THE RETURNS TO EARLY ADOPTION OF INFORMATION TECHNOLOGIES: ORDER OF ADOPTION OR LEVEL OF ADOPTION ADVANTAGES?

Abstract

We study whether the early adoption of information technologies provides competitive advantages, and the source of these advantages. Existing research in this area has failed to differentiate between advantages that arise from the timing of adoption (e.g., asset preemption) and those that are related to higher levels of adoption (e.g., organizational learning). In this paper, we break down these advantages into two components: order of adoption (early vs. late adopters) and level of adoption (high vs. low internal diffusion levels). The empirical analysis examines the adoption and subsequent intrafirm diffusion of the automated teller machine in a sample of Spanish savings banks, using a long data panel, from 1988 to 2004. The results show that advantages associated to the order of adoption outweigh those associated to the level of adoption, and they have the potential to be long lasting. Our findings are consistent across various estimation methods, and we assess different performance dimensions (ROA, income, and efficiency). Our modelling and findings have implications for managers and can be applied to the study of the early adoption of modern technologies.

Key words:

First-mover advantage, early adoption, intrafirm diffusion, firm performance, banking sector, ATM

THE RETURNS TO EARLY ADOPTION OF INFORMATION TECHNOLOGIES: ORDER OF ADOPTION OR LEVEL OF ADOPTION ADVANTAGES?

1. INTRODUCTION

In the Information Systems (IS) literature there has been significant interest in understanding how Information Technology (IT) affects firm performance (Bharadwaj, 2000; Santhanam and Hartono, 2003; Chae, Koh and Prybutok, 2014; Kim et al., 2011; Mithas, Whitaker and Tafti, 2017). Substantial advances have been made, leading to improved comprehension of how IT should be analyzed (Bharadwaj, 2000; Wade and Hulland, 2004), its potential to generate and sustain competitive advantages (Piccoli and Ives, 2005; Mithas et al., 2012), and its impact on various aspects of firm performance (Chae et al., 2014).

In this paper, we aim to contribute to the IS literature by exploring how the early adoption of IT might generate competitive advantage. First Mover Advantages (FMA) theory examines the ability of pioneering firms to earn economic rents in excess of followers (Lieberman and Montgomery, 1988, 2013). Most studies on FMA focus on new product introductions and entry into new markets (Fosfuri, Lanzolla and Suarez, 2013, Lieberman and Montgomery, 2013, Zachary et al., 2015, Domínguez, Gómez and Maícas, 2021, Gómez, Pérez-Aradros and Salazar, 2022). However, firms may also achieve first movers' status by being the first in adopting a new technology (Bukchin and Kerret, 2020).

Previous research has recognized this by studying the relative performance of early vs, late adopters of automated teller machines (Pfeffers and Dos Santos, 1996, Dos Santos and Pfeffers, 1995) and digital technologies (Nafizah, Roper and Mole, 2023), for example. But research on early adoption advantages has focused solely on comparing firms based on the first adoption of a new technology (i.e., inter-firm diffusion) and it has not considered the differences between firms that may arise in the internal diffusion process (i.e., intra-firm diffusion). This distinction between the two processes is important in the context of FMA, both, theoretically and empirically. First, Porter (1985) defines FMA in the context of new technologies as those that occur because of an initial technological gap that turns into competitive advantages that persist even after the gap has closed. The mere distinction between adopters and non-adopters (i.e., inter-firm diffusion) of the technology is not a good indicator of whether the gap has closed or not. Inter- and intra-firm diffusion are two different dimensions of the technological gap: "intra-firm diffusion tends to lag behind inter-firm diffusion over the whole of the diffusion process and inter-firm diffusion patterns … may thus be poor indicators of overall diffusion" (Battisti and Stoneman, 2003).

Second, the evidence reveals significant differences in the drivers of inter- and intra-firm diffusion (Battisti and Stoneman, 2003; 2005; Hollenstein and Woerter, 2008, Arvanitis and Ley, 2013, Waters, 2017; Khalifa, 2022) with learning from the use of the technology being a main determinant of the intra-firm diffusion process (Stoneman, 1981; Mcwilliams and Zilbermanfr, 1996; Kalirajan, and Shand, 2001; Battisti and Stoneman, 2003; Battisti and Stoneman, 2005; Hollenstein and Woerter, 2008). Not distinguishing between the two processes would amount to conflating the potential sources of FMA.

Despite these two reasons that justify the analysis of intra-firm diffusion, research on the advantages of the early adoption of new technology (Pfeffers and Dos Santos, 1996, Dos Santos and Pfeffers, 1995; Sinha and Noble, 1997; Nafizah, Roper and Mole, 2023) has not considered that the superior profitability of an early adopter could be attributed to differences in the level of internal diffusion of a technology (i.e., a technological gap not closed). This is surprising, as it is well known that the uneven diffusion of technologies is a key determinant of the productivity gap between frontier and laggard firms (OECD, 2021). "To fully assess the impact of a new technology … assessing the extent of use and the dynamics of the adoption and of the replacement process of the old with the new technology by the adopting firms (the intra-firm diffusion) is as important as observing the number of users (inter-firm diffusion), with the former being the main driver of the realization of the benefits from adoption" (Cave, Waterson and Battisti: 2023: 20). Essentially, if early adopters usually possess higher levels of adoption during internal diffusion, omitting technology implementation levels in early adoption studies could mistakenly attribute FMA to what are differences in firms' internal diffusion processes.

Additionally, research on the advantages of the early adoption of new technology argues that learning from the use of the technology may be behind FMA (Pfeffers and Dos Santos, 1996, Dos Santos and Pfeffers, 1995). However, this research does not recognize that organizational learning is associated with the intra-firm diffusion process (Stoneman, 1981; Mcwilliams and Zilbermanfr, 1996; Kalirajan, and Shand, 2001; Battisti and Stoneman, 2003; Battisti and Stoneman, 2005; Hollenstein and Woerter, 2008). Not controlling for differences in intrafirm diffusion among firms would be failing to control for differences in learning among firms. We argue that if learning takes place *"*…the greatest impact of a new application will only occur sometime after the application is first implemented" (Pfeffers and Dos Santos, 1996: 383). However, the long duration of intra-firm diffusion processes casts doubts on whether differences in learning could be the source of long-term FMA. In fact, "… the time profiles that we observe in the diffusion process stretch to a number of decades, and it is difficult to believe that these learning processes are so slow as to lead to such elongated time scales for adoption." (Battisti and Stoneman, 2003: 1647).

In this context, our article analyses the role of intrafirm diffusion in early adopter advantages. This process is widely observed in IT implementation (Fichman, 2001; Swanson and Ramiller, 2004). One of the stylized facts from the intrafirm diffusion literature is that firms gradually incorporate new technologies into their operations for experimentation, learning purposes, and due to various complexities involved in implementation. This intrafirm diffusion process follows a distinctive time pattern that can help us to better understand the effect of IT on firm performance. As a result of considering the intrafirm diffusion process, the model we propose distinguishes between two different types of advantages for early adopters: *order of adoption advantages*, related to the FMA that early adopters can obtain because of the timing of adoption, and *level of adoption advantages*, related to their level of implementation of the technology. These two advantages are distinct, and their delineation is important to correctly identify the returns to early adoption and the underlying mechanisms.

We apply the model to study the profitability, revenues and efficiency associated with the diffusion of the Automated Teller Machine (ATM) in the Spanish savings banks between 1988 and 2004. This time horizon starts from the year in which each firm had already installed its first ATM terminal and captures most of its evolution until the ATM became an essential technology in the savings bank sector (Bátiz-Lazo, 2007; Consoli, 2005). The ATM has been identified as "The most important financial innovation…" in the financial sector (Volcker, 2009; Shepherd-Barron, 2017). As of 2022, with 3.0 million ATMs worldwide (RBR London, 2023a), this technology remains a pivotal innovation, currently evolving from mere cash dispensers to a multifunctional device that challenges the need for physical bank branches.

Our study contributes to the literature on the consequences of early technology adoption in several ways. First, we propose a model that innovatively examines the returns of early technology adoption, focusing on advantages that persist even after the technology adoption gap among competing firms is closed (Porter, 1985). Unlike prior research on FMA, which overlooks the intra-firm diffusion process, our approach enables a precise determination of whether advantages in the long-run stem from first-mover actions or differential adoption levels. We provide both a novel analysis and a method to quantify the impact of adoption level and adoption order on superior performance. Second, our model contributes to a deeper understanding of the mechanisms behind FMA. While existing literature on FMA (Pfeffers and Dos Santos, 1996; Dos Santos and Pfeffers, 1995; Sinha and Noble, 1997) links organizational learning to FMA, our findings challenge this view. We suggest that in the long run organizational learning may not be the primary driver of profitability differences between early and late adopters, especially once the intra-firm diffusion process concludes. Third, the model sheds light on the potential longevity of early adopter benefits. Advantages linked to intra-firm diffusion diminish as the technology gap closes (Porter, 1985), while those tied to the order of adoption, can persist post-diffusion. Our findings suggest enduring advantages for early adopters even after the internal diffusion process has ended. Finally, our decomposition and findings not only have managerial implications, but we also suggest how they should be used for the analysis of more modern technologies subject to a process of intrafirm diffusion.

2. LITERATURE

2.1 IT and Firm Performance

The impact of IT on organizational performance has been a significant research subject in the IS literature. Scholars have adopted various viewpoints in conceptualizing IT, elucidated different mechanisms to explain its effect on performance, and investigated distinct dimensions of organizational performance influenced by IT. Online Appendix A.1 summarizes a selection of relevant studies on the effect of IT on firm performance.

The resource-based view (RBV) (Barney, 1991; Peteraf, 1993) is a common approach to analyze the impact of IT initiatives on firm performance (Bharadwaj, 2000; Wade and Hulland, 2004; Oh and Pinsonneault, 2007). The RBV examines the essential resources constituting IT initiatives. According to this perspective, IT initiatives comprise three critical resource groups: *IT infrastructure*, which includes the assets forming the material core of the IT on which the initiative is based (e.g., hardware, software, data); *IT human resources*, encompassing the personnel who use the IT and implement the initiative, along with their capabilities, knowledge, and skills; and *IT-enabled intangibles*, referring to various intangible assets developed through the IT initiative, such as organizational culture, relationships with value chain participants, experience, and know-how.

For an IT initiative to be a source of sustainable competitive advantage and superior performance, it must differentiate the firm along relevant competitive dimensions. From the perspective of the RBV, this means that the resources on which it is based must be different from those controlled by other firms (Barney, 1991; Wade and Hulland, 2004). If competitors have similar resources, or resources offering equivalent productive services, they can replicate the IT initiative, negating any competitive advantage and undermining potential superior performance. Thus, in order to achieve and sustain superior performance, it is imperative that (i) the firm creates some form of competitive asymmetry (e.g., by developing its own technology internally or adopting it before competitors) and that (ii) there are ex-post limits to competition that prevent rivals from imitating the resources necessary to implement the IT initiative (Barney, 1991; Peteraf, 2003; Wade and Hulland, 2004).

The IS literature has identified several mechanisms that generate ex-post limits to competition, including *social complexity, erosion barriers, path dependence, and organizational learning* (Mithas et al., 2012; Piccoli and Ives, 2005). Notably, among the three constituents of an IT initiative, the IT infrastructure has weaker mechanisms to prevent imitation. During the early IT development stages (from the 1960s to the early 1990s), IT initiatives were typically built upon proprietary IT infrastructures, generating heterogeneity vis-à-vis competitors. However, with the onset of the Network Computing era in the mid-1990s (Sabherwal and Jeyaraj, 2015), specialized IT providers proliferated, and IT infrastructures became highly commoditized. Consequently, from the perspective of the RBV, the technological core of the IT initiative (i.e., hardware, software, data) cannot confer competitive advantages or superior performance by itself. Instead, firms must create that advantage through developing superior human IT resources and/or IT-enabled intangibles, which, in turn, must be safeguarded by the aforementioned mechanisms (Mata, Fuerst, and Barney, 1995; Wade and Hulland, 2004).

2.2 Early adoption and the profits of IT

Despite the commoditization of IT infrastructures, the literature has argued that firms are able to create competitive asymmetry by being the first to adopt a new technology (Pfeffers and Dos Santos, 1996, Dos Santos and Pfeffers, 1995; Nafizah, Roper and Mole, 2023). Online Appendix A.2 lists a selection of previous studies that have examined the consequences of the early adoption of IT on performance and their characteristics. From these papers, we can conclude that (i) there were early adoption benefits, (ii) the impact of the technology on performance was not immediate, but it took several years for the early adopters to reap the benefits, indicating the need to perform longitudinal studies, and (iii) the main mechanism that the authors propose to explain early adopter advantages is organizational learning, with followers going through a learning process similar to early adopters, which gives early adopters a differential advantage (Peffers and Dos Santos, 1996).

In this paper, we argue that previous research has not considered an essential element of the process of diffusion of new technologies when analyzing early adopter advantages, namely its internal diffusion. We contend that considering intra-firm diffusion is crucial for understanding both whether there are advantages associated with early adoption and to elucidate whether organizational learning (as previous research suggests) or other mechanisms are behind these advantages. We elaborate on this idea based on two sets of arguments.

The first set of arguments relates to the definition of FMAs. Porter (1985: 71) points out that pioneer advantages "allow a leader to translate a technology gap into other competitive advantages that persist even if the technology gap closes." However, previous studies on early adoption advantages have treated early adoption as a single event (Pfeffers and Dos Santos, 1996, Dos Santos and Pfeffers, 1995; Sinha and Noble, 1997; Nafizah, Roper and Mole, 2023), without distinguishing between the adoption and the implementation of the technology (i.e., acquisition and deployment, in the terminology of Fichman and Kemerer, 1999). This is surprising, because the literature on diffusion has long distinguished between inter-firm and intra-firm diffusion (Fichman and Kemerer, 1999; Battisti and Stoneman, 2003).

Intrafirm diffusion refers to the process by which a divisible technology is accumulated over time within a firm, gradually substituting the old technology (Battisti and Stoneman, 2003). Technologies are first implemented in selected organizational subunits (e.g., departments, factories, subsidiaries, offices), and later progressively spread throughout the organization. As a result, the level of adoption varies over time. Intrafirm diffusion processes have been observed for various IT, such as optical scanners (Levin, Levin and Meisel, 1992), electronic mail systems (Astebro, 1995), automated teller machines (Fuentelsaz, Gómez and Polo, 2003, Fuentelsaz, Gómez and Palomas, 2009, 2012), flexible production systems (Battisti and Stoneman, 2005), e-business activities (Battisti et al., 2009), energy-saving technologies (Arvanitis and Ley, 2013), green energy technologies (Stucki and Woerter, 2016) and smart-meters (Strong, 2019).

Considering the process of intra-firm diffusion of the technology is crucial, because inter-firm diffusion is a poor indicator of overall diffusion (Battisti and Stoneman, 2003). It is well known that the acquisition of a technology does not necessarily imply its deployment, leading to what the literature has referred to as assimilation gaps (Fichman and Kemerer, 1999). Furthermore, it is important to highlight that diffusion within firms is the primary determinant of the benefits of adopting new technologies (Fichman and Kemerer, 1999; Battisti and Stoneman, 2003). As a result, failing to consider the intrafirm diffusion process in the analysis of early adoption advantages not only overlooks the main driver of technology returns, but also impedes us to ascertain whether the technological gap has closed and, therefore, it doesn't align with the definition of FMA (Porter, 1985).

The second set of arguments relates to the mechanisms explaining interfirm and intrafirm diffusion. Research has shown that these mechanisms are very different (Battisti and Stoneman, 2003; Battisti and Stoneman, 2005; Hollenstein and Woerter, 2008, Arvanitis and Ley, 2013, Waters, 2017; Khalifa, 2022), with intrafirm diffusion being consistently linked to organizational learning (Stoneman, 1981; Mcwilliams and Zilbermanfr, 1996; Kalirajan, and Shand, 2001; Battisti and Stoneman, 2003; 2005; Hollenstein and Woerter, 2008). An example of this is Stoneman (1981)'s model of intrafirm diffusion, in which Bayesian learning is the main driver of the internal diffusion of the technology and it produces the sigmoid curve that characterizes diffusion studies. In fact, the economics literature argues that to implement IT, organizations must adapt their internal structures, routines, processes, and information systems (McElheran 2015), requiring significant co-invention (Bresnahan and Greenstein, 1996). To effectively address these challenges, firms require substantial doses of organizational learning. Relatedly, the Information Systems literature has also argued that the assimilation gap can be explained by the existence of knowledge barriers (Atewell, 1992) or, in other words, the "effort of organizational learning required to obtain necessary knowledge and skills" (Fichman and Kemerer, 1999: 261).

Previous research on FMA in the context of IT has argued that organizational learning is also the main mechanisms behind the advantages of pioneers (Pfeffers and Dos Santos, 1996, Dos Santos and Pfeffers, 1995; Sinha and Noble, 1997). For example, Pfeffers and Dos Santos (1996) argue that investing in IT involves two forms of learning: (1) 'learning by doing', or first-order learning, where users gradually learn to use an IT to its fullest potential, and (2) second-order learning, which involves changes to the IT itself or to its organizational context.

In line with previous research, we argue that organizational learning is a main driver of intrafirm diffusion and a determinant of performance differences related to the process of internal diffusion of a technology. However, even if as a result of learning "…the greatest impact of a new application will only occur sometime after the application is first implemented" (Pfeffers and Dos Santos, 1996: 383), intrafirm diffusion processes tend to be very long and

"… it is difficult to believe that these learning processes are so slow as to lead to such elongated time scales for adoption." (Battisti and Stoneman, 2003: 1647). Therefore, we also contend that learning processes associated to the diffusion of the technology come to an end. Once the learning process has been exhausted, which likely happens before the end of the intra-firm diffusion process, it's improbable that learning differences between companies would justify early adopter advantages. Since the learning we refer to is associated with internal diffusion, it is crucial to consider the intrafirm diffusion process. If organizational learning is the driver of early adoption advantages, it is improbable that it remains a persistent determinant of these advantages, as it will likely be exhausted as the intrafirm diffusion process progresses.

3. BREAKING DOWN THE RETURNS TO EARLY ADOPTION

3.1 Theoretical Approach

Our model proposes that early adoption can result in two types of advantages: (1) order of adoption advantages and, (2) level of adoption advantages. *Order of adoption advantages* are possible because early adopters face limited or no competition for some time at the beginning of the diffusion process. For example, early adopters could establish strong mechanisms to prevent imitation, such as erosion barriers based on resource pre-emption at the start of the new activity (Piccoli and Ives, 2005; Mithas et al., 2012). These resources may be, for instance, employees with specific IT skills, or preferred locations (Peffers and Dos Santos, 1996). If the technology allows the firm to improve customer service, early adopters may also acquire customers that are subsequently locked in, giving them greater market share and scale economies. "Once this market share position is obtained, customers may incur "switching costs" if they decide to move to another firm" (Dos Santos and Peffers, 1995: 244). Early adopters may also gain a reputation as technology leaders, which creates a favorable opinion among customers (Peffers and Dos Santos, 1996). In terms of the RBV, this means that early adopters could benefit from superior human IT resources and IT-enabled intangibles. Importantly, order of adoption advantages might resist imitation of the IT initiative by rivals and fit with Porter (1985)'s definition of FMA.

Level of adoption advantages represent the second potential source of competitive advantage for early adopters. They are related to inter-firm differences in the process of intrafirm diffusion. A stylized fact in studies of intrafirm diffusion is that it is highly time-dependent (Fuentelsaz et al., 2003). Early adopters enjoy a head start in the deployment of the technology. As a result, they benefit from a technology gap in the form of higher levels of adoption, because their intrafirm diffusion process has been advancing for longer. However, the existence of an intrafirm diffusion process, heavily dependent on organizational learning, opens up an opportunity for early adopters to gain an advantage based on the level of deployment of IT infrastructures. This allows early adopters to significantly advance processes that take a long time to complete, such as assimilation and deployment of the IT infrastructure (Fichman and Kemerer, 1999; Fichman, 2001; Swanson and Ramiller, 2004). Research on IT adoption and implementation has examined several post-adoption processes, including routinization, infusion, assimilation, and intrafirm diffusion (Fichman, 2001; Swanson and Ramiller, 2004). Evidence suggests that firms may differ in the extent to which they successfully complete these processes, leading to assimilation gaps and performance differences among competitors (e.g., Astebro, 1995; Fichman and Kemerer, 1999, 2001).

Therefore, *level of adoption advantages* are related to the time profile of the intrafirm diffusion process, which generates the opportunity for advantages (i.e., leading in the adoption of an IT may bestow temporary advantages based on the IT infrastructure). The wide availability of IT infrastructure providers makes this type of resources easily imitable (Mata et al., 1995). Superior outcomes arise because early adopters progress ahead in the technology implementation process, having started earlier; however, as we discuss below, these differences diminish over time. If, as studies argue, "early followers go through a learning process similar to the innovators" when implementing the technology (Peffers and Dos Santos, 1996: 382), these disparities will have dissipated by the end of the intrafirm diffusion process, rendering the advantage temporary. Furthermore, followers may expedite the learning process, narrowing the gap with early adopters.

In summary, a central idea in our model is that *order of adoption* advantages differ from *level of adoption* advantages. Order of adoption advantages correspond to Porter (1985)'s definition of FMA as those that persist when the technology gap has closed. *Level of adoption advantages* result from the *lead time* that early adopters enjoy deploying the IT infrastructure, i.e., from intrafirm diffusion. These latter advantages result from differences in technology implementation and end once the internal diffusion process has been completed in firms.

Figure 1 illustrates a typical intrafirm diffusion process. This process usually follows a sigmoidal pattern (Fuentelsaz et al., 2003). At the beginning of the process, intrafirm diffusion advances slowly. After a certain point, it takes off, increasing at a higher rate until it approaches the maximum level, when it slows again. Therefore, firms that adopt earlier are expected to show superior adoption levels. However, the typical diffusion path shows that the technology gap narrows over time. For instance, in Figure 1, a one-year lead time results in a greater gap at t_1 (t_1, t_{1+1}) than at t_2 (t_2, t_{2+1}). Consequently, these differences in level of adoption will be transitory.

Figure 1: Standard intrafirm diffusion pattern

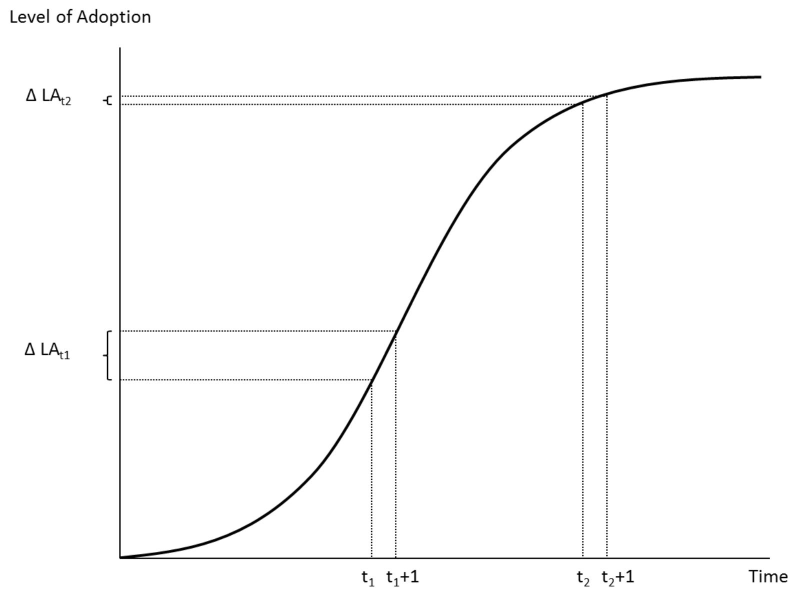

Figure 1 illustrates a typical intrafirm diffusion process. The Y-axis represents the level of adoption, or the extent to which the new IT has been integrated into the operations of the adopting firm. It is measured as units of technology per unit of operation (e.g., percentage of employees using the application, percentage of products incorporating the software, or percentage of units of equipment based on the new technology). This process typically follows a sigmoidal pattern characterized by a slow start, followed by rapid internal diffusion, and ending with a new period of slow diffusion as firms approach saturation. The X-axis represents time. The figure shows how, because of this intrafirm diffusion pattern, a one-year lead time results in a different gap in the level of adoption depending on the stage of the intrafirm diffusion process.

This time profile reveals that early adoption is conducive to higher levels of adoption, resulting in heterogeneous levels of technology deployment among firms at different stages of the diffusion process. In order to understand superior performance among early adopters of an IT, we have to distinguish between the two types of advantages. *Order of adoption advantages* accrue to the few first adopters. Subsequent adopters may still have access to some of them, but the size of these advantages decreases with the number of previous adopters (e.g., specialized employees may be still available, and some markets segment may be still untapped, but greater competition will lead to worse conditions than those enjoyed by the very first adopters). *Level of adoption advantages* stem from differences in the timing of the incorporation of the technology to productive activities. These advantages are not exclusive to the early adopters. Any adopter enjoys these advantages in relation to subsequent adopters, and these advantages are associated with the difference in the deployment of technology between firms.

3.2 Disentangling the components of the returns to early adoption

Figure 2 shows a graphical representation of our conceptual model. On the Y axis, we represent the value generated by each unit of the level of adoption. The distinctive conditions faced by early adopters increase the value that they obtain from the technology (e.g., superior human IT resources or IT enabled intangibles). The position on this axis represents order effects that correspond to conventional FMA (Porter, 1985): Early adopters obtain a higher value per unit of level of adoption than late adopters because of their *order of adoption advantage*s. On the X axis, we represent the level of adoption, which is approximated as units of technology per unit of activity[1]. This dimension is linked to intrafirm diffusion (Battisti and Stoneman, 2003). Early adopters have higher levels of adoption because of their lead time in the intrafirm diffusion process. Therefore, firms may obtain *level of adoption advantages* simply as a result of having begun the intrafirm diffusion process earlier than subsequent adopters.

The model considers the two sources of advantages of early adopters. The returns that a firm obtains from the technology are the product of the value of each unit of the level of adoption (Y axis) multiplied by the level of adoption (X axis). An early adopter (point A) obtains returns equal to the combination of areas I, II, and III (β_{EA} x LA_{EA}). Assuming the presence of FMA and a standard intrafirm diffusion process, point B represents a late adopter (lower returns from each unit and lower level of adoption). Returns to late adoption are equal to area I (β_{LA} x LA_{LA}). The advantage that the early adopter enjoys is the return in excess of the late adopter and corresponds to areas II and III. Importantly, these areas represent different components of the excess returns of the early adopter.

[1] Note that by *level of adoption* we mean *the extent to which IT has been integrated into the operations of the adopting firm*. In our empirical context, it is measured as "*units of technology per unit of operation*". However, previous research on intrafirm diffusion has also used other measures, such as the percentage of production equipment incorporating the technology, the percentage of output derived from the technology, the percentage of departments using the technology, or the number of activities performed using the technology (Battisti and Stoneman, 2003; 2005; Battisti et al., 2009).

Figure 2: Components of returns to early adoption

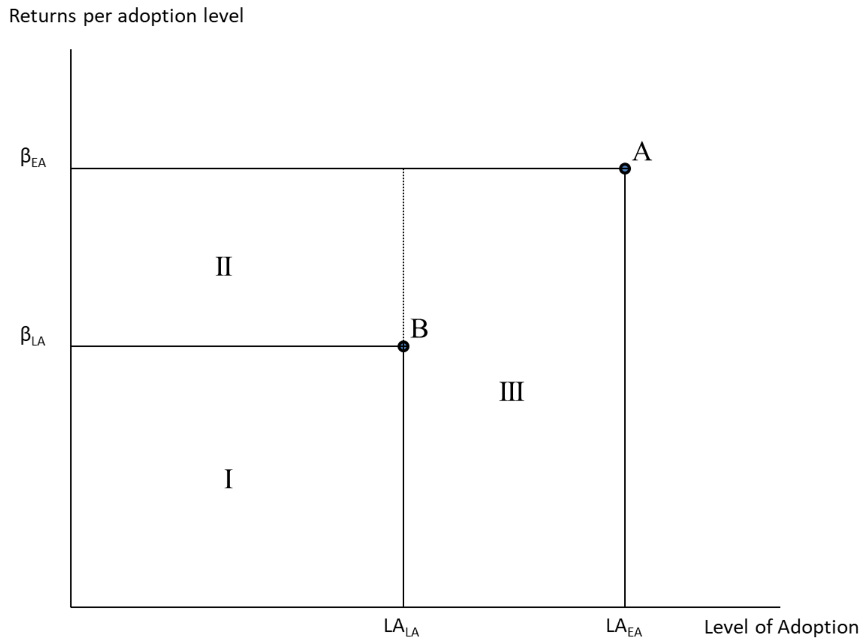

Figure 2 shows a graphical representation of our conceptual model. The Y-axis represents the value generated by each unit of the level of adoption. The X-axis represents the level of adoption. Early adopters are expected to experience both higher levels of adoption, and higher returns per level of adoption. Therefore, A represents an early adopter, and B represents a late adopter. The early adopter has a total return equal to areas I+II+III. The late adopter has a return equal to area I. The early adopter obtains a superior performance equal to areas II + III, which are the order of adoption component and the level of adoption component, respectively.

Area II, calculated as (β_{EA} - β_{LA}) x LA_{LA}, represents the excess returns that accrue to firm A that are obtained strictly from conventional FMA (Porter, 1985). These advantages persist even when the technological gap is closed (i.e., imitators have matched the level of adoption of the early adopters). In other words, returns stemming from conventional FMA are interpreted as the differential in gains that the early adopter would still obtain if its level of adoption were the same as that of the late adopter. The difference β_{EA} - β_{LA} represents the advantage that the early adopter enjoys from each unit of adoption. Area II is the component that does not depend on differences in level of adoption. We label area II as the order of adoption component of the excess returns of the early adopter.

Area III also represents excess returns that accrue to the early adopter. However, they do not fit into a strict definition of FMA (Porter, 1985) because they depend on the technological gap generated by the differences in the level of adoption. It is calculated as β_{EA} x (LA_{EA} –

LA_{LA}). Area III depends on differences in level of adoption, and it is therefore transitory. As late adopters close the technology gap, i.e., as LA_{LA} approaches LA_{EA}, this component will be eroded, and eventually disappear. However, even in the absence of FMA (i.e., $\beta_{EA} = \beta_{LA}$), early adopters would still benefit from this type of excess returns during the time it takes late adopters to close the technological gap. We label area III as the level of adoption component of the excess returns of the early adopter.

To sum up, in this model the advantages are represented by considering the location of the firm on the axes. Order of adoption advantages imply higher positions on the Y axis, while level of adoption advantages imply higher positions on the X axis. Returns obtained from different adoption timing strategies are the result of the product of the two axes. Excess returns depend on the relative positions of different adopters on these axes. The different areas represent the different types of returns to early adoption. Therefore, it is possible to identify the two different components of returns to early adoption and quantify them empirically. More importantly, the model also allows us to distinguish conventional FMA returns (Porter, 1985) from other returns from early adoption whose omission would confound the empirical analysis. In the remainder of the article, we illustrate the model by analyzing FMA in the context of the ATM. We also offer suggestions on how an analysis of pioneer advantages should be conducted in other contexts and with modern technologies.

4. RESEARCH SETTING

4.1. Data and research context

We illustrate the model using a dataset that depicts the diffusion of the ATM.[2] The ATM has been regarded as "…the most important financial innovation …" (Volcker, 2009) and, undoubtedly, it is a technology that has dramatically transformed the operations of banks.

[2] To be consistent with the literature, we distinguish between cash machines (or cash dispensers) and ATMs. "The most common confusion has been to use the terms cash machine and ATM interchangeably, blurring the distinction in their functionality" (Bátiz-Lazo and Reid, 2008: 110). Please see below for more information.

Worldwide, the number of installed ATMs reached 3.0 million in 2022 (RBR London, 2023a). Predicting the evolution of the installed base is challenging, as the COVID-19 pandemic resulted in numerous temporary closures and promoted the adoption of cashless payment methods. Despite this, recent reports indicate that cash usage remains resilient, and the number of ATMs is growing in many countries (RBR London, 2023b).

The dataset employed in this research focuses on the diffusion of the ATM in the Spanish banking sector. The data is publicly accessible and delineate the number of ATM terminals installed by each savings bank.[3] Our analysis encompasses the years 1988 to 2004. Although most data are available since 1986, we started our research in 1988, as this is the first year in which all Spanish savings banks already had adopted the ATM. In this way, we focus on comparing profitability between firms that have adopted the technology and avoid any influence on the results stemming from comparing adopters and non-adopters. In any case, the removal of those two periods makes our results more conservative, since pioneer advantages tend to decrease over time. It is important to clarify that in some of the analyses that we conduct below, we use data from the years 1986 and 1987 on some explanatory variables that are used as internal instruments or as lagged variables, in a way that allows us to lose fewer observations of the dependent variable. Additionally, we gathered information on the adoption of the ATM from 1981 and this allows us to distinguish between early and late adopters.

We conclude our study in 2004, given that the subsequent year witnessed shifts in accounting regulations within the sector, potentially diminishing data comparability. Furthermore, incorporating more recent years would necessitate accounting for the period of

[3] Firm level data are collected by CECA, the Spanish professional association that includes all the savings banks. This organization publishes a yearly report about the sector that provides the financial statements of all the savings banks in Spain and the data on the number of ATMs installed by each bank, among other information. These reports can be accessed online from their website (https://fondohistorico.ceca.es/fondohis/fondos.nsf/WAnuariosF?ReadForm). Data about market and national level variables (e.g., population) are collected from the Spanish National Statistics Institute (http://www.ine.es). The data used in this research are freely available from these sources.

profound consolidation and operational alterations that transpired during the financial crisis, which began in 2007 (Adrian and Shin, 2010). By constraining our observation window to 1988-2004, we assess nearly two decades of ATM diffusion, while circumventing analytical challenges arising from comparing adopters and non-adopters and from atypical economic circumstances. Importantly, as shown by Table 2 (below), this observation window is sufficient to capture the process of intrafirm diffusion up to the point in which there are no statistically significant differences in the level of adoption between early and late adopters.

The ATM "is an expensive, industry-specific piece of capital equipment which also embodies the first step in multi-channel delivery strategies for the provision of retail banking services" (Bátiz-Lazo, 2018: 1). The ATM connects the users with the bank where their checking account reside. Through an ATM terminal, users can execute various fundamental tasks, such as cash withdrawal and account balance inquiries. This technology's origins date back to the late 60's, with the advent of the cash machine (Bátiz-Lazo, 2007). Unlike modern ATMs, the cash machine was stand-alone, analog device, unconnected to the databases of the adopting firm, and consequently did not provide the typical benefits of contemporary ATMs. These early machines functioned akin to vending machines, exchanging previously obtained tokens for fixed, unmodifiable sums of cash (Bátiz-Lazo and Maixe-Altes, 2011; Bátiz-Lazo and Reid, 2008). Database-connected cash machines emerged in the mid-1970s, though they lacked real-time connectivity. Rather, withdrawal information was conveyed after cash delivery had occurred. The first devices offering real-time banking services appeared in the late 1970s (Bátiz-Lazo, 2007: 23; Barras, 1990). Integration with information systems and data repositories proved necessary for advanced service provision. In this study, the term 'ATM' strictly denotes this latter category of devices.

ATMs enhance the delivery of routine banking services to customers, increasing the value provided by the bank. First, ATMs offer 24-hour access, enabling customers to obtain basic services at any time. Second, the technology complements human tellers' activities, boosting branch productivity (Fuentelsaz et al., 2009). Third, ATMs significantly reduce operating costs, with costs per transaction ranging between 28% and 40% of traditional branch costs (European Central Bank, 1999). Fourth, ATM display screens enable adopting banks to present timely, personalized offers to users during transactions. Consequently, this technology delivers notable operational enhancements and facilitates a broad array of strategic initiatives aimed at improving service quality and customer proximity (Consoli, 2005; Bátiz-Lazo, 2007).

We chose this technology for our analysis for several reasons. First, it is a technology with a well-documented intrafirm diffusion process (Fuentelsaz et al., 2003). Second, previous evidence suggests early adopter advantages for this technology, mainly justified in terms of organizational learning (Dos Santos and Peffers, 1993; 1995; Peffers and Dos Santos, 1996; Sinha and Noble, 1997). However, this evidence is based on a conceptualization of early mover advantages that does not account for differences between *order of adoption advantages* and *level of adoption advantages*. Third, the technology has been widely adopted by banks worldwide (RBR London, 2023a), and has been available for an extended period, as required by this type of analysis (Peffers and Dos Santos, 1996).

The first operating ATMs with complete information-based functionalities were installed in the mid-1970s in the United Kingdom, a few years before its introduction in Spain (Bátiz-Lazo, 2007)[4]. Similarly, Sinha and Noble (2005) place the first decade of development of the ATM in the US in the 1970s (see also Hannah and McDowell, 1984, 1987 or Saloner and Shepard, 1995). This means our observation window, 1988-2004, encompasses the technology's second and third decades. By doing so, we reduce the possibility of mistakenly identifying short-run advantages based on the lack of imitation efforts as *order of adoption* or *level of adoption*

[4] Following (Bátiz-Lazo, 2007:16), "Cash dispensers remained stand-alone machines until Lloyds Bank introduced the first 'on-line" dispensers in December 1972. This development marked the move from the cash dispenser to the first ATM in the UK." Nonetheless, "on-line" in this context implied that withdrawal data were sent to the bank after the cash had been withdrawn. This data restriction also implied certain restrictions in operations.

advantages.[5] Crucially, the availability of a lengthy observation window allows for a better evaluation of FMA, as it "is dynamic, not static" (Lieberman and Montgomery, 2013: 316; Peffers and Dos Santos, 1996).

The integration of the technology into the activities of the Spanish savings banks proceeded relatively slowly. In Spain, ATMs are almost exclusively located on-site, meaning that the ratio between the number of ATMs and the number of branches is a good proxy for the intensity of technology adoption. In 1981, Spanish savings banks operated a mere 169 ATM terminals (one per every 52 branches). One year later, this figure had risen to 522 (one for every 18 branches) - still quite low. In 1988, the year in which every savings bank was already operating at least its first ATM terminal, there were a total of 5,609 terminals, nearly one for every two branches. The intensity with which each adopter integrated the technology into its activities increased progressively over the observation window. In 1988, the average level of adoption among savings banks stood at 0.46 ATMs per branch. Sixteen years later, in 2004, the industry's level of adoption had reached 1.41 ATMs per branch.

The technology exhibits relatively minimal variation in its operational principles during the period analyzed in this research. The most used services (i.e., cash withdrawal and account balance inquiries) have been available during the entire observation window, adhering to the same operating principles. Although several technical improvements have expanded the services provided from ATMs and the technology's capabilities, the majority of these changes, which some authors have labelled "ATM 2.0", occurred after our research window concluded (Sodhu, 2011). Finally, it is essential to highlight that international specialized providers market the technology, rendering it accessible to any potential adopter. Therefore, the IT infrastructure

[5] In this sense, our approach is similar to Scott, van Reenen and Zachariadis (2017), who analyze the performance effects of SWIFT adoption between 1998 and 2005. Note that in their case, data on SWIFT adopters is available from 1977 to 2006. However, we have two advantages: (1) information on the intensity of adoption, and (2) a much longer panel.

is highly imitable, making the analysis of intrafirm diffusion and its interrelationship with the returns to adoption particularly relevant.

5. METHODOLOGY

5.1 Identification of level of adoption and order of adoption advantages

Level of adoption advantages stem from greater penetration of the technology into organizational operations. The level of adoption is observable, allowing for empirical verification of these advantages by comparing the level of adoption of early and late adopters. Such advantages exist when early adopters have consistently higher levels of adoption than late adopters. We anticipate these differences to diminish gradually as intrafirm diffusion approaches the technology's saturation level.

Order of adoption advantages pertain to greater returns (e.g., profitability) achieved from the technology at a given level of adoption by early adopters. Profitability derived from a technology is not directly observable, necessitating a statistical analysis to estimate the magnitude of these returns. Order of adoption advantages are confirmed if each category of adopter obtains different returns, with early adopters reaping greater benefits. As with standard statistical analyses of profitability, controlling for various confounding factors is essential.

Our empirical model assumes the form presented in equation (1):

$$ROA_{it} = \theta_1 \left(\frac{ATMs}{Branches}\right)_{it} + \theta_2 \left(\frac{ATMs}{Branches}\right)_{it} \times 1\langle early\ adopter\rangle_i + \sum_{n=1}^{k} \beta_n X_{nit} + \gamma_i + \gamma_t + \varepsilon_{it} \qquad (1)$$

Where ROA represents Return on Assets, the dependent variable (Hannan, 1991), $\left(\frac{ATMs}{Branches}\right)_{it}$ measures the level of adoption, $1\langle early\ adopter\rangle_i$ is a dummy variable identifying early adopters, X represents other explanatory variables, with n denoting the k different variables in the model, i identifying banks and t specifying time periods. Following the structure-conduct-performance paradigm for banking firms (Hannan, 1991; Molyneux and Forbes, 1995), X encompasses variables capturing market power. To differentiate between the effects of market power and efficiency on ROA (Demsetz, 1973; Molyneux and Forbes, 1995), we also include a control for banking firm efficiency (Hannan, 1991). Lastly, X incorporates a control for bank risk-taking (Molyneux and Forbes, 1995; Fuentelsaz et al., 2012). We elaborate on the model by adding ATM and IT specific controls, as discussed below.

The components γ_i and γ_t identify firm and year unobservable fixed components, and ε_{it} represents a random, time varying effect. In subsequent sections, we define the variables, justify the model's specification, and discuss its interpretation.

5.2 Definition of variables

Dependent variable. As our primary dependent variable, we employ an accounting-based financial performance measure: Return on Assets (ROA). This metric has frequently been utilized in previous research on IT returns (Bharadwaj, 2000; Santhanam and Hartono, 2003; Chae et al., 2014). ROA is calculated as the ratio of pre-tax profits to total assets, represented in percentage points.

Independent variables. A key theoretical variable in this analysis is *Level of Adoption*, which measures the intensity of intrafirm diffusion via the ratio of ATMs to branches for each firm. The level of adoption captures the diffusion process within the firm and is preferable to other approaches, such as using a dummy variable that simply distinguishes between early and late adopters, as previous research has done. As discussed above, we anticipate higher adoption levels for early adopters. Utilizing dummies for early adopters (as the papers presented in Online Appendix A.2 have done) would conflate order and level of adoption advantages. To estimate the differential value for early adopters, we interact *Level of adoption* with a dummy variable, $1\langle early\ adopter\rangle_i$, which takes the value of 1 for early adopters and of 0 for late adopters. The parameter on the interaction term represents the excess returns early adopters attain compared to late adopters, aligning with our order of adoption advantage definition.

We establish the adoption date as the earliest instance when the focal savings bank reports an operating ATM terminal, starting in 1981. Early adopters are savings banks already operating ATM terminals by 1981, while subsequent adopters are designated as late adopters[6]. In 1981, Spanish savings banks operated only 169 ATM terminals, or 1 per every 223.000 inhabitants. These firms adopted the technology well before its widespread adoption. At the beginning of our observation window in 1988, this classification reveals 14% of firms as early adopters (11 firms out of 76), and 86% of late adopters (65 firms).

Control Variables. The estimation includes several control variables, namely, *Concentration*, *Firm Size, Branch Size*, *Risk* propensity, *Inefficiency*, *Rural* orientation, *ATM terminals* operated by the focal firm, the *Density of ATMs in the Market* in which the focal firm operates, and the percentage of *IT employees* within each bank.

Concentration is calculated as the Herfindahl Index of the provinces where each savings bank operates, with market shares proxied via branch shares. For banks operating in multiple provinces, the weighted average of each province's Herfindahl Index is calculated, using the number of branches operated by the focal firm on each province as weights. Concentration influences both financial performance (Scherer and Ross, 1990) and ATM adoption propensity (Hannan and McDowell, 1984), necessitating control for this effect.

Firm Size is calculated as the logarithm of total assets, while *Branch Size* is the logarithm of the ratio total assets to total branches. Size is a crucial factor in technology adoption (Astebro, 2004). Larger firms can benefit from cost-spreading and better access to external funding and are more likely to have obsolete equipment to be replaced. Size may impact returns on assets through scale economies and coordination costs, making it essential to control for this variable. Both firm and branch size are introduced, as technical unit's size is more relevant for some factors than the organization's overall size (Astebro, 2004).

[6] 1981 is the earliest period in which we can identify adopters and non-adopters. Using it as a cut-off point we identify 14% of firms as early adopters. If we take 1982 as the cut-off point, we get 54% of firms as early adopters, which seems to be an overestimate. In analyses not shown here, we distinguish three cohorts: early adopters (adopting by 1981, 14% of firms), early followers (adopting in 1982, 40%), and late adopters (adopting from 1983, 46%). These results also confirm that early adopters perform better than other adopters, in line with the model.

Risk is the ratio of total loans to total assets. Higher values indicate greater exposure to potential losses, such as customer bankruptcy (Molyneux and Forbes, 1995). Riskier firms may also adopt new technologies and processes more readily, generating correlation between level of adoption and this variable. *Inefficiency* is the ratio of operating costs to operating margin. Less efficient firms may have greater need for innovation and lower profitability, introducing a negative correlation between financial performance and level of adoption (Hannan, 1991). *Rural* represents the percentage of branches of each bank in towns with less than 10,000 inhabitants, capturing the extent of operations in geographical areas with potentially less developed economic and technological infrastructures. A negative impact on ROA is expected.

ATM Terminals is the number of ATMs operated by the focal firm, and controls for firm-specific network effects. ATMs are subject to network economies (Saloner and Shepard, 1995), and could provide firms with larger installed bases a competitive advantage. In Spain, savings banks' ATMs were interoperable, allowing customers of any savings bank to access other banks' ATMs, sometimes for a small fee. Although interoperability makes network effects at the firm level unlikely, the purpose of this variable is to control for network effects.

Density of ATMs in the Market is calculated as the number of ATM terminals per 1000 inhabitants in the provinces where the focal firm operates. There is no branch or province level information on the distribution of ATM terminals. To calculate this variable, we consider that each savings bank homogeneously distributes its ATM terminals across its branches. This variable captures potential market saturation, where higher density implies lower returns for adopters (Fuentelsaz et al., 2012). It also captures potential positive network spillovers from a high density of interoperable ATMs in markets where savings banks place their branches.

IT employees is the ratio of IT-related employees to total employees. This variable is commonly used in the IS literature to measure firms' investments in IT and assess their impact on productivity. For example, Wu, Jin and Hitt (2017) used this variable to measure the effects of such investments on the sales of over 6000 US firms in the context of a Cobb-Douglas production function. We anticipate a positive effect on the profitability of savings banks.

Lastly, we incorporate year and firm fixed effects. Year fixed effects are introduced to parameterize industry-wide shocks, such as interest rates, regulatory and accounting norm changes, or economic cycles. Technological improvements are treated as industry-wide shocks since they are developed by suppliers and made available to savings banks under similar conditions. We expect year fixed effects to absorb these improvements. Furthermore, as ATMs are interoperable in the Spanish retail banking, industry year fixed effects would capture network effects common to the firms operating in the industry. Firm fixed effects are introduced to capture firm specific non observable characteristics, such as R&D skills or learning capabilities. [7] Consequently, we estimate a two-way fixed effects model, as displayed in equation 1. Descriptive statistics and correlations are shown in Table 1.

[7] We conducted a series of tests to select the appropriate specification of the model. The Breusch-Pagan Lagrange multiplier test, which tests for unobserved heterogeneity at the firm level, rejected the null hypothesis, which is interpreted as evidence of firm-specific effects and the need for panel data techniques (Wooldridge, 2002). Heterogeneity can be modeled as a fixed or random effect. The Hausman test rejected the null hypothesis, supporting a fixed effects specification. There were several mergers during the period. Whenever a merger takes place, we consider the resulting savings bank as a new firm with its own fixed effect. The resulting firm is considered an early adopter only if the largest firm in the merger was an early adopter.

Table 1. Correlations and descriptive statistics

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
Mean	1.09	0.90	0.16	14.17	9.01	0.56	0.57	0.43	339.85	0.46	0.03
Standard Deviation	0.52	0.47	0.08	1.30	0.43	0.13	0.08	0.20	745.79	0.25	0.02
Minimum	-2.32	0.09	0.07	10.21	7.93	0.23	0.33	0.00	2.00	0.04	0.00
Maximum	4.35	3.21	0.53	17.93	10.34	0.89	0.97	0.96	6922.00	1.06	0.10
1. Returns on Assets	1.00										
2. Level of Adoption	-0.02	1.00									
3. Concentration	0.03	-0.07	1.00								
4. Size	-0.07	0.47	0.02	1.00							
5. Branch Size	0.06	0.68	0.05	0.67	1.00						
6. Risk	-0.08	0.42	-0.12	0.06	0.11	1.00					
7. Inefficiency	-0.49	-0.11	-0.03	-0.35	-0.46	0.12	1.00				
8. Rural	0.14	-0.58	0.25	-0.28	-0.35	-0.37	-0.07	1.00			
9. ATM teminals	-0.16	0.38	-0.07	0.61	0.31	0.08	-0.01	-0.32	1.00		
10. Density of ATMs in the market	-0.10	0.56	-0.07	0.26	0.35	0.46	-0.05	-0.28	0.22	1.00	
11. IT Employees	0.04	-0.14	0.09	-0.17	-0.03	-0.17	-0.01	0.10	-0.21	-0.20	1.00

6. RESULTS

Our model investigates the existence of level of adoption advantages and order of adoption advantages. We first examine if early adopters exhibit higher levels of adoption than subsequent adopters. Second, we determine whether, for a given level of adoption, early adopters enjoy greater returns from the technology than subsequent adopters. We then use our model to divide the (expected) superior performance of early adopters into order of adoption and level of adoption advantages.

6.1 Level of adoption differences

Table 2 displays the average level of adoption of each cohort of adopters throughout the observation window and when divided into four different sub-periods. During the observation window, the difference first slightly increases and then decreases, moving from 0.09 (1988-1991) to 0.10 (1992-1999), and finally to 0.07 (2000-2004). Mean comparison tests reveal that differences in the level of adoption between early adopters and late adopters were statistically significant from 1988 to 1999, becoming statistically insignificant by the end of the observation window (period 2000-2004). This suggests that late adopters closed the technological gap by the end of the observation window. These findings align with our predicted evolution of the level of adoption.

Table 2: Levels of adoption

	Early Adopters (1)	Late Adopters (2)	Difference (3)
1988-2004	1.00	0.87	0.13 (0.00)***
1988-1991	0.60	0.51	0.09 (0.03)**
1992-1995	0.91	0.81	0.10 (0.06)*
1996-1999	1.13	1.03	0.10 (0.09)*
2000-2004	1.33	1.26	0.07 (0.16)

Levels of adoption measured as average number of ATM terminals per branch.
In parentheses, the p-value of mean-comparison tests. Two-tailed test of significance: * p<.10; ** p<.05; *** p<.01

Table 3: Estimation results

Return on Assets	(1) OLS-FE	(2) OLS-FE	(3) OLS-FE	(4) 2SLS	(5) 2SLS-GMM	(6) Garen	(7) H-T
Level of Adoption	-	0.192** (0.085)	-0.745 (0.771)	0.036 (0.139)	0.055 (0.137)	0.075 (0.119)	0.211*** (0.076)
Level of Adoption X 1< Early Adopters >	-	0.239** (0.102)	0.184* (0.109)	0.293** 0.114	0.285** (0.111)	0.184* (0.108)	0.247*** (0.092)
Level of Adoption X Firm Size	-	-	0.062 (0.051)	-	-	-	-
Level of Adoption X Rural	-	-	-0.141 (0.252)	-	-	-	-
Level of Adoption X IT employees	-	-	1.669 (1.489)	-	-	-	-
Concentration	-0.501 (0.583)	-0.594 (0.579)	-0.603 (0.580)	0.188 (0.440)	0.224 (0.435)	-0.056 (0.437)	-0.797 (0.551)
Firm Size	-0.729*** (0.167)	-0.513*** (0.175)	-0.609*** (0.194)	-0.458** (0.215)	-0.352* (0.202)	-0.468** (0.198)	-0.318** (0.119)
Branch Size	0.060 (0.162)	-0.131 (0.173)	-0.108 (0.188)	-0.177 (0.210)	-0.239 (0.206)	-0.131 (0.206)	-0.354** (0.149)
Risk	-0.494* (0.285)	-0.394 (0.287)	-0.296 (0.292)	-0.499 (0.304)	-0.516* (0.299)	-0.336 (0.293)	-0.306 (0.257)
Inefficiency	-3.000*** (0.261)	-2.990*** (0.259)	-2.935*** (0.261)	-3.245*** (0.304)	-3.234*** (0.303)	-3.033*** (0.293)	-2.962*** (0.244)
Rural	-0.472** (0.239)	-0.433* (0.236)	-0.296 (0.293)	-0.256 (0.241)	-0.259 (0.238)	-0.441* (0.238)	-0.272 (0.216)
ATM terminals	-0.000 (0.000)	0.000 (0.000)	-0.000 (0.000)	0.000 (0.000)	-0.000 (0.000)	-0.000 (0.000)	-0.000* (0.000)
Density of ATMs in the market	-1.518*** (0.313)	-1.648*** (0.318)	-1.586*** (0.341)	-0.875** (0.392)	-0.771** (0.382)	-1.508*** (0.358)	-1.634*** (0.274)
IT employees	-0.122 (0.686)	0.027 (0.679)	-2.054 (0.202)	0.119 (0.413)	0.005 (0.391)	0.045 (0.523)	-0.070 (0.645)
Year dummies	Yes***	Yes***	Yes***	Yes***	Yes***	Yes***	Yes***
$\eta_{levelofadoption}$	-	-	-	-	-	0.304 (0.326)	-
$\eta_{levelofadoption}$ **X Level of Adoption**	-	-	-	-	-	0.026 (0.217)	-
Return on Assets (1970)	-	-	-	-	-	-	9.181 (0.296)
Firm Size (1970)	-	-	-	-	-	-	0.000 (0.000)
Accounts per branch (1970)	-	-	-	-	-	-	-0.000 (0.000)
Founding date	-	-	-	-	-	-	0.000 (0.002)
1<Early adopter>	-	-	-	-	-	-	0.724 (0.757)
Constant	-	-	-	-	-	-	1.023** (4.913)
F (or $\chi 2$)	15.89***	15.50***	14.03***	12.86***	12.92***	13.88***	481.11***
R2	0.35	0.36	0.37	0.35	0.34	0.36	-
Observations	902	902	902	748	748	878	883
Kleibergen-Paap rk Wald F statistic	-	-	-	89.82	89.82	-	-
Hansen's J (p-value)	-	-	-	0.278	0.278	-	(Sargan test) 0.98
Hausman test (p-value)	-	-	-	-	-	-	0.98

Robust standard errors in parentheses. Two-tailed test of significance: * p<.10; ** p<.05; *** p<.01

6.2 Order of adoption advantages

Order of adoption advantages entail differences in profitability between early and late adopters. This profitability cannot be directly observed, necessitating estimation. Table 3 presents the results of our two-way fixed effects estimations.

The first column displays the baseline model containing only control variables. Column 2 introduces the main theoretical variables. In this model, the variable *Level of Adoption* positively impacts firm performance (θ_1=0.192, $p<0.05$). The effect of its interaction term with the early adopter dummy variable is positive and significant (θ_2=0.239, $p<0.05$).[8] It is important to remember that this parameter represents the profitability bonus obtained by early adopters compared to late adopters. This result implies that early adopters achieve statistically significantly higher returns from each unit of adoption than late adopters, consistent with the existence of order of adoption advantages, as outlined by Porter (1985). The total benefit attained by early adopters by adding one ATM in every one of their offices is the combination of these two parameters, equating to 0.431 percentage points (0.192 + 0.239 = 0.431).

6.3 Components of the returns to early adoption

In this section, we quantify the components of return to early adoption and analyze their evolution between 1988 and 2004 using the predicted values of the models presented in Table 3 (Column 2). These returns are measured as percent points of ROA. Table 4 presents the breakdown, displaying total returns obtained by early adopters from the ATM, their excess returns, and the two components discussed in our model. The order of adoption component corresponds to area II in Figure 2, representing excess returns independent of level of adoption differences. The level of adoption component corresponds to area III in Figure 2. Each cell in Table 4 reveals the absolute value in percentage points and the relative importance of each

[8] Note that we do not include the dummy for early adopters as a direct effect because this is a type of firm-specific time constant effect. The inclusion of fixed effects in the estimation already controls for this effect. We explore other sources of endogeneity in section 6.5 below.

component with respect to the total excess returns (in parenthesis). Total returns are calculated as the product (β_{EA} x LA_{EA})[9]. Excess returns are calculated as the difference between the early adopter returns and late adopter returns (β_{LA} x LA_{LA}). This analysis is conducted for the entire observation window and each sub-period, using the values from Sections 6.1 (Table 2) and 6.2 (Column 2 in Table 3).

As shown in Table 4, the excess returns of early adopters grew over time. This is because early adopters enjoy a greater return from each unit of the level of adoption, and the level of adoption grows constantly during the observation window. These excess returns grew from 0.160 percent points in 1988-1991 [(0.431 x 0.60) - (0.192 x 0.51)] to 0.331 in 2000-2004 [(0.431 x 1.33) - (0.192 x 1.26)].[10]

The order of adoption component of the returns is calculated as the excess return that early adopters would obtain if they had the same level of adoption as late adopters [(β_{EA}-β_{LA}) x LA_{LA}]. We can "predict" or "simulate" such a scenario by taking the values we obtain in our estimates of the order of adoption advantages (Table 3, Column 2), and the actual level of adoption of late adopters (Table 2). As can be seen in Table 4, the absolute value of this component rose during the observation window. The order of adoption component grew from 0.122 percent points in 1988-1991 [(0.431-0.192) x 0.51] to 0.301 in 2000-2004 [(0.431-0.192) x 1.26]. The relative weight of the order of adoption component on excess returns increased its importance from 75.8% to 90.9% of total excess returns.

The level of adoption component is calculated as excess returns that depend on differences in level of adoption, that is, area III, [β_{EA} x (LA_{EA}-LA_{LA})]. This component first increased from 0.039 to 0.043 percent points, and then fell to 0.030 percent points. In relative terms, the component varied from 24.2% to 9.1% of excess returns. This is because, as firms reach the

[9] β_{EA} refers to the returns that the early adopter obtains from each level of adoption unit, according to our estimations in Section 6.2. LA_{EA} refers to the level of adoption of the early adopter. Subscript LA refers to late adopters.

[10] To predict, we take the data from Table 3 (column 2) and Table 2.

upper limit of their intrafirm diffusion processes, differences in level of adoption tend to dissipate, reducing the importance of the level of adoption component.

Table 4: Decomposition of the returns to early adoption

	Total Returns	Excess Returns	Order of Adoption Component	Level of Adoption Component
1988-2004	0.433	0.266 (100%)	0.208 (78.1%)	0.058 (21.9%)
1988-1991	0.258	0.160 (100%)	0.122 (75.8%)	0.039 (24.2%)
1992-1995	0.392	0.236 (100%)	0.193 (81.8%)	0.043 (18.2%)
1996-1999	0.487	0.289 (100%)	0.246 (85.1%)	0.043 (14.9%)
2000-2004	0.573	0.331 (100%)	0.301 (90.9%)	0.030 (9.1%)

6.4. Further analyses

There are a number of issues that could affect the value that the savings banks derive from ATMs and affect our results[11]. First, *Firm Siz*e is often associated with reputation and the presence of other complementary assets that might enhance the value of the technology. Secondly, in Spain ATMs are predominantly located in bank branches, so firms with more urban branches (i.e., lower values in the *Rural* variable) could have more profitable locations than those with branches in rural areas. Third, *IT employees* have frequently been used as a proxy for IT investments in the past, allowing us to assess the presence of other IT investments within the firm, which may act as specialized assets. Consequently, these three factors may increase the value of the ATM for early adopters, without being actually a consequence of early adoption. This could artificially inflate our estimations of the effect of early adoption.

[11] We thank an anonymous reviewer for bringing this to our attention.

We consider these three effects in a new specification. With this new specification we can explore whether the effect of early adoption (θ_2) captures the influence of other relevant firm characteristics that are not the result of early adoption but increase its value. In particular, we propose the following equation:

$$ROA_{it} = \theta_1 \left(\frac{ATMs}{Branches}\right)_{it} + \theta_2 \left(\frac{ATMs}{Branches}\right)_{it} \times 1\langle early\ adopter \rangle_i + \theta_3 \left(\frac{ATMs}{Branches}\right)_{it} \times Firm\ Size + \theta_4 \left(\frac{ATMs}{Branches}\right)_{it} \times Rural + \theta_5 \left(\frac{ATMs}{Branches}\right)_{it} \times IT\ employees + \sum_{n=1}^{k} \beta_n X_{nit} + \gamma_i + \gamma_t + \varepsilon_{it} \quad (2)$$

In Equation 2, the total effect of level of adoption is the sum of a direct effect (θ_1) that is obtained by any adopter, the effect of early adoption (θ_2), and the additional effects of a set of variables not explicitly modeled as components of the effect of level of adoption in Equation 1. While these variables were controlled for in the baseline model, we now investigate whether they also influence the profitability of each unit of level of adoption, potentially confounding our analysis of order of adoption advantages.

The results are presented in Table 3, column 3. As can be observed, none of the interactions are statistically significant. The parameter capturing order of adoption advantages remains positive and statistically significant, and its change respecting the model in column 2 is not statistically significant. Additionally, the inclusion of these interactions does not improve the explanatory power of the model. This finding rules out that the advantages derived from the order of entry are due to factors related to reputation (associated with *Firm Size*), location preemption (associated with *Rural*), or complementary and specialized assets (associated with *Firm Size* and *IT employees*).

Another effect to be further explored is the existence of network effects (Saloner and Shepard, 1995). As already mentioned, in the case of Spain all ATMs have been part of the same network since their introduction and have been interoperable. Therefore, savings banks should not enjoy individual network effects. Instead, there should be an industry-level effect, and as such, be captured by year fixed effects. It could also be argued that aspects such as customer loyalty or the existence of a small fee when using an ATM from another bank allow individual-level network effects to be generated. This effect could then be modeled through the number of ATMs controlled by each bank (variable *ATM terminals*). These two possibilities are already accounted for in the main estimations. There is, however, an additional way in which network effects could benefit savings banks: the value of the technology for a firm could be greater the larger its installed network of ATMs. Thus, for the same level of adoption, firms with more ATMs installed would be more profitable[12]. To analyze networks effects of this form, we interacted *Level of Adoption* with the number of ATMs of each firm (*ATM terminals*). The results of this estimation are shown in Online Appendix A.3. As can be seen, no significant effect is found for the interaction. The conclusions of the main model hold.

6.5 Analyses of endogeneity

Our analysis of order of adoption advantages may be subject to endogeneity in two distinct ways. First, there could be variables not explicitly included in the model that affect both the decision to adopt an IT and its effect on firm performance, creating an artificial relationship or introducing a bias in our estimates. This is known as omitted variable bias. Second, the decision to be a pioneer is exposed to self-selection (Cirik and Makadok, 2021). This situation also generates omitted variable bias, although there are methods to assess its existence and correct for its effect. Therefore, we also explore this potential source of endogeneity.

We start by estimating the model through two-stage least squares (2SLS). This method allows to control for endogeneity of unknown form, so it is suitable for any of the possible sources of endogeneity that we have pointed out. We use internal instruments (Chung et al., 2019). We use three lags of the potentially endogenous variable *Level of Adoption*, and of its interaction with the early adopter dummy. Lagged values are appropriate instruments because they cannot be associated with unanticipated shocks in the dependent variable in the current

[12] We thank an anonymous reviewer for bringing this to our attention

period, which is (theoretically) sufficient to satisfy the orthogonality condition. The results of the first stage are available in Online Appendix A.4. We perform two tests to formally check for the appropriateness of the instruments. First, we obtain the Kleibergen-Paap rk Wald F statistic. It takes a value of 89.82, well above the thresholds recommended by Stock and Yogo (2005). Therefore, the instruments are sufficiently correlated with the potentially endogenous variables, i.e., they fulfil the relevance condition. Second, we calculate the Hansen's J statistic. Under the null hypothesis, the instruments are not correlated to the error term in the main equation, i.e., they fulfil the orthogonality condition. The null hypothesis cannot be rejected (p-value = 0.278). According to these tests, the instruments can be considered appropriate. The results (Table 3, column 4) confirm the existence of order of adoption advantages.

To reduce the concern that the choice of instruments is conditioning the results, we present two alternative approaches in Online Appendix A.5. In the first, we use the second and third lags of the *Level of Adoption*. In the second, we instrument the *Level of Adoption* with the average level of adoption of the industry, excluding the focal firm (Chung et al., 2019). The results are similar to those presented in the main estimations.

We find heteroskedasticity in the 2SLS model (Pagan and Hall= 49.304, $p<0.05$). In this case, the Generalized Method of Moments (GMM) estimator is more efficient than 2SLS (Baum, Schaffer and Stillman, 2003). Therefore, we use a GMM estimator, the two-stage heteroskedastic least squares (H2SLS) described by Davidson and McKinnon (1993). It is a version of the 2SLS which incorporates the GMM. In this estimation, the first stage is the same as in the 2SLS and, therefore, we utilize internal instruments (Chung et al. 2019). As in the 2SLS estimation, the instruments also satisfy the relevance and orthogonality conditions. The results of the GMM estimation are shown in Table 3, column 5. These results again support the existence of an order of adoption advantage for early adopters.

It is important to note that 2SLS and GMM estimates trade consistency for efficiency, i.e. these estimates are expected to be unbiased, but less precise than OLS estimates. Therefore, it is advisable to test for endogeneity. To do this, we use the C-statistic to test for the endogeneity of the two potentially endogenous regressors (i.e., *Level of Adoption*, and its interaction with the early adopter dummy). Under the null hypothesis, OLS estimates, which are more efficient, are preferred, and endogeneity is not a concern. We cannot reject the null hypothesis of the test (χ^2=2.42, p=0.30). We therefore conclude that endogeneity does not bias our estimates, and the OLS estimations are preferred (those in column 2).

These methods (2SLS and GMM) correct the endogeneity problem by replacing the potentially endogenous variables in the main equation with an estimate of these variables that is uncorrelated with the error term of the main equation. An alternative way of interpreting the endogeneity problem is through self-selection bias. Although this problem should theoretically be corrected by the above estimations, we apply Garen's (1984) method, which is specifically designed to correct for this form of endogeneity (see, Saldanha et al. 2017; Chung et al. 2019 or Chen et al., 2021, for recent applications).

Garen (1984) proposed a two-stage approach, which is a generalized version of Heckman's selection model that can be applied to continuous selection variables. We treat *Level of Adoption* as the selection variable. Following previous research using panel data (Chung et al, 2019:1089), in the first stage (i.e., the selection equation) we use the lagged value of the potentially endogenous variable as an instrument, along with the contemporaneous values of all other independent variables. The lagged value is used to fulfil the exclusion restriction and is therefore not included in the second stage. The results of this first stage are reported in Online Appendix A.6. We then calculate the residuals of this first stage ($\eta_{\text{levelofadoption}}$). In the second stage, these residuals are introduced as a control variable to account for endogeneity due to selection bias. Also, their product with the potentially endogenous variable ($\eta_{\text{levelofadoption}}$ x *Level* *of Adoption*) is included to account for unobserved heterogeneity across the range of values of the potentially endogenous variable (Saldanha et al. 2017; Cheng et al., 2021). The results of this second stage are reported in Table 3, Column 6. The results confirm that early adopters enjoyed higher returns from their level of adoption. Interestingly, both the residuals and their product with the potentially endogenous variable are not statistically significant in the second stage, suggesting that endogeneity is not a concern.

We also tested the robustness of our results to self-selection bias through the Hausman and Taylor (1981) estimator (Pan et al., 2023). This method has been previously used to analyze the advantages of early movers in the context of new product introductions (see, for example, Boulding and Christen, 2008; 2009). This approach has three main features: (1) it is able to include time invariant variables because it is not a fixed effects model; (2) it provides the coefficients associated with time invariant variables (i.e. the early adoption dummy); and (3) it allows considering both $1\langle early\ adopter\rangle_i$ and *Level of Adoption* endogenous. Column 7 of Table 3 shows the results of this estimation.[13] The results are similar to those in the main estimations. In all the models, both the Hausman and Sargan tests fail to reject the null hypothesis ($p > 0.10$), suggesting model consistency and instrument validity (Greene, 2003; Hausman and Taylor, 1981; Sargan, 1958).

Finally, we assess the extent to which our results are sensitive to omitted variables bias. We use the Impact Threshold of Confounding Variables (ITCV) to calculate the maximum value of partial correlation that an omitted variable should have to both our dependent variable,

[13] To estimate the Hausman-Taylor model we need variables that are correlated with the decisions made by savings banks regarding the adoption of the ATM. We use the *Founding Date* of the Savings Bank and three other exogenous variables: *Return of Assets (1970)*, which measures ROA in 1970, *Firm Size (1970)*, measured as total assets in 1970, and *Accounts per branch (1970)*, calculated as number of total accounts in 1970 divided by the number of branches in the same year. These variables are extracted from "*Anuario Español de los Bancos, Bolsas y Cajas de Ahorros*", published in 1971. We choose these variables because they predate 1972, the year in which Lloyds Bank introduced the first "on-line" cash dispensers in UK, and consequently are completely exogenous to the adoption of the ATM. These variables identify characteristics of savings banks which may influence their decisions about the timing and intensity of adoption, but should have no impact on their profitability during our observation window. The estimation considers both the condition of early adopter and *Level of Adoption* as endogenous variables.

ROA, and our main theoretical variable, the level of adoption of early adopters, to generate a spuriously significant result (Busenbark et al., 2022; Xu et al., 2019). We obtain an ICTV of 0.168. Following the advice provided by Busenbark et al., (2022), we compare this threshold to our set of control variables, and we identify that only one of them shows enough partial correlation (*Inefficiency*). One control variable over the threshold implies moderate risk of omitted variables bias. In practical terms, for an omitted variable to change our conclusions, it should have a greater combined effect on firm profitability and the level of adoption than variables such as *Firm Size*, *IT employees*, or *Rural*. Given that our various endogeneity robust estimation methods are consistent with our main estimates, that the C-statistic suggests that no correction for endogeneity is necessary, and that in Garen's approach the results suggest that endogeneity is not biasing our results, we consider that our main estimates are reliable.

6.6 Performance dimensions

Research on the impact of IT on performance has been interested in the sources of the advantage created by the adopted technology, paying attention not only to profitability, but also to revenues and cost (e.g., Bharadwaj, 2000; Santhanam and Hartono, 2003; Kim et al., 2011; Chae et al., 2014; Mithas et al., 2012, 2017). The results are mixed, with some predominance of findings that suggest that IT initiatives have a more consistent effect on revenues and profitability than on costs (Mithas et al., 2012, 2017). We explore the origins of the competitive advantage created by the early adoption of ATM by estimating its effect on revenue generation (operating income) and cost savings (inefficiency). Operating revenues may increase because ATM allow firms to increase maintenance fees for credit account holders, charge fees to users of other banks, and increase the reach of products taken out by their customers. Efficiency may improve because the ATM allows banks to reduce the operating costs they incur when providing retail banking services.

To study the effect of early adoption on revenues and efficiency we use two-way fixed effects models analogous to those in our main estimates, taking as dependent variable a measure of operating income (operating income normalized by firm size) and a measure of inefficiency (the one described in section 5.2). We also test simple models without controls, but with year and firm fixed effects to check the robustness of our results. In addition, we explore level-level and log-level relationships between the theoretical variables and the operating income and inefficiency measures. Our results, shown in Table 5, suggest that early adopters enjoyed an operating income advantage compared to late adopters. This advantage is robust to the different model specifications we tested (see Columns 1 to 4 of Table 5). However, we do not observe an advantage in the case of efficiency, as the *Level of Adoption* is not significant in any of the estimations (see, in this case, Columns 5 to 9 of Table 5). It should be noted that the overall fit of the income models is better than that of the efficiency models.

7. CONCLUSIONS

Discussion of findings. In this paper, we present a model that identifies the advantages of early adoption of new technologies, as defined by Porter (1985). Our results show the existence of advantages derived from the early adoption of the ATM, aligning with previous research. However, unlike prior studies, we distinguish between the order of adoption and level of adoption advantages, isolating and quantifying the former.

Table 5: The effect of early adoption in incomes and cost

	Operating Incomes/size		Ln(Operating Incomes/size)		Inefficiency		Ln(Inefficiency)	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Level of Adoption	-0.074 (0.145)	0.020 (0.138)	0.013 (0.024)	0.009 (0.023)	0.016 (0.013)	0.012 (0.012)	0.030 (0.023)	0.022 (0.021)
Level of Adoption x 1< Early Adopters >	0.344** (0.174)	0.409** (0.176)	0.083*** (0.029)	0.085*** (0.029)	-0.023 (0.015)	-0.021 (0.015)	-0.039 (0.027)	-0.032 (0.027)
Concentration	-2.894*** (1.057)	-	-0.577*** (0.180)	-	-0.188** (0.082)	-	-0.310** (0.145)	-
Firm Size	-1.414*** (0.301)	-	-0.125** (0.051)	-	0.015 (0.026)	-	0.033 (0.046)	-
Branch Size	0.476 (0.299)	-	0.022 (0.050)	-	-0.081*** (0.026)	-	-0.153*** (0.045)	-
Risk	1.812*** (0.495)	-	0.276*** (0.084)	-	-0.097** (0.042)	-	-0.199*** (0.075)	-
Inefficiency	-2.315*** (0.457)	-	-0.356*** (0.078)	-	-	-	-	-
Rural	-0.499 (0.416)	-	0.005 (0.071)	-	-0.058* (0.034)	-	-0.116* (0.061)	-
ATM Terminals	0.000 (0.000)	-	-0.000 (0.000)	-	0.000 (0.000)	-	0.000 (0.000)	-
Density of ATMs in the market	0.086 (0.548)	-	-0.079 (0.093)	-	0.155*** (0,047)	-	0.298*** (0.083)	-
IT employees	1.092 (1.220)	-	0.268 (0.208)	-	0.107 (0.098)	-	0.204 (0.173)	-
Year dummies	Yes***	Yes***	Yes***	Yes***	Yes***	Yes***	Yes***	Yes***
F	468.72***	620.51***	290.49***	391.24***	5.50***	5.95***	5.72***	5.93***
R2	0.94	0.93	0.91	0.90	0.16	0.12	0.17	0.12
Observations	902	902	902	902	902	902	902	902

Robust standard errors in parentheses
Two-tailed test of significance: * p<.10; ** p<.05; *** p<.01

Our findings are novel, revealing that early adopters benefit from order of adoption advantages; that is, they derive higher value from each ATM adopted compared to later adopters. We demonstrate that the benefits of early adopters are not linked to the technology's intrafirm diffusion process. The durability of these advantages, even after the intrafirm diffusion process concluded, refutes the notion that they arise from differences in levels of technology deployment among companies, and suggest that they do not stem from organizational learning. Instead, our analysis indicates that the observed advantages stem from other factors. In the empirical analysis conducted in Section 6, we explored the role of reputation, location, specialized assets and network effects. None of these mechanisms seem to be the driving force behind the advantages associated with the order of adoption.

Our results also reveal the existence of advantages related to the level of adoption. Since intrafirm diffusion proceeds slowly, early adopters have a head start in the diffusion process. Differences in the level of adoption are significant for understanding the returns from early adoption during the diffusion process, as they correlate with disparities in operational and financial performance (Devaraj and Kohli, 2003; Fuentelsaz et al., 2009; 2012). In previous research, these differences have been linked to processes associated with technology implementation, such as organizational learning. The extent of this advantage depends on the S-shaped intrafirm diffusion process. We expected it to be temporary and to diminish as intrafirm diffusion progresses and adoption levels near the upper limit of the sigmoidal pattern, as it is the case. By the end of our observation period, late adopters catch up with early adopters, neutralizing this advantage.

In decomposing the excess returns to early adoption, we have found out that the order of adoption component significantly surpasses the level of adoption component. Not only is the latter smaller, but it also declines in importance over time. This finding is intriguing, as conceptually, FMA are expected to wane over time (Makadok, 1998; Lieberman and

Montgomery, 2013; Zachary et al., 2015). We had foreseen a decrease in both the size and relative significance of the level of adoption component. However, we had no initial reason to assume that the level of adoption component would be small compared to the order of adoption component. In different contexts (different IT initiatives), the relative importance of these components may be reversed, with the level of adoption component being the main determinant of returns to early adoption.

It is important to note that, in addition to profitability, our paper also explores the effects of early adoption on efficiency and revenue. In this regard, our findings are also consistent with previous literature (e.g., Mithas et al., 2012, 2017), which observes an effect of information technologies on revenue, but not on efficiency.

Overall, our results demonstrate the importance of technology adoption strategy on firm performance. It should be noted that in our sample, the effects of early adoption are still evident many years later, even though we are dealing with a technology that has now been widely accepted and extensively used by both banks and consumers. Our findings suggest enduring advantages for early adopters even after the internal diffusion process has ended.

Contributions. Our first contribution is the proposal of a model measuring the returns to early adoption of IT, specifically the excess returns that early adopters obtain compared to late adopters. In the model, early adopters receive i) *order of adoption advantages*, and ii) *level of adoption advantages*. While *order of adoption advantages* accrue only to early adopters, *level of adoption advantages* are obtained by every firm in relation to subsequent adopters, and therefore they do not fit into a strict definition of FMA (Porter, 1985). We show how the model can be used to analyze both types of advantages independently and their combined effect on early adoption strategy returns. It is important to highlight that the model can be applied to the analysis of the early adopter advantages of more recent technologies, as we explain below.

Secondly, it is important to highlight that the use of the model and the decomposition of early adopter advantages into two components allow us to advance in identifying the mechanisms that might be behind these advantages. This distinction has not been made in previous papers (Pfeffers and Dos Santos, 1996; Dos Santos and Pfeffers, 1995; Sinha and Noble, 1997) and is significant because it can help future research to place more or less emphasis on the study of specific mechanisms. For instance, in our case, it suggests that the advantages of early adopters are not associated with organizational learning or other processes related to intrafirm diffusion. On the contrary, our results suggest that the advantages derive from the firm's position in the order of adoption because it enjoys limited or no competition in the use of technology for a period of time. Some factors that might explain such advantages could include preemption of critical resources, such as employees with specific IT skills or preferred locations or customer switching costs (Lieberman and Montgomery, 1988; Gómez and Maícas, 2011). Early adopters may also gain a reputation as technology leaders, which creates a favorable opinion among customers (Peffers and Dos Santos, 1996).

Limitations. Our research has some limitations. Firstly, our analysis provides clues about the mechanisms that may explain the benefits derived from the early adoption of new technologies. However, we must acknowledge that it has not enabled us to identify specific mechanisms. We cannot rule out that organizational learning has been significant during the time that the internal diffusion process of the technology lasts, although, as explained, its effects should disappear in the long run. Our results suggest that factors not related to intrafirm diffusion, such as reputation or switching costs, may have been important in explaining the order of entry advantages we observe. Nonetheless, the mechanisms we have studied, namely reputation, location, investments in specialized assets and network effects were not significant. The analyses on efficiency and revenue that we have conducted indicate that, in our case, the advantages seem to come from the revenue side and not from efficiency. Future research should aim to identify such mechanisms. The literature on pioneer advantages provides clues on how this could be achieved by identifying mediating mechanisms between the order of entry and the attainment of a superior return by pioneers (please, see Gómez and Maícas, 2011).

Second, we have considered the deployment of technology, but not its use by the firms. As we have previously mentioned, some articles distinguish between different concepts related to the intra-firm diffusion process, including routinization, infusion, assimilation, and intrafirm diffusion (Fichman, 2001; Swanson and Ramiller, 2004). This could be a problem, for instance, if the customers of some banks tended to use ATMs more than those of another bank.

Third, companies choose the timing of adoption. Therefore, the early adopter condition is potentially subject to self-selection bias. If this were the case, we would face a endogeneity problem, common in studies that have analyzed pioneer advantages. In our case, the potential endogeneity problem has been addressed using various methods (2SLS, GMM, Garen's method, and Hausman-Taylor estimator). Naturally, the ideal scenario would involve companies being randomly assigned to early adopter and follower groups. Although not impossible, it is difficult to imagine real-life examples where this could occur (Cirik and Makadok, 2021). Another limitation is the possibility that our model may be incorrectly specified, leading to biases from omitted variables. Our efforts have focused on including firm-specific variables that could minimize this issue. Furthermore, the methods we have used (e.g., fixed effects) also account for the possibility of omitted variables.

Implications and future research. Our paper has managerial implications. First, the paper highlights how the level and timing of adoption impact performance, offering insights for executives. In particular, it emphasizes that closing the technology gap does not mean closing the profitability gap between pioneers and followers and shows the relevance of the timing of technology adoption. Second, it suggests that relying solely on organizational learning may not suffice to maintain long-term advantages gained from early adoption and that there are other mechanisms involved. Therefore, executives should focus on other factors, such as the ones highlighted before (e.g., preemption of scarce assets or switching costs), and that are available at the beginning of the diffusion process. Third, it shows that the advantages of early adoption may endure and have an impact on performance over the long run.

The model we present here is useful for future research because it provides a framework that can be used to develop similar studies in other IT. However, a clarification is needed on the concept of intrafirm diffusion, which is central to our model. Some of the early IT systems were usually embedded in divisible production equipment with a physical presence. This would be the case, for example, with computers, optical scanners, ATMs, or CNC machines. In these cases, the concept of intrafirm diffusion applies naturally. In fact, this type of IT has been a common context for research on intrafirm diffusion (Battisti and Stoneman, 2003; 2005; Fuentelsaz et al., 2003; Levin et al., 1992). In recent years, there are still abundant cases where IT comes embedded in productive equipment, such as autonomous cleaning robots, autonomous internal transportation systems, or self-service kiosks (see, for instance, Scherer, Wunderlich, and von Wangenheim, 2015). In these cases, the application of our model based on order and level of adoption level advantages is straightforward. The physical support (i.e., IT infrastructure) is easily obtainable, but the need to deploy the technology, together with the possibility of obtaining advantages based on IT-enabled intangibles and IT human resources, generate a context in which our proposal should be applicable to analyze the type of advantages obtained.

However, in some of the most relevant IT in recent years, the physical element is not so clearly divisible. This would be the case with cloud computing, artificial intelligence or blockchain, to name but a few. This makes the concept of intra-firm diffusion somewhat more confusing. However, there is research that has applied the concept of intrafirm diffusion in similar contexts. For example, Battisti, Canepa and Stoneman (2009), analyzing e-business, approximate the concept of intrafirm diffusion through the adoption of different IT functionalities. Astebro (1995), analyzing electronic mail systems, approximates intrafirm diffusion as the number of users within the firm. Thus, even in cases where IT is implemented as a unit for the entire organization, it is possible to analyze the process by which the technology is progressively implemented, which would make it possible to extend our model.

The question is then to identify the appropriate way to approach this implementation process. In the literature, there are concepts related to IT implementation such as routinization, infusion, assimilation, or intrafirm diffusion (Fichman, 2001; Swanson and Ramiller, 2004). These processes have been shown to be highly correlated, and it is common for progress in one to require high levels in another. These processes depend on organizational learning and adaptation, which tend to result in assimilation gaps (e.g., Astebro, 1995; Fichman and Kemerer, 1999, 2001). Moreover, they are associated with the impact of IT on firm performance (Devaraj and Kohli, 2003).

Once the appropriate concept is chosen, the problem is to identify an appropriate measure. It is important to note that these measures should be conceived as stock measures, and it is not sufficient to consider only the order of adoption or the effort to implement the technology in each period. Furthermore, it is important to emphasize that we should make sure that we observe the entire process of assimilation of the technology so that we can determine whether the technological gap (Porter, 1985) is closed. The fact that intra-firm diffusion processes tend to be long, suggests that empirical studies of early adopter advantages need to use long panel data sets that allow for the correct identification of order-of-entry advantages. This information is rarely available, and this is what makes this type of study so difficult and unique.

Additionally, it is essential to consider that the mechanisms potentially driving the advantages of early adoption might differ depending on the technology analyzed. For example, the advantages associated with organizational learning that occur during the internal diffusion process may not be very relevant in cases where the technology is simple and not subject to significant knowledge barriers. However, learning may be more important in the case of more complex technologies. Regarding the mechanisms associated with the order of entry, we will encounter technologies that are only used internally. In other cases, besides being part of a business process, the technology may interact with the consumer, generating switching costs, or influencing the firm's reputation.

REFERENCES

Adrian, T., Shin, H. S. 2010. The Changing Nature of Financial Intermediation and the Financial Crisis of 2007–2009. ***Annual Review of Economics*** 2(1):603-618.

Anuario español de los bancos bolsas y cajas de ahorros. 1971. Editorial Maestre. Madrid.

Arvanitis, S. and Ley, M. 2013. Factors determining the adoption of energy-saving technologies in Swiss firms: An analysis based on micro data. ***Environmental and Resource Economics*** 54:389-417

Astebro, T. 1995. The effect of management and social interaction on the intrafirm diffusion of electronic mail systems. ***IEEE Transactions on Engineering Management*** 42:319-331

Astebro, T. B. 2004. Sunk costs and the depth and probability of technology adoption. ***Journal of Industrial Economic*** 52:381-399.

Attewell, P. (1992). Technology diffusion and organizational learning: The case of business computing. ***Organization science***, 3(1), 1-19.

Barney, J. 1991. Firm resources and sustained competitive advantage. ***Journal of Management***, 17(1), 99-120.

Barras, R. 1990. Interactive innovation in financial and business services: The vanguard of the service revolution. ***Research Policy*** 19: 215-237

Bátiz-Lazo, B. 2007. Emergence and evolution of proprietary ATM networks in the UK. ***MPRA Paper No. 3689*** (Online at http://mpra.ub.uni-muenchen.de/3689/).

Bátiz-Lazo, B. 2018. *Cash and Dash: How ATMs and Computers Changed Banking*. Oxford University Press. Oxford, United Kingdom.

Bátiz-Lazo, B. Reid, R. J. 2008. Evidence from the patent record on the development of cash dispensing technology. In **History of Telecommunications Conference, 2008**. HISTELCON 2008. IEEE, pages 110–114. IEEE, 2008

Bátiz-Lazo, B., Maixé-Altés, J. C. 2011. Organisational Change and the Computerisation of British and Spanish Savings Banks, circa 1950-1985. **Technological Innovation in Retail Finance: International Historical Perspectives**. Bernardo Bátiz-Lazo, J. Carles Maixé-Altés and Paul Thomes (eds.). London and New York, Routledge, pp. 137-154

Battisti, G., Canepa, A., Stoneman P. 2009. E-Business usage across and within firms in the UK: profitability, externalities and policy. ***Research Policy*** 38:133-143

Battisti, G., Stoneman, P. 2003. Inter- and intra-firm effects in the diffusion of new process technology ***Research Policy*** 32:1641-1655

Battisti, G., Stoneman, P. 2005. The intra-firm diffusion of new process technologies ***International Journal of Industrial Organization*** 23:1-22.

Baum, C. F., Schaffer, M. E., & Stillman, S. (2007). Enhanced routines for instrumental variables/generalized method of moments estimation and testing. ***The Stata Journal***, 7(4), 465-506.

Bharadwaj, A. S., Bharadwaj, S. G., Konsynski, B. R. 1999. Information technology effects on firm performance as measured by Tobin's q. ***Management Science***, 45(7), 1008-1024.

Bharadwaj, A. S. 2000. A resource-based perspective on information technology capability and firm performance: an empirical investigation. ***MIS Quarterly***, 169-196.

Boulding, W., Christen, M. 2008. Disentangling pioneering cost advantages and disadvantages. ***Marketing Science*** 27: 699–716

Boulding, W., Christen, M. 2009. Pioneering plus a broad product line strategy: higher profits or deeper losses? ***Management Science*** 55(6):958–967

Bresnahan, T., Greenstein, S. M., 1996. Technical Progress and Co-Invention in Computing and in the Use of Computer. ***Brookings Papers on Economics Activity: Microeconomics***,1-83

Bukchin, S., & Kerret, D. (2020). Once you choose hope: early adoption of green technology. ***Environmental Science and Pollution Research***, 27, 3271-3280.

Busenbark, J. R., Yoon, H., Gamache, D. L., Withers, M. C. 2022. Omitted variable bias: Examining management research with the impact threshold of a confounding variable (ITCV). ***Journal of Management***, 48(1), 17-48.

Cave, J. A., Waterson, M., & Battisti, G. (2023). Economic Impact of Significant New Deployment of Infrastructure: Historical examples and links to potential high impact outcomes for 5G.

Chae, H. C., Koh, C. E., Prybutok, V. R. 2014. Information technology capability and firm performance: contradictory findings and their possible causes. ***MIS quarterly***, 38(1), 305-326.

Chen, L., Hsieh, J. J. P.-A., Rai, A., and Xu, S. X. (2021). The Impact of Service Employees' Infusion Use of CRM Systems on Customer Satisfaction in Face-to-Face and Virtual Channels. ***MIS Quarterly*** (45:2), pp 719-754.

Chung, S., Animesh, A., Han, K., & Pinsonneault, A. (2019). Software patents and firm value: A real options perspective on the role of innovation orientation and environmental uncertainty. ***Information Systems Research***, 30(3), 1073-1097.

Cirik, K., Makadok, R. 2021. First-mover advantages versus first-mover benefits: What's the difference and why does it matter?. ***Academy of Management Review***, (ja).

Consoli, D. 2005. The dynamics of technological change in UK retail banking services: an evolutionary perspective. ***Research Policy*** 34(4):461-480.

Davidson, R., & MacKinnon, J. G. 1993. Estimation and inference in econometrics (Vol. 63). New York: Oxford.

Demsetz, H. (1973). Industry structure, market rivalry, and public policy. ***The Journal of Law and Economics***, 16(1), 1-9

Devaraj S., Kohli R. 2003. Performance impacts of information technology: Is actual usage the missing link. ***Management Science*** 49(3):273-289

Domínguez, B., Gómez, J., & Maícas, J. P. (2021). Multimarket pioneers: Does multimarket contact improve their performance?. ***Long Range Planning***, 54(2), 101993.

Dos Santos, B. L., Peffers, K. 1993. The effects of early adoption of information technology: An empirical study. ***Journal of Information Technology Management*** 4(1):1-13

Dos Santos, B. L., Peffers, K. 1995. Rewards to investors in innovative information technology applications: first movers and early followers in ATMs. ***Organization Science*** 6(3):241-259

European Central Bank. 1999. The effects of technology on the EU banking systems. Available at *www.ecb.eu/pub/pdf/other/techbnken.pdf*

Fichman, R. G. 2001. The role of aggregation in the measurement of IT-related organizational innovation. ***MIS Quarterly***, 427-455.

Fichman, R. G., Kemerer, C. F. 1999. The illusory diffusion of innovation: An examination of assimilation gaps. ***Information Systems Research***, 10(3), 255-275.

Fichman, R. G., Kemerer, C. F. 2001. Incentive compatibility and systematic software reuse. ***Journal of Systems and Software***, 57(1), 45-60.

Fosfuri, A., Lanzolla, G., Suarez, F. F. 2013. Entry-timing strategies: the road ahead. ***Long Range Planning*** 46(6):300–311.

Fuentelsaz, L., Gómez, J., Palomas, S. 2009. The effects of new technologies on productivity: An intrafirm diffusion-based assessment. ***Research Policy*** 38:1172-1180

Fuentelsaz, L., Gómez, J., Palomas, S. 2012. Production technologies and financial performance. The effect of uneven diffusion among competitors. ***Research Policy*** 41(2):401-413

Fuentelsaz, L., Gómez, J., Polo Y. 2003. Intrafirm diffusion of new technologies: An empirical application ***Research Policy*** 32:533-551.

Garen, J. (1984). The returns to schooling: A selectivity bias approach with a continuous choice variable. ***Econometrica: Journal of the Econometric Society***, 1199-1218.

Gómez, J., & Maícas, J. P. (2011). Do switching costs mediate the relationship between entry timing and performance?. ***Strategic Management Journal***, 32(12), 1251-1269.

Gomez, J., Pérez-Aradros, B., Salazar, I. 2022. How to beat early movers: The role of competitive strategy and industry dynamism on followers' performance in the telecommunications industry. ***Long Range Planning***, 55(5), 102244.

Greene, W. H. (2003). Econometric analysis. Pearson Education India.

Hannan, T. H. 1991. Foundations of the structure-conduct-performance paradigm in banking. ***Journal of Money, Credit and Banking***, 23(1), 68-84

Hannan T. H., McDowell J. M. 1984. The determinants of technology adoption: The case of the banking firm. ***RAND Journal of Economics*** 15(3):328-335

Hannan, T. H., McDowell. J. M. 1987. Rival Precedence and the Dynamics of Technology Adoption: An Empirical Analysis. ***Economica*** 54:155–71

Hausman, J., Taylor, W. 1981. Panel Data and Unobservable Individual Effects. ***Econometrica*** 49(6):1377–1398.

Hollenstein, H., & Woerter, M. (2008). Inter-and intra-firm diffusion of technology: The example of E-commerce: An analysis based on Swiss firm-level data. ***Research Policy***, 37(3), 545-564.

Kalirajan, K. P., & Shand, R. T. (2001). Technology and farm performance: paths of productive efficiencies over time. ***Agricultural Economics***, 24(3), 297-306.

Khalifa, A. B. (2022). Inter-and intra-firm diffusion of technology: the example of software, hardware, and network communications empirical evidence for Tunisian manufacturing firms. ***Journal of the Knowledge Economy***, 1-28.

Kim, G., Shin, B., Kim, K. K., ¡Lee, H. G. 2011. IT capabilities, process-oriented dynamic capabilities, and firm financial performance***. Journal of the Association for Information Systems***, 12(7), 1.

Levin, S. G., Levin S. L., Meisel J. B. 1992. Market structure, uncertainty, and intrafirm diffusion: The case of optical scanners in grocery stores. ***Review of economics and Statistics,*** 74:345-350

Lieberman, M. B., Montgomery, D. B. 1988. First-mover advantages. ***Strategic Management Journal*** 9:41-58

Lieberman, M. B., Montgomery, D. B. 2013.Conundra and Progress: Research on Entry Order and Performance. ***Long Range Planning*** 46(4-5):312-324

Makadok, R. 1998. Can first-mover and early-mover advantages be sustained in an industry with low barriers to entry/imitation? ***Strategic Management Journal*** 19: 683–696.

Mata, F. J., Fuerst, W. L., Barney, J. B. 1995. Information technology and sustained competitive advantage: A resource-based analysis. ***MIS Quarterly*** 19(4): 487-505

McElheran, K. 2015. Do Market Leaders Lead in Business Process Innovation? The Case(s) of E-Business Adoption. ***Management Science*** 61(6):1197-1216.

McWilliams, B., & Zilbermanfr, D. (1996). Time of technology adoption and learning by using. ***Economics of Innovation and New technology***, 4(2), 139-154.

Mithas, S., Krishnan, M. S., Fornell, C. 2016. Research note—Information technology, customer satisfaction, and profit: Theory and evidence. ***Information Systems Research***, 27(1), 166-181.

Mithas, S., Ramasubbu, N., Sambamurthy, V. 2011. How information management capability influences firm performance. ***MIS Quarterly***, 237-256.

Mithas, S., Rust, R. T. 2016. How information technology strategy and investments influence firm performance. ***Mis Quarterly***, 40(1), 223-246.

Mithas, S., Tafti, A., Bardhan, I., Goh, J.M. 2012. Information technology and firm profitability: Mechanisms and empirical evidence. ***MIS Quarterly*** 36(1): 205-224

Mithas, S., Whitaker, J., Tafti, A. 2017. Information technology, revenues, and profits: Exploring the role of foreign and domestic operations. ***Information Systems Research***, 28(2), 430-444.

Molyneux, P., & Forbes, W. (1995). Market structure and performance in European banking. ***Applied Economics***, 27(2), 155-159.

Nafizah, U. Y., Roper, S., & Mole, K. (2023). Estimating the innovation benefits of first-mover and second-mover strategies when micro-businesses adopt artificial intelligence and machine learning. ***Small Business Economics***, 1-24.

"Paul Volcker: Think More Boldly", ***The Wall Street Journal***, 14 December 2009, accessed 25 July 2020, <https://www.wsj.com/articles/SB10001424052748704825504574586330960597134>

Oh, W., Pinsonneault, A. 2007. On the assessment of the strategic value of information technologies: conceptual and analytical approaches. ***MIS quarterly***, 239-265.

Pan, Y., Mithas, S., Po-An Hsieh, J. J., & Liu, C. W. (2023). Do Risk Preferences Shape the Effect of Online Trading on Trading Frequency, Volume, and Portfolio Performance?. ***Journal of Management Information Systems***, 40(2), 440-469.

Peffers, K., & Dos Santos, B. L. 1996. Performance effects of innovative IT applications over time. ***IEEE Transactions on Engineering Management*** 43(4):381-392

Peteraf, M. A. 1993. The cornerstones of competitive advantage. A resource based view. ***Strategic Management Journal*** 14:179-191

Piccoli, G., & Ives, B. 2005. IT-dependent strategic initiatives and sustained competitive advantage: a review and synthesis of the literature. ***MIS Quarterly*** 29(4): 747-776.

Porter, M. E. (1985). The Competitive Advantage: Creating and Sustaining Superior Performance 1985 New York.

RBR London (2023a) "Developing ATM markets see strong growth as post-COVID recovery continues", accessed 28th February 2024 < https://www.rbrlondon.com/wp-content/uploads/2023/06/GAIS23_Press_Release_270623.pdf>

RBR London (2023b) "Retailers' demand for cash automation solutions surges", accessed 28th February 2024 < https://www.rbrlondon.com/wp-content/uploads/2023/01/RCA23_Press_Release_110123.pdf>

Sabherwal, R., Jeyaraj, A. (2015). Information technology impacts on firm performance. ***MIS Quarterly***, 39(4), 809-836.

Saldanha, T. J., Lee, D., Mithas, S. 2020. Aligning information technology and business: The differential effects of alignment during investment planning, delivery, and change. ***Information Systems Research***, 31(4), 1260-1281.

Saldanha, T. J., Mithas, S., & Krishnan, M. S. (2017). Leveraging customer involvement for fueling innovation. ***MIS Quarterly***, 41(1), 267-286.

Saloner, G., Shepard, A. 1995. Adoption of technologies with network effects: an empirical examination of the adoption of automated teller machines. ***RAND Journal of Economics***, 26:479-501

Santhanam, R., Hartono, E. 2003. Issues in linking information technology capability to firm performance. ***MIS Quarterly***, 125-153.

Sargan, J. D. (1958). The estimation of economic relationships using instrumental variables. ***Econometrica: Journal of the Econometric Society***, 393-415.

Scherer, F. M., Ross, D. 1990. ***Industrial Market Structure and Economic Performance***. Houghton Mifflin, Boston, MA.

Scherer, A., Wunderlich, N. V., & von Wangenheim, F. (2015). The Value of Self-Service: Long-Term Effects of Technology-Based Self-Service Usage on Customer Retention. ***MIS Quarterly***, 39(1), 177-200.

Scott, S. V., Van Reenen, J., & Zachariadis, M. (2017). The long-term effect of digital innovation on bank performance: An empirical study of SWIFT adoption in financial services. ***Research Policy***, 46(5), 984-1004.

Setia, P., Setia, M., Krishnan, R., Sambamurthy, V. 2011. The effects of the assimilation and use of IT applications on financial performance in healthcare organizations. ***Journal of the Association for Information Systems***, 12(3), 1.

Shepherd-Barron, J. (2017) “Meet the true star of financial innovation — the humble ATM”, Financial Times, accessed 25 July 2020, < https://www.ft.com/content/052f9310-5738-11e7-80b6-9bfa4c1f83d2>

Sinha, R. K., & Noble, C. H. (1997). The performance consequences of subfield entry. ***Strategic Management Journal***, 18(6), 465-481.

Sinha, R. K., Noble, C. H. 2005. A model of market entry in an emerging technology market. ***IEEE Transactions on Engineering Management*** 52(2): 186–198.

Sodhu, I. (2011) “ATM 2.0: What the Cash Machine's Second Act Can Teach Business Leaders” Forbes. April 1st, 2011, <https://www.forbes.com/sites/indersidhu/2011/04/01/atm-2-0-what-the-cash-machines-second-act-can-teach-business-leaders/?sh=5607dc3f4f89>

Stock, J., & Yogo, M. 2005. Asymptotic distributions of instrumental variables statistics with many instruments. Identification and inference for econometric models: Essays in honor of Thomas Rothenberg, 6, 109-120.

Stoneman, P. 1981. Intra-firm diffusion, Bayesian learning and profitability. ***The Economic Journal***, 91(362), 375-388.

Strong, D.R. 2019. Impacts of diffusion policy: determinants of early smart meter diffusion in the US electric power industry. ***Industrial and Corporate Change*** 28(5) 1343-1363.

Stucki, T. and Woerter, M. 2016. Intra-firm diffusion of green energy technologies and the choice of policy instruments. ***Journal of Cleaner Production*** 131: 545-560.

Swanson, E. B., Ramiller, N. C. 2004. Innovating mindfully with information technology. ***MIS Quarterly***, 553-583.

Wade, M., Hulland, J. 2004. The resource-based view and information systems research: Review, extension, and suggestions for future research. ***MIS Quarterly***, 107-142.

Waters, J. (2017). Determinants of initial technology adoption and intensification: evidence from Latin America and the Caribbean. ***Economics of Innovation and New Technology***, 26(4), 334-352.

Wu, L., Jin, F., & Hitt, L. M. (2018). Are all spillovers created equal? A network perspective on information technology labor movements. ***Management Science***, 64(7), 3168-3186

Wooldridge J.M. 2002. ***Econometric analysis of cross section and panel data***. MIT Press

Xu, R., Frank, K. A., Maroulis, S. J., Rosenberg, J. M. 2019. konfound: Command to quantify robustness of causal inferences. ***The Stata Journal***, 19(3), 523-550.

Zachary, M., Gianiodis, P. T., Payne, G. T., Markman, G., D. 2015. Entry Timing: Enduring Lessons and Future Directions. ***Journal of Management*** 41(5):13 88-1415.

APPENDIX A1: Selected studies on the effect of IT on firm performance

Table A1: Selected studies on the effect of IT on firm performance

Study	Sample	IT measure	Performance measure	Main results	Observations
Bharadwaj et. al (1999)	A set of US publicly traded firms (1989-1993), 53% manufacturing sector and 47% services sector	IT budget	*Market performance* : Tobin's q	IT has a consistent positive effect on market value.	One of the few papers that assess the extent to which the stock markets values IT investment
Bharadwaj (2000)	A set of firms identified as IT leaders by *Information Week* (1991-1994), and a matched comparison group	IT leaders identified by IT specialists.	*Profitability* : Return on Assets (ROA), Return on Equity (ROE), *Revenue* : Operating income to Assets (OI/A), operating income to sales (OI/S), operating income to employees (OI/E) *Cost* : total operating expenses to sales (OEXP/S), cost of goods sold to sales (COGS/S), selling and general administrative expenses to sales (SG&A/S)	Profitability and revenues are higher for IT leaders. Mixed results for costs	Proposes a Resource Based View (RBV) analysis.
Santhanam and Hartono (2003)	A set firms identified as IT leaders by *Information Week* (1991-1994), and other firms in their industries.	IT leaders identified by IT specialists.	*Profitability* : ROA, ROE *Revenue* : OI/A, OI/S, OI/E *Cost* : OEXP/S, COGS/S, SG&A/S	Profitability and revenues are higher for IT leaders. Costs tend to be lower for IT leaders.	Replicates and elaborates the analysis by Bharadwaj (2000). Compares IT leaders with the industry average. Results are robust to the consideration of past performance. Advantages are sustainable.
Oh and Pinsonneault (2007)	A set of small and medium sized Canadian firms in the manufacturing industry (year not specified)	Reserchers' evaluation of IT portfolio of the firm to identify the purpose of each IT application	Survey-based and perceptual measures of revenue, cost and profitability	IT can reduce cost and increase revenue and profitability. The effect depends on the type of IT application and the theoretical approach used to define the research design.	The article compares the two dominant approaches ot the analysis of the impact of IT: RBV and Contingency theory
Mithas, Ramasubbu and Sambamurthy (2011)	Firms within a large business group (1999-2003).	Assessment by independent examiners of information management capabilities of each firm in the business group	Assessment by independent examiners of management capabilities (performance, customer, process) and results (customer focused, financial, human resource, organizational effectiveness)	Information management capabilities reinforce other critical management capabilities of the firm, which, in turn, have robust effects on firm performance	Uses qualitative measures which allow the measurement of complex dimensions of organizational capabilities and performance
Kim et al., (2011)	A sample of Korean firms from a wide range of industries (year not specified)	Survey and perceptual measures of IT personnel expertise, intrastructure flexibility, and management capabilities	*Profitability* : overall financial performance over the three previous years	IT has a positive effect on performance. The effect is mediated by the development of process oriented dynamic capabilities	The article analyzes the effect of IT on firm performance applying a Dynamic Capabilities perspective
Setia et al. (2011)	A set of US hospitals (2004)	Range of IT applications adopted (spread), and number of years using them (longevity)	*Profitability* : net income by patient-day	Mixed results. Greater IT longevity in the clinical domain improves performance, but in the business domain decreases performance. Complex patterns of complementarity and substitution between spread and longevity	One of the few papers which cosiders explicit measures of actual assimilation of the IT
Mithas et al (2012)	A set of IT intensive firms around the world (1998 - 2003)	IT investments normalized by employee	*Profitability* : net income per employee *Revenue* : revenue per employee *Cost* : operating expenses before depretiation per employee	IT has a robust effect on revenues and profitability, but not on costs	The article discusses the conditions under which IT can generate competitive advantages
Chae, Koh and Prybutok (2014)	A set of firms identified as IT leaders by *Information Week* (2001-2004), and a matched comparison group	IT leaders identified by IT specialists.	*Profitability* : ROA, ROE *Revenue* : OI/A, OI/S, OI/E *Cost* : OEXP/S, COGS/S, SG&A/S	In the long term, only cost advantages are sustained	Replicates and elaborates the analysis by Bharadwaj (2000). Explores whether results are robust to the commoditization of IT
Mithas and Rust (2016)	A set of IT intensive US firms (2003-2004)	IT investment as a percentage of firm sales	*Profitability* : operating income before depretiation divided by sales *Market performance* : Tobin's q	IT investments have a positive ffect o profitability and market performance, and this effect is stronger under the appropriate strategic emphasis.	The article analyzes whether a strategic emphasis on cost reduction, revenue increase, or a dual emphasis is the best option for different levels of IT spending

Table A1: Selected studies on the effect of IT on firm performance (Continued)

Study	Sample	IT measure	Performance measure	Main results	Observations
Mithas, Krishnan and Fornell (2016)	A set of IT intensive US firms (1994-1996 and 1999-2006)	IT investment as a percentage of firm sales	*Profitability* : operating income before depretiation and taxes *Operational performance* : customer satisfaction	IT investments have a positive effect on cusomer satisfaction. In the case of profitability, the effect is negative in 1994-1996, and positive in 1999-2006	The article elaborates on how over time the effect of IT on performance have changed
Mithas, Whitaker and Tafti (2017)	A set of IT intensive US firms operating both domestically and internationally (1999-2006)	IT expenditures	*Profitability* : pretax incomes *Revenue* : net sales *Cost* : difference between net sales and pretax incomes	IT spending has a positive effect on profitabiity and revenues, both domestic and foreign. In the case of cost, the effect is null (foreign) or even harmful (higher domestic costs)	The article explores how IT can support an internationalization strategy, with diferential effects on domestic and foreign performance
Saldanha, Lee and Mithas (2020)	A set of IT intensive Indian firms (2008)	IT expenditures	*Revenue* : total turnover	IT investment has a positive effet on firm revenue. This effect can be greater depending on the IT alignment strategy.	The article explores how IT alignment improve the effectiveness of IT investments depending on the volume of investment, and the stage of the IT lifecycle in which the alignment effort is made

APPENDIX A2: Empirical studies on the effect of early adoption of IT on firm performance

Table A2: Empirical studies on the effect of early adoption of IT on firm performance[14]

Study	Sample	Technology	Dependent variable	Observation window	Methodology	Measurement of IT	Observations
Peffers and Dos Santos (1996)	2534 US banks	Automated teller machine	Market share, income	1971-1984	Ordinary Least Squares, Nonlinear Least Squares	Dummy variables, one for each of the four years in which a bank in the sample may have first adopted ATMs (1971-1974)	Learning is the main mechanism explaining early adopter advantages. It is not explicitly modelled. No distinction between order of entry and level of adoption advantages
Dos Santos and Peffers (1995)	2534 US banks	Automated teller machine	Market share, income	1971-1983	Ordinary Least Squares	Dummy variables, one for each year in the adoption period from 1971 to 1979	Learning, switching costs and preemption of assets are the mechanisms proposed to explain early adopter advantages. None of them is explicitly modelled. No distinction between order of entry and level of adoption advantages
Dos Santos and Peffers (1993)	3838 US banks	Automated teller machine	Market share, employee efficiency	Investments in ATMs for 1971-1979; bank performance data for 1972, 1980, 1982 and 1984	Ordinary Least Squares	Dummy variable, taking a value of 1 if bank j adopted ATMs between 1971 and 1979, and zero otherwise	ATMs can reduce costs or improve quality of service. No clear isolating mechanism is proposed. No distinction between order of entry and level of adoption advantages
Sinha and Noble (1997)	3500 US banks	Automated teller machine	Return on assets, net income	1971-1979	Fixed and Random Effects Panel Data	Years since adoption of ATM; dummy variable taking a value of 1 if the bank has adopted in the first two years, and zero otherwise	Learning curve advantages (they are not measured). No distinction between order of entry and level of adoption advantages
This paper	Spanish saving banks	Automated teller machine	Return on assets, income, efficiency	1988-2004	Fixed Effects Panel Data, Two Stage Least Squares, Generalized Method of Moments, Hausman Taylor	*Order of adoption* : dummy variable taking a value of 1 for firms having adopted by 1981, and zero otherwise; *Level of adoption* : number of ATMs per branch from 1988 to 2004	We distinguish between order of entry and level of adoption advantages. We explicitly model reputation, location preemption and investments in other IT.

[14] Table A2 lists a selection of previous studies that have examined the consequences of the early adoption of IT on performance and their characteristics. The number of empirical papers that we could find in our review is very small, likely due to the difficulties researchers face in finding data on the exact time of adoption.

APPENDIX A3: Analysis of network externalities

Table A3: Analysis of network externalities

	Return on Assets
Level of Adoption	0.196** (0.090)
Level of Adoption X 1< Early Adopters >	0.241** (0.103)
Level of Adoption X ATM terminals	-0.000 (0.000)
Concentration	-0.594 (0.578)
Firm Size	-0.513*** (0.174)
Branch Size	-0.131 (0.172)
Risk	-0.395 (0.285)
Inefficiency	-2.991*** (0.258)
Rural	-0.431* (0.236)
ATM terminals	0.000 (0.000)
Density of ATMs in the market	-1.647*** (0.316)
IT employees	0.026 (0.678)
Year dummies	Yes***
F	15.01***
R2	0.36
Observations	902

Robust standard errors in parentheses.
Two-tailed test of significance: * p<.10; ** p<.05; *** p<.01

APPENDIX A4: First stage results of 2SLS estimations

Table A4: First stage results of 2SLS estimations

	(1)	(2)
	Level of Adoption	**Level of Adoption X 1< Early Adopters >**
Concentration	0.099 (0.160)	0.075 (0.087)
Firm Size	-0.218*** (0.049)	-0.091*** (0.027)
Branch Size	0.262*** (0.049)	0.023 (0.027)
Risk	0.024 (0.077)	-0.128*** (0.042)
Inefficiency	0.068 (0.072)	0.003 (0.040)
Rural	-0.043 (0.066)	0.007 (0.036)
ATM terminals	0.000*** (0.000)	0.000*** (0.000)
Density of ATMs in the market	0.302*** (0.089)	0.138*** (0.049)
IT employees	-0.267 (0.175)	-0.168* (0.097)
Year dummies	Yes***	Yes***
Excluded instruments:		
Level of Adoption (t-1)	0.919*** (0.044)	-0.017 (0.024)
Level of Adoption (t-2)	-0.005 (0.056)	-0.003 (0.031)
Level of Adoption (t-3)	-0.122*** (0.041)	-0.016 (0.023)
Level of Adoption X 1< Early Adopters > (t-1)	-0.366*** (0.081)	0.586*** (0.045)
Level of Adoption X 1< Early Adopters > (t-2)	0.209** (0.094)	0.235*** (0.053)
Level of Adoption X 1< Early Adopters > (t-3)	0.124 (0.076)	0.037 (0.042)
F	269.71***	146.70***
R2	0.93	0.87
Observations	748	748

Robust standard errors in parentheses.
Two-tailed test of significance: * p<.10; ** p<.05; *** p<.01

APPENDIX A5: alternative sets of instruments for 2SLS

In the main estimations we propose the use of internal instruments, in particular the first three lags of the potentially endogenous variables (i.e., *Level of Adoption* and its interaction with the early adopter dummy) as instruments. However, it should be noted that we find autocorrelation in our independent variable, *ROA*, which means that lagged *ROA* is correlated with contemporaneous *ROA*. This can cause problems when using internal instruments. In particular, explanatory variables that are lagged by one period could have an effect on the dependent variable that is mediated by their effect on the lagged dependent variable, leading to a correlation between the lagged explanatory variable and the contemporaneous dependent variable in the main equation (and, consequently, with the error term). In this case, this instrument would not be valid, as it would not fulfil the orthogonality condition[15].

It can be argued that we use tests to check whether the instruments are significantly correlated with the error term of the main equation, such as Hansen's J-statistic. However, these tests take as their null hypothesis that the instruments are not correlated with the error term (and hence with the dependent variable), which is not sufficiently conservative given the suspicion of autocorrelation. Therefore, we tested two alternative instrument proposals.

First, we again use internal instruments, but only from the second lag onwards. In this case, we assume that the explanatory variables from two or three years ago have no direct effect on the contemporaneous dependent variable. The results of this estimation are shown in the first three columns of Table A5. The first two columns show the first stage, and the third column shows the second stage. The tests shown at the bottom of column 3 confirm that these instruments satisfy the conditions of relevance (Kleibergen-Paap rk Wald F statistic above the critical values of Stock and Yogo [2005]) and orthogonality (Hansen's J does not reject the null hypothesis). As can be seen, the results remain qualitatively the same.

[15] We are grateful to an anonymous reviewer for bringing this to our attention.

Second, we use external instruments. In particular, we instrument the adoption level variable through the average adoption level of the industry excluding the focal firm (Chung et al., 2019). This average level may affect the firm's adoption level through inter-organizational learning and imitation, while it is the firm's own level that should affect its performance. We include both the contemporaneous value and the first lag of the instruments so that we can compute tests to check whether the instruments are appropriate. The results of this second alternative are shown in the columns 4 to 6 of Table A5. Columns 4 and 5 show the first stage, and column 6 shows the second stage. The tests shown at the bottom of column 6 confirm that the instruments are valid. The results remain, once again, qualitatively the same.

Table A5: Alternative sets of instruments for 2SLS

	(1)	(2)	(3)	(4)	(5)	(6)
	Level of adoption	Level of adoption x 1<early adopter>	ROA	Level of adoption	Level of adoption x 1<early adopter>	ROA
Level of Adoption	-	-	-0.044 (0.151)	-	-	0.163 (0.112)
Level of Adoption X 1< Early Adopters >	-	-	0.269** 0.114	-	-	0.315*** (0.112)
Concentration	0.367* (0.188)	0.140 (0.090)	0.254 (0.447)	-0.052 (0.044)	-0.293* (0.161)	-0.098 (0.446)
Firm Size	-0.385*** (0.076)	-0.107*** (0.035)	-0.518** (0.219)	-0.095*** (0.019)	-0.101 (0.066)	-0.367* (0.190)
Branch Size	0.477*** (0.085)	0.020 (0.044)	-0.109 (0.222)	0.104*** (0.019)	0.025 (0.071)	-0.216 (0.199)
Risk	0.131 (0.152)	-0.179* (0.096)	-0.493 (0.314)	0.012 (0.025)	-0.159* (0.092)	-0.290 (0.298)
Inefficiency	0.156 (0.113)	0.020 (0.035)	-3.240*** (0.305)	0.048** (0.022)	0.009 (0.061)	-3.028*** (0.297)
Rural	-0.119 (0.081)	0.005 (0.034)	-0.272 (0.240)	-0.056*** (0.020)	0.048 (0.045)	-0.429 (0.239)
ATM terminals	0.000*** (0.000)	0.000 (0.000)	0.000 (0.000)	0.000 (0.000)	0.000** (0.000)	-0.000 (0.000)
Density of ATMs in the market	0.500*** (0.169)	0.156 (0.097)	-0.801** (0.395)	0.135*** (0.036)	-0.034 (0.094)	-1.569*** (0.363)
IT employees	-0.225 (0.287)	-0.190 (0.318)	0.053 (0.414)	0.126** (0.060)	-0.064 (0.218)	-0.058 (0.509)
Year dummies	Yes***	Yes***	Yes***	Yes***	Yes***	Yes***
Excluded instruments:						
Level of Adoption (t-2)	0.809*** (0.110)	-0.024 (0.018)	-	-	-	-
Level of Adoption (t-3)	-0.178** (0.089)	-0.026** (0.013)	-	-	-	-
Level of Adoption X 1< Early Adopters > (t-2)	-0.323** (0.132)	0.585*** (0.087)	-	-	-	-
Level of Adoption X 1< Early Adopters > (t-3)	0.322** (0.126)	0.208** (0.097)	-	-	-	-
Industry level of adoption	-	-	-	-47.803*** (0.669)	-11.247*** (3.372)	-
Industry level of adoption (t-1)	-	-	-	4.218*** (0.630)	4.472 (2.775)	-
Industry Level of Adoption X 1< Early Adopters >	-	-	-	0.026 (0.187)	-0.243 (0.825)	-
Industry Level of Adoption X 1< Early Adopters >(t-1)	-	-	-	-0.055 (0.173)	1.141 (0.793)	-
						-
F	93.23***	35.64***	12.53***	2946.34***	10.36***	14.17***
R2	0.87	0.83	0.34	0.99	0.80	0.35
Observations	748	748	748	878	878	878
Kleibergen-Paap rk Wald F statistic	-	-	36.76	-	-	51.39
Hansen test (p-value)	-	-	0.43	-	-	0.47

Robust standard errors in parentheses. Two-tailed test of significance: * p<.10; ** p<.05; *** p<.01

APPENDIX A6: Garen's Method. Selection equation.

Table A6: Garen's Method. Selection equation.

	Level of adoption
Concentration	-0.036 (0.138)
Firm Size	-0.264*** (0.072)
Branch Size	0.341*** (0.081)
Risk	0.092 (0.129)
Inefficiency	0.064 (0.100)
Rural	-0.042 (0.065)
ATM terminals	0.000** (0.000)
Density of ATMs in the market	0.290** (0.130)
IT employees	-0.117 (0.214)
Level of Adoption (t-1)	0.788*** (0.054)
Year dummies	Yes***
F	382.57***
R2	0.67
Observations	878

Robust standard errors in parentheses.
Two-tailed test of significance: * p<.10; ** p<.05; *** p<.01