, joint ownership of the ATM network is usually correlated with monopoly power of the network (i.e. there is only one network or each network has a substantial market share).

The key public policy issue involved here is whether the existence of ATM networks with monopoly power need to be regulated. The existing response of the US authorities has been that because the ATM networks are jointly owned by a large number of banks mitigates the concern over the fact that the ATM networks have monopoly power, thus reducing the need for antitrust regulation. McAndrews and Rob (1996) provide an analysis that questions this conclusion.

They provide a model which shows that the view that joint ownership is benign in an antitrust context, only focuses on transactions between the ATM network and the banks (i.e. the upstream relationship). They argue however, that this view ignores the downstream relationship between networks and consumers. The key issue driving this result is the existence of network externalities, which can be more fully exploited in concentrated ATM markets. Thus a system of joint ownership of a single ATM network, rather than curtailing market power, will actually exacerbate it because of the existence of network externalities. The policy conclusion that follows from this argument is that the antitrust authorities should weigh the network benefits versus the costs of monopoly power when evaluating jointly owned ATM networks, rather than simply assuming that assuming that joint ownership is an adequate protection against monopoly power as is currently the case.

While McAndrews and Rob (1996) examine the issue of ATMs and antitrust regulation using a theoretical model, Prager (1999) examines the issue empirically. Specifically, she uses data to examine the impact of mergers in ATM networks on the prices and quantities of ATM transactions. She argues that if ATM mergers result in cost savings (as argued by the banks) then prices would be expected to decline and/or quantities of transactions would be expected to increase. On the other hand, the opposite would happen if the ATM mergers created additional market power.

Prager (1999) uses data from surveys of network prices over time including data on transaction volume, switch fees and interchange fees for both merging and non-merging networks. Her key finding is that she could not find any significant differences in both ATM prices and quantities between merging and non-merging networks. These findings imply that either the market power and efficiency impacts from ATM mergers were offsetting or they were both very small. The policy implication of this is that consumers were not harmed by the ATM mergers in the study.

### 3.4. ATMs as new technology

The paper by Hannan and McDowell (1990) examines ATMs in the context of the introduction of new technologies by banks. Hannan and McDowell (1990) argue that
while much of the literature has examined how market structure can impact new technologies, the introduction of ATMs by the banks is a case of the opposite process – how new technologies can in turn impact market structure. In particular they examine how the introduction of ATMs impacts the concentration of local banking markets. They argue that the introduction of ATMs will impact market structure if there are differences in adoption behavior by larger relative to smaller firms (or vice versa). If for example, ATMs are mostly introduced by larger firms, and this increases the market share of those firms, then this will tend to increase concentration in individual banking markets. On the other hand, if smaller firms are able to increase their market share by increasing ATMs then this will tend to reduce concentration.

The main findings of Hannan and McDowell (1990) is that banks have been able to attract customers by adopting ATM technologies. However, they could not detect any significant impact of ATMs on market concentration, which implies that large banks and smaller banks have tended to implement ATM technologies on a proportionately equal basis. They do find however that if larger firms are more efficient than smaller firms, then the level of concentration tends to go up.

4. Debit cards (electronic fund transfer at point of sale)

This section examines the literature on electronic fund transfer at point of sale (EFTPOS) networks, which are also sometime referred to as Debit Card networks. While there has been a very large increase in these networks, the academic research on this topic still remains relatively small. EFTPOS networks are facilities provided by merchants for consumers to directly withdraw cash at the point of sale in order to pay the merchant for goods. This means that consumers do not have to access cash though ATMs or use their credit cards. We argue here however that some of the issues involved in EFTPOS networks are just as complex and interesting as those found in credit cards and ATMs. For this reason we argue that there are many opportunities for research in this area.

4.1. The choice between debit cards and ATM cash

An important recent paper on EFTPOS networks is that of Markose and Loke (2003). They use a Nash game to model the relationship between a consumer using cash withdrawn from ATM machines and debit cards used at the point of sale (EFTPOS). The aim of this paper is to attempt to explain why the use of cash has been declining in many countries, and is being replaced by the use of cards at EFTPOS networks. The central question of their theory is under what conditions a consumer will use cash and under what conditions a consumer will use a debit card. The develop a Nash game where the consumer makes predictions on the proportion of merchants in the economy who are linked to the EFTPOS system. The greater the proportion of merchants linked to EFTPOS the lower will be their demand for cash from ATM networks. On the other hand, the probability of the merchant installing the EFTPOS system depends on the proportion of retail consumption that will be paid for using this system. The authors develop a Nash equilibrium where both cash (from ATMs) as well as debit card use both coexist.

The problem of the consumer in deciding between ATM cash and EFTPOS debit cards, involves issues such as the costs of using cash from an ATM (which include both shoe leather costs of finding an ATM as well as the costs of using an ATM from another bank network); the interest forgone from holding cash as well as the probability that a merchant will accept EFTPOS transactions. The implication of this model is that if all the merchants in the economy provide EFTPOS service, then the system will produce a result with zero ATM cash usage. A key finding of the model is that a corner solution where there is complete EFTPOS coverage as well as zero cash use will exist if the interest rate is above zero (e.g. 2%). At this point the incentives for consumers to economize on cash ceases.

In terms of the policy implications of this paper, Markose and Loke (2003) argue that their model can explain why money demand functions began to break down in the late 1970s as new technologies such as EFTPOS and ATMs began to be introduced. Most money demand models only used interest rates and income to measure the consumers demand for cash and did not consider the fact that the existence of EFTPOS could significantly reduce the demand for cash.

4.2. Empirical estimates of consumer payment choices

Humphrey et al., 1996 attempt to estimate how consumer payments break down between different payment mechanisms in different countries. They examine a variety of types of payments including cash, checks, other paper based payments as well as electronic payments which include both credit and debit cards that are used at the point of sale (EFTPOS). They also attempt to explain the determinants of these choices between different countries. The data they use runs from 1987 to 1993.

The main conclusion of Humphrey et al., 1996 is that over time and across countries electronic forms of payment are rising significantly because they are generally cheaper than paper forms. Humphrey et al., 1996 also attempt to explain why the mix of payment systems is so different across different countries. The main factors influencing this mix are factors that reflect the availability of different options to consumers as well as institutional and cultural differences between countries. Cultural and institutional factors are particularly strong at explaining differences across countries. For example the use of non-cash payment systems are related to per capita income, the availability of new payment systems and also the prevalence of violent crime within countries.
One of their key findings is that the United States differs from many of the other developing countries when it comes to the use of electronic vs. non-electronic forms of payment. They find that Japan has high cash holdings per person, low non-cash use, but a high percentage of electronic payments. The United States on the other hand has low cash holdings, high non-cash use and a low percentage of electronic payments. Humphrey et al., 1996 argue that electronic payments (including debit and credit cards at point of sale) in the US until 1993 has been relatively slow in the US compared to other countries. The reasons they provide for this include the banks reliance on credit cards for loan revenue as well as the US historical dependence on using checks for payments.

Humphrey (2004) updates this study by examining how cards have replaced cash in the US. He finds that the share of cash in consumer payments fell by a third from 1974 to 2000, while at the same time the share of cards in consumer payments doubled. By 2000, the share of cards was actually greater (27%) than the share of cash (20%). The rest was made up of checks etc.

An important element in examining consumer choices of payment system is to estimate the impact that prices have on consumer choices between payment alternatives. Humphrey et al. (2001) attempt to do this, using data from Norway. The aim of their paper is to examine whether differences in pricing between paper means of transactions and electronic means of transactions will provide an incentive for consumers to shift away from paper towards electronic payments. Humphrey et al. (2001) argue that Norway provides a good case study, because there has been a deliberate attempt to increase the prices of paper transactions relative to electronic transactions in order to encourage consumers to shift to more efficient electronic transactions.

Humphrey et al. (2001) estimate the consumer’s demand for three payment choices – checks, ATMs and EFTPOS. In their estimation of the consumer’s demand they applied the Christensen et al. (1975) approximation logarithm technique to estimate the indirect utility function. To do this they invoked the assumption of separability between POS1 and other uses of the three payment instruments. Their data source is a survey on the quantity and price for the three payment transaction for Norwegian saving banks and commercial banks over the period 1989–1995. The data is aggregated into two groups, saving banks and commercial banks. The total sample size is 26 observations.

They find evidence of net substitution between ATMs and checks (both ways). However, they find a one way substitution from checks to POS in response to partial increase in the check fees. On the other hand, they found no evidence of substitution between ATMs and EFTPOS.

5. New estimates on the substitution between ATMs and EFTPOS

In this paper we provide new estimates on the question posed by Humphrey et al. (2000) by examining the substitution between ATMs and EFTPOS, using a new database consisting of detailed bank level data on ATM and EFTPOS transactions. While the data used by Humphrey et al. (2001) is aggregate country level data, our data is bank level data taken from Spain. Our new database consists of 1242 panel data observations of bank specific data from Spain (compared to the 26 observations used by Humphrey et al. (2000)). Our data contains both prices and quantities of EFTPOS and ATM transactions. Furthermore, our data enables us to distinguish between transactions by consumers who are members of the ATM network. Our large and detailed database allows us to use regression based models that control for endogeneity (i.e. 3SLS) in our estimates. In particular, we estimate the substitution between ATMs and EFTPOS in a model that allowed us to endogenize the strategic choices of consumers demand for the two types of payments (ATM and EFTPOS) and banks strategic choice of prices and quantities of ATMs and branches.

In Spain customers can either withdraw cash directly from an own-bank ATM (where there are no ATM fees), a foreign bank ATM (i.e. a bank where a consumer is not a member and thus has to pay ATM surcharges) or purchase a commodity directly in a store using point of sale machine (EFTPOS) technology. Humphrey et al. (2001) argue that Norway provides a good case study, because there has been a deliberate attempt to increase the prices of paper transactions relative to electronic transactions in order to encourage consumers to shift to more efficient electronic transactions. In addition, EFTPOS allows customers to withdraw “cash back” up to specific limits. From the consumer’s perspective, however, there is an asymmetry in the way ATM fees and EFTPOS fees are charged. Customers pay ATM surcharges for the use of a foreign ATM while such a fee is absent if a EFTPOS machine is used. Accordingly, a customers’ cost of using foreign ATMs is much higher (non-zero) than that for foreign non-bank EFTPOS. Humphrey et al. (2000) estimated the consumer's demand for the two types of payment using a new database consisting of detailed bank level data on ATM and EFTPOS transactions. While the data used by Humphrey et al. (2000) is aggregate country level data, our data is bank level data taken from Spain. Our new database consists of 1242 panel data observations of bank specific data from Spain (compared to the 26 observations used by Humphrey et al. (2000)).

5.1. Hypothesis

We test two hypotheses:

H1: EFTPOS transactions and ATMs transactions are substitutes.

H2: There is a cross substitution between the ATM surcharge and EFTPOS transaction volume.

5.2. Data and variable definition

We use proprietary data provided by the Spanish Confederation of Savings Banks (CECA). The database contains quarterly bank level information on ATM and

\footnote{POS (Point of Sale) includes both paper based (checks), electronic based (EFTPOS, Electronic Funds Transfer at Point of Sale) and cash payments.}
EFTPOS machines, number and value of transactions and sources of bank income of 46 Spanish savings banks from 1997:1 to 2003:3. In total there are 1,242 panel observations. These institutions correspond to all savings banks operating within the payment systems in Spain. These data are adjusted to reflect mergers over the period. These saving banks represent around 60% of total card payment transactions in Spain. In Spain, the number of ATMs increased from 511 in 1996 to 55,399 in 2004 and the number of EFTPOS increased from 30,437 in 1996 to 1,055,103 in 2004. Detailed variable definitions used in this study are presented in Table 1. Table 2 provides summary statistics for our variables.

5.3. Empirical methodology

To test our main hypotheses (H1 and H2) of consumer’s use of ATMs and EFTPOS we estimate a system of five equations that endogenize the strategic choices made by the banks as well as their customers (members and foreign customers). Here foreign customers are defined as the customers with EFTPOS/ATM cards that belong to a different network. There are four networks in Spain. Banks choose ATM surcharges, and their market share of ATMs. Bank members (i.e. customers) and foreign customers choose their use of ATMs and EFTPOS. This results in the following five endogenous variables: ATM surcharge, Market share of ATMs, ATM transactions of Members, ATM transactions of foreign customers and EFTPOS transactions. The following system of five equations is estimated using three stage least squares (3SLS) with fixed effects to account for bank individual heterogeneity, as well as time fixed effects:

\[
\begin{align*}
\text{ATM surcharge},_i &= \theta_0 + \theta_1, \text{Market share of ATMs},_i \\
&+ \theta_2, \text{ATM transactions members},_i \\
&+ \theta_3, \text{EFTPOS transactions members},_i \\
&+ \theta_4, \text{ATM transactions foreign customers},_i \\
&+ \theta_5, \text{EFTPOS transactions foreign customers},_i \\
&+ \theta_6, C_{i,t} + \mu_i + \varepsilon_{i,t}, \\
\end{align*}
\]

\[
\begin{align*}
\text{Market share of ATMs},_i &= \phi_0 + \phi_1, \text{ATM surcharge},_i \\
&+ \phi_2, \text{ATM transactions members},_i \\
&+ \phi_3, \text{EFTPOS transactions members},_i \\
&+ \phi_4, \text{ATM transactions foreign customers},_i \\
&+ \phi_5, \text{EFTPOS transactions foreign customers},_i \\
&+ \phi_6, C_{i,t} + \kappa_i + \varepsilon_{i,t}, \\
\end{align*}
\]

<table>
<thead>
<tr>
<th>Table 1</th>
</tr>
</thead>
<tbody>
<tr>
<td>Variable</td>
</tr>
<tr>
<td>ATM surcharge</td>
</tr>
<tr>
<td>Market share of ATMs</td>
</tr>
<tr>
<td>ATM transactions Members</td>
</tr>
<tr>
<td>ATM transactions foreign customers</td>
</tr>
<tr>
<td>EFTPOS transactions</td>
</tr>
<tr>
<td>EFTPOS Market share</td>
</tr>
<tr>
<td>Average ATM transaction value for foreign customers</td>
</tr>
<tr>
<td>Average ATM transaction value for own customers</td>
</tr>
<tr>
<td>Average EFTPOS transaction value</td>
</tr>
<tr>
<td>Log (National ATMs)</td>
</tr>
<tr>
<td>Log (National EFTPOS)</td>
</tr>
<tr>
<td>GDP per capita</td>
</tr>
<tr>
<td>Population growth</td>
</tr>
</tbody>
</table>
Table 2
Descriptive statistics

<table>
<thead>
<tr>
<th>Variable</th>
<th>Mean</th>
<th>Std. dev.</th>
</tr>
</thead>
<tbody>
<tr>
<td>ATM surcharge</td>
<td>0.0223</td>
<td>0.0221</td>
</tr>
<tr>
<td>Market share of ATMs</td>
<td>0.0223</td>
<td>0.0398</td>
</tr>
<tr>
<td>ATM transactions for members</td>
<td>9.2234</td>
<td>1.3726</td>
</tr>
<tr>
<td>EFTPOS transactions</td>
<td>15.6880</td>
<td>1.9253</td>
</tr>
<tr>
<td>EFTPOS market share</td>
<td>0.0202</td>
<td>0.0372</td>
</tr>
<tr>
<td>ATM transactions for foreign customers</td>
<td>15.8757</td>
<td>2.0231</td>
</tr>
<tr>
<td>EFTPOS transactions</td>
<td>11.2556</td>
<td>2.3567</td>
</tr>
<tr>
<td>ATM transactions for foreign customers</td>
<td>3.0811</td>
<td>1.2303</td>
</tr>
<tr>
<td>EFTPOS transactions for foreign customers</td>
<td>4.1123</td>
<td>1.0545</td>
</tr>
<tr>
<td>Average ATM transaction value for members (€)</td>
<td>9.2234</td>
<td>0.0725</td>
</tr>
<tr>
<td>Average ATM transaction value for foreign customers (€)</td>
<td>4.3624</td>
<td>1.2306</td>
</tr>
<tr>
<td>Average EFTPOS transaction volume (€)</td>
<td>42.3490</td>
<td>2.6101</td>
</tr>
<tr>
<td>Log (National ATMs)</td>
<td>4.3624</td>
<td>0.0725</td>
</tr>
<tr>
<td>Log (National EFTPOS)</td>
<td>8.0335</td>
<td>0.2829</td>
</tr>
<tr>
<td>GDP per capita (€)</td>
<td>85.2355</td>
<td>5.6328</td>
</tr>
<tr>
<td>Population growth</td>
<td>21458.02</td>
<td>3157.19</td>
</tr>
<tr>
<td>Market share of ATMs</td>
<td>0.0732</td>
<td>0.0652</td>
</tr>
</tbody>
</table>

This table includes descriptive statistics for the variables used in our analysis. These variables are taken from the following data sources: (1). The Spanish Confederation of Savings Banks, (2). Spain’s National Statistical Office (INE). The data covers a period from 1997:1 to 2003:3.

ATM transactions members $i,t$

$$= \psi_0 + \psi_{1,t} \text{ATM surcharge}_{i,t} + \psi_{2,t} \text{Market share of ATMs}_{i,t} + \psi_{3,t} \text{EFTPOS transactions members}_{i,t} + \psi_{4,t} \text{ATM transactions foreign customers}_{i,t} + \psi_{5,t} \text{EFTPOS transactions foreign customers}_{i,t} + \psi_{6,t} \text{C}_{i,t} + \rho_i + \nu_{i,t},$$  

(3)

ATM transactions foreign customers $i,t$

$$= \gamma_0 + \gamma_{1,t} \text{ATM surcharge}_{i,t} + \gamma_{2,t} \text{Market share of ATMs}_{i,t} + \gamma_{3,t} \text{EFTPOS transactions members}_{i,t} + \gamma_{4,t} \text{ATM transactions members}_{i,t} + \gamma_{5,t} \text{EFTPOS transactions foreign customers}_{i,t} + \gamma_{6,t} \text{C}_{i,t} + \delta_i + \omega_{i,t},$$  

(4)

Total EFTPOS transactions members $i,t$

$$= \lambda_0 + \lambda_{1,t} \text{ATM surcharge}_{i,t} + \lambda_{2,t} \text{Market share of ATMs}_{i,t} + \lambda_{3,t} \text{ATM transactions foreign customers}_{i,t} + \lambda_{4,t} \text{ATM transactions members}_{i,t} + \lambda_{5,t} \text{C}_{i,t} + \zeta_i + \xi_{i,t},$$  

(5)

where $C$ is a vector of control variables, $\theta$, $\phi$, $\psi$, $\gamma$ and $\lambda$ are coefficients and $\mu$, $k$, $r$, $\theta$, $\delta$, $\epsilon$, $\nu$, $\omega$ and $\zeta$ are error terms.

In the above system, Eqs. (1) and (2) explain the strategic choices by banks: i.e. ATM surcharges and the market share of ATMs. The last three equation ((3)–(5)) explain the demand of ATM services (by bank customers and foreign customers) as well as the demand of EFTPOS services.

Our control variables ($c$) include the size of ATMs and EFTPOS networks at the national level, the influence of economic development (GDP per capita) and the potential growth in the demand for cards (population growth), average ATM transaction value for members and foreign customers and average EFTPOS transaction value. Our choice of the instrumental variables is based on their relevance to the dependent variables. For example, in the EFTPOS transactions equation (Column iv), our instrument is EFTPOS market share.

5.4. Main results

Table 3 shows the results of estimating the five equations system using 3SLS. From this equation system we can examine the two hypotheses proposed in Section 5.1.

5.4.1. H1: EFTPOS transactions and ATMs transactions are substitutes

Here the key column (equation) in Table 3 are columns (iv) and (v) which show in the context of the five equation model the effects of use of ATM (EFTPOS) transactions on EFTPOS (ATM) transactions. If the two payment mechanisms are substitutes we would expect to see the coefficient on ATM in the EFTPOS transaction being negative. If the coefficient were positive this would be consistent with the two payment mechanisms being complements. As can be seen in column (v) of Table 3, the sign on the volume of ATM transactions in the EFTPOS regression is negative and significant at the 1% level.

This substitution is further confirmed by the sign of the coefficient on EFTPOS transactions for foreign customers in the ATM transaction regression (column (iv)). Since foreign users of EFTPOS are not subject to a fee for their use of a non-bank terminal we would expect an increase in the
Table 3
Tests of consumer’s rationality in the use of ATMs and EFTPOS (1997:1–2003:3)

<table>
<thead>
<tr>
<th>Dependent variable</th>
<th>ATM surcharge</th>
<th>Market share of ATMs</th>
<th>ATM transactions for members</th>
<th>ATM transactions for foreign customers</th>
<th>EFTPOS transactions</th>
</tr>
</thead>
<tbody>
<tr>
<td>(i)</td>
<td>(ii)</td>
<td>(iii)</td>
<td>(iv)</td>
<td>(v)</td>
<td></td>
</tr>
<tr>
<td>Constant</td>
<td>0.35E–14***</td>
<td>–1.6E–15†</td>
<td>0.15E–12</td>
<td>0.26E–13</td>
<td>–0.30E–15</td>
</tr>
<tr>
<td></td>
<td>(0.01)</td>
<td>(0.01)</td>
<td>(0.01)</td>
<td>(0.21)</td>
<td>(0.01)</td>
</tr>
<tr>
<td>ATM surcharge</td>
<td>–</td>
<td>0.011†</td>
<td>–7.70</td>
<td>–15.54†</td>
<td>0.13***</td>
</tr>
<tr>
<td></td>
<td>(0.06)</td>
<td>(7.04)</td>
<td>(10.99)</td>
<td>(0.01)</td>
<td>(0.02)</td>
</tr>
<tr>
<td>Market share of ATMs</td>
<td>61.35***</td>
<td>–</td>
<td>904.06***</td>
<td>1466.86***</td>
<td>–</td>
</tr>
<tr>
<td></td>
<td>(8.95)</td>
<td>(40.15)</td>
<td>(49.61)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Market share of POS</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>0.60</td>
</tr>
<tr>
<td>ATM transactions for members</td>
<td>–0.03***</td>
<td>0.004***</td>
<td>–</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td></td>
<td>(0.03)</td>
<td>(0.01)</td>
<td>(0.01)</td>
<td>(0.01)</td>
<td>(0.79)</td>
</tr>
<tr>
<td>EFTPOS transactions for members</td>
<td>–0.09</td>
<td>0.10435</td>
<td>–1.15</td>
<td>–1.82</td>
<td>–</td>
</tr>
<tr>
<td></td>
<td>(0.56)</td>
<td>(21.41)</td>
<td>(34.17)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>ATM transactions for foreign customers</td>
<td>–0.82***</td>
<td>0.010***</td>
<td>–</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td></td>
<td>(0.03)</td>
<td>(0.01)</td>
<td>(0.01)</td>
<td>(0.01)</td>
<td>(0.01)</td>
</tr>
<tr>
<td>EFTPOS transactions for foreign customers</td>
<td>2.53***</td>
<td>0.11***</td>
<td>–13.64***</td>
<td>–194.35***</td>
<td>–</td>
</tr>
<tr>
<td></td>
<td>(0.72)</td>
<td>(21.24)</td>
<td>(33.73)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total annual transactions per ATM</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–0.01***</td>
<td>(0.01)</td>
</tr>
<tr>
<td>Average ATM transaction value for foreign customers</td>
<td>–7.97</td>
<td>0.48</td>
<td>–51.90</td>
<td>–797.66</td>
<td>–3.06</td>
</tr>
<tr>
<td></td>
<td>(5.34)</td>
<td>(0.31)</td>
<td>(33.30)</td>
<td>(540.35)</td>
<td>(2.40)</td>
</tr>
<tr>
<td>Average ATM transaction value for members</td>
<td>8.82</td>
<td>–0.06</td>
<td>56.53</td>
<td>105.87</td>
<td>–0.40</td>
</tr>
<tr>
<td></td>
<td>(2.21)</td>
<td>(0.14)</td>
<td>(152.70)</td>
<td>(240.43)</td>
<td>(1.06)</td>
</tr>
<tr>
<td>Average EFTPOS transaction value</td>
<td>2.01***</td>
<td>–0.02</td>
<td>19.87</td>
<td>31.91</td>
<td>0.03</td>
</tr>
<tr>
<td></td>
<td>(0.43)</td>
<td>(0.02)</td>
<td>(25.98)</td>
<td>(40.91)</td>
<td>(0.18)</td>
</tr>
<tr>
<td>Log (National ATMs)</td>
<td>–5.06***</td>
<td>–0.03</td>
<td>–46.45</td>
<td>65.87</td>
<td>0.63***</td>
</tr>
<tr>
<td></td>
<td>(0.88)</td>
<td>(0.04)</td>
<td>(44.61)</td>
<td>(70.40)</td>
<td>(0.30)</td>
</tr>
<tr>
<td>Log (National EFTPOS)</td>
<td>3.25***</td>
<td>–0.11***</td>
<td>112.18***</td>
<td>178.21***</td>
<td>0.73***</td>
</tr>
<tr>
<td></td>
<td>(0.51)</td>
<td>(0.02)</td>
<td>(20.63)</td>
<td>(32.60)</td>
<td>(0.15)</td>
</tr>
<tr>
<td>GDP per capita</td>
<td>–0.13E–05**</td>
<td>0.47E–08</td>
<td>–0.23E–05</td>
<td>–0.41E–06</td>
<td>–0.14E–06</td>
</tr>
<tr>
<td></td>
<td>(0.01)</td>
<td>(0.01)</td>
<td>(0.01)</td>
<td>(0.01)</td>
<td>(0.01)</td>
</tr>
<tr>
<td>Population growth</td>
<td>0.024</td>
<td>–0.004</td>
<td>1.06</td>
<td>–0.07</td>
<td>–0.02</td>
</tr>
<tr>
<td></td>
<td>(0.03)</td>
<td>(0.02)</td>
<td>(2.85)</td>
<td>(4.49)</td>
<td>(0.01)</td>
</tr>
<tr>
<td>$\chi^2$</td>
<td>34.68</td>
<td>2.34</td>
<td>4.09**</td>
<td>5.46**</td>
<td>42.73***</td>
</tr>
</tbody>
</table>

This table presents the results from a 3SLS five equation system with the five endogenous variables (ATM surcharge, Market share of ATM, ATM transactions for members, ATM transactions for foreign customers and EFTPOS transactions). Our two hypothesis of consumer’s rationality in the use of ATMs and EFTPOS are tested using estimated coefficients in the ATM transaction for foreign customers and EFTPOS transaction (Columns iv and v). If consumers are rational then we expect to find a negative relationship between ATM and EFTPOS transactions for foreign customers, a negative relationship between total EFTPOS and total ATM transaction and a positive relationship between ATM surcharge and EFTPOS transaction. Time effects included in all the equations. Standard errors in parenthesis.

* Indicates $p$-value of 10%.
** Indicates $p$-value of 5%.
*** Indicates $p$-value of 1%.
use of EFTPOS by foreign customers to reduce the need for cash and thus ATM usage. As can be seen from column (iv) this is confirmed by the negative statistically significant sign on the EFTPOS transaction for foreign customers variable. Thus, allowing for endogeneity, the results using these Spanish data are consistent with this hypothesis.

5.4.2. H2: There is a cross substitution between the ATM surcharge and EFTPOS transaction volume

The results from our five equation system also supports the second hypothesis. From column (v) it can be seen that the higher the ATM surcharge – which is the cost of ATM use by non-bank customers (foreign) – the greater is the volume of EFTPOS transactions. This is also consistent with rational behavior, i.e. a higher price of one payment mechanism results in greater use of a low (no) cost alternative.

Overall, the results using the Spanish saving bank data, recognizing issues of endogeneity and other external effects on payment use and pricing, support consumer rationality in their use of retail payment mechanism.

6. Conclusions and future research agenda

The main conclusion of this survey is that research in the area of credit cards, debit cards and ATMs is still inconclusive and unsettled, and that additional research remains to be conducted. The topic of credit card pricing provides one example of the unsettled nature of this research. For example, as we describe in detail above, theoretical explanations for the behavior of credit card interest rates in the existing literature include (1) adverse selection, (2) search costs, (3) switching costs, (4) rational consumers, (5) time inconsistency on the part of consumers, (6), fixed and variable interest rates, (7) tacit collusion, (8) the option value of card debt, and (9) risk and return. This list shows the complexity of this seemingly simple question, as well as the fact that no single paradigm has emerged to provide a definitive answer among the competing theories. Clearly, significant further research is required, both theoretical as well as empirical, to confirm or refute these alternative hypotheses.

At the moment, the main constraint for additional research progress in the payment system area is in the availability of useful data. It seems evident that researchers who have access to detailed data can make significant progress in furthering our understanding of credit cards, debit cards and ATMs. Of particular importance will be large and detailed datasets at either the individual consumer level or alternatively at the individual bank level. The two key elements in much of the theoretical literature discussed above are decision making by individual consumers and decision making by individual banks. Researchers who are able to generate large datasets in order to examine empirically how such decisions are made will make very significant progress in this area.

Another important conclusion from this literature survey is the diverse fields of finance and economics that form a basis for much of the research on credit cards, debit cards and ATMs. In the papers described above, the relevant economic sub-fields have included financial economics, banking, monetary economics, macroeconomics, industrial organization, regulatory economics, consumer behavior, and network economics. Once again, this list illustrates the complexity of seemingly simple – but still unresolved – issues, such as pricing and substitutability among payment vehicles. It can be argued however, that successful future research on these topics will draw from a variety of these sub-fields simultaneously to develop richer and more useful models and explanations. The fact that the issues of credit cards, debit cards and ATMs can be analyzed from so many different perspectives provides an important advantage to future researchers on these topics in the development of new explanations that are alternative to those already in the literature.

Finally, with respect to finance and financial behavior, the area of consumer finance research generally is very small relative to the quantity of financial research on corporate finance for example. Clearly, much more work can be done on understanding how consumers make financial decisions to develop a more general equilibrium understanding of the behavior of credit markets in particular and financial markets in general.

References


