TILBURG ◆ UNIVERSITY

TILEC

TILEC Discussion Paper

DP 2007-035

Investment and usage of new technologies: evidence from a shared ATM network

By

Stijn Ferrari, Frank Verboven, Hans Degryse

November 2007

ISSN 1572-4042

Investment and usage of new technologies: evidence from a shared ATM network

Stijn Ferrari, Frank Verboven, Hans Degryse*

November 2007

Abstract

When new technologies become available, it is not only essential that firms have the correct investment incentives, but often also that consumers make the proper usage decisions. This paper studies investment and usage in a shared ATM network. Because all banks coordinate their ATM investment decisions, there is no strategic but only a pure cost-saving incentive to invest. At the same time, because retail fees for cash withdrawals are regulated to zero at both branches and ATMs, consumers may not have the proper incentives to substitute their transactions from branches to the available ATMs. We develop an empirical model of coordinated investment and cash withdrawal demand, where banks choose the number of ATMs and consumers decide whether to withdraw cash at ATMs or branches. We find that banks substantially underinvested in the shared ATM network and thus provided too little geographic coverage. This contrasts with earlier findings of strategic overinvestment in networks with partial incompatibility. Furthermore, we find that consumer usage of the available ATM network is too low because of the zero retail fees for cash withdrawals at branches. A direct promotion of investment (through subsidies or other means) can improve welfare, but the introduction of retail fees on cash withdrawals at branches would be more effective, even if this does not encourage investment per se.

**Stijn Ferrari*: Catholic University of Leuven. Email: Stijn.Ferrari@econ.kuleuven.be. *Frank Verboven*: Catholic University of Leuven and C.E.P.R. (London). E-mail: Frank.Verboven@econ.kuleuven.be. *Hans Degryse*: Tilburg University, TILEC, Catholic University of Leuven and CESIfo. Email: H.Degryse@uvt.nl.
The authors thank Lapo Filistrucchi, Javier Suarez, Patrick Waelbroeck and seminar participants at Leuven, Louvain-la-Neuve, TILEC, the 2007 EARIE conference in Valencia, the IDEI/Bruegel conference on "Regulation, Competition and Investment in Network Industries" in Brussels, the Telecom Paris-ENST conference on "The Economics of Payment Systems", and the 2007 CEPR-ESSFM in Gerzensee.

1 Introduction

The incentives to invest in new technologies are not necessarily in line with social welfare. On the one hand, firms may underinvest because they are not able to appropriate all consumer surplus. On the other hand, they may strategically overinvest because they do not account for the business stealing effects on their competitors. While the reasons for under- and overinvestment in new technologies have become reasonably well-understood,[1] empirical evidence remains limited. Furthermore, the focus on investment incentives often provides an incomplete picture; in many settings it is also essential to understand how consumers respond and are willing to use the new technology. The joint importance of both investment and usage of new technologies has been stressed by policy makers, but it has received little attention in academic research.[2]

Automated teller machines (ATMs) provide a particularly interesting case to study the interplay between the firms' investment and the consumers' usage decisions of a new technology. The technology became available in the seventies and provided important opportunities to the banks to save on the high variable costs from branch transactions by inducing consumers to substitute to lower cost ATM transactions. Since substantial fixed investment costs were required to create sufficient geographical coverage, banks quickly joined forces to build large compatible, or shared, ATM networks. By the mid-nineties many European countries and U.S. states effectively had a single or a dominant shared ATM network, accessible to most consumers. However, in several countries this trend reversed with the introduction of surcharges. The partial incompatibility resulting from these and related retail fees provided the possibility of strategic overinvestment, as has become well-documented in several recent studies on ATMs.[3] In contrast, very little is known about potential underinvestment in compatible or shared ATM networks. Furthermore, even less is known on how effective these investments in infrastructure have been in inducing consumers to substitute and use the new cost-saving technology.

This paper aims to shed light on the firms' investment incentives and the consumers' usage decisions in a shared ATM network. We consider the case of Belgium, where until

[1] See in particular Arrow (1962), Gilbert and Newbery (1982) and many related contributions regarding the role of market structure on the incentives to invest in new technologies.

[2] For example, the most recent OECD (2007) Communications Outlook uses both infrastructure investment and consumer usage criteria in evaluating the performance of new information and communication technologies such as broadband. In contrast, the economics literature on technology adoption tends to treat firm investment and consumer adoption separately, as illustrated by Stoneman's (2002) review of the literature.

[3] See, for example, Gowrisankaran and Kraner (2007), Hannan and Borzekowki (2007), Ishii (2005) and several other studies, as reviewed below.

recently all banks jointly owned a single shared ATM network and coordinated their ATM investment decisions. At the same time, regulation prevented the banks from charging retail fees for cash withdrawals at both their own branches and at their shared ATMs. In this environment, which is representative for many countries at various points in time, ATM investment was not driven by revenue or strategic incentives. It was instead essentially motivated by a variable cost-saving incentive, i.e. the prospect that increased ATM availability would induce consumers to substitute out of high-cost branch transactions towards low-cost ATM transactions. We ask the following two related questions. First, to which extent was this environment of coordinated investment responsible for underinvestment in the ATM network? Second, to which extent did the zero retail fees on cash withdrawals provide the wrong signal to consumers and did it induce them to use branches too often relative to the available ATMs? To address these questions, we assess how a direct promotion of investment (through subsidies or other means) and the introduction of retail fees can contribute to improving social welfare.

We develop an empirical model of consumer cash withdrawal demand and coordinated ATM investment in local markets. Consumers' demand for branch and ATM cash withdrawals depends on local ATM availability. The banks' ATM investment decisions involve a trade-off between variable transaction cost savings and additional fixed costs from expanding the network. The model generates the following insights. First, it measures how increased ATM availability induces consumers to substitute out of cash withdrawals at branches to ATMs. Second, it allows us to infer the relative importance of fixed costs per ATM and the variable cost savings from increased ATM usage. To estimate the model we have collected a unique data set on ATM cash withdrawal demand and the number of ATMs, covering the entire network across Belgian local markets.

We find evidence of substantial underinvestment in the provision of ATMs: the total number of ATMs is only about half of the socially optimal number and the number of markets without an ATM is three times higher than in the social optimum. These findings stem from the fact that the coordinating banks cannot appropriate consumer surplus (in the form of improved convenience from increased ATM availability) and do not have strategic motives to invest. However, while the limited ATM investments and especially the perceived lack of geographic coverage have been highly sensitive political issues, this is only part of the welfare story. We find that the welfare losses also stem from the fact that cash withdrawal fees on both branches and ATMs have been regulated to zero, so that consumers make too limited use of the existing ATM network. To achieve the maximum welfare gains the direct promotion of ATM investment (through subsidies or other means) should be combined with the introduction of cost-based cash withdrawal fees for branch transactions. In fact, we find that a second-best "fees-only" policy is more effective than an alternative "subsidies-only" policy that directly promotes ATM investment while keeping fees regulated to zero. This is because a fees-only policy also accomplishes desirable cost-saving substitution to ATMs without requiring large additional fixed cost investments in a dense ATM network. At a more general level, our findings imply that economic analysis and policy may often be too preoccupied with stimulating investment per se, and should be more concerned with providing the correct price signals to achieve an efficient usage of the cost-reducing investments.

Apart from the general interest question on investment and usage of cost-reducing technologies, the paper contributes to the growing empirical literature on ATMs. Most of this literature has been motivated by the recent move to partial incompatibility after the introduction of surcharges in the U.S. In particular, Gowrisankaran and Krainer (2007) and Ishii (2005) develop structural models of ATM investment, enabling a welfare analysis.[4] They focus on respectively the stand-alone revenue motives and strategic motives from investing in ATMs, and their results indicate a tendency towards overinvestment in ATMs relative to the social optimum. They do not take into account the cost-saving incentives for ATM investment. In contrast, based on a unique data set on cash withdrawals, we focus on this pure cost-saving incentive: we find evidence of substantial underinvestment in ATM network coverage, combined with an insufficient usage of the existing investments due to incorrect price signals. While several studies have measured the variable cost savings from ATM cash withdrawals based on aggregate cost information, no work has attempted to integrate this in a model to study the investment incentives and consumer usage responses.[5]

From a methodological perspective our empirical model of ATM investment closely relates to the empirical entry literature. Bresnahan and Reiss (1990, 1991) and Berry (1992) introduced models of free entry. Berry and Waldfogel (1999) added a demand side to the free entry model enabling them to draw inferences on fixed costs. Ishii (2005) models ATM investment and deposit demand sequentially, using Pakes et al.'s (2006) moment inequalities method for the investment part of the model. In contrast with Berry and Waldfogel (1999) and Ishii (2005) we consider coordinated investment. Furthermore, we allow ATM demand and investment to be simultaneously determined, i.e. we account for the fact that they may depend on common unobserved local market characteristics. Intuitively, banks tend to invest

[4] Other empirical contributions on the effects of greater incompatibility on ATM investment include Knittel and Stango (2004, 2006) and Hannan and Borzekowski (2007). Theoretical contributions on ATM deployment and efficiency include Matutes and Padilla (1994), Bernhardt and Massoud (2005) and Donze and Dubec (2006).

[5] Studies measuring cost savings from ATM withdrawals link bank accounting cost data with the number of ATMs and the number of bank branches, see e.g. Felgran (1984), Berger (1985), Humphrey (1994), or Humphrey et al. (2003).

especially in markets where they expect a high ATM demand. We therefore account for both selection and endogeneity issues in measuring the causal impact of ATM investment on ATM demand.[6]

The paper is organized as follows. Section 2 discusses the relevant industry background in an international context and takes a first look at our data set. Section 3 presents the model of coordinated investment, and compares it with socially optimal investment. Sections 4, 5 and 6 respectively discuss the econometric specification, the empirical results, and the policy counterfactuals. Section 7 presents a robustness analysis from two extensions of the model, and Section 8 concludes.

2 Industry background and data set

We study ATM investment and cash withdrawal demand (usage) based on the shared network in Belgium in 1994. Before developing the econometric model, we discuss the relevant industry background in an international context and have a first look at our data set.

2.1 Industry background

The evolution to a shared ATM network Banks traditionally used their own branch networks to provide cash withdrawal services to its customers. In the late sixties and early seventies the first ATMs emerged, providing the banks with opportunities to reduce labour costs at their branches. In both the U.S. and Europe, the banks initially developed private ATM networks, accessible to their own customers only. However, to cut costs banks quickly started to cooperate, resulting in the development of shared networks, accessible to all customers of the participating banks. In the U.S., the interconnection of smaller regional networks was followed by a process of consolidation of many shared networks and the introduction of national networks.[7] In Europe, there was a similar trend towards shared ATM networks. This resulted in single or dominant shared networks in large countries such as France and Italy, as well as in several smaller countries such as Belgium, Finland, the Netherlands, Sweden or Switzerland; see B.I.S. (1999, 2003) and Snellman (2006). However,

[6] Because our model treats investment and demand simultaneously, we can make equilibrium investment and demand predictions under alternative policies. In contrast, Ishii's conclusions about overinvestment in incompatible networks are not based on equilibrium predictions because of the complexities with multiple equilibria in her framework. She therefore only looks at the direction of each bank's investment decisions.

[7] See McAndrews (1991) for a discussion of the evolution to shared ATM networks in the early years in the U.S.; and Ishii (2005) for a review of the recent U.S. evolution and the introduction of surcharges.

parallel to these shared ATM networks, banks continued to provide cash withdrawal services to their own depositors through their traditional branch networks.

Against this background we analyze the shared ATM network in Belgium in 1994. In the late seventies cooperation between the large banks resulted in the emergence of two competing ATM networks. Consumers could withdraw cash from any ATM of their own network, but had no access to the competing network. Because of cost considerations and public pressure to increase user convenience the two networks were made compatible in 1987, enabling all Belgian debit card holders to withdraw cash from ATMs of either network. A few years later, in 1990, the two networks merged completely to create a common network operator, Banksys, co-owned by all the banks. Banksys managed the shared ATM network and the emerging electronic services with debit cards.[8]

During 1990-2005, an ATM-committee within Banksys made the decisions to invest in additional ATMs, and replace or remove existing ones. This ATM-committee consisted of representatives of the larger banks, a representative of the smaller banks and a representative of the network operator. The committee decided on the number and location of ATMs for each local market. The ATMs were always installed at one of the banks' branches, hence never "off-premise" (e.g. in shopping malls). The banks had to bear the costs of the ATMs that were located at their branches, including the fixed investment and maintenance costs and the variable costs of cash withdrawals (e.g. refilling ATMs). There were a number of mechanisms to ensure cooperation among all banks. First, there was a mutual understanding that banks should host ATMs in proportion to their market shares. In practice, most banks indeed had a market share in the ATM network close to their market share in terms of branches or deposits.[9] Second, banks received compensation for the ATM services through cost-based interchange fees.

Figure 1 shows the evolution of the total number of ATMs in Belgium since 1979. The shared ATM network has grown nearly linearly during the eighties to reach maturity in the early nineties. Our data set covers a cross-section of local markets in 1994. This year is well-suited for studying ATM investment and demand. First, Figure 1 shows that 1994 represents a mature long-term situation, making it reasonable to abstract from dynamic considerations. Second, in 1994 consumers still made only limited use of electronic payment services and of

[8] As in some other countries, there was one other very small network, Postomat, accessible only to the customers of the Belgian Postal Bank. This network joined forces with Banksys in 2000.

[9] In 1994, the seven large banks were ASLK, Generale Bank, Gemeentekrediet, BBL, Kredietbank, Cera, and BACOB. They had a nation-wide presence and their market shares in terms of branches (deposits) of respectively 15 (12), 14 (13), 12 (15), 12 (10), 10 (10), 12 (5) and 8 (5) percent. Their market shares in the ATM network were respectively 21, 21, 16, 13, 9, 5 and 7 percent.

incompatible private ATMs, installed within the banks' own branches.[10]

Retail fees and costs of cash withdrawals Banks have a long tradition of low or zero variable retail fees for providing payment services to their *own customers*, including cash withdrawal services at own branches or shared ATMs. Cash withdrawal services to *customers of other banks* are typically not available at branches, but they are available at the banks' shared ATMs, possibly at a retail fee in the form of a surcharge. McAndrews (2003) provides an overview of the various retail fees for ATM cash withdrawal services to the banks' own customers (on-us fees and on-other's fees) and to non-customers (surcharges).[11]

Government regulation in Belgium has for a long time completely prevented the banks from charging retail fees for any payment related services, including cash withdrawals at branches or ATMs. Decreasing margins following intensified competition, a drop in the interbanking rates, and public demand for more transparency increased the banks' needs for charging retail fees. Intensive lobbying eventually resulted in stepwise liberalizations in 1991 and 1993, enabling the banks to charge variable retail fees for cash withdrawal services. In practice, however, a universal service obligation kept the variable fees equal to zero until the late nineties.[12] In sum, Belgian banks have generally charged zero variable retail fees, both for branch cash withdrawal services to their own customers and for the shared ATM cash withdrawals services to all debit card holders. This situation is similar to that in many other European countries, as surveyed in a study by Retail Banking Ltd. (2005) for the European Commission.

The absence of retail fees does clearly not reflect the banks' costs. While precise estimates are difficult to find, it is well-known that the variable costs for cash withdrawal services are considerably higher at branches than at ATMs. Berger (1985) and Humphrey (1994) find that the variable costs are about twice as high at branches than at ATMs. According to Kimball and Gregor (1995), the per transaction cost is \$0.27 at ATMs, compared to \$1.07 at branches, while Fasig (2001) states that transaction costs vary between \$0.15 to \$0.50 at ATMs and \$1 to \$2 at branches. The variable cost savings at ATMs should however

[10] The incompatible private ATMs were exclusively for the bank's own customers, allowing for cash withdrawals and other traditional branch transactions, such as the ordering of documents and the transfering of funds. In 1994, less than one fifth of the branches were equipped with an incompatible private ATM, and we will simply treat them as an integrated part of the banks' branch networks.

[11] In addition to the retail fees, there are also wholesale fees, i.e. the interchange fees that banks charge to each other, and switch fees charged by the network operator to the banks.

[12] The universal service obligation forced banks to offer a minimum amount of payment services, including cash withdrawals at branches or ATMs at a low cost. In practice, banks only charge a small annual fixed fee for payment services and no variable retail fees.

be balanced against the fixed costs of investing and maintaining ATMs. According to the Belgian network operator Banksys, the fixed cost per ATM amounts to about €2,300 per month. This is similar to estimates quoted by Ishii (2005) for the U.S.[13] We will come back to these cost estimates in our empirical analysis, where we will infer the ratio of fixed ATM costs over variable cost savings from our econometric model of ATM investment.

Summary Belgian banks have for a long time coordinated their investment decisions in the shared ATM network. The year 1994 is well-suited for an empirical analysis of ATM investment and demand since the network had matured and competing electronic payment services were still of limited importance. Banks charged no retail fees for cash withdrawal services, whether at incompatible branches or at shared ATMs. However, banks could realize potentially important variable cost savings from cash withdrawals at ATMs instead of branches, to be traded off against the fixed costs from setting-up and maintaining the shared ATM network. These observations will motivate our empirical model of coordinated investment, developed and estimated in the next sections.

2.2 A first look at the data

2.2.1 The data set

Our main data set consists of ATM cash withdrawal demand and the number of ATMs for a cross-section of local markets in Belgium in 1994. The markets are defined by postal codes, which are part of administrative municipalities and typically consist of about one or two traditional towns. To reduce potential problems with overlapping markets, we focus on a subsample of 659 non-urban markets (defined as markets with a population density of less than 800 per km^2), having on average about 8,700 inhabitants. But we also considered a robustness analysis based on the full sample of all 842 markets including the cities. For each market, we observe the total number of ATM cash withdrawals and their nominal monetary value, both expressed as 1994 monthly averages. In addition, we observe the number of shared ATMs, defined as the number of distinct ATM locations per market. We also collected data on the banks' branch locations in 1994, and on various demographic characteristics such as population size.[14]

[13] Ishii (2005) quotes 2003 American Bankers' Association numbers, according to which the cost of buying an ATM machine is \$50,000, and annual maintenance cost is between \$12,000 and \$15,000. Using her five-year linear depreciation period this amounts to a monthly fixed cost of between \$1,833 and \$2,083.

[14] The data set on the ATMs was provided to us by the ATM network operator Banksys. The data on the branch locations is from B.V.B., the Belgian Banking Federation. The demographic characteristics were obtained from the N.I.S. (National Institute of Statistics), Ecodata (Federal Government Agency for

Table 1 provides precise definitions of our variables, and Table 2 presents summary statistics for the cross-section of 659 non-urban markets, and the subsample of 310 markets with at least one ATM. The per capita number of ATM cash withdrawals Q_A is on average 0.56 per month for all markets, and 0.78 for the markets with at least one ATM.[15] The average value per cash withdrawal V_A/Q_A is €101 (average across the markets with at least one ATM). The availability of ATMs across the local markets is rather limited. There are no ATMs in 349 out of 659 markets, and in those markets with at least one ATM the average number of ATMs N is only 1.57.

Consumers can also withdraw cash from their own bank branches rather than from the shared ATMs. We do not have rich data on branch cash withdrawals at the local market level, but at the national level we estimate that consumers make about 2.07 cash withdrawals per month.[16] Hence, ATM usage is relatively limited: only about one third of the cash withdrawals take place at the shared ATMs and the remaining two-thirds are at the branches. Branch availability to consumers in need for cash can be measured since we observe the number of branches per market for each bank. Since branches of rival banks are not compatible, a crude aggregate measure of branch availability is the average number of branches per bank in each market. Table 2 shows that there are on average 0.86 branches per bank across all markets, and on average 1.25 branches per bank in the sample of markets with at least one shared ATM. Consumers thus tend to find about the same amount of branches of their own bank as shared ATMs within a local market.

The remaining variables are the market demographics. In our empirical analysis these may affect both ATM cash withdrawal demand and the profitability of investing in ATMs. The demographics include population (number of inhabitants per market), the market surface (in km^2), the number of enterprises, the fraction of foreigners, the fraction of young (under the age of 18) and elderly (over the age of 65), the unemployment rate, and a dummy variable for the region of Flanders (Dutch-speaking part of Belgium). Table 2 shows that several of the demographics may differ depending on whether the full sample or the subsample of markets with at least one ATM is considered. In particular, the average population size is 8,738 across all markets, but up to 13,445 in markets where banks invested in at least one ATM.

Economics), and the R.S.Z. (the National Institute of Social Security).

[15] These averages become slightly larger when city markets are included, i.e. 0.80 for all markets, and 0.98 for the markets with at least one ATM.

[16] The estimate of 2.07 cash withdrawals per month is based on recent 2004 information at the national level on cash withdrawals. Note that the government also used an estimate of 2 cash withdrawals per month in its universal service obligation proposal for the banks.

Preliminary relationships Table 3 shows the relationship between our main variables of interest, i.e. the per capita number of ATM cash withdrawals Q_A and the number of shared ATMs, N. The average of Q_A is 0.63 across markets with only one ATM and this gradually increases as N increases, to reach an average of 1.13 in markets with 5 available ATMs. Table 3 also shows that the average value per cash withdrawal V_A/Q_A decreases in N, but only weakly from €102 in markets with one ATM to €98 in markets with 5 ATMs.

To gain further insights in the relationship between ATM demand or usage and ATM availability, we estimate two simple OLS regressions, based on the sample of markets with at least one ATM. The first regression takes $\ln Q_A$ as the dependent variable and relates this to $\ln N$ and the log of the number of branches per bank, after controlling for the market demographics. In the second regression $\ln V_A/Q_A$ is the dependent variable, and includes the same explanatory variables. The regressions should be interpreted with care, as $\ln N$ may be correlated with the error term because of both sample selection and endogeneity issues: banks tend to invest in no or few ATMs in markets where they expect a low ATM demand, and vice versa.

Table 4 shows the results. We focus mainly on the regression for $\ln Q_A$ in the first part of the table. The elasticity of ATM cash withdrawal demand Q_A with respect to the number of ATMs N is 0.63, which is positive and highly significant. This may describe a causal effect of ATM availability on demand or usage, or simply reflect the fact that banks invest in many ATMs when they expect high demand. Furthermore, the elasticity of ATM cash withdrawal demand with respect to the number of branches per bank is -0.51. In absolute value this is not significantly different from the coefficient on the number of ATMs (p-value of 0.10). Hence, ATM usage increases by about the same amount when ATM availability increases as when per bank branch availability decreases. This indicates that consumers use ATMs as a substitute for branches to withdraw cash.

The second part of Table 4 shows the regression for $\ln V_A/Q_A$. The elasticity of the average value per cash withdrawal V_A/Q_A with respect to the number of ATMs N is negative and significant, but its magnitude is quite small (-0.03). Hence, while Q_A increases substantially with ATM availability, V_A/Q_A decreases only to a small extent. This suggests that the positive relationship between ATM withdrawals and availability is not due to the fact that consumers withdraw a lower value per transaction, but rather because they substitute out of cash withdrawals at branches.

We emphasize again that the regressions on ATM availability should be interpreted with care, because N is an endogenous variable implying both selection and endogeneity issues with simple OLS estimation. The next sections develop and estimate a model of ATM demand and coordinated investment in ATMs that take these issues into account. This will enable us to obtain more reliable conclusions on the causal effect of ATM availability on demand, and to perform a welfare analysis regarding the optimality of ATM investment and demand under alternative scenarios.

3 The model

3.1 Overview

When banks can charge retail fees for cash withdrawal services, they have at least two broad profit motives for adopting ATMs. First, there is the pure stand-alone profit motive associated with the fee revenues from ATM cash withdrawals. Second, there is a strategic motive when the fees come in the form of on-other's fees and/or surcharges, i.e. additional fees for consumers using ATMs from banks other than their own. These fees result in partial incompatibility between different ATM networks, providing banks with larger networks a strategic advantage over their rivals, as they can more easily attract new customers, or raise their rivals' costs. The recent ATM literature has largely focused on these two profit incentives for adopting ATMs, see e.g. McAndrews (2003) for an overview of the theoretical literature and Hannan et al. (2003), Knittel and Stango (2004), Ishii (2005), Gowrisankaran and Krainer (2007), and Hannan and Borzekowski (2007) for recent empirical contributions.

There is, however, also a third profit incentive for adopting ATMs, the pure cost-saving incentive, which is present even if banks cannot charge retail fees. An ATM network with a broad geographic coverage induces customers to switch from branch to ATM cash withdrawals. This implies potentially important variable cost savings, but these need to be balanced against the fixed costs of setting up the ATM network. The cost-saving incentive is therefore larger if firms coordinate their ATM investment decisions and set up a shared ATM network.

The cost-saving incentive was already highlighted in the early literature as an important motive for ATM investment, but data limitations prevented a proper identification. Our analysis models and identifies precisely this cost-saving incentive in an environment where the two other profit incentives are absent because of zero retail fees, as in many countries. Consistent with our industry background we first develop a model of coordinated ATM investment and demand in the absence of retail fees for cash withdrawals. This model will form the basis of our econometric analysis. We then consider the socially optimal outcome, and show how a social planner can intervene by regulating fees (at branches and ATMs), and/or providing subsidies per installed ATM. This will be used to perform a counterfactual policy analysis.

Our empirical model builds on earlier models of free entry, originating from Bresnahan and Reiss (1990, 1991) and Berry (1992). Berry and Waldfogel (1999) added a demand side to a free entry model, which enabled them to separately identify the demand and fixed cost parameters. Recent related work that incorporates both an entry and a demand equation can be found in e.g. Abraham et al. (2005), Ishii (2005), and Smith (2007). Our own model also consists of an entry and demand equation, but the entry equation comes from a model of coordinated entry rather than one of free entry.

3.2 Coordinated investment

For a cross-section of local markets we observe the monthly number of ATM cash withdrawals per capita Q_A and the number of shared ATMs N. For each market our model of coordinated ATM investment specifies how Q_A and N (or usage and investment) are simultaneously determined and depend on observed and unobserved market characteristics.

A market consists of L consumers. Each consumer may withdraw cash at the branch of its bank or at a shared ATM, and the demands depend on the availability of ATMs. Let ATM cash withdrawal demand, $Q_A = Q_A(N)$, be increasing in the number of ATMs N: as N increases, the average distance to an ATM in the local market decreases so that demand for cash withdrawals at ATMs increases. Similarly, let cash withdrawal demand at branches, $Q_B(N)$, be decreasing in N: as the availability of ATMs increases, it becomes relatively less attractive to withdraw cash at branches and consumers substitute to ATMs. Total cash withdrawal demand is $Q(N) = Q_A(N) + Q_B(N)$. Let $Q(N)$ be nondecreasing in N, i.e. an increase in the availability of ATMs leads to an increase in the total number of withdrawals, unless total cash withdrawal demand is inelastic with respect to N. In sum, increasing the number of ATMs leads to substitution from branch to ATM cash withdrawals, and to an overall expansion of cash withdrawals unless total cash demand is inelastic.

Banks coordinate their ATM investment (or entry) decisions, in line with our industry background discussed in Section 2. In each market, they choose the number of shared ATMs N to maximize their joint profits $\Pi(N)$. The joint profits consist of a stand-alone component Π_0, independent of N, and of several other components that depend on N. There is a constant variable cost per ATM cash withdrawal of c_A, and a constant variable cost per branch cash withdrawal of $c_B > c_A$. The fixed cost of an ATM is F, which consists of both investment and maintenance costs. Banks do not charge retail fees for cash withdrawal at either ATMs or branches. They also do not obtain subsidies per ATM. The banks' joint profits in a given market are then given by:

$$\Pi(N) = \Pi_0 - c_A Q_A(N) L - c_B Q_B(N) L - FN. \tag{1}$$

This is simply the stand-alone profit component Π_0, minus the total variable costs from ATM and branch cash withdrawals, minus the fixed costs of all shared ATMs in the market. Note that the joint profits do not depend on the interchange fees, which banks pay to each other through the network operator. These interchange fees are simply transfers between banks and cancel out when adding up the banks' individual profits to obtain joint profits. If firms would choose ATMs in an uncoordinated way to maximize their individual profits, then the interchange fees become potentially relevant and may serve as a mechanism to soften competition for depositors; see Matutes and Padilla (1994) and Donze and Dubec (2006) for analyses of the strategic use of interchange fees when banks do not coordinate their ATM investment decisions.

The banks' marginal joint profits from investing in N ATMs are:

$$\Pi(N) - \Pi(N-1) = -c_A \left(Q_A(N) - Q_A(N-1)\right) L - c_B \left(Q_B(N) - Q_B(N-1)\right) L - F.$$

To interpret this economically, substitute out $Q_B(N)$ using $Q(N) = Q_A(N) + Q_B(N)$. The marginal joint profits can then be rewritten as:

$$\Pi(N) - \Pi(N-1) = \underbrace{(c_B - c_A)\left(Q_A(N) - Q_A(N-1)\right) L}_{\substack{\text{variable cost saving} \\ \text{due to substitution}}} - \underbrace{c_B \left(Q(N) - Q(N-1)\right) L}_{\substack{\text{variable cost increase} \\ \text{due to market expansion}}} - F. \tag{2}$$

This says that the change in banks' joint profits from one additional ATM consists of three components. First, an additional ATM induces consumers to substitute from high variable cost cash withdrawals at branches to low variable cost cash withdrawals at ATMs. Second, an additional ATM may increase the total number of cash withdrawals, which generates additional variable costs. Third, there is a fixed cost involved in installing an additional ATM. If total cash withdrawal demand $Q(N)$ is inelastic, the second term cancels so that an increase in the number of ATMs reduces to a simple trade-off between variable cost savings and an additional fixed cost.

The banks choose the number of shared ATMs N to maximize their joint profits. The optimal number of ATMs is $N = 0$ if:

$$\Pi(1) - \Pi(0) < 0 \tag{3}$$

and $N = n > 0$ if:

$$\Pi(n+1) - \Pi(n) < 0 \leq \Pi(n) - \Pi(n-1), \tag{4}$$

i.e. the marginal joint profits from investing in n ATMs should be positive, and the marginal joint profits from investing in $n+1$ ATMs should be negative. These are necessary conditions for joint profit maximization. They are also sufficient if the joint profits $\Pi(N)$ are concave in N, or equivalently if the marginal joint profits are decreasing in N. Note how the requirement of decreasing marginal joint profits in our model of coordinated entry parallels the common requirement of decreasing individual profits in traditional empirical models of free entry (as in e.g. Bresnahan and Reiss (1990).

3.3 Socially optimal investment

The above model describes ATM investment when banks coordinate and cannot charge retail fees on either ATM or branch cash withdrawals. This describes the "status quo situation" and forms the basis for our empirical analysis. In our policy counterfactuals presented in Section 6, we will compare the status quo with the social optimum, and assess how a regulator can set subsidies and/or retail fees to induce banks to implement the social optimum. Subsidies should be viewed as an instrument to directly promote ATM investment, but other means such as tax deductions may obviously also be possible. Retail fees mainly serve to influence ATM demand or usage, i.e. they may induce consumers to use ATMs given the available ATM network.

Suppose that the banks can charge a retail fee t_A per ATM cash withdrawal and a retail fee t_B per branch cash withdrawal.[17] Consumer surplus $CS(N, t_A, t_B)$ is increasing in N and decreasing in both fees t_A and t_B. The per capita demand for cash withdrawals at ATMs is $Q_A(N, t_A, t_B)$. This is increasing in N (as in the status quo situation where $t_A = t_B = 0$), decreasing in the own retail ATM fee t_A and increasing in the branch retail fee t_B. Similarly, per capita demand for cash withdrawals at branches is $Q_B(N, t_A, t_B)$, decreasing in N, increasing in t_A and decreasing in t_B. The earlier status quo demands with zero fees are defined as $Q_A(N) \equiv Q_A(N, 0, 0)$ and $Q_B(N) \equiv Q_B(N, 0, 0)$. In section 4 we provide a utility-consistent specification that relates $CS(N, t_A, t_B)$ to $Q_A(N, t_A, t_B)$ and $Q_B(N, t_A, t_B)$.

Producer surplus is equal to the banks' joint profits. These now also include retail fee revenues and a subsidy S per ATM:

$$\Pi(N, t_A, t_B, S) = \Pi_0 + (t_A - c_A)\, Q_A(N, t_A, t_B)L + (t_B - c_B)\, Q_B(N, t_A, t_B)L + (S - F)N, \quad (5)$$

[17] The ATM retail fee t_A applies to all consumers regardless their bank affiliation. This rules out surcharges and on-other's fees, so that ATMs remain fully compatible. The branch retail fee t_B only applies to the banks' own customers since branches are incompatible.

assumed to be concave in N. This extends the status quo profit function (1) to include the fees and subsidies, so $\Pi(N) \equiv \Pi(N, 0, 0, 0)$.

Total welfare in the presence of retail fees and subsidies, $W(N, t_A, t_B)$, is then the sum of producer surplus (5), consumer surplus and government revenues $-SN$, i.e.

$$W(N, t_A, t_B) = \Pi(N, t_A, t_B, S) + CS(N, t_A, t_B) - SN. \tag{6}$$

Note that total welfare is independent of the subsidy S, since SN is also part of $\Pi(N, t_A, t_B, S)$, and cancels out as it is just a transfer from the social planner to the banks. The social optimum or first-best solution then maximizes $W(N, t_A, t_B)$ with respect to N, t_A and t_B.

The status quo situation may not be socially optimal for two reasons. First, banks choose the number of ATMs N to maximize their own joint profits, and they do not take into account the effects on consumer surplus. Since consumer surplus $CS(N, t_A, t_B)$ is increasing in N and $\Pi(N, t_A, t_B, S)$ is concave in N, the banks will underinvest in N if the subsidy S is equal to zero.[18] Second, the retail cash withdrawal fees t_A and t_B are below variable costs and in fact regulated to zero. This implies that the demand for ATM withdrawals and especially for the high variable cost branch withdrawals may be distorted.

The social planner can induce the banks to implement the social optimum in a decentralized way, by first setting S (instead of N), t_A and t_B, and subsequently letting banks coordinate on N, given S, t_A and t_B. Formally, use (5) to compute the banks' marginal joint profits with fees and subsidies, and obtain inequality conditions analogous to (4). These define the banks' joint profit maximizing number of ATMs $n^*(t_A, t_B, S)$, given the fees t_A and t_B and the subsidy S. The social planner then maximizes total welfare $W(n^*(t_A, t_B, S), t_A, t_B)$ with respect to the fees t_A and t_B and the subsidy S. Note that while the subsidy has no direct effect on welfare (as SN is a transfer that cancels out), it has an indirect impact by influencing the banks' coordinated investment $n^*(t_A, t_B, S)$.

In our counterfactual policy analysis we will compare the status quo situation with the social optimum or first-best, as implemented through optimal fees t_A and t_B and a subsidy S. We will also consider two second-best solutions. In the "fees-only" case, the social planner keeps the subsidy at $S = 0$, and chooses t_A and t_B to maximize $W(n^*(t_A, t_B, 0), t_A, t_B)$. In the "subsidies-only" case, the social planner keeps the fees at $t_A = t_B = 0$, and chooses S to maximize $W(n^*(0, 0, S), 0, 0)$. We will assess to which extent the fees-only and subsidies-only cases improve over the status quo situation and come close to the first-best.

[18] To see this, set $S = 0$ and suppress the retail fees as arguments. When N is continuous, the coordinated optimum N^C solves $\Pi'(N^C) = 0$ and the social optimum N^S solves $\Pi'(N^S) + CS'(N^S) = 0$. The second term is positive by assumption, so that the first term is negative. Hence, $\Pi'(N^S) < \Pi'(N^C)$, so that $N^S > N^C$ by concavity of $\Pi(N)$. This argument still holds if S is positive and sufficiently close to zero.

4 Econometric specification

We now apply the model of coordinated investment and present the econometric specification. For a cross-section of markets we observe per capita ATM cash withdrawal demand Q_A and the number of ATMs N. They are simultaneously determined and depend on observed and unobserved market characteristics. We develop a specification that will enable estimation by standard joint maximum likelihood.

4.1 ATM demand or usage

Section 3 allowed total demand for cash withdrawals $Q(N)$ to be increasing in N. We now assume that additional ATMs mainly involve substitution from branches to ATMs without raising the total number of cash withdrawals. So total demand is inelastic and equal to a constant $Q(N) = \overline{Q}$. This is not unreasonable here, since our reduced form evidence, presented in Section 2, suggested that consumers do not withdraw lower values per withdrawal as ATM availability increases. We can then write the ATM and branch demands as shares in total cash withdrawal demand, i.e. $Q_A(N) = s_A(N)\overline{Q}$ and $Q_B(N) = (1 - s_A(N))\,\overline{Q}$, where $s_A(N) \in [0, 1]$ is the ATM cash withdrawal share. We will now specify this share, based on a model of consumers affiliated to the different banks.

Each consumer is affiliated to a bank i, and decides to make $\overline{Q}$ cash withdrawals at either one of the N shared ATMs in the market, or at one of the B^i branches of bank i to which she is affiliated. She incurs a total price p_A for an ATM cash withdrawal and p_B^i for a branch cash withdrawal. Let $p_A = p_A(N)$ be decreasing in N, capturing the fact that the consumer's expected travel cost decreases as the number of ATMs increases. More specifically, the expected travel cost is equal to the travel cost per unit of distance k, times the expected distance to the nearest ATM. Assuming a spatial Poisson process for the consumers' and ATMs' locations, this expected distance is equal to $\frac{1}{2}\sqrt{M/N}$, where M is the surface of the market where the consumers and ATMs are located.[19] The total price p_A for an ATM cash withdrawal is the sum of the expected travel cost and the retail cash withdrawal fee t_A (zero under the status quo), so $p_A = p_A(N) = \frac{k}{2}\sqrt{M/N} + t_A$. Similarly, the price for a branch cash withdrawal is $p_B^i = p_B(B^i) = \frac{k}{2}\sqrt{M/B^i} + t_B$.

Indirect utility or consumer surplus of a depositor affiliated to bank i takes the following

[19] The expected distance between a consumer and the nearest ATM is thus inversely proportional to the square root of the number of ATM locations, which is known as the square root law. See for example Kolesar and Blum (1973) for a derivation.

logit form:

$$CS^i\left(p_A\right) = y + \frac{1}{\alpha}\ln\left(\exp\left(v_A - \alpha p_A\right) + \exp\left(v_B - \alpha p_B^i\right)\right)\overline{Q}, \tag{7}$$

where v_A and v_B are the intrinsic utilities for withdrawing cash at ATMs and branches, respectively. This specification can be derived from either a logit discrete choice or a representative consumer model; see Anderson, de Palma and Thisse (1992). We will consider an alternative semi-log specification for consumer surplus in Section 7.

Applying Roy's identity to (7), bank i 's consumers have the following share of ATM cash withdrawals in their total cash withdrawals:

$$s_A^i\left(p_A\right) = \frac{1}{1 + \exp\left(v_B - v_A - \alpha\left(p_B^i - p_A\right)\right)}. \tag{8}$$

The aggregate ATM cash withdrawal share, as a function of the number of ATMs N in the market, is then

$$s_A(N) = \sum_i w^i s_A^i\left(p_A(N)\right), \tag{9}$$

where w^i is the market share of bank i. We assume that the banks' market shares w^i are independent of the number of ATMs, since the ATMs are shared and unlike incompatible ATMs do not provide a strategic advantage (Matutes and Padilla (1994)).[20]

The aggregate ATM cash withdrawal share $s_A(N)$ is the deterministic part of demand. Total per capita demand for cash withdrawals $\overline{Q}$ is the random part and is specified as:

$$\ln\overline{Q} = X\beta + \eta_1, \tag{10}$$

where X is a vector of observed market characteristics influencing $\overline{Q}$ and η_1 is an unobserved error term affecting total demand in the market.

Using $Q_A = s_A(N)\overline{Q}$ and (10), we obtain the following equation for ATM cash withdrawal demand:

$$\ln Q_A = \ln s_A(N) + X\beta + \eta_1, \tag{11}$$

where $s_A(N)$ is given by (8) and (9). This is the ATM demand equation to be taken to the data, for $t_A = t_B = 0$. Intuitively, the market characteristics X influence ATM demand through the parameter vector β. The number of ATMs and branches, entering p_A and p_B, influence ATM demand through the parameter α. The remaining parameter to be estimated is $v_B - v_A$, the intrinsic utility difference from withdrawing cash at a branch rather than an ATM. While $v_B - v_A$ can be made a function of market characteristics, it is not well identified

[20] We do not directly observe the market shares w^i at the local market level. As a proxy we take the market share according to the number of branches and suitably rescale so that the national market shares according to our proxy equal the observed national market shares.

from β since total cash withdrawal demand $\overline{Q}$ is unobserved. We therefore estimate $v_B - v_A$ as a constant, and assess identification by comparing total cash demand $\overline{Q}$ as predicted by our model with our estimate from an external source ($\overline{Q} = 2.07$ withdrawals per month, as discussed in Section 2.2). Furthermore, in Section 7 we will consider an alternative functional form for $s_A(N)$ to show the robustness of our results.

4.2 ATM investment

Banks coordinate their ATM investment (or entry) decisions to maximize their joint profits. With inelastic demand of total cash withdrawals, i.e. $Q(N) = \overline{Q}$, and $Q_A(N) = s_A(N)\overline{Q}$, the marginal joint profits (2) from investing in an additional ATM simplify to:

$$\Pi(N) - \Pi(N-1) = (c_B - c_A)\left(s_A(N) - s_A(N-1)\right)\overline{Q}L - F. \tag{12}$$

Intuitively, investing in one more ATM involves a simple trade-off between an additional fixed cost F against the variable cost savings from consumers substituting from branch to ATM cash withdrawals, as reflected in the higher ATM cash withdrawal share.

Substituting the marginal joint profits (12) in the necessary inequality conditions for optimality (3) and (4), the joint-profit maximizing number of ATMs is $N = 0$ if

$$\left(s_A(1) - s_A(0)\right)\overline{Q}L < \frac{F}{c_B - c_A} \tag{13}$$

and $N = n > 0$ if:

$$\left(s_A(n+1) - s_A(n)\right)\overline{Q}L < \frac{F}{c_B - c_A} < \left(s_A(n) - s_A(n-1)\right)\overline{Q}L. \tag{14}$$

These inequality conditions for joint profit maximization are also sufficient if $s_A(N)$ is concave in N. In the empirical analysis we will verify whether this is indeed the case at our obtained parameter estimates.

The investment model does not separately identify the fixed costs from the variable cost savings, but only the ratio.[21] We specify this ratio as:

$$\ln \frac{F}{c_B - c_A} = W\gamma + \eta_2, \tag{15}$$

where W is a vector of observed market characteristics and η_2 is an unobserved error term. In Section 7 we will extend this specification and allow F to depend on N, thereby allowing for economies of density.

[21] Ishii (2005) and Gowrisankaran and Krainer (2007) identify fixed costs by making assumptions on the variable costs of cash withdrawals. We do not make these assumptions at the estimation stage. In our policy counterfactuals, we also make identifying assumptions, but rather on fixed costs for which we have better information than on variable costs.

Substituting (15) in the inequality conditions (13) and (14), the number of ATMs is $N = 0$ if

$$\ln\left(s_A(1) - s_A(0)\right) + X\beta + \eta_1 + \ln L < W\gamma + \eta_2 \tag{16}$$

and $N = n > 0$ if

$$\ln\left(s_A(n+1) - s_A(n)\right) + X\beta + \eta_1 + \ln L < W\gamma + \eta_2 < \ln\left(s_A(n) - s_A(n-1)\right) + X\beta + \eta_1 + \ln L. \tag{17}$$

These investment conditions are similar to the inequalities in an ordered probit model. They can be taken to the data, together with the demand equation (11). Note that the demand error term η_1 also enters the investment conditions (16)–(17). Intuitively, a high demand shock does not only imply a high ATM demand Q_A, but also high marginal joint profits, inducing banks to invest in many shared ATMs N. This emphasizes the importance of properly accounting for the fact that Q_A and N are simultaneously determined and may depend on the same unobserved factors. We turn to estimation next.

4.3 Estimation

For our cross-section of markets we observe the number of shared ATMs N and ATM cash withdrawal demand Q_A unless $N = 0$. Defining

$$\begin{aligned} \varepsilon_1 &\equiv \eta_1 \\ \varepsilon_2 &\equiv \eta_2 - \eta_1 \\ Z\theta &\equiv X\beta + \ln L - W\gamma \\ \tau_n &\equiv \ln\left(s_A(n) - s_A(n-1)\right), \end{aligned} \tag{18}$$

we can write the demand equation (11) and the investment inequalities (16)–(17) more compactly as follows:

$$\begin{aligned} &\text{For } N = 0: && Q_A \quad \text{unobserved} \\ & && Z\theta + \tau_1 < \varepsilon_2 \\ &\text{For } N = n > 0: && \ln Q_A = \ln s_A(n) + X\beta + \varepsilon_1 \\ & && Z\theta + \tau_{n+1} < \varepsilon_2 < Z\theta + \tau_n. \end{aligned} \tag{19}$$

The model thus essentially consists of a demand or usage equation, and investment inequalities as in an ordered probit model.

If one is not interested in the parameters determining ATM investment, one may in principle estimate the demand equation separately to learn about the causal impact of ATM availability on ATM cash withdrawal demand. However, OLS estimation would be unwarranted because of the endogeneity and selection problems stemming from the simultaneous determination of Q_A and N. Intuitively, Q_A and N tend to be strongly correlated even in the absence of a causal relationship, because banks tend to invest in many ATMs under high demand shocks and in few ATMs under low demand shocks. For very low demand shocks, banks decide to invest in no ATMs, the traditional selection problem. Econometrically, the error terms ε_1 and ε_2 will be correlated since the demand error term η_1 enters both error terms through $\varepsilon_1 \equiv \eta_1$ and $\varepsilon_2 \equiv \eta_2 - \eta_1$. This correlation arises here for economic reasons, i.e. the fact that the unobserved demand term η_1 influences the banks' investment decisions.[22]

One solution to deal with the simultaneity of Q_A and N is to include a correction term in the demand equation in the spirit of Heckman's (1978) and Amemiya's (1984) binary response selection models. Several recent papers extend these models to a non-binary response framework; see e.g. Mazzeo (2002), Manuszak and Moul (2006), Watson (2007), or Cohen and Mazzeo (2007). Our econometric specification enables a more efficient approach, i.e. estimate the demand and investment model jointly using maximum likelihood. Since we are interested in both the demand and cost side parameters, we follow this approach here.[23]

Let $f_{12}(\cdot)$, $f_1(\cdot)$ and $f_2(\cdot)$ be the joint and marginal density functions of ε_1 and ε_2. We can then write the likelihood contributions for our sample of markets. For markets with $N = 0$ we have

$$P(N = 0) = \int_{Z\theta+\tau_1}^{\infty} f_2(u_2)du_2,$$

and for markets with $N = n > 0$, we have

$$f(\ln Q_A | N = n)P(N = n) = \int_{Z\theta+\tau_{n+1}}^{Z\theta+\tau_n} f_{12}(\varepsilon_1, u_2)du_2,$$

where $\varepsilon_1 = \ln Q_A - \ln s_A(n) - X\beta$ from (19).

Assume that ε_1 and ε_2 have a bivariate normal distribution, with means of zero, variances of σ_1^2 and σ_2^2 and a covariance of σ_{12}. Following standard practice in simpler Tobit II models

[22] The econometric model can be compared with Gronau's (1974) model of wage determination: wages are only observed for individuals who decide to participate in the labour market, and this participation decision may depend on the same unobserved factors ("skills") as the wages. The difference with our framework is that the participation decision in Gronau's model is a binary decision that matters for selection but does not directly influence wages. In contrast, the investment (or entry) decision is an ordered variable that matters for selection and in addition directly influences demand.

[23] To estimate the investment model, one could alternatively consider Pakes et al.'s (2006) moment inequality approach, which achieves partial identification in a more general setting (e.g. allowing for multi-agent strategic interactions and more general functional forms). One advantage of maximum likelihood in our application is that it enables simultaneous estimation of the demand and investment model, thereby accounting for common unobservables affecting both demand and investment.

with normal errors, this enables us to write the second likelihood contribution as a product of (conditional) univariate normals. Denoting the standard normal distribution and density functions by $\Phi(\cdot)$ and $\phi(\cdot)$, respectively, we can thus rewrite the likelihood contributions as

$$P(N=0) = 1 - \Phi\left(\frac{Z\theta + \tau_1}{\sigma_2}\right)$$

and

$$\begin{aligned} f(\ln Q_A)P(N &= n|\ln Q_A) = \\ &\frac{1}{\sigma_1}\phi\left(\frac{\varepsilon_1}{\sigma_1}\right)\left(\Phi\left(\frac{Z\theta + \tau_n - (\sigma_{12}/\sigma_1^2)\,\varepsilon_1}{\sqrt{\sigma_2^2 - \sigma_{12}^2/\sigma_1^2}}\right) - \Phi\left(\frac{Z\theta + \tau_{n+1} - (\sigma_{12}/\sigma_1^2)\,\varepsilon_1}{\sqrt{\sigma_2^2 - \sigma_{12}^2/\sigma_1^2}}\right)\right). \end{aligned}$$

In many latent variable models the standard deviation σ_2 is not identified. In our application, however, σ_2 is identified since one parameter of the variables in Z is restricted, i.e. the parameter for $\ln L$ (the log of the number of consumers) entering Z is equal to one; see (18). This restriction is based on the reasonable assumption that per capita cash withdrawal demand does not depend on the number of consumers.

5 Empirical results

The empirical model consists of the ATM demand or usage equation (11) and the investment or entry equation (16)-(17), as also summarized by (19). To estimate this model we observe Q_A and N, and a set of market characteristics for a cross-section of 659 markets, as discussed earlier in Section 2 and Tables 1 and 2. The market characteristics enter the demand equation (11) through X, and the entry equation (16)-(17) through W. Intuitively, X and W contain the market-level determinants of respectively total per capita cash withdrawals $\ln \overline{Q}$, and the ratio of fixed costs over variable cost savings $\ln F/(c_B - c_A)$. We will set $X = W$, hence we allow $\ln \overline{Q}$ and $\ln F/(c_B - c_A)$ to be affected by the same determinants. A first specification includes an intercept only, so that $\ln \overline{Q}$ and $\ln F/\left(c_B - c_A\right)$ are assumed to be uniform across markets. A second specification allows $\overline{Q}$ and $F/\left(c_B - c_A\right)$ to vary according to the following market demographics: the number of enterprises, the percentage of foreigners, young, elderly and unemployed, and a region dummy for Flanders.

Demand model only We first present the results from estimating the demand equation (11) only. This equation contains the ATM cash withdrawal share function $s_A(N)$, as given by (9), which is nonlinear in the parameters $v_B - v_A$ and α. We therefore estimate the demand equation by maximum likelihood, but this does not take into account that Q_A and N are simultaneously determined. The results from the demand equation thus serve to highlight the endogeneity and selection issues associated with the number of ATMs N.

Table 5 shows the parameter estimates. The parameter α is estimated to be highly significant in both the specification without and with market characteristics influencing $\overline{Q}$. Recall that α measures how ATM usage is affected by the implicit price of an ATM withdrawal p_A. Since this implicit price is inversely proportional to $\sqrt{N}$, the significant estimate of α means that consumers withdraw significantly more cash at ATMs than at branches in markets where N is high. The parameter estimate of α allows us to compute the demand elasticity with respect to the number of ATMs, $E_A = \frac{\partial Q_A(N)}{\partial N}\frac{N}{Q_A}$.[24] This elasticity estimate (evaluated at the sample mean) is quite high, i.e. 1.09 and 0.89 in the two respective specifications. This may however not describe the causal effect of N on Q_A, but only a correlation since banks may have an incentive to invest in many ATMs when they observe a high ATM demand shock, and vice versa. The simultaneous model of ATM demand and investment will take this into account.

The second specification in Table 5 shows how market demographics affect the total number of per capita cash withdrawals $\overline{Q}$ (at ATMs and branches). Cash withdrawals tend to be significantly higher in markets with many elderly, which may indicate that this demographic group does not make use of electronic payments to the same extent. Cash withdrawal demand is significantly lower in the region of Flanders, and the other demographics do not play a statistically significant role. The implied value of $\overline{Q}$ (at sample means) is precisely estimated at 1.28 in the first and 1.21 in the second specification.[25] Note that this is significantly lower than the estimate of 2.07 cash withdrawals per capita and per month, available from our external sources.

Simultaneous demand and investment model Table 6 displays the maximum likelihood estimates from the full simultaneous equations model (19), consisting of the demand or usage equation (11) and the investment equation (16)-(17). As discussed earlier, the model allows for correlation between the unobserved shocks affecting both demand and investment, thereby accounting for endogeneity and selection issues associated with N. We first verified that demand is concave at the estimated parameters in all markets. Hence, the inequalities (16)-(17) are both necessary and sufficient for optimal investment.

First, consider the *demand parameters* (α, $v_B - v_A$ and β). Several parameters differ substantially from the estimates of the single equation model. Most notably, for both the specification without and with market characteristics, the estimate of α is almost three

[24]Using (11), (9) and (8), this elasticity is given by $E_A = \frac{\partial Q_A(N)}{\partial N}\frac{N}{q_A} = \frac{\alpha}{2}\frac{p_A}{s_A}\sum_i w^i s_A^i(1 - s_A^i)$.

[25]Using (10), this is simply computed as $\overline{Q} = \exp(X\beta)$, evaluated at the sample means for X.

times smaller (and it is also more precise). This translates into lower elasticities of ATM cash demand with respect to the number of ATMs, i.e. estimates (at sample means) of 0.65 in both specifications. Intuitively, the elasticity estimates of 1.09 and 0.89 in the single equation demand model only capture the correlation between N and Q_A, which may be high if unobserved demand shocks induce banks to invest in many ATMs. The lower elasticity estimates of the simultaneous equation model capture the causal effect of N on Q_A, which is what we are interested in when making welfare comparisons.

To further appreciate how the simultaneous equations model corrects for the endogeneity of N, consider the estimated correlation between the structural demand and cost errors η_1 and η_2, as computed from σ_1, σ_2 and σ_{12}.[26] These are relatively low in both specifications (respectively, -0.14 and -0.06) and insignificant in the second specification. However, they translate into highly significant negative correlations between our econometric errors $\varepsilon_1 = \eta_1$ and $\varepsilon_2 = \eta_2 - \eta_1$, of respectively -0.73 and -0.69. Intuitively, the demand error η_1 enters both the demand and the investment equation, so banks take into account that a high ATM demand shock η_1 also implies high marginal profits from investing in ATMs. Hence, both demand Q_A and N will tend to be high, which is properly accounted for through the covariance parameter σ_{12}. The single demand equation ignored this covariance, resulting in an overestimation of α.

The parameter $v_B - v_A$ is positive and significant in both specifications: other things equal, consumers prefer a cash transaction at a branch over one at an ATM. This is as expected since consumers can combine a branch visit with several other services that are not necessarily available at shared ATMs. Finally, the second specification in Table 6 again shows that the total number of cash withdrawals $\overline{Q}$ is especially high among elderly and outside Flanders. The implied estimate of $\overline{Q}$ (at sample means) is now equal to 1.71 and 1.93 in the two respective specifications. While these numbers are still somewhat lower than the estimate of 2.07 from our external sources, the underestimation is no longer statistically significant in contrast with the single equation demand estimates.

Second, consider the *investment (or entry) parameters* (γ), as shown in the second part of Table 6. The first specification without demographics includes a highly significant intercept $\gamma_0 = 8.26$, which translates in a precise estimate of the ratio of the fixed cost over variable cost savings, i.e. $F/(c_B - c_A) = \exp(\gamma_0) = 3,876$, with a 95% confidence interval of $[3,556; 4,197]$. The second specification with demographics implies a very similar ratio of $3,932$ (at sample

[26] Given the mean-zero bivariate normal distribution of ε_1 and ε_2, the structural errors $\eta_1 = \varepsilon_1$ and $\eta_2 = \varepsilon_1 + \varepsilon_2$ have a mean-zero bivariate normal distribution with variances of σ_1^2 and $\sigma_1^2 + \sigma_2^2 + 2\sigma_{12}$, and a covariance of $\sigma_1^2 + \sigma_{12}$. The correlation between η_1 and η_2 then equals $\dfrac{\sigma_1^2 + \sigma_{12}}{\sigma_1 \sqrt{\sigma_1^2 + \sigma_2^2 + 2\sigma_{12}}}$.

means). To assess whether this ratio is reasonable, we can use fixed and variable cost information from our external sources, discussed earlier in Section 2.1. According to the network operator Banksys the monthly fixed costs of an ATM are €2,300. Our estimated ratio $F/(c_B - c_A)$ then implies that the variable cost savings from cash withdrawals at ATMs instead of branches amount to €0.59 and €0.58 in the two respective specifications. This is of a similar order of magnitude as Kimball and Gregor's (1995) estimated variable cost savings of $0.80.

To evaluate the fit of the full simultaneous equations model, we compare the model's predicted number of ATMs with the observed number in each market. To predict N, we take a large number of draws of ε_1 and ε_2 for each market (100 draws). For each market and each draw, we compute the joint profit maximizing N based on the parameter estimates and equilibrium condition (17). For each market, we then compute the average of N over all draws, and take this as the predicted N for the given market. Similar to Berry and Waldfogel (1999), we then compute the correlation between the predicted and the observed number of ATMs for the markets. This is equal to 0.78 and 0.81 in the models without and with demographics, implying an R^2 of, respectively 0.60 and 0.65.

Sensitivity analysis Our discussion focused on two specifications, based on the sample of 659 non-urban markets. We also estimated the model using the complete sample of 842 markets, i.e. including the urban areas. Tables A1 and A2 in the Appendix show that most of the parameter estimates are similar, and they do not generate different qualitative conclusions. We have also considered a sensitivity analysis regarding some specification assumptions, which will be discussed in Section 7.

6 Policy counterfactuals

We use our parameter estimates to compare the status quo situation of coordinated entry and zero retail fees with several alternative scenarios with positive retail fees and/or subsidies per ATM. We ask how these alternative scenarios influence ATM investment (a highly sensitive political issue because of a perceived lack of geographic coverage), ATM demand or usage, and total welfare.

6.1 Approach

We compute the predicted number of ATMs, the number markets without an ATM and the various welfare components (consumer surplus, producer surplus and government revenues) under various scenarios. The first one is simply the status quo scenario, in which banks coordinate ATM investment, charge zero retail cash withdrawal fees at branches and ATMs and obtain no subsidies per ATM. The second scenario is the social optimum or first-best solution. As discussed in Section 3, the first-best can be obtained in two ways: in a centralized way by maximizing total welfare (6) with respect to the number of ATMs N and cash withdrawal fees t_A and t_B; or in a decentralized way through welfare-maximizing retail fees and subsidies while allowing banks to coordinate on N given these fees and subsidies. The third and fourth scenarios are the second-best "fees-only" and "subsidies-only" scenarios, where the social planner optimally chooses either fees or subsidies but not both, and banks subsequently coordinate on N.

The status quo predictions follow the approach used for computing the model's fit as described in the previous section. The predictions of the other three scenarios are similar but slightly more involved. To illustrate, we explain the approach of the third scenario, where the social planner optimally sets the retail cash withdrawal fees, but maintains zero subsidies. For each market we take a large number of draws of ε_1 and ε_2 (i.e. 100 draws).[27] For each market and each draw we take a possible fee structure (t_A, t_B), we compute the joint profit maximizing number of ATMs $n^*(t_A, t_B, 0)$, based on the equilibrium condition (17), and compute total welfare $W(n^*(t_A, t_B, 0), t_A, t_B, 0)$. We then search over alternative (t_A, t_B) to find the fees that maximize total welfare. We repeat this approach for each market and each draw to obtain optimal fees and the implied welfare components for each market and each draw. We subsequently compute summary information across markets on the optimal retail fees and/or subsidies, on ATM investment and usage, and welfare. We present both the means and the standard errors from our 100 draws.

This approach assumes that the social planner can set optimal fees specific to each market. In reality, it would be more reasonable to assume that a regulator sets a uniform fee for all markets. We also followed that approach and obtained very similar results. However, we prefer to present the results from the optimal market-specific fees since this provides a sharper economic intuition and a clear-cut benchmark for first-best.

Since our empirical model assumes that total cash withdrawal demand $\overline{Q}$ is inelastic, optimal welfare only depends on the difference $t_B - t_A$ and not on the levels of t_A and t_B separately. This facilitates the exposition and especially the calculations as it is only necessary to search over the difference $t_B - t_A$.[28]

[27] We take these draws from the normal distribution. This differs from Berry and Waldfogel (1999), who take draws from a truncated normal distribution such that the status quo is perfectly predicted. We also followed their approach and obtained similar results.

[28] Concretely, in our third scenario with optimal fees and no subsidies we search over 201 possible values

To perform our counterfactuals we need some identification assumptions. The empirical model only identified the ratio $F/(c_B - c_A)$. Based on our external estimate, we assume that $F =$€2,300, implying that $c_B - c_A =$€0.59. The results were similar for higher values of fixed costs with correspondingly higher variable cost savings. The empirical model does also not identify the travel cost per unit of distance k (in €/km) from the price parameter α. In our welfare analysis, a high assumed value of k implies a high weight to consumer surplus relative to producer surplus. We therefore use two alternative values $k =$€0.1 and $k =$€0.25. The higher value is a commonly used by companies and tax authorities to reimburse travel costs. The lower value roughly corresponds to Gowrisankaran and Krainer's (2007) estimate of ATM travel costs (using a different model and data). We focus our discussion on the results for k=€0.25, and present the results for k=€0.1 as a robustness check in Table A3 of the Appendix.

6.2 Results

Status quo The first column of Table 7 shows the 1994 status quo predictions, when fees and subsidies are zero. The predicted total number of ATMs across all markets is 490 (standard error of 19), which is close to the actually observed number of ATMs of 486. The predicted number of markets without an ATM is 330, again not significantly different from the actual number of unserved markets of 349. Hence, under the status quo over one half of the non-urban markets are unserved by shared ATMs.

The low density of the ATM network is reflected in a low number of per capita ATM cash withdrawals under the status quo. The model predicts monthly per capita ATM cash withdrawals of 0.59, which is close to and not significantly different from the observed number of 0.56. This is only one third of total cash withdrawals at ATMs and branches, hence ATM usage is rather low.

First-best The second column of Table 7 shows the first-best predictions. As discussed, these can be obtained either in a centralized way by choosing N and $t_B - t_A$, or in a decentralized way by setting S and $t_B - t_A$, and subsequently allowing banks to continue to coordinate on N. Table 7 shows that a regulator would like to invest in a much larger shared ATM network: the total number of ATMs across markets under the social optimum is 1018

$t_B - t_A$ in the interval $(-5, 5)$ for each market and each draw. Since there are 659 markets and 100 draws per market, the equilibrium number of ATMs $n^*(t_A, t_B, 0)$ has to be computed over 13 million times. In the first-best scenario, for each market and each draw we in addition consider 53 possible subsidy values S in the interval $(0, 2600)$ for each price difference $t_B - t_A$. This amounts to about 702 million computations of the socially optimal number of ATMs $n^*(t_A, t_B, S)$.

(standard error of 20), which is more than twice as much as the status quo number of 490. The number of unserved markets is almost three times lower, dropping from 330 under the status quo to 114 under the social optimum. Finally, ATM usage is substantially larger in the social optimum: the number of cash withdrawals at ATMs increases from a monthly per capita average of 0.59 to 1.42, amounting to more than two thirds of total cash transactions (at ATMs and branches).

One may therefore conclude that ATM investment and usage have been considerably lower than socially optimal. As discussed in the theoretical framework, the undervestment is due to the fact that the coordinating banks do not take into account the effects of their investments on consumer surplus. The suboptimal ATM usage may be either due to the too low ATM network size or due to the regulatory context with zero fees for cash withdrawals. It is therefore of interest to look at the subsidies and retail fees that implement the social optimum. Table 7 shows that the optimal extra retail fee for cash withdrawals at branches $t_B - t_A$ is on average €0.62, which is essentially cost-based (close to the extra variable costs at branches of $c_B - c_A =$€0.59). The accompanying optimal ATM subsidy S is €2,236, which is also essentially cost-based (close to the fixed costs of an ATM of €2,300).[29]

Total welfare for our sample of non-urban areas in Belgium increases by €2.16 million per month. Banks capture the largest share of the welfare gains: they would receive an additional €6.4 million per month, because of the subsidies, the fee revenues from cash withdrawals at branches and the variable cost savings from consumers substituting to ATMs. The government loses €2.27 million per month, due to the subsidies paid to the banks. Perhaps surprisingly, consumers lose €1.97 million per month despite the much more dense shared ATM network. Their benefits in the form of reduced travel costs to ATMs are overwhelmed by the losses from the fees they have to pay for branch withdrawals.

Second-best How close can one reach to the first-best through either "subsidies-only" or "fees-only"? Consider first the subsidies-only case, shown on the third column of Table 7. This should be interpreted as a policy to directly promote ATM investment, so other instruments such as tax deductions may achieve the same outcome. The optimal subsidy per ATM is on average €1,545, which is about two thirds of the fixed cost of an ATM. This results in a substantial increase in the number of ATMs, from 490 to 1022, very close to the socially optimal number. Similarly, there is a considerable drop in the number of markets

[29] When N is a continuous variable, it can easily be shown that the social optimum requires $t_B - t_A = c_B - c_A$ and $S = F$. This is not necessarily true when N can only take integer values as in our set-up, but the simulation results are nevertheless close. This tendency to cost-based optimal fees and subsidies also explains why the standard errors of the predicted fees an subsidies are so low.

without an ATM, from 330 to 148 markets, which is again close to geographic coverage in the social optimum. However, the monthly number of cash withdrawals at ATMs increases only moderately from 0.59 to 0.84, which is still far below the social optimum of 1.42. As a result, the total welfare gains amount to only €0.59 million per month, far less than the maximum attainable welfare gains of €2.16 million per month under the first-best. Intuitively, this is because the pure promotion of ATM investment is a rather expensive way to promote usage of the cost-saving technology. Note finally that consumers gain €1.03 million per month (in contrast with the consumer losses in first-best): they save on travel costs because of the more dense ATM network, and they do not have to pay retail fees.

Now consider the fees-only case, shown on the final column of Table 7. The optimal extra retail fee for cash withdrawals at branches is on average €0.47 per transaction, slightly below the extra variable cost of about €0.60 for cash withdrawals at branches. The extra retail fee has only minor effects on the banks' investment decisions, hence geographic coverage remains suboptimal. This is because the retail fees make it relatively more profitable to serve customers at branches, thus reducing the variable cost saving incentive from adopting ATMs. However, the fees induce consumers to substitute out of branches and use ATMs more often: the number of cash withdrawals at ATMs increases from 0.59 to 0.88 per capita and per month. These changes result in a monthly increase of total welfare by €1.12 million, largely because of profit increases at the expense of consumers. Interestingly, a policy of raising retail fees without subsidies is thus more effective in improving welfare than a policy of introducing subsidies while keeping fees at zero, despite the fact that the latter policy brings ATM investment closely in line with the first-best. This is because a fees-only policy also induces consumers to substitute to ATM withdrawals, hence realizing variable cost savings without the need of extra fixed cost investments in ATMs.

Note that a fees-only policy is even more effective than a subsidies-only policy if we assume a lower cost per km, k=€0.10, as shown in Table A3 of the Appendix. Intuitively, under a lower k consumers receive less weight in total welfare, so that fees become an even more effective instrument to improve welfare.

Summary The policy counterfactuals show that there is substantial underinvestment in the shared ATM network, implying a too large number of unserved markets. Furthermore, ATM demand or usage is too low; consumers use branches too often to withdraw cash. A policy that combines cost-based cash withdrawal fees and ATM subsidies can achieve the social optimum and raise welfare to a significant extent. A second-best fees-only policy is more effective in raising welfare than the direct promotion of investment through a subsidies-only policy, since it induces consumers to substitute to ATM withdrawals without requiring expensive extra investments in ATMs. However, if geographic coverage is a policy objective per se (because of distributional considerations), a subsidies-only policy may still be preferable to a fees-only policy.

7 Extensions

To assess the robustness of our results, we extended our model in two ways. First, we considered an alternative functional form for our logit ATM market share specification (8). Second, we introduced the possibility of economies of density by allowing the fixed costs per ATM to depend on the number of ATMs in the market.

7.1 Alternative demand specification

Our logit ATM market share specification (8) contained the parameter $v_B - v_A$, the intrinsic utility from withdrawing cash at branches instead of ATMs. As explained in Section 4.1, $v_B - v_A$ is not well identified from β, the parameter vector entering total cash withdrawal demand $\overline{Q}$. The reason is that we only observed ATM demand Q_A and not total cash withdrawal demand $\overline{Q}$. Our approach to this identification problem was to identify the $v_B - v_A$ from β through the non-linearity of the market share specification (8), and subsequently assess whether $\overline{Q}$ as predicted from (10) was close to our country-level estimate of $\overline{Q}$ from an external source. The estimates of our simultaneous demand and investment model showed that this was indeed the case.

To shed further light on the identification issue, we now consider an alternative ATM market share specification. As an alternative to (7), let indirect utility or consumer surplus of a depositor affiliated to bank i take the following form:

$$CS^i(p_A) = y + \frac{1}{\alpha}\left(v_B - \alpha p_B^i + \exp\left(v_A - \alpha p_A - \left(v_B - \alpha p_B^i\right)\right)\right). \tag{20}$$

Applying Roy's identity to (20), a bank i consumer's share of ATM cash withdrawals in total cash withdrawals is

$$s_A^i(p_A) = \exp\left(v_B - v_A + \alpha\left(p_B^i - p_A\right)\right). \tag{21}$$

We refer to this as our semilog specification. Substituting (21) in (9) and (11), we obtain the following specification for ATM transaction demand

$$\ln Q_A = v_B - v_A + \ln \sum\nolimits_i w^i \exp\left(\alpha\left(p_B^i - p_A\right)\right) + X\beta + \eta_1.$$

This shows that $v_B - v_A$ now enters linearly, so that it is clearly not identified from the intercept β_0 in β, not even through the functional form. After estimating the model, we therefore set $v_B - v_A$ and β_0 such that the predicted $\overline{Q}$ is equal to 2.07 for the representative market.

The empirical results for the full simultaneous demand and investment model are shown in Table 8. The first column repeats the earlier results for the logit specification, and the second column shows the results for the semilog specification. A comparison shows that the results are very similar; the same is evidently true for the policy counterfactuals (not shown).

7.2 Economies of density

Our investment specification assumed a fixed cost F per ATM, independent of the number of ATMs in the market. In practice, it is possible that there are economies of density in setting up an ATM network. For example, the network operator's fixed maintenance costs may be lower when there are many nearby ATMs in the same market. Holmes (2007) provides a thorough analysis of economies of density based Walmart's location decisions. To account for economies of density we extend our specification of the ratio of fixed cost over variable cost savings (15) to

$$\ln \frac{F}{c_B - c_A} = W\gamma + \lambda \ln N + \eta_2.$$

If $\lambda < 0$, there are economies of density since an increase in N lowers the fixed cost F per ATM (assuming that $c_B - c_A$ is independent of N).

The empirical results are shown in the third column of Table 8. We indeed find evidence of economies of density, since $\lambda = -0.36$ (standard error of 0.10). Most other parameter estimates are close to those in the first column, where $\lambda = 0$. Since the ratio of fixed cost over variable cost savings is no longer constant, we present the ratio for markets with $N = 1$ and $N = 2$ (covering 90% of the markets with an ATM). The ratio is equal to $4,176$ when $N = 1$ and $3,264$ when $N = 2$, compared with our earlier constant estimate of $3,932$. We also considered policy counterfactuals, continuing to assume a constant subsidy S per ATM. Because of the economies of density, the optimal subsidy per ATM was on average lower than in our baseline case without economies of density, but most other results remained similar.

8 Conclusion

We have analyzed investment and usage in a shared ATM network. Because ATMs are compatible and there are no retail fees, banks have no strategic or revenue motives but only a pure cost-saving incentive for investing in ATMs. Furthermore, because retail fees for cash withdrawals are regulated to zero, consumers may have insufficient incentives to use the available cost-saving ATMs. We developed an empirical model of coordinated investment and ATM cash withdrawal demand, and applied it to the Belgian market in the early nineties. Our results showed that banks substantially underinvested in the shared ATM network because they cannot appropriate all consumer surplus. This contrasts with earlier findings of overinvestment in ATM networks with partial incompatibility due to surcharges. Furthermore, we found that usage of the ATM network is too low because of the zero retail fees for cash withdrawals at branches. A direct promotion of investment (through subsidies or other means) can improve efficiency, but the introduction of proper retail fees on cash withdrawals at branches would be more effective in raising welfare, even if it does not encourage investment per se. Our results stress the importance of both the correct investment incentives to firms and price incentives to consumers.

Our analysis is based on the institutional context of Belgium, with a fully shared network, coordinated investment, and no retail fees for cash withdrawals at branches or ATMs. Our analysis is however also relevant for understanding the situation in many U.S. states before the introduction of surcharges in the mid-nineties. Combining our results with the recent U.S. findings, one may conclude that there has been a shift from a substantial underinvestment to an overinvestment due to the introduction of surcharges. Our analysis is also relevant for the current or recent situation in many other European countries, including larger countries such as France and Italy and smaller countries such as Finland, the Netherlands, Sweden or Switzerland. As discussed, these countries have in common the presence of a single or dominant shared network. However, some of these countries may still differ in specific institutional details, e.g. the level of (non-discriminatory) fees or the extent of coordination of the investment decisions. It would therefore be interesting in future work to apply or extend our framework to learn whether our results of underinvestment in compatible networks can be generalized. More generally, we hope our work will stimulate further research that jointly considers investment and usage of new technologies.

References

Abraham, J., Gaynor, M., Vogt, W. (2005), “Entry and Competition in Local Hospital Markets,” NBER Working paper 11649, 55 p.

Amemiya, T. (1984), “Tobit models: A Survey,” *Journal of Econometrics*, 24, pp. 3-61

Anderson S., De Palma A., Thisse, J. (1992), "Discrete choice theory of product differentiation," MIT press, 422 p.

Arrow, K. (1962), "Economic Welfare and the Allocation of Resources for Inventions," in *The Rate and Direction of Inventive Activity*, ed. R. Nelson, Princeton University Press, 635 p.

Berger, A. (1985), "The Economics of Electronic Fund Transfers," Working paper Board of Governors of the Federal Reserve System.

Bernhardt, D., Massoud, N. (2005), "Endogenous ATM Networks and Pricing," mimeo, 30 p.

Berry, S. (1992), "Estimation of a Model of Entry in the Airline Industry," *Econometrica*, 60(4), pp. 889-917

Berry, S., Waldfogel, J. (1999), "Free entry and social inefficiency in radio broadcasting," *RAND Journal of Economics*, 30(3), pp. 397-420

Bank for International Settlements (1999, 2003), "Payment and settlement systems in selected countries," Committee on Payment and Settlement Systems Red Book

Bresnahan, T., Reiss, P. (1990), "Entry in monopoly markets," *Review of Economic Studies*, 57(4), pp. 531-553

Bresnahan, T., Reiss, P. (1991), "Entry and Competition in Concentrated Markets," *Journal of Political Economy*, 99(5), pp. 977-1009

Cohen, A., Mazzeo, M. (2007), "Investment Strategies and Market Structure: An Empirical Analysis of Bank Branching Decisions," mimeo, Northwestern University, 37 p.

Donze, J., Dubec, I. (2006), "The role of interchange fees in ATM networks," *International Journal of Industrial Organization*, 24, pp. 29-43

Fasig, L. (2001), "This window closed - Fleet touts replacement ATMs as always open," *Providence Journal-Bulletin*, March 17, Business Section, 1B

Felgran, S. (1984), "Shared ATM Networks: Market Structure and Public Policy," *New England Economic Review*, Jan/Feb 1984, pp. 23-38

Gilbert, R., Newbery, D. (1982), "Preemtive Patenting and the Persistence of Monopoly," *American Economic Review*, 72, pp. 514-526

Gowrisankaran, G., Krainer, J. (2007), "The Welfare Consequences of ATM Surcharges: Evidence from a Structural Entry Model," mimeo, 50 p.

Gronau, R. (1974), "Wage Comparisons: A Selectivity Bias," *Journal of Political Economy*, 82, pp. 1119-1143

Hannan, T., Borzekowski, R. (2007), "Incompatibility and Investment in ATM networks," *Review of Network Economics*, 6(1), pp. 1-15

Hannan, T., Kiser, E., Prager, R., McAndrews, J. (2003), "To Surcharge or Not to Surcharge: An Empirical Investigation of ATM pricing," *The Review of Economics and Statistics*, 85(4), pp. 990-1002

Hayashi, F., Sullivan, R., Weiner, S. (2003), "A Guide to the ATM and Debit Card Industry," Federal Reserve Bank of Kansas City.

Heckman, J. (1978), "Dummy Endogenous Variables in a Simultaneous Equation System," *Econometrica*, 46(4), pp. 931-959

Holmes, T. (2007), "The Diffusion of Wal-Mart and Economies of Density," mimeo, University of Minnesota, 36 p.

Humphrey, D. (1994), "Delivering Deposit Services: ATMs versus Branches," *Federal Reserve Bank of Richmond Economic Quarterly*, pp. 59-81

Humphrey, D., Willeson, M., Bergendahl, G., Lindblom, T. (2003), "Cost Savings from Electronic Payments and ATMs in Europe," Working paper Federal Reserve Bank of Philadelphia.

Ishii, J. (2005), "Interconnection Pricing, Compatibility, and Investment in Network Industries: ATM Networks in the Banking Industry," mimeo, Stanford University, 50 p.

Kimball, R., Gregor, W. (1995), "How Distribution Is Transforming Retail Banking: Changes Leading Banks Are Making," *Journal of Retail Banking Services*, 17(3), pp. 1-9

Knittel, C., Stango, V. (2004), "Incompatibility, Product Attributes and Consumer Welfare: Evidence from ATMs," NBER Working paper 10962, 43 p.

Knittel, C., Stango, V. (2006), "Strategic Incompatibility in ATM Markets," NBER Working paper 12604, 31 p.

Kolesar, P., Blum, E. (1973), “Square Root Laws for Fire Engine Response Distances,” *Management Science*, 19(12), pp. 1368-1378

Mankiw, N., Whinston, M. (1986), “Free Entry and Social Inefficiency,” *RAND Journal of Economics*, 17, pp. 48-58

Manuszak, M., Moul, C. (2006), “Prices and Endogenous Market Structure in Office Supply Superstores,” forthcoming in *Journal of Industrial Economics*

Matutes, C., Padilla, J. (1994), “Shared ATM networks and banking competition,” *European Economic Review*, 38, pp. 1113-1138.

Mazzeo, M. (2002), “Competitive Outcomes in Product Differentiated Oligopoly,” *Review of Economics and Statistics*, 84, pp. 716-728

McAndrews, J. (1991), “The Evolution of Automatic Teller Machine Networks,” *Federal Reserve Bank of Philadelphia Business Review*, May 1991

McAndrews, J. (2003), “Automated Teller Machine Network Pricing: A Review of the Literature,” *Review of Network Economics*, 2(2), pp. 146-158

OECD (2007), “OECD Communications Outlook 2007,” OECD, 315 p.

Pakes, A., Porter, J., Ho, K., Ishii, J. (2006), “Moment Inequalities, and Their Application,” mimeo, 79 p.

Retail Banking Research Ltd. (2005), “Study of the Impact of Regulation 2560/2001 on Bank Charges for National Payments,” Retail Banking Research London, 66 p.

Smith, H. (2007), “Efficiency Gain from Ownership Deregulation: Estimates for the Radio Industry,” mimeo, Oxford University, 29 p.

Snellman, H. (2006), “Automated Teller Machine network market structure and cash usage,” Bank of Finland Scientific Monographs E38, 105 p.

Stoneman, P. (2002), “The Economics of Technological Diffusion”, Blackwell, 317p.

Watson, R. (2007), “Product Variety and Competition in the Retail Market for Eyeglasses,” mimeo, University of Texas, 49 p.

Table 1: Variable description (referring to the sample of markets)

ATM withdrawals (Q_A)	monthly per capita number of cash withdrawals at shared ATMs
withdrawal value (V_A/Q_A)	value per cash withdrawal at shared ATMs
number of ATMs (N)	number of shared ATMs
number of branches per bank	number of branches per bank
population (L)	population
surface (M)	surface (in km^2)
enterprises	number of enterprises
foreign	fraction of foreigners in the population
young	fraction of population under 18
elderly	fraction of population over 65
unemployment rate	unemployment rate
Flanders	indicator variable for Dutch-speaking part of Belgium

Table 2: Summary statistics

	all markets		markets with at least one ATM	
	mean	st. dev.	mean	st. dev.
ATM withdrawals (Q_A)	0.56	0.47	0.78	0.37
withdrawal value (V_A/Q_A)	73.21	45.67	101.15	7.14
number of ATMs (N)	0.74	0.97	1.57	0.84
number of branches per bank	0.86	0.64	1.25	0.65
population (L)	8738	7314	13445	7884
surface (M)	36.78	29.25	45.41	32.55
enterprises	1329	2366	1466	2104
foreign	0.04	0.06	0.05	0.06
young	0.22	0.02	0.21	0.02
elderly	0.16	0.02	0.16	0.02
unemployment rate	0.03	0.02	0.03	0.02
Flanders	0.58	0.49	0.63	0.48
number of observations	659		310	

Notes: For a description of the variables, see Table 1. The means for Q_A and V_A $/Q_A$ are population-weighted. Sources: Banksys, N.I.S., B.V.B. and R.S.Z..

Table 3: ATM cash withdrawal demand and number of ATMs

	number of observations	ATM withdrawals		withdrawal value	
		mean	st. dev.	mean	st. dev.
N=0	349	0	0	0	0
N=1	185	0.63	0.31	102.22	7.38
N=2	90	0.83	0.37	101.27	6.97
N=3	23	0.96	0.41	100.35	6.67
N=4	9	1.01	0.32	97.08	6.47
N=5	2	1.13	0.35	97.65	2.68
N=6	1	0.74		92.73	
total	659	0.56	0.47	73.21	45.67

Notes: For a description and summary statistics of the variables, see Tables 1 and 2. The means for ATM withdrawals and withdrawal value are population-weighted. Source: Banksys.

Table 4: Reduced-form demand regressions

	param.	st. err.	param.	st. err.
	ATM withdrawals		withdrawal value	
(log) number of ATMs ($\ln N$)	0.63	(0.07)	−0.03	(0.01)
(log) number of branches per bank	−0.51	(0.07)	0.02	(0.01)
constant	−0.97	(0.69)	4.73	(0.08)
enterprises	−0.44	(1.32)	−0.08	(0.14)
foreign	−0.57	(0.51)	0.09	(0.06)
young	−0.29	(2.10)	−0.50	(0.23)
elderly	4.84	(1.50)	0.03	(0.17)
unemployment rate	−2.63	(2.56)	−1.44	(0.28)
Flanders	−0.14	(0.11)	0.05	(0.01)
R^2	0.37		0.56	
number of observations	310		310	

Notes: For a description and summary statistics of the variables, see Tables 1 and 2. Dependent variables are log of ATM withdrawals ($\ln Q_A$) respectively log of per transaction withdrawal value ($\ln(V_A/Q_A)$). Enterprises is the number of enterprises in the market, divided by 100000.

Table 5: Parameter estimates and predictions from demand model only

	param.	st. err.	param.	st. err.
	demand equation (11)			
α	6.18	(1.05)	5.47	(1.19)
$v_B - v_A$	−0.59	(0.38)	−0.75	(0.48)
constant	0.25	(0.12)	0.08	(0.58)
enterprises			0.08	(0.83)
foreign			0.13	(0.33)
young			−1.07	(1.68)
elderly			3.97	(1.26)
unemployment rate			−3.70	(2.26)
Flanders			−0.33	(0.08)
σ_1	0.39	(0.01)	0.36	(0.01)
	implied demand predictions			
total cash withdrawals ($\overline{Q}$)	1.28	(0.15)	1.21	(0.17)
share of ATM cash withdrawal ($s_A(N)$)	0.56	(0.07)	0.60	(0.09)
elasticity (E_A)	1.09	(0.09)	0.89	(0.09)
number of observations	310		310	

Notes: For a description and summary statistics of the variables, see Tables 1 and 2. Enterprises is the number of enterprises in the market, divided by 100000. Implied demand predictions are at sample means.

Table 6: Parameter estimates and predictions from simultaneous demand and entry model

	param.	st. err.	param.	st. err.
	demand equation (11)			
α	2.30	(0.22)	2.18	(0.22)
$v_B - v_A$	0.51	(0.25)	0.65	(0.27)
constant	0.54	(0.17)	0.28	(0.56)
enterprises			−1.43	(0.85)
foreign			−0.11	(0.29)
young			−0.75	(1.68)
elderly			4.31	(1.19)
unemployment rate			−0.23	(2.32)
Flanders			−0.21	(0.09)
σ_1	0.44	(0.02)	0.40	(0.02)
	implied demand predictions			
cash withdrawals ($\overline{Q}$)	1.71	(0.28)	1.93	(0.35)
share of ATM cash withdrawal ($s_A(N)$)	0.35	(0.05)	0.32	(0.06)
elasticity (E_A)	0.65	(0.02)	0.65	(0.03)
	investment equation (16)–(17)			
constant	8.26	(0.04)	9.22	(0.69)
enterprises			−0.23	(0.96)
foreign			−1.85	(0.49)
young			1.14	(2.08)
elderly			−5.13	(1.60)
unemployment rate			−4.68	(2.50)
Flanders			−0.23	(0.10)
σ_2	0.70	(0.03)	0.61	(0.03)
σ_{12}	−0.22	(0.03)	−0.17	(0.02)
	implied cost predictions			
$F/(c_B - c_A)$	3876	(164)	3932	(158)
R^2	0.60		0.65	
number of observations	659		659	

Notes: For a description and summary statistics of the variables, see Tables 1 and 2. Enterprises is the number of enterprises in the market, divided by 100000. Implied demand and cost predictions are at sample means.

Table 7: Policy counterfactuals

	status quo	first-best	subsidy only	fees only
	optimal subsidies and fees			
average subsidy per ATM S	0	2236	1545	0
		(2.60)	(13.05)	
average fee $t_B - t_A$	0	0.62	0	0.47
		(0.01)		(0.01)
	ATM investment and demand			
total number of ATMs	490	1018	1022	463
	(18.59)	(20.41)	(22.20)	(14.84)
total number of markets without ATM	330	114	148	312
	(10.18)	(7.08)	(7.67)	(10.51)
average share of ATM cash withdrawals	0.25	0.66	0.38	0.39
	(0.01)	(0.01)	(0.00)	(0.01)
average number of ATM cash withdrawals	0.59	1.42	0.84	0.89
	(0.03)	(0.04)	(0.03)	(0.04)
	welfare (in millions of euro)			
change in producer surplus	0	6.40	1.14	4.06
		(0.20)	(0.02)	(0.17)
change in consumer surplus	0	−1.97	1.03	−2.94
		(0.08)	(0.04)	(0.10)
change in government revenues	0	−2.27	−1.58	0
		(0.05)	(0.03)	
change in total welfare	0	2.16	0.59	1.12
		(0.10)	(0.03)	(0.08)

Notes: Number of observations is 659 markets. 100 simulations draws per market. Standard errors are in parentheses. Cost per unit of distance k=€0.25, ATM fixed costs F=€2300, as discussed in the text.

Table 8: Parameter estimates and predictions from simultaneous demand and entry model: extensions

	param.	st. err.	param.	st. err.	param.	st. err.
	base model		semilog		ec. of density	
	demand equation (11)					
α	2.18	(0.22)	1.33	(0.06)	1.94	(0.22)
$v_B - v_A$	0.65	(0.27)	1.17		−0.18	(0.42)
constant	0.28	(0.56)	0.31	(0.55)	−0.08	(0.57)
enterprises	−1.43	(0.85)	−1.64	(0.83)	−1.69	(0.88)
foreign	−0.11	(0.29)	−0.17	(0.29)	−0.07	(0.30)
young	−0.75	(1.68)	−0.71	(1.70)	−1.57	(1.74)
elderly	4.31	(1.19)	4.46	(1.19)	4.57	(1.23)
unemployment rate	−0.23	(2.32)	0.19	(2.33)	0.85	(2.34)
Flanders	−0.21	(0.09)	−0.21	(0.09)	−0.23	(0.09)
σ_1	0.40	(0.02)	0.41	(0.02)	0.41	(0.02)
	implied demand predictions					
cash withdrawals ($\overline{Q}$)	1.93	(0.35)	2.08	(0.07)	1.21	(0.25)
share of ATM cash withdrawal (s_A)	0.32	(0.06)	0.29	(0.00)	0.52	(0.10)
elasticity (E_A)	0.65	(0.03)	0.59	(0.03)	0.41	(0.07)
	investment equation (16)–(17)					
constant	9.22	(0.69)	9.20	(0.66)	9.56	(0.64)
enterprises	−0.23	(0.96)	−0.46	(0.91)	−0.46	(0.90)
foreign	−1.85	(0.49)	−2.03	(0.47)	−1.68	(0.47)
young	1.14	(2.08)	1.04	(1.99)	−0.72	(1.97)
elderly	−5.13	(1.60)	−4.80	(1.53)	−5.07	(1.46)
unemployment rate	−4.68	(2.50)	−4.21	(2.41)	−2.38	2.42)
Flanders	−0.23	(0.10)	−0.26	(0.10)	−0.15	(0.09)
economies of density (λ)	0		0		−0.36	(0.10)
σ_2	0.61	(0.03)	0.61	(0.03)	0.59	(0.03)
σ_{12}	−0.17	(0.02)	−0.18	(0.02)	−0.18	(0.02)
	implied cost predictions					
$F/(c_B - c_A)$ for $N = 1$	3932	(158)	3894	(164)	4176	(157)
$F/(c_B - c_A)$ for $N = 2$	3932	(158)	3894	(164)	3264	(123)
R^2	0.65		0.64		0.65	
number of observations	659		659		659	

Notes: For a description and summary statistics of the variables, see Tables 1 and 2. Enterprises is the number of enterprises in the market, divided by 100000. Implied demand and cost predictions are at sample means.

Appendix: Sensitivity analysis

Table A1: Parameter estimates and predictions from demand model only: full sample of markets

	param.	st. err.	param.	st. err.
	demand equation (11)			
a	6.36	(0.99)	5.50	(1.11)
$v_B - v_A$	−0.30	(0.33)	−0.48	(0.44)
constant	0.39	(0.12)	1.65	(0.38)
enterprises			0.75	(0.37)
foreign			0.62	(0.33)
young			−5.16	(0.94)
elderly			0.33	(1.02)
unemployment rate			−3.85	(1.72)
Flanders			−0.38	(0.07)
σ_1	0.44	(0.01)	0.42	(0.01)
	implied demand predictions			
cash withdrawals ($\overline{Q}$)	1.47	(0.18)	1.35	(0.20)
share of ATM cash withdrawal (s_A)	0.62	(0.07)	0.66	(0.10)
elasticity (E_A)	0.73	(0.07)	0.58	(0.07)
number of observations	467		467	

Notes: For a description and summary statistics of the variables, see Tables 1 and 2. Enterprises is the number of enterprises in the market, divided by 100000. Implied demand predictions are at sample means.

Table A2: Parameter estimates and predictions from simultaneous demand and entry model: full sample of markets

	param.	st. err.	param.	st. err.
	demand equation (11)			
α	2.99	(0.27)	2.70	(0.24)
$v_B - v_A$	−0.08	(0.20)	0.15	(0.21)
constant	0.25	(0.10)	1.49	(0.41)
enterprises			0.93	(0.40)
foreign			0.69	(0.32)
young			−5.55	(1.03)
elderly			1.39	(1.06)
unemployment rate			−1.43	(1.87)
Flanders			−0.30	(0.08)
σ_1	0.51	(0.01)	0.47	(0.01)
	implied demand predictions			
cash withdrawals ($\overline{Q}$)	1.28	(0.13)	1.45	(0.17)
share of ATM cash withdrawal (s_A)	0.55	(0.05)	0.49	(0.05)
elasticity (E_A)	0.43	(0.01)	0.44	(0.01)
	investment equation (16)–(17)			
constant	8.10	(0.04)	9.67	(0.55)
enterprises			−1.77	(0.57)
foreign			−2.06	(0.38)
young			1.38	(1.65)
elderly			−8.25	(1.31)
unemployment rate			−5.54	(1.91)
Flanders			−0.24	(0.09)
σ_2	0.99	(0.04)	0.80	(0.03)
σ_{12}	−0.39	(0.03)	−0.28	(0.03)
	implied cost predictions			
$F/(c_B - c_A)$	3284	(134)	3337	(126)
R^2	0.56		0.62	
number of observations	842		842	

Notes: For a description and summary statistics of the variables, see Tables 1 and 2. Enterprises is the number of enterprises in the market, divided by 100000. Implied demand and cost predictions are at sample means.

Table A3: Policy counterfactuals (alternative cost per unit of distance)

	status quo	first best	subsidy only	fees only
	optimal subsidies and fees			
average subsidy per ATM S	0	2243	1208	0
		(1.18)	(17.39)	
average fee $t_B - t_A$	0	0.62	0	0.53
		(0.01)		(0.01)
	ATM investment and demand			
total number of ATMs	490	700	711	409
	(18.59)	(13.47)	(20.77)	(11.81)
total number of markets without ATM	330	131	234	284
	(10.18)	(8.21)	(10.09)	(10.15)
average share of ATM cash withdrawals	0.25	0.83	0.32	0.62
	(0.01)	(0.01)	(0.01)	(0.01)
average number of ATM cash withdrawals	0.59	1.75	0.72	1.36
	(0.03)	(0.05)	(0.03)	(0.05)
	welfare (in millions of euro)			
change in producer surplus	0	6.54	0.76	5.29
		(0.20)	(0.02)	(0.19)
change in consumer surplus	0	−1.96	0.21	−2.74
		(0.05)	(0.01)	(0.07)
change in government revenues	0	−1.57	−0.86	0
		(0.03)	(0.02)	
change in total welfare	0	3.02	0.11	2.55
		(0.15)	(0.01)	(0.14)

Notes: Number of observations is 842 markets. 100 simulations draws per market. Standard errors are in parentheses. Cost per unit of distance k=€0.10, ATM fixed costs F=€2300, as discussed in the text.

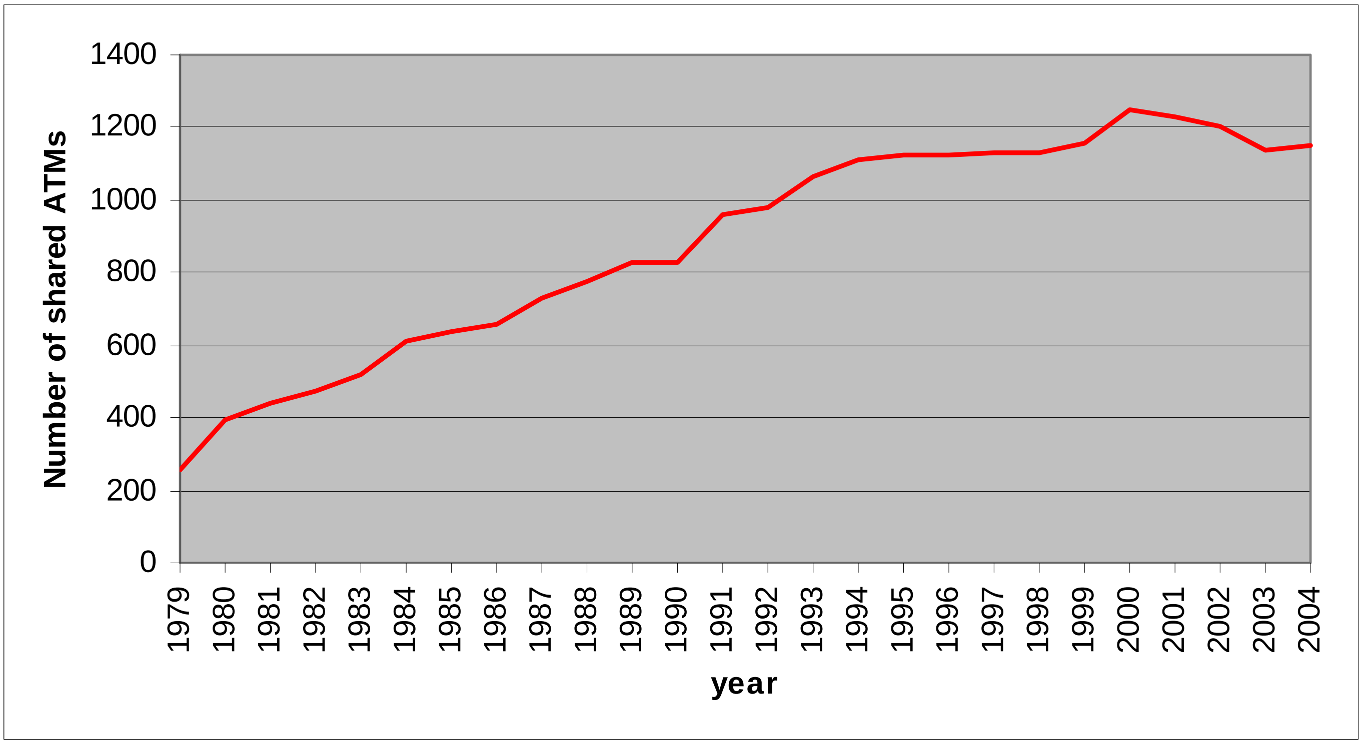

Figure 1: Evolution of the shared ATM network in Belgium (1979-2004)