The Value of Information in Competitive Markets: The Impact of Big Data on Small and Medium Enterprises[1]

Jose Enrique Galdon-Sanchez, UPNA
Ricard Gil, Queen's University
Guillermo Uriz-Uharte, Compass Lexecon

July 2021

Abstract

A firm may gain competitive advantage over its rivals through access to market information. Yet, evidence thus far suggests only large firms invest in technology that facilitates access to information, potentially increasing their leverage over smaller competitors. This paper aims to fill a gap in the literature by empirically investigating how the performance and decision-making of small and medium-size enterprises change when gaining access to strategically valuable market information. To do so, we evaluate the impact of an unprecedented Big Data information service diffused at zero cost by a large European bank among its small and medium-size business customers. Upon program adoption, adopting firms had monthly access to reports elaborated by the bank with rich information about each firm's clientele portfolio and that of its competitors analyzing Big Data credit card transactions. Using first-differences, we find adoption is associated with a 4.5% increase in establishment revenue, whereas IV estimation results show that adoption causally increases revenue by 9% for those establishments whose adoption decision is most strongly affected by the instrument. The main mechanism behind this result appears to be the information technology prompting establishments to target existing, yet unexploited, business opportunities. Consistent with this mechanism, we find that adopting establishments increase their sales to customer segments from underrepresented gender-age groups in their customer portfolio prior to adoption. Our evidence also suggests that adopting establishments improved their resource allocation efficiency between weekly peak and off-peak times. These findings suggest that small and medium enterprises obtain substantial returns from information access, and therefore, managerial inattention and high adoption costs are likely to be key barriers preventing small firms from investing in resources to acquire and analyze market information.

JEL Codes: G20, L20, L80, M15, O32, O33
Keywords: information, competitive markets, Big Data, data sharing, small and medium enterprises.

---

[1] We thank Kristina McElheran, Antonio Cabrales, Aureo de Paula, Imran Rasul, James MacKinnon, Lars Nesheim, Michelle Sovinsky, Paul Belleflamme, Riccardo D'Adamo, and seminar participants at the NBER Summer Institute Meetings of the Economics of Digitization, UCL, Amazon, EARIE, European University Institute, Queen's University, APIOC – Tokyo, the LED YES conference, and Washington University Olin School of Business for their comments. We thank Fundacion BBVA for funding this research project. Jose Enrique Galdon-Sanchez also thanks financial support from the Ministerio de Ciencia, Innovacion y Universidades (Project DGCT, No. PGC2018-093542-B-I00). The usual disclaimer applies.

"The world's most valuable resource is no longer oil, but data."

The Economist May 6th 2017

1. Introduction

While neoclassical economics implicitly assumes that [perfect] information is widely available to firms and decision makers, the reality is that imperfect and asymmetric information is ubiquitous in markets and organizations. In fact, economists have showed that information plays a central role in understanding the development and functioning of markets in a wide variety of contexts such as monetary policy and financial markets (Hayek, 1945; Fama, 1970; Lucas, 1972; Grossman and Stiglitz, 1980), labor and education markets (Stigler, 1962; Spence, 1973), healthcare and insurance markets (Rothschild and Stiglitz, 1976), or product markets where quality and reputation are key determinants of competitive advantage (Akerloff, 1970).

A key mechanism through which information affects the economy is decision-making. Not only consumers make purchasing decisions based on information available to them through advertising and consumer reports, but also information is a key input for firms in their day-to-day production and marketing strategies (Bergemann and Bonatti, 2019). In fact, in a competitive market environment, information can be a source of competitive advantage through a variety of channels, for instance, by lowering costs (more efficient resource allocation and production processes) or enhancing the understanding of business opportunities (better product customization and improved forecasting demand). While the literature on firm dynamics has recognized the role of productivity in shaping firm growth, Foster et al. (2008, 2016) have shown that information on customer demand may be also an important determinant of firm growth.

Over the last few decades, the rise of information technology (IT hereafter) has lowered the marginal cost of collecting, processing and using information for decision-making (Brynjolfsson et al., 2011; Brynjolfsson and McElheran, 2016a and 2016b; Agrawal et al., 2018). Most recently, this has originated the eruption of the Big Data revolution and the spread of data analytics, consequently, enabling data-driven decision-making (DDD hereafter) over traditional decision-making based on intuition. In the words of Jim Barksdale, the former CEO of Netscape, a good maxim for modern management practice is: "If we have data, let's look at data. If all we have are opinions, let's go with mine."[2]

However, because the adoption of IT has concentrated in large corporations,[3] the growing literature studying adoption patterns and consequences of IT has focused on large firms. Large firms are more likely to adopt because they benefit the most from these technologies by improving their internal processes (Brynjolfsson and Hitt, 2000; Bresnahan et al., 2002; Bartel et al., 2007) and gaining better access to markets (Jensen, 2007; Foster et al., 2008 & 2016; Goyal, 2010; Einav and Levin, 2014).

In contrast, adoption of these technologies has been anecdotal among small and medium enterprises (SMEs hereafter).[4] Identifying how market information, through Big Data adoption for instance, affects SMEs' performance and decision-making would help answering at least two relevant questions. First, it will inform us on whether SMEs are deterred from investing in market information through the adoption of Big Data technologies because *(a)* returns from adoption are too low for them, *(b)* adoption costs are too high, or (c) due to lack of managerial awareness and understanding of the returns and costs of the technologies. Finding positive returns for SMEs can open the door for public (and private) interventions intended to decrease adoption costs and facilitate data sharing initiatives (Jones and Tonetti, 2020). This would increase adoption rates among SMEs, which in turn may reduce the current performance differences between firms with "intuition-driven" and "data-driven" decision-making practices, and contribute to decrease market concentration with all its consequences on market outcomes such as prices, quality, and innovation.

Second, this exercise will improve our understanding on how access to market information affects firms' competitive strategies and market equilibrium outcomes. Recent years have witnessed dramatic increases in the availability of data, and the pace is accelerating. Therefore, it is of first-order importance to understand how this improvement in information access is likely to affect firm strategic decision-making and market competition in order to stay ahead of events and regulate markets accordingly. In this paper, we aim to contribute to

[2] See the link here, https://casestudies.storetrials.com/we-have-data-lets-look-at-data-e8a06e2e3331.

[3] Brynjolfson and McHeleran (2016b) show that data-driven decision-making is concentrated in plants with three key advantages: size, high levels of potential complements such as information technology and educated workers, and "awareness."

[4] In this paper, when referring to SMEs, we use the terms enterprises, firms, and establishments interchangeably.

this debate by accomplishing two goals. On the one hand, we estimate the distribution of returns to market information access through the adoption of a Big Data information-sharing technology that facilitates the implementation of DDD practices among SMEs. On the other hand, we provide evidence of which mechanisms operate behind the estimated effect of market information access on SME performance.

To achieve our goals, we use information from a large European bank on the deployment of a Big Data information-sharing program in Spain among its SMEs customers. Upon voluntarily signing up to this unprecedented and unique free Big Data information-sharing program, SMEs receive every month a monthly report on their sales profiles relative to other neighboring establishments in their same sector based on all credit and debit card transactions occurred during the previous month in all the establishments with a bank point-of-sale (hereafter POS). Therefore, the program reduces the costs of access of SMEs to market information through a double channel. On the one hand, given their limited number of customers, SMEs usually do not have the capacity to generate large volumes of data on consumer behavior. On the other hand, SMEs often lack the capacity to analyze large volumes of disaggregated data and draw conclusions from it. This program processes the raw data and offers SMEs a report that, despite having very rich information, is easier to understand than unstructured data. Following an earlier pilot release in 2014, the program was officially launched in the spring of 2016 for the whole country, targeting all establishments with a bank POS. While our initial data comprises comprehensive information on credit and debit card transactions for nearly all POS in Spain between 2014 and 2018, our final working data contains quarterly information for 487,610 establishments, out of which 7,110 adopted the technology across all provinces, 18 sectors and 77 subsectors in the country.[5]

Our empirical methodology uses first-differences OLS regressions of quarterly revenue data on adoption with sector-zip code-quarter fixed effects and establishment-specific time trends

[5] The final number of adopters approximately represents 1.5% of all potential adopters of the program. Adoption was free and did not require high specific investment or high user sophistication levels, so we are confident a priori that the low adoption rate of this technology is not driven by high adoption costs. Therefore, low adoption rates can be explained by either low returns of adoption, strategic motives or managerial inattention, and lack of awareness. Our findings will show that adoption returns are sizeable, and that information received did not depend on adoption of others. Consequently, lack of awareness seems to be the most plausible reason behind the low adoption rates observed in the data.

as baseline specification. We complement this analysis with an instrumental variable approach (IV hereafter) where we take advantage of the fact that different establishments within the same sector-zip code dyad are affiliated to different bank branches. Our IV is the number of adopters (other than the focal establishment) in the establishment's bank branch in a given quarter. This strategy allows to use exogenous variation across establishments within the same sector, zip code and quarter. The rationale for the instrument comes from detailed conversations with bank managers in that the bank did not compensate its employees for the diffusion of the program, and therefore differences in program diffusion across branches were explained by idiosyncratic preferences and affinity of branch employees with the program. High affinity branches' employees put a considerable amount of effort in the promotion of the program, increasing adoption among its customers. Low affinity branches' employees put no effort in the promotion and, as a result, their customers were unlikely to adopt because they were not made aware of the existence and benefits of the program.

We find that adoption is associated with a 4.5% increase in revenue from credit and debit card transactions, whereas our IV strategy shows that adoption causally increases establishments' revenue by 9% for those establishments whose adoption decision is most strongly affected by our instrument. This finding is robust to several falsification and placebo tests. Moreover, our heterogeneity analysis shows that smaller adopters and adopters operating in zip codes with more competitors earn higher returns, while differences in establishment sophistication do not appear to drive differences in the returns of adoption.

We investigate the role of various potential mechanisms behind these findings. First, adoption may prompt establishments to target existing, yet unexploited, business opportunities such as new age-gender client groups previously underserved. Second, even in the absence of discovering new business opportunities, adoption may help establishments improve the efficiency of their internal resource allocation through better scheduling hours and sale effort allocation across peak and off-peak times. Our analysis shows direct support for the former mechanism in that adopting establishments increase their sales to underserved customer segments. Our evidence first points out that the increase in revenue after adoption comes from an increase in both the number of transactions and the number of customers, while keeping the average number of transactions per customer and the average value per transaction constant. Not only they increase their number of customers, their new customers also come from underrepresented geographic and gender-age groups in their customer portfolio prior to adoption. We also find evidence consistent with the latter mechanism, that is, adopting establishments reshuffle their sales towards idle times of the week even when the demographics of their clientele portfolio does not change.

It is worth highlighting that adopters not only change their portfolio of customers when they discover new business opportunities, but also when they choose to broaden their customer base into a more diverse portfolio of customers. Consequently, we find that non-adopters revenue goes down when a competitor adopts in their same zip code and business sector, suggesting that more information seems to increase competition among incumbent establishments. While some theories may predict that more information may drive establishments to become more specialized, we find the opposite result in our analysis. Establishments with more information start serving more customer types and mimicking the customer portfolios of their closest competitors. This finding has direct and important consequences for our understanding of the impact of information on the degree of competition in a market and, ultimately, on consumer surplus and total welfare.[6] If access to more information makes establishments specialize in serving narrow market segments, the factual degree of competition would indeed decrease, prices would increase and welfare could potentially decrease. Instead, our findings suggest a positive association between more information in a market, and the degree of competition and total welfare.

Our findings and their implications contribute to three main streams of literature. First, our paper contributes to the literature that analyses the impact of market information on a firm's strategic decision-making. It is customary in the industrial organization literature, and more generally in Economics, to assume firms' full knowledge on market fundamentals when making optimal strategic decisions. However, there is abundant evidence that firm's information is usually far from perfect (e.g., Cyert and March, 1963; Baum and Lant, 2003; Li et al., 2017; Kim, 2019).[7] Therefore, understanding whether firms are able to benefit from

[6] While more transparency on the consumer side is associated with more competition (Brown and Goolsbee, 2002; Jin and Leslie, 2003; Liberti et al., 2019), more transparency on the producer side is thought to have opposite effects as it facilitates tacit collusion among incumbent firms (Stigler, 1964; Tirole, 1988; Pettengill, 1979; Choi et al., 1990; Bertoletti and Poletti, 1997; Carlin et al., 2012).

[7] Relatedly, our findings also have implications for the literature on inattention in organizations to the extent that information technology attenuates inattention and information gaps within organizations and their market interactions. Our findings are consistent with theories of organizational slack (Cyert and March, 1963; Cohen et al., 1972) and absorptive capacity (Cohen and Levinthal, 1990), or most recently, rational inattention on

more and better information, and how they react to it, should be of first-order importance to comprehend and regulate competition dynamics in a world shaped by an increasing availability of data (Foster et al, 2008 & 2016; Bergemann and Bonatti, 2019; Bergemann et al., 2021).

We also contribute to a second stream of literature that focuses on the study of persistent performance differences (PPDs hereafter) among otherwise-equal firms within an industry. While traditional explanations for the dispersion in productivity have pointed out competition (Syverson, 2004 and 2011; Hsieh and Klenow, 2009; Galdon-Sanchez and Schmitz, 2002) or search costs (Hortacsu and Syverson, 2004) as main driving factors, Gibbons and Henderson (2013) highlight the importance of management practices to explain the observed distribution of PPDs in an economy. Furthermore, Bloom and Van Reenen (2007) have provided consistent evidence that certain managerial practices are more likely to be associated with high productivity levels, and that information technologies are important enablers of such managerial practices (Sadun and Van Reenen, 2005; Bloom et al., 2012). Our paper identifies access to market information, through the adoption of Big Data IT, as an input of production that facilitates changes in behavior and strategies, which translates into changes in performance in SMEs.[8] Thus, access (or lack thereof) to market information among SMEs can reduce (exacerbate) the productivity gap between large and small firms. Interestingly, our evidence suggests that high adoption costs and managerial inattention (instead of low adoption returns) may be responsible for the lack of use of market information, and low adoption rate of this program in particular and information technology more generally, among SMEs.

Finally, our third contribution is to the growing literature studying the role of IT in enabling DDDs (Brynjolfsson and McElheran, 2016a and 2016b; Brynjolfsson et al., 2011). McAfee and Brynjolfsson (2012) argue that Big Data allows managers to evaluate and measure precisely the impact of their decisions through DDDs. Einav et al. (2017) assess gains from e-commerce, Farboodi et al. (2019) present data as a valuable intangible asset driving the

organizational focus (Dessein et al., 2016), inattentive sellers and price rigidity (Matějka, 2016; Levitt, 2006), and retail outlet competition for consumer attention (Anderson and De Palma, 2012).

[8] Consequently, our paper also contributes to the literature that studies adoption patterns of IT. This literature has focused on the impact of IT in local wages (Forman et al., 2012), firms' organization (Brynjolfsson and Hitt, 2000; Bresnahan et al., 2002; Bloom et al, 2013), R&D and innovation (Mohnen et al., 2018; Uriz-Uharte, 2019), and productivity (Hauswald and Marquez, 2003; Sadun and Van Reenen, 2005; Bloom et al., 2012).

skewness of firm size and productivity distribution, and Bajari et al. (2019) show that Big Data allows firms to lower forecasting errors and therefore enables better decision making.[9]

To the best of our knowledge, ours is the first paper to evaluate empirically the gains of access to Big Data market information among SMEs in the retail and customer service sectors. While a limited number of other papers have studied the impact of data analytics on online retailers (Muller et al, 2018; Jin and Sun, 2019 & 2020; Berman and Israeli, 2020), our study investigates the impact of information on largely unsophisticated, brick-and-mortar retail and customer service establishments. To the extent that our sample is representative of the average establishment in a representative economy, our findings contribute to an ongoing debate regarding the complementarities between a firm's scale and the adoption of information technologies that may enable the implementation of DDDs in organizations. Brynjolfsson and McElheran (2016b) document an increase in productivity among large manufacturing plants upon adoption of IT that facilitates the switch towards DDDs practices throughout their organizations. Angle and Forman (2018) use a different sample of manufacturing plants to establish that productivity gains from IT adoption are only present in larger plants. Our paper here differs from these studies, and others in this literature, in a number of ways. First, our sample is composed by small and medium-sized downstream establishments. Second, the data-sharing program adopted is homogenous across establishments and it provides information not only about the adopting establishment but also about its competitors' strategies and market opportunities. Third, our IV approach allows us to provide causal estimates of the impact of adoption on productivity.

This program provides a unique opportunity to study how access to market information might affect firms' strategies and decision-making. A closely related paper to ours is Kim (2019) in that it provides evidence that small firms may lack knowledge of competitors' decisions even when this information is readily accessible.[10] She shows that, upon receiving concrete information about their closest rivals' prices, small firms change their strategies to align closer to their competitors' strategies. Moreover, she finds suggestive evidence that

[9] Goldfarb and Tucker (2019) survey the literature on the economics of digitalization and IT.

[10] Other recent papers with a similar informational context are Nagaraj (2020) and Fabregas et al (2019). The former explores the impact of public data infrastructure and shows how better information increases market entry. The latter provides empirical evidence of smallholder farmers' valuation for neighboring agricultural information.

managerial inattention plays an important role in explaining the firms' lack of awareness. Our paper differs from hers in two noticeable ways. First, our data contains the universe of establishments using POS across all retail sectors and, therefore, our results are richer and applicable to policy design. Second, we show firms are able to react to access to rich multidimensional information comparing their own client portfolio to that of their competitors. Our findings support the view that it is managerial inattention and lack of awareness, and not a lack of sophistication or cognitive capacity, driving the change of behaviour upon information access. In our setting, information allows firms to increase their revenues by becoming aware of existing, unexploited business opportunities. Moreover, in line with findings in previous research, such as Bloom et al. (2013), Bruhn et al. (2018) or Giorcelli (2019), we find evidence consistent with an improvement in the internal resource allocation of firms, SMEs in our case, after receiving information. Interestingly, the impact of information is not larger for more sophisticated establishments. This finding is likely due to the combination of two factors. First, firms that are more sophisticated were probably already using a considerable amount of information in their decision-making process prior to adoption. Second, the information-sharing program in our study processes the information facilitating its understanding and use by less sophisticated managers.

While managerial implications of our findings are clear for managers of SMEs, policy implications are even more relevant. In our setting (an average OECD economy like Spain), large firms (more than 50 employees) account only for less than 1% of all firms in the country and 48% of employment whereas SMEs account for more than 50% of employment and almost 99% of firms.[11] These patterns in the size distribution of firms and employment are representative for all industrialized and OECD countries. To the extent that our results provide estimates of the private returns of access to market information for SMEs, intervention and government policy aiming to correct for socially inefficient adoption is desirable.

The structure of the paper is as follows. We describe the empirical setting and our data in section 2. Section 3 lays out the methodology and discusses identification. In section 4, we describe our main results and explore the role of demand-discovery mechanisms. Section 5

[11] See information on the size distribution of firms in Spain here, http://www.ipyme.org/Publicaciones/Retrato-PYME-DIRCE-1-enero-2019.pdf.

provides additional results regarding other mechanisms and aggregate effects of adoption. Finally, section 6 concludes.

2. Institutional Detail and Data Description

2.1. The Bank

Our empirical setting is the market for SMEs in Spain. Our data come from one of the largest European banks with a global presence, and with a high market share presence in Spain. Hereafter, we refer to the data provider as "the bank". The bank is also a major player in the credit card market both as credit (and debit) card issuer and as credit card POS provider.

Amidst its prevalence and salience in the marketplace, the bank launched a pilot program for its POS clients in one region of Spain in the fall of 2014 and went national in the spring of 2016. The program aimed to bring the benefits of Big Data technology to SMEs using the bank's credit card POS.[12] The bank provided this program for free, and adoption was voluntary. It is also important to note that the bank did not compensate its employees for the diffusion of this program, which explains the ex-post scant adoption rates of the technology. If anything, bank employees would offer the adoption of the program to a client as a source of value added to an already existing business relationship.

To join the program, a POS client would follow a two-step process. First, the client would physically visit a bank branch and meet with a branch employee that would facilitate signing up for the program. Once the client had signed up, the bank would send her an email with setting up information for accessing the incoming monthly reports. Second, the client would need to follow the indications in the email received. These instructions would prompt the client to answer a few questions regarding her analytical and marketing savviness. At this point in the process, the newly signed up customer became familiar with the online platform that the bank used to deliver its monthly report. This platform contained different tools and orientation videos to familiarize the client with the report information and therefore maximize the understanding, accessibility and customer experience from this service. Finally, note that when clients signed up for this service online, they had to acknowledge a

[12] A Bank manager supervising the program went on public record to describe the program as "This program brings data technology available only for big firms to SMEs. Through this tool, retailers can get to know better their sector and customers. This allows them to improve their decision making."

waiver on their liability with the program. Regardless of when a customer signed up for the program, the signee would receive its first report during the first week of the following calendar month.

Upon opting in for this service, the bank generated a monthly report for each adopter, which became available through the program's online platform. It is necessary to highlight that the nature of the information in the report is descriptive and not predictive. Accordingly, the report contained summary statistics regarding the number and value of credit card transactions in the previous month. The report disaggregated this information on credit card transactions by client demographic groups such as age, gender and zip code, as well as other classifications such as new vs. returning customers or the time and day of transactions. The report also contained the same set of aggregated information for business competitors in the same zip code. This set of information on each store's direct competitors provided a reference point and allowed program participants discover differences between their own performance and client portfolio and those of their closest competitors. In other words, each monthly report effectively provided precise market research information on the local market in which each program participant operated.[13]

To understand further the program, we need to describe the nature of the information used to generate the monthly reports. The reports originated from credit and debit card transactions made by both bank-issued cards in all POS in the country (both POS from the bank and from other financial institutions) and other bank-issued cards in the POS of the bank. Because the bank of our study holds a substantial market share in the credit card market in the country, the report information issued by the program and received by the adopters was representative of the population of credit card transactions in the market for both the adopter and her competitors.

2.2. Data Description

Our data is the universe of all transactions from credit cards issued by the bank from January 2014 to December 2018. The data is unique in that it details, for each transaction,

[13] Figures A1 and A2 provide samples of some of the information contained in the monthly reports as well as the presentation of the information. The content of these figures is not exhaustive of all the information in the reports.

establishment-specific and card-specific identifiers. On the one hand, it is important to note that we observe any establishment in the country as long as this establishment has an active POS. The data set also contains information on the establishment location, sector and subsector. On the other hand, the data contains cardholder information at the card level such as age, gender and residence zip code. A zip code in our context is equivalent to a 5-digit zip code in the US.

Overall, the raw data contains transaction-level information for nearly 2.5 million establishments distributed across all provinces, 18 sectors and 77 subsectors.[14] Because of our confidentiality agreement with the bank, we aggregate transaction information at the establishment-quarter level. Additionally, we make two other changes to our initial data set. First, we drop all establishments with less than 5 transactions on average per quarter. Second, we focus our analysis on all establishments in sector-zip code pairs where we observe, at least, one adopter during our sample period. These changes decrease computational burden while preserving all the within-zip code-sector variation in technology adoption from the original data. This variation is precisely what will allow us to achieve our goal of estimating the impact of technology adoption at the establishment level.

Our final working data set contains information from a total of 487,610 establishments, including all 7,100 program adopters in the universe. Figure 1 shows the evolution of the number of adopters from July 2014 to end of 2018. While the bank first launched the technology as a pilot program in a few locations, its official launching took place in mid-2016 where the number of adopters increased rapidly to a level right around 7,100 in late 2018. This number represents approximately 1.5% of the total number of clients of the bank with a point of sale and 0.3% of establishments with a point of sale in the country.[15] Table 1 shows that our data set accounts for a total of 4,610,085 establishment-quarter observations. In our sample, the average establishment collects 4,715 Euros per quarter spread across 120 transactions. These distributions are clearly skewed, as the average transaction value is 64

[14] See Table A1 for a list of sectors and subsectors.

[15] See our description of the technology adoption in section 2.1. and the introduction of our IV strategy later in the paper to understand why the adoption rate was as low as 1.5%. Namely, bank employees were not compensated directly for its diffusion. Moreover, it is worth noting that, although a previous literature has raised concerns about the possibility that an increase of transparency on the supply side can facilitate collusion (Stigler, 1964; Tirole, 1988), this technology in particular does not present any serious threat of collusion due to its low adoption rate.

euros. Finally, it is important to note that the average store sells to 74 customers in a quarter and the average value per customer is 85 Euros.

The bottom half of Table 1 describes these variables and other characteristics that we use to explore impact heterogeneity for the subsample of 7,100 adopters.[16] The average adopter collects 6,200 Euros per quarter in 153 transactions with an average transaction of 80 Euros. Each adopter serves 92 customers per quarter, each of which spends 102 Euros on average. Finally, adopters have on average of 75 competitors of the same sector in their same zip code. Finally, we use the fact that adopters may answer three different questions regarding their analytical, marketing and digital capabilities when registering onto the online platform that will grant them access to the monthly reports.[17] Each one of these questions provide Likert scales from 1 to 5. We create a measure of analytical savviness by averaging all three answers from all adopters who answer all three questions. Not all adopters respond to this questionnaire. In fact, only 3,495 adopters out of the total 7,100 responded (49.2%). The average sophistication score following this measure is 3.53, with a median of 3.67 and a standard deviation of 0.89. Once we have described our data, we proceed to present our empirical methodology in the following section.

[16] Among subsectors, those with most adopters were restaurants (952), clothing stores (726), grocery stores (604), hairdressers and beauty stores (411), bars and cafeterias (404), jewelry stores (333), motor garages and car dealerships (287), shoe shops (270), hotels (245) and pharmacies (236). The other 2632 adopters (37% of all adopters) were distributed among the remaining 60 subsectors.

[17] The three questions and potential answers are as follows. First question: How digital are you? (1) I do not use computers often or internet in my daily file; (2) I have an email account. I use internet to see the news, search for information, etc.; (3) I have personal social media. I use internet daily. I use internet to communicate with my customers/providers; (4) I have social media and business webpage. I have hired a product online at least once. I use internet daily to communicate with my customers/providers; (5) I make internet-based marketing campaigns and analyze the traffic in my webpage. I use online tools for management.

Second question: Do you use data for management? (1) I only use intuition-driven management practices. I think measuring and analyzing data has no value for my business; (2) I think there is a value in data, but I do not know where to find data or what I could use it for; (3) I analyze my sales periodically. I read news articles with information about my sector, and think how to apply this to my business; (4) I measure my sales and analyze the data in order to improve my performance. I have a database with my customers' contact. I search on the internet information about my sector; (5) I have a database /CRM with detailed information about my customers, and I use this to make promotions. I analyze my sales margins by product. I buy market studies to plan my activity.

Third question: What is your relation with marketing? (1) I never do marketing campaigns; (2) I take care of my shop window and my service to attract and increase customer loyalty, but I never do marketing campaigns out of my establishment; (3) I make promotions, 2x1, gifts, etc. Sometimes I have made mail campaigns or bought advertising space; (4) I frequently make marketing campaigns, advertising and discounts. I use email and social media to cultivate customer relations; (5) I have a marketing plan in which I design campaigns and events. I inform my clients about customer-specific promotions. I count with a loyalty program. I advertise my business in the media (physical advertising, press, or the internet).

3. Empirical Methodology and Identification

3.1. Baseline Regression Specification

First and foremost, we are primarily interested in estimating the returns to adoption of the Big Data IT program in our sample. For that purpose, our baseline specification is such that,

$$Y_{isjt} = \mu + \beta Adoption_{isjt} + \gamma X_{isjt} + \alpha_i + \theta_{sjt} + u_{isjt} \qquad (1)$$

where Y_{isjt} is the log of the outcome variable such as revenues, number of transactions, or number of new customers for establishment i in sector s located in zip code j and quarter t. Our main variable of interest is $Adoption_{isjt}$, which is a dummy variable that takes value 1 if establishment i has adopted the technology before quarter t, and 0 otherwise. This variable varies within establishment over time for adopters, and remains at 0 for non-adopters. See Figure 2 for a representation of the timeline between the time when an establishment signs up for the program, the delivery of its first report and our variable $Adoption_{isjt}$ taking value 1. In this example, the establishment signs up in the middle of the second quarter (month 4) and only starts receiving a report on May 1st. Our adoption variable takes value 1 in the quarter following adoption and all quarters after that.

Our regression specification also includes time-varying controls X_{isjt} such as dummies for the first four quarters an establishment enters our sample (new establishments may display different growth at the beginning of their existence) as well as establishment fixed effects α_i and sector-zip code-quarter-specific fixed effects θ_{sjt}. Finally, u_{isjt} is our residual.

Our working specification will take first differences from specification (1) above,

$$\Delta Y_{isjt} = \beta \Delta Adoption_{isjt} + \gamma \Delta X_{isjt} + \theta_{sjt} + \Delta u_{isjt} \qquad (2)$$

where ΔY_{isjt} is first differences in our dependent variable, and $\Delta Adoption_{isjt}$ is first differences in the technology adoption dummy variable in specification (1). It is important to note that $\Delta Adoption_{isjt}$ takes value 1 in the quarter right after adoption and value 0 in all other quarters. This specification in first-differences also contains controls X_{isjt} such as dummies for the first four quarters after an establishment enters our sample, sector-zip code-quarter-specific fixed effects θ_{sjt}, and a residual Δu_{isjt}.

Before coping with endogeneity concerns in the next subsection, we argue here that estimating our parameter of interest β with first-differences allows us to tackle several potential issues of identification. On the one hand, first-differences are equivalent to introducing establishment fixed effects and therefore controls for any time-invariant correlation between the error term and the probability of adoption at the establishment level. On the other hand, our context is far from being stationary, and therefore first-differences estimation partially addresses issues of autocorrelation in the error term. Finally, this regression specification relaxes the requirement of strict exogeneity in the regressors only requiring weak exogeneity for the consistency of estimates.

3.2. Instrumental Variables and Identification

A pervasive concern in the technology adoption literature, and elsewhere in any empirical study in economics, is the endogeneity and self-selection of establishments into adoption of a technology. In our context, this concern is problematic if the establishment-specific idiosyncratic error terms are correlated with adoption. First, we can think that adoption may be more likely in high performance establishments. To address this concern, we include a dummy variable in our specification and estimate our main specification in first differences. Second, adoption may also be more intense in sectors or areas receiving positive temporary shocks. We control for this by including sector-zipcode-quarter fixed effects. Third, adoption may be more likely in establishments growing more intensively. To control for this issue, we will show robustness results including establishment specific trends. Finally, other issues that can compromise identification of the impact of adoption on outcome variables include sporadic episodes of positive or negative growth that coincide with the timing of adoption (e.g., adoption may coincide with the implementation of other investments), or an increase in the incentives of an adopting establishment to sell more by credit card vs. cash in order to obtain more information about its customers. In these cases, the first-differences regression specification (2) with OLS will erroneously attribute changes in productivity to technology adoption. To address this concern of endogeneity in the adoption decision, we use an IV identification strategy.

To this end, we look for changes in an establishment's environment that may exogenously change the probability of adoption across establishment within sector-zip code-quarter triads while being orthogonal to establishment-specific productivity and demand shocks. With this goal in mind, we derive an IV strategy that exploits the fact that different establishments in the same sector-zip code dyad may hold their business bank account in different bank branches located in different zip codes. Hereafter, we call the bank branch where an establishment has its business bank account the establishment branch.

Our conversations with bank managers provide a strong foundation for our IV strategy below. As explained in our institutional detail section, the bank did not compensate its employees for the diffusion and adoption of this technology. If anything, HQ paid for brochures and advertising boards and distributed them equally among bank branches. The variation in adoption across branches was rooted in the affinity of their employees with the program. The larger the affinity of an employee, the higher the level of her promotional effort despite not being compensated for it. In other words, the distribution of employees' affinity to the program across bank branches, and therefore the distribution of promotional effort across bank branches, is orthogonal to the distribution of potential gains from adoption of the program across establishments' branches.

Our IV is the number of adopters per quarter (across sectors and zip codes) in the establishment branch, other than the focal establishment. Figure 3 sheds light on the rationale behind our IV. Assume two zip codes, A and B. Each zip code has a bank branch. There are two bakeries in zip code A (bakery 1 and bakery 2) and one pharmacy in zip code B. Our instrument highlights the variation in establishment branch for each of the establishments' location. While bakery 1 located in zip code A uses the bank branch in zip code A, bakery 2 also located in zip code A uses the bank branch in zip code B. The pharmacy in zip code B uses the bank branch in zip code B. Per our IV strategy, the adoption of the pharmacy located in zip code B (using the bank branch located in zip code B) rises the probability of adoption of bakery 2, but it does not affect the probability of adoption of bakery 1 in zip code A.

Specifically, and through the lens of our example in Figure 3, the increase in the probability of adoption of bakery 2 may come from two different channels. On the one hand, branch employees in zip code B may exert larger promotional effort on the diffusion of the program, and therefore increase the probability of adoption of all its corporate clients, including the pharmacy and bakery 2 (as explained before). On the other hand, the pharmacy's adoption also increases the probability of adoption of bakery 2 through peer effects at the establishment branch level. In our empirical application, we do not observe promotional effort of the program at the branch employee level. Therefore, our instrument relies on variation across bank branches in the number of adopters over time to proxy for differences in promotional effort.

Our identification strategy posits that the number of adopters at the establishment branch (as opposed to the branch in the same zip code of the focal establishment) increases the probability of adoption because that is proportional to the promotional effort of the employees. In our example of Figure 3, the pharmacy adopts the technology and that increases the probability of adoption of bakery 2 because they share the same establishment branch. In contrast, the probability of adoption of bakery 1 does not change due to the pharmacy's adoption despite being in the same sector and zip code as bakery 2 because bakery 1 does not share establishment branch with the pharmacy. Therefore, our instrument provides variation in the probability of adoption across establishment in the same sector-zip code dyad.

Reached this point, our identification strategy needs to address the validity of our exclusion restriction. Our strategy exploits differences in probability of adoption across establishments within the same sector-business-quarter triad, which in fact takes into account all sector-zip code-quarter level productivity and demand shocks. Then, our exclusion restriction assumption would fail if a correlation exists between establishment-specific shocks and promotional effort of the establishment branch within a quarter. Equivalently, heterogeneous trends in performance within sector-zip codes across different establishments affiliated to different establishment branches would also violate our exclusion restriction.[18]

Moreover, our identification strategy does not rest on the assumption that different establishments within a sector-zip code dyad with different establishment branches are alike. Even if there is self-selection of establishments into different branches of different

[18] Note that we include bank-branch time trends to control for this possible concern in a robustness specification (displayed in our main results in Table 2).

characteristics (perhaps located in different zip codes), our identification strategy exploits differences in promotional effort of the program over time within branch and mostly relies on the timing of promotional effort being orthogonal to the timing of program introduction to market.

Note that even if there exists peer-effects between establishments of a same sector-zip code dyad that do not share establishment branch, this alternative mechanism would work against the variation provided by our instrument utilized by our identification strategy. Nevertheless, it is also important to point out that (1) the introduction of sector-zip code-quarter perfectly captures this type of peer effects between establishments in the same sector and zip code, and (2) this second order effects should not be a concern for our exclusion restriction. Ultimately, a necessary exclusion restriction for the plausibility of our IV is that sharing the same establishment branch only affects the probability of adoption, but it does not directly affect performance.[19] Finally, it is paramount to emphasize the fact that the bank did not introduce any other program with [partially or fully] overlapping characteristics during our sample period.

4. Results and Mechanisms

We describe the results of our empirical analysis in three different steps. First, we show our main results of running regression specification (2), and follow up with exploring heterogeneity in the impact of adoption of the technology. Second, we continue our analysis by investigating mechanisms driving the main results. Third, we conclude this section by evaluating the overall impact of the technology at the sector level, and discussing how our results link back to the existing literature.

4.1. Main Results and Heterogeneity

Table 2 shows the results of running our baseline specification where the dependent variable is the log of quarterly credit card revenue per establishment. Columns 1-3 run OLS regressions in first-differences under alternative deviations of the baseline specification. Column 1 shows that adoption is associated with an increase of 4.5% in revenue. Columns 2

[19] As a robustness check, we produce evidence in Table A2 in the Appendix where the IV does not include peers in the same sector.

and 3 are the result of running leads and lags dummies of the adoption quarter. On the one hand, column 2 shows that the increase in revenues is concentrated in the quarter after adoption and we do not observe further increases or decreases in subsequent quarters. This means that the gains in revenues are permanent up to a year after adoption. In fact, although not shown here, we do not observe a reversion to the mean when including further leads up to $t+7$.[20] On the other hand, column 3 runs a placebo test by including as a regressor a dummy that takes the value of one in the quarter prior to adoption (the quarter before the first-difference of the adoption variable takes the value of one). Results show there are no increases in revenue preceding the quarter of adoption.

Columns (4) and (5) of Table 2 show results of implementing our IV strategy. Column 4 shows estimates of the first stage results. Column 5 shows the results of running our IV strategy on the baseline specification of Column 1. We find that the effect jumps from 4.5% to 9.0%. We carry out the Hausman test to check for endogeneity, and we cannot reject the null hypothesis that adoption is exogenous (p-value=0.24).[21]

A potential concern is that bank branches with more adopters are also branches that attract better firms, perhaps because they have better account managers, and these better firms are more likely to both sign up for the program and grow. For this reason, in columns (6) to (8) of Table 2, we run OLS and IV specifications with bank-branch specific time trends capturing differences over time across bank branches and the respective firms selecting into those bank branches. Our results are robust and similar to those using earlier specifications. OLS results show that adoption is associated with a 4.4% increase in revenues and an 8.4% casual increase.

Following Forman et al. (2012), we believe the estimate magnitude is larger than in the baseline regressions due to the existence of heterogeneous returns to technology adoption. If bank branches that have customer establishments with higher potential returns of technology adoption made more promotional effort, then we are likely to observe a jump in their estimate of returns from adoption when applying our IV strategy. In the same manner, if bank branches

[20] Results available upon request.

[21] While Table 2 presents our baseline results, appendix Table A2 includes regression specifications with establishment-specific time trends and subsector-zip code-quarter FE. All our findings are robust to changes in the specification. Table A2 also shows results where the IV does not include peers in the same sector. In addition, appendix Table A3 includes an event-study exercise in which the sample is limited to only adopting establishments.

with a higher affinity of its employees to the technology not only exert higher promotional effort, but also do a better job in explaining the characteristics and functionalities of the technology, it is likely the case that adopters in those branches obtain higher returns from adoption. In other words, there are reasons to believe the local average treatment effect may be larger than the average treatment effect. This implies that although the instrument affects revenue only through its impact on technology adoption, the returns to technology adoption are larger for those establishments whose adoption decisions are most strongly affected by our instrument.

Once we have determined that technology adoption causally increases establishment revenue by 9%, we investigate the presence of heterogeneity in this effect. We explore heterogeneous effects in two different ways. First, we investigate heterogeneous effects across sectors, subsectors and geographical regions. We plot the distribution of effects across these three dimensions in Figure 4. Note that all three distributions of effects are centered around zero, and that the heterogeneity across sectors shows the lowest variance with range between -0.25 and +0.25. The distribution with largest variance is across subsectors ranging from -0.5 to 1, and the distribution across regions is in between those as it reflects different distributions of sector and subsectors across regions.[22]

Second, we investigate heterogeneous effects across different establishment characteristics. For this purpose, we split our sample of adopters into three different dimensions: analytical savviness of the adopter, establishment size prior to adoption, and degree of local market competition. We report our heterogeneity results for both OLS and IV control function in Table 3. Columns 1 and 2 investigate how analytical savviness drives the impact of technology adoption. For this matter, we take advantage of the fact that the adopters must answer three different questions regarding their analytical, marketing and digital capabilities when registering onto the online platform.[23] Answering these questions will grant them access to the monthly reports. We use their answers to compute our measure of sophistication

[22] Retail sectors benefitting more from adoption are technologies, home wellness and beauty, and accommodation. Retail sectors benefitting less are sports and toys, and supermarkets. A closer look into subsectors shows positive returns of adoption (other than the above mentioned sectors) for tobacco stores, car rental shops, musical instruments, photography, fast-food restaurants, and gardening and floristry. Subsectors with negative returns are pubs and discos, press, optician shops, and gas stations. So far as geographical regions are concerned, those with a higher number of inhabitants (and adopters) make up for most of the centered distributions of returns around 5-8%, while positive and negative outliers correspond to small regions.

[23] See Footnote 16 for a detailed description of all three questions.

in this table. Hereafter and for simplicity, we call this variable *level of sophistication*, and we create dummies for adopters above and below the median level of sophistication among adopters. Column 1 runs OLS first-difference regressions and shows that adoption is associated with increases of 4.4% and 4.6% in revenue for adopters above and below the median level of sophistication, respectively. Column 2 applies our IV control function approach and shows that the returns are now 8.7% and 9.7% for adopters above and below the median level of sophistication, respectively. Note that Table 3 reports the p-value of the test for equal returns for both firm types. According to the reported p-values of 0.95 and 0.75, we cannot reject that these rates of return are statistically the same. This suggests the role of cognitive capacity does not seem to drive returns to technology adoption.[24]

Next, we explore how establishment size correlates with the impact of technology adoption. We measure size by the average quarterly revenue of an establishment in all observed quarters prior to adoption.[25] We then create a dummy variable "*Large*" that gives value 1 to an establishment if its size is above the median size of adopters in the same sector, and 0 otherwise. We also create a dummy variable "*Small*" that gives value 1 to an establishment if its size is below the median size of adopters in the same sector, and 0 otherwise. Column 3 shows that the impact of technology adoption in large establishments is not statistically different from zero, and it is 7.96% in small establishments. When applying our IV strategy, the estimate for large establishments becomes statistically significant at 7.3% and the estimate for small establishments increases to 14.6%. These findings point out that the returns to access this technology vary greatly with establishment size. We are able to reject that these returns are the same, so we can safely conclude that smaller establishments benefit more from technology adoption.[26]

---

[24] The program provides, processes and analyzes data for the adopter and, therefore, the information provided is easy to understand. This is consistent with the fact that we observe an impact even for less sophisticated adopters. This finding is important when considering policy implications regarding access to Big Data IT technology of less sophisticated and smaller establishments. Appendix Table A4 includes results disaggregated for establishments above and below the median level of sophistication for each of the categories of analytical, marketing and digital capabilities. Interestingly, we find that establishment with low analytical sophistication seem to be obtaining higher returns from adoption than those with high analytical sophistication. By contrast, establishments with higher digital sophistication obtain higher returns of adoption than those with lower digital sophistication.

[25] The results remain qualitatively unchanged when measuring size by market share within a sector-zip code or subsector-zip code.

[26] Later in the paper, our analysis of mechanisms will show that information increases establishment revenues by highlighting untapped business opportunities and streamlining resource allocation. Larger establishments

Finally, Columns 5 and 6 explore the heterogeneity of the results along the dimension of the degree of local market competition. For this purpose, we calculate the average number of competitors in the same sector and zip code for each adopting establishment over the sample period.[27] The number of competitors averages 74 with a median of 45 and a standard deviation of 90 (with a highly skewed distribution ranging from three to 967)**.** Once again, we create dummies that divide the adopters into those above and below the median number of competitors. Results in column 5 show that the association between adoption and revenue increases is statistically significant for establishments in highly competitive markets, and it is not statistically significant for establishments in less competitive environments. Column 6 reports our IV results and shows that adoption increases revenue by 11% in more competitive markets. These results across more and less competitive sector-zip code dyads are statistically different from each other at the 11% level.[28]

In summary, our heterogeneity results are insightful in depicting scenarios where technology adoption yields higher returns. Our findings in Table 3 show that those establishments of smaller size and those operating in markets that are more competitive derive higher returns from adoption. Sophistication and digital experience do not seem to matter for the returns to technology adoption in our context.

4.2. Mechanisms

Our findings in the previous section establish that technology adoption increases establishment revenue by 9%. Moreover, we also find that this effect is heterogeneous. In fact, smaller establishments and those in more competitive markets seem to benefit more from adoption. In this subsection, we aim to understand the mechanisms behind our findings. In our empirical setting, establishments adopt a technology that provides information on their performance relative to others in their local market. Using the fact that the reports gather

---

most likely performed well in both dimensions already before adoption, thus their larger size. Therefore, we expect the scope for improvement of big establishments in both margins to be smaller and, accordingly, adoption returns less significant.

[27] The results remain qualitatively unchanged when measuring competition by the average number of competitors in the same subsector and zip code.

[28] Appendix Table A5 shows results on the heterogeneous returns of adoption for early vs. late adopters. Interestingly, we find that *(i)* returns, in general, do not differ between early and late adopters of the program, *(ii)* returns, however, seem to be negligible for establishments adopting during the pilot period, *(iii)* returns are higher for the first adopter in each sector-zip code than for later adopters, and *(iv)* sign-up time of adopters in the quarter previous to our adoption dummy taking value 1 does not impact returns.

information of different nature, we separate the new information obtained by adopters into different dimensions that point out distinct mechanism with different direct effects and implications on adopter behavior. The report received may highlight business opportunities that the establishment was not aware of or did not pay much attention to in the past. The receipt and processing of this information may drive an adopter to serve different customer profiles, that is, different age-gender groups, customers from nearby zip codes, or customers that purchase their goods and services during different times of the week. Because this mechanism requires exploiting new business opportunities, hereafter and for simplicity, we call these "demand-driven" mechanisms.[29]

4.2.1. Number of Transactions and Customers

We start our analysis of mechanisms by investigating how the increase in revenue relates to the number of transactions and customers. Table 4 shows results using three different dependent variables. While Column 1 shows that the adoption of the technology is associated with an increase of 3.9% in the number of customers, Column 2 uses our IV strategy and reports that the causal effect of technology adoption in the number of customers is a 12% increase. Parallelly, Columns 3 and 4 investigate whether the increase in the number of customers comes paired with changes in the revenue per customer. This would happen if new customers were spending more or less than original customers, or if old customers were changing their spending levels. We find no changes in the average revenue per customer. Moreover, in Columns 5 and 6, we find no changes in the number of transactions per customer. These results suggest that adopting firms are able to attract new customers but that new customers purchasing patterns are not statistically different from old customers in two important dimensions as amounts spent and number of transactions.

[29] While providing evidence on the specific actions that adopting establishments implemented after adoption would be interesting, our data does not provide such categorical and qualitative information that would capture changes in marketing and promotional strategies or internal changes in the organization of the workplace. In conversations with the bank, we had discussed the possibility of designing a survey for adopting establishments to gain a better understanding of these aspects, but changes in EU privacy law and potential exposure to legal liabilities on the bank side prevented us from following down that path. Our findings show that mean transaction values do not change due to adoption, which is consistent with prices remaining stable. Our results also show that adopting establishments are able to increase revenues by discovering new business opportunities, attracting new customers and diversifying their client portfolios, but we are not able to show detailed evidence on the specific actions implemented to that end.

Table 5 turns to the study of the impact of technology adoption on the number of transactions and the average transaction value. We find that, consistently with our results on Table 4, technology adoption increases the number of transactions by 4.4% in the OLS specification and 13% in the IV specification. Moreover, although the change in revenue per transaction is not statistically different from zero in the OLS specification, it shows a drop of 3.9% in the IV regression. These findings suggest that adopting establishments are not increasing (perhaps even decreasing) their prices post-adoption, consistent with a potential increase in head-to-head competition of adopters with their sector competitors in their zip codes.

4.2.2. Customer Demographics and Geographical Reach

Once we have established that technology adoption facilitates the discovery of new business opportunities through increases in the number of customers and the number of transactions, we continue our analysis by examining whether the average demographic profile of the customers of an establishment changes upon adoption. Because we show in Tables 4 and 5 that there are differences between the average customer pre-adoption and average customer post-adoption, we now examine changes in the customer profile by age, gender and zip code. Importantly for us, the report identifies customers according to two gender groups (male/female), six age groups (<25, 25-34, 35-44, 45-54, 55-64, 65>), and customer dwelling zip code. Moreover, the report highlights the most important customer profile of the adopting establishment and its sector-zip code independently. This allows establishments to identify differences between their own main customer type and the main customer type of their closest competitors in the same sector and in the same zip code.

For this purpose, we identify the main customer type (one of the 12 gender-age groups described above) for each sector-province dyad and calculate the share of revenues from each establishment's main customer type according to their sector-province dyad. Then we create two dummy variables: "*Large Share*" that equals 1 if the share of revenues from the main customer type is above the median among all adopters, and 0 otherwise; and "*Small Share*" that gives value 1 to an adopter if the share of revenues from the main customer type is below the median among all adopters, and 0 otherwise. Columns 1 and 2 in Table 6 show that technology adoption does not significantly change the share of revenues from the main customer type. Yet, in Columns 3 and 4 we investigate whether the no-effect is a true no-effect or is the result of compositional issues. Indeed, findings in Columns 3 and 4 show that those establishments with larger shares of the main customer type pre-adoption are likely to decrease the share of revenues from the main customer type upon adoption. Conversely, establishments with smaller shares increase their share of revenues from the main customer type. Column 4 uses the control function IV and so we include the error term of the first stage as control in the regression specification.

In appendix Table A6, we investigate whether our results on the changes in the share of revenue from the main customer type are driven by the numerator (more or less sales to this customer type) or the denominator (more or less sales to other customer types and in total). Our findings in appendix Table A6 show that findings in Columns 1 to 4 of Table 6 are driven by: (1) establishments with small share of sales to the prime customer group increasing their sales to this group; and (2) establishments with a high share of sales to the prime customer group decreasing the share of their sales to this group as a result of selling more to other groups but not reducing their sales to the prime customer group.

In columns 5 to 8 of Table 6, we examine how the diversity of the customer profile per establishment changes with adoption. Our dependent variable uses information of the shares of revenue per each of the 12 age-gender groups in each establishment and computes a Herfindahl-Hirschman Index (HHI hereafter) of customer diversity. Our HHI measure would take value 1 if an establishment sold 100% of their goods and services to only one of the 12 groups, and would take value 0.083 if it sold equally to all 12 age-gender groups. Columns 5 and 6 show that technology adoption decreases the concentration of sales by 3.4%.

Columns 7 and 8 explore whether the decrease in concentration comes from establishments with high or low degrees of concentration pre-adoption. For this purpose, we compute the HHI of customer type concentration for each adopter pre-adoption. Then we create two dummy variables: "*High Concentration*" that equals 1 if the HHI of the establishment is above the median among all adopters, and 0 otherwise; and "*Low Concentration*" that gives value 1 to an adopter if its HHI is below the median among all adopters, and 0 otherwise. Results in Columns 7 and 8 show that the decreases in concentration in Columns 5 and 6 are entirely coming from establishments with high concentration rates pre-adoption. Those establishments in the upper half of the concentration distribution decrease concentration by 8.7% after adoption. These results indicate that upon receiving information about competitor strategies and market opportunities, establishments tend to diversify more their client portfolio instead of focusing on a narrower customer niche. Just like we did in column 4, column 8 uses the control function IV and so we include the error term of the first stage as control in the regression specification.

Finally, Table 7 investigates whether adopters change the spatial composition of their customer base. For simplicity, we compute for each establishment the share of revenue from customers residing in other zip codes. Columns 1 and 2 show that technology adoption does not seem to have an effect on the average share of revenue from customers in other zip codes. Building from this finding, columns 3 and 4 decompose the main effect into adopters with large and small shares of revenue from customers in other zip codes pre-adoption. We create two dummy variables: "*Large Share*" that equals 1 if the revenue share coming from customers in other zip codes is above the median among all adopters, and 0 otherwise; and "*Small Share*" that gives value 1 to an adopter if its revenue share coming from customers in other zip codes is below the median among all adopters, and 0 otherwise. Our results provide evidence that the increase in revenue from customers in other sectors is concentrated in establishments with low share of such type of customer pre-adoption. Column 4 uses the control function IV including the error term of the first stage as control in the regression specification, and shows that those establishments with a lower-than-median share of customers from other zip codes increase their share of customers from other areas in 2.6 percentage points or 3.1% over the mean in the sample.[30]

As far as the demand-driven mechanism is concerned, evidence in Tables 6 and 7 is consistent with a mechanism where establishments discover new business opportunities and implement new marketing strategies to take advantage of the new (to them) information. Given the size of the SMEs in our sample, the new marketing strategies ought to be as simple as paying more attention to customers in age-gender groups underserved prior to adoption, or taking an interest on those first-time customers that are likely to reside in another zip code. We do not expect the SMEs in our sample to engage in ambitious advertising campaigns or grandiose promotional events to attract more customers.

[30] A potential concern with our IV strategy is that those bank branches with more adopters may have also been located in zip codes where establishments were more likely to have higher rates of out-of-zip code customers. We find no statistically significant correlation between our IV and the share of out-of-zip code customers in our data.

In summary, our findings show that the increase in revenue comes as a direct consequence of establishments expanding their customer portfolio in a variety of ways. Adopters do not just increase their number of customers, but they target new age-gender profiles and look for customers beyond their zip code.[31]

5. Additional Results

Following the previous section, we consider different additional effects of the program adoption on the internal operations of adopters, as well as the impact of adoption on non-adopting competitors and the impact on aggregate sector revenues.

5.1. Allocation of Effort among Peak and Off-peak week times

Alternatively, we consider the impact of the newly revealed information contained in the report received by adopters on their time management and internal operations. In addition to the discovery of new business opportunities, establishments may also learn that competing establishments organize their sales in different days of the week and times of the day. In some cases, a reorganization of their time schedule during the week may help establishment managers improve the logistical efficiency of their allocation of resources such as personnel, time and effort. Upon receiving information from the monthly report, an establishment may reallocate clients and sales to different parts of the week (days or hours), improving the distribution of workload during the week.

Following this logic, we study changes in the distribution of revenue across different days and hours upon technology adoption. To do so, we divide the week in 4 time slots, namely, weekday morning (until 3 pm), weekend morning, weekday evening (after 3 pm) and

[31] An interesting question is whether the 9% increase in revenue translates into a 9% increase in profits or whether input consumption (number of employees, open hours, etc.) also increases potentially making profit increase null. Although we do not have information on profit margins, we can abstract from any improvement in resource allocation about what can happen when establishments adopt. On the one hand, if establishments were operating with excess capacity, discovering new opportunities would increase profits by increasing revenue without increasing input consumption. On the other hand, if establishments were operating at full capacity and optimal scale, they would not be willing to increase production, which is at odds with the increase in revenue we see in the data. If they were not operating at an optimal scale, discovering new business opportunities would imply increasing input consumption to increase revenues and increasing profits. Therefore, we can conclude that adoption results in an increase in profits, albeit perhaps smaller or larger than the 9% increase in revenue. The improvement in resource allocation would reinforce the profit increase.

weekend evening. We identify the peak shopping time for each sector-province dyad and calculate the share of revenues from each establishment's peak shopping time according to their sector-province dyad. The average establishment had 37.3% of its sales taking place during its shopping peak time. Table 8 regresses the log of revenues at the peak and off-peak (the other three time slots) times of the week on adoption. While columns 1 and 2 show no change in the sales at peak time, columns 3 and 4 show an increase in the sales at off-peak time. This relative shift of revenue among business hours could be driven by *(i)* a supply-side gain in efficiency, or *(ii)* shifting business to serve new demographics with different shopping schedules – this would be encompassed in what we have called demand-driven mechanism. To distinguish between these two explanations, columns 5 and 6 control for changes in demographics of the clientele. Once we control for changes in demand demographics (log of sales for each of the 12 age-gender customer categories and log of sales for out-of-zip code customers), the magnitude of the effect goes down from 17% to 8%. While not shown here, this finding is robust to controlling for changes in the HHI of customer types, the share of out-of-zip code sales, or total sales. These findings suggest that technology adoption triggers changes in business hours not explained by changes in customer demographics and, therefore, those changes may be due to improvements in supply-side efficiency.[32]

An alternative way to investigate this same issue is parallel to our analysis in the customer portfolio above, and implies the use of the concentration measure HHI for the distribution of revenues among all four time slots. Our HHI measure would take value 1 if an establishment sold 100% of their goods and services during only one of the four time slots, and would take value 0.25 if it sold equally in all four time slots. Columns 1 and 2 of Table 9 show that technology adoption decreases concentration by 4.4%. When controlling for changes in demand demographics in columns 3 and 4, the magnitude decreases slightly to 3.1%. Note that these results are consistent with our findings in Table 8 above. Technology adoption both discovers business opportunities and improves logistical efficiency in adopting establishments.

[32] Appendix Table A7 runs the same specifications as Table 8 with the average value per transaction as dependent variable. We do not observe any changes in average transaction value in neither peak times nor off-peak times.

Before concluding this section, we want to note that it is empirically challenging to separate reshuffling resources across time slots during the week from the discovery of new business opportunities as these may come hand-in-hand. We attempt to disentangle these two channels with a different set of empirical evidence that aims to estimate whether establishments reshuffle resources across different time slots while holding constant their customer demographic portfolio.

In summary, we have investigated the role of demand-driven mechanisms and changes in the operational behavior of adopters. On the one hand, we find compelling evidence consistent with the existence of a demand-driven mechanism, that is, technology adopters are able to identify new business opportunities and tilt their customer portfolio in response to the monthly information received. On the other hand, we also find evidence that increases in sales due to technology adoption are also coming from improving processes and workload distribution.

5.2. The impact of adoption on non-adopters and aggregate revenues

Let us qualify our results so far. Our findings contribute to a well-established literature in economics studying how information and the extent of its availability influence market outcomes. Following recent developments of this literature, we investigate this question through the lens of IT adoption, and particularly, the adoption of an information-sharing program that provides access to the benefits of Big Data technology without bearing the financial costs of adoption. Our paper differs from the existing literature in that it focuses on the impact of access to market information, through technology, on SMEs. With the obvious exception of online SMEs (SMEs exclusively engaging in e-commerce), IT adoption by SMEs is lagging, and so the puzzle of why SMEs are not investing in information access through existing technologies remains. Our findings show that SMEs can and do incorporate information provided by Big Data and data-sharing in their decision making. In fact, we find that program adoption increases establishment revenues by 9% with large heterogeneity in the effects of adoption. Most importantly, our study of mechanisms shows that the increase in revenues is driven by both an improvement of strategies targeting more lucrative customer types, and a more efficient internal organization. Consequently, our findings suggest that SMEs can obtain high returns from access to market information. Thus, low adoption rates are likely due to high adoption costs or managerial inattention. Our evidence may justify government intervention through either direct provision or regulation allowing businesses to share information and resources.

In addition to the mechanisms described above, it is important to investigate whether the increase in revenues due to technology adoption comes from business stealing (potentially from non-adopters) or net value generated from better service.[33] Table 10 investigates the effect of adopters on non-adopters' revenues. We first define non-adopter competitors as the rest of the establishments in the same zip code-sector dyad. Column 1 estimates the impact of adoption on adopters and non-adopters including sector-quarter fixed effects (we cannot control for zip code-sector-quarter fixed effects given our definition of non-adopter). Column 1 also includes sector-zip code trend and trend squared. In Column 2, we define competitors as establishments in the same subsector-zip code so that we can introduce zip code-sector-quarter fixed effects but also subsector-zip code specific trend and trend squared. We note that adoption is associated with decreases in revenues of non-adopters and that the impact is stronger in closer competitors – those in your sector (1.5% decrease when subsector competitor adopts the technology). When we instrument for adoption, the impact on non-adopters remains qualitatively unchanged. Column 4 runs the same specification as of column 2 dropping all adopters and so comparing performance of non-adopters before and after adoption in their sector-quarter fixed effects and finding the same exact finding of a drop of revenue of 1.4% upon adoption of a competitor. Following that, columns 5 to 7 account for a different definition of adoption where it only has an impact on non-adopters the first time it occurs within a sector and zip code. Our OLS findings are robust to this definition change, while our result when implementing IV becomes statistically non-significant (although still negative and close to 1%). Finally, column 8 examines impact heterogeneity departing from the specification in column 2 and splitting the sample of competitors of adopters into those in high competition and low competition subsectors. In this specification, a subsector is high competition or low competition if it is above or below

[33] A potential concern regarding our main finding is that adopters are not really experiencing an increase in revenue, and are just merely increasing the share of POS transactions relative to cash transactions in their establishment. We use data from Esselink and Hernandez (2017) in their report of the European Central Bank detailing the use of cash versus credit and debit cards in different sectors, to separate sectors in our data that are more cash intensive from those more credit card intensive in their transactions. Appendix Table A8 shows the impact of adoption is the same across sectors with different intensity of cash transactions.

the median of number of competitors, respectively (same definition we employed in Table 3 earlier in the paper). We find no difference on the impact of adoption on a competitor's revenue depending on the degree of competition in its subsector and zip code.

These findings seem to suggest that some of the gains in revenues following adoption are coming from business stealing effects from competitors. Therefore, with the evidence in Table 10, we cannot reject that adoption has no effect in total welfare. To shed light on this matter, we aggregate data at the sector-zip code level and run specifications in first differences in Table 11. Our adoption dummy here takes value 1 if an establishment of a given sector and zip code adopts, and 0 otherwise. Therefore, our dummy in first differences in the right-hand side of the specification does not take into account the incidence of the second adopter, third adopter, and so on. The findings of our OLS regressions show a positive association between sector-zip code revenues and adoption. In particular, column 2 shows an increase of 1.6% in revenues at the sector-zip code level when introducing sector-quarter and sector-zip code fixed effects in the specification. As for our IV strategy, we use the number of other sectors with adopting establishments in the same zip code and the same quarter as IV. Column 3 shows the IV is positively correlated with adoption. Finally, column 4 shows no statistically significant causal effect of adoption on revenues at the sector and zip code levels. This result is consistent with those in Table 10 and shows that the positive impact of adoption on establishment revenues is likely to be caused by business stealing due to the development of a competitive advantage over other close-by competitors, and not due to an increase of overall sales in the adopters' respective local markets.[34]

Even though we cannot use the findings in Table 11 to conclude that this technology is welfare improving, we must also consider the fact that (by revealed preference) consumers switching from one establishment to another must be enjoying net increases in their utility. If so, switching behavior should be an indication of a welfare-improving technology.[35]

[34] Borusyak and Jaravel (2018) point out potential concerns in the identification of treatment effects in event-study settings estimated with individual and time fixed-effects. OLS does not recover a reasonable weighted average of the treatment effects as long-run effects are weighted negatively. In our framework, we estimate the baseline specification in first-differences, precluding the problem of assigning incorrect weights to long-run effects.

[35] Switching behavior would be associated with welfare decreases in extraordinary circumstances such as: (1) firms with lower marginal costs are losing market share (probably unusual in retail); or (2) there is firm exit combined with an increase in competition where small establishments are gaining more than mid and big–size establishments (rather implausible).

Additionally, our findings also suggest that adopters become more efficient, which translates into lower costs for the same level of revenue and surplus generated. If so, welfare gains from widespread technology adoption may come from efficiency gains and not so much from sales and consumer surplus.[36]

6. Conclusions

This paper investigates the role of information in competitive markets through the evaluation of the impact of an information-sharing program, through Big Data technology, on small and medium-sized enterprises. In our empirical context, small and medium-size establishments were invited by their credit card POS provider to register free of charge in a program that would deliver monthly reports of their performance and their competitors' performance, as well as demographic and geographic characteristics of their customers and those of the customers of their competitors. Using first-difference OLS regressions we find adoption is associated with a 4.5% increase in revenue from credit and debit card transaction, and our IV strategy shows a causal impact of adoption on revenues of 9% for those establishments whose adoption decision is most strongly affected by the instrument. We find heterogeneous effects of technology adoption such that smaller adopters and adopters in markets with more competitors benefit more from the information provided by the technology. We find no differences across levels of entrepreneurial sophistication and digital experience among adopters. We also investigate mechanisms through which new or better structured information delivered by the monthly reports may have triggered the observed increase in revenues. We find that adopting establishments increase their revenues from both targeting underserved market segments and reshuffling resources and effort to off-peak times that were underutilized prior to adoption.

Our findings have managerial and policy implications for the understanding of the impact of market information on the performance of SMEs as well as the adoption patterns and the economic impact of new technologies. Departing from the existence of PPDs coupled with

[36] Because of the low adoption rate of this technology, we cannot say much about the potential effects of scalability of the adoption of the technology at the market level. We have run similar specifications to those in Table 11 with the share of adopters as explanatory variable, instead of our adoption dummy, and find consistent results with those in Table 11. Higher adoption rates in each local market are associated with increases in market sales, but we find no statistically significant causal effect of increases in market adoption rates on sales at the sector and zip code levels.

increases in market power of large firms (De Loecker et al., 2020) and decreases in business dynamism (Akcigit and Ates, 2019), it is important to understand how the arrival of the new information technologies such as Big Data and data analytics can affect these trends. The adoption of first-generation IT was mainly concentrated among large firms contributing to increase the gap between large and small firms. However, these patterns of adoption could be expected as these technologies were mainly intended to improve internal coordination and these gains are lower in small firms. By contrast, second-generation IT not only focuses on offering firms opportunities for better internal organization, but also offers them information about their competitive environment (consumers' preferences and their competitors' actions). Thus, there is a large scope for small firms to benefit from this new generation of information technologies. However, if high adoption costs prevent the adoption of these new information technologies and bar access to market information by small firms, it is likely the case that the disparities between large and small firms will grow even larger. As a result, private adoption decisions and investments in access to market information may be socially inefficient, thus opening the door for government intervention or data sharing initiatives to mitigate adoption costs.

Our findings show that small enterprises strategically use the information provided by the information-sharing program to their advantage yielding a sizable average return of adoption and heterogeneity across establishments of different sizes. A cautious interpretation of the evidence presented in this paper calls for an estimate of the lower bound of the cost of adoption of information technology that would facilitate access and analysis of market information to managers of SMEs. In fact, managers and owners of small and medium-sized enterprises may perceive the cost of adoption to be at least 9% of their revenues if we want to rationalize their behavior prior to the free supply of information provided by the technology evaluated in this study. Consequently, future research should investigate the nature and behavior of the costs of adoption under different competitive environments. On the one hand, it is important to understand whether the scarce adoption patterns observed in small and medium-sized enterprises can be addressed with mere awareness campaigns, the provision of other technologies that exhibit complementarities with market information data, or the provision of adequate skilled human capital to operate such technology and process its information to be used as valid input in decision-making. On the other hand, future research should also aim to *(i)* enhance our understanding of the potential market-level effects of scalability of the adoption of these types of information-sharing programs and information technologies, and *(ii)* provide further evidence of the specific actions taken by firms as a result of an increase in market information such as changes in pricing strategies, introduction of new products, or marketing and advertising campaigns.

References

Akcigit, Ufuk, and Sina T. Ates (2019). "What Happened to U.S. Business Dynamism?" NBER Working Paper #25756.

Akerlof, George (1970). "The Market for 'Lemons': Quality Uncertainty and the Market Mechanism," *Quarterly Journal of Economics*, 84: 488-500.

Anderson, Simon P., and André De Palma (2012). "Competition for Attention in the Information (Overload) Age." *The RAND Journal of Economics* 43 (1): 1–25.

Angle, Patricia, and Chris Forman (2018). "Does IT Level the Playing Field for Small Establishments? Evidence from Manufacturing." SSRN Electronic Journal.

Agrawal, Ajay, Joshua Gans, and Avi Goldfarb (2018). Prediction Machines: The Simple Economics of Artificial Intelligence. Harvard Business School Press.

Bajari, Patrick, Victor Chernozhukov, Ali Hortaçsu and Junichi Suzuki (2019). "The Impact of Big Data on Firm Performance: An Empirical Investigation," *AEA Papers and Proceedings*, vol 109, pp. 33-37.

Bartel, Ann, Casey Ichniowski, and Kathryn Shaw (2007). "How Does Information Technology Affect Productivity? Plant-Level Comparisons of Product Innovation, Process Improvement, and Worker Skills." *Quarterly Journal of Economics* 122 (4): 1721–58.

Baum, J. A., & Lant, T. K. (2003). Hits and misses: Managers'(mis) categorization of competitors in the Manhattan hotel industry. In Geography and Strategy (pp. 119-156). Emerald Group Publishing Limited.

Bergemann, Dirk, and Alessandro Bonatti (2019). "Markets for information: An introduction." *Annual Review of Economics* 11: 85-107.

Bergemann, Dirk, Alessandro Bonatti, and Tan Gan (2021). "The Economics of Social Data." Cowles Foundation Discussion Paper No. 2203R3.

Berman, Ron, and Ayelet Israeli (2020). "The Value of Descriptive Analytics: Evidence from Online Retailers." Harvard Business School Marketing Unit Working Paper 21-067.

Bertoletti, Paolo, and Clara Poletti (1997). "X-Inefficiency, Competition and Market Information." *The Journal of Industrial Economics*, 45(4): 359–375.

Bloom, Nicholas, Benn Eifert, Aprajit Mahajan, David McKenzie, and John Roberts (2013). "Does Management Matter? Evidence from India," *Quarterly Journal of Economics* 128(1): 1-51.

Bloom, Nicholas, Luis Garicano, Raffaella Sadun, and John Van Reenen. (2014). "The distinct effects of Information Technology and Communication Technology on firm organization," *Management Science* 60 (12): 2859-2885.

Bloom, Nicholas, Raffaella Sadun, and John Van Reenen. (2012). "Americans do I.T. Better: US Multinationals and the Productivity Miracle," *American Economic Review* 102 (1): 167-201.

Bloom, Nicholas, and John Van Reenen (2007). "Measuring and Explaining Management Practices across Firms and Countries," *Quarterly Journal of Economics* 122(4): 1351-1408.

Borusyak, Kirill, and Xavier Jaravel (2018). "Revisiting Event Study Designs." Working Paper LSE. Available at SSRN: https://ssrn.com/abstract=2826228.

Bresnahan, Timothy F., Erik Brynjolfsson, and Lorin M. Hitt (2002). "Information Technology, Workplace Organization, and the Demand for Skilled Labor: Firm-Level Evidence." *Quarterly Journal of Economics* 117 (1): 339–76.

Bruhn, Miriam, Dean Karlan, and Antoinette Schoar (2018). "The Impact of Consulting Services on Small and Medium Enterprises: Evidence from a Randomized Trial in Mexico," *Journal of Political Economy* 126 (2) :635-687.

Brynjolfsson, Erik, and Lorin M. Hitt (2000). "Beyond Computation: Information Technology, Organizational Transformation and Business Performance." *Journal of Economic Perspectives* 14 (4): 23–48.

Brynjolfsson, Erik, Lorin M. Hitt, and Heekyung Hellen Kim. (2011). "Strength in Numbers: How Does Data-Driven Decision Making Affect Firm Performance?" Unpublished.

Brynjolfsson, Erik, and Kristina McElheran (2016a). "The Rapid Rise of Data-Driven Decision Making," *American Economic Review Papers and Proceedings* 106 (5): 133-139.

Brynjolfsson, Erik, and Kristina McElheran (2016b). "Data in Action: Data-Driven Decision Making in U.S. Manufacturing." Center for Economic Studies Working Paper 16–06.

Brown, Jeffrey R., and Austan Goolsbee (2002). "Does the Internet Make Markets More Competitive? Evidence from the Life Insurance Industry." *Journal of Political Economy* 110 (3): 481-507.

Carlin, Bruce I., Shaun W. Davies, and Andrew Iannaccone (2012). "Competition, Comparative Performance, and Market Transparency." *American Economic Journal: Microeconomics*. 4 (4): 202–237.

Choi, S. Chan, Wayne S. Desarbo, and Patrick T. Harker (1990). "Product Positioning under Price Competition." *Management Science* 36 (2): 175–199.

Cohen, Wesley, and Daniel Levinthal (1990). "Absorptive capacity: A new perspective on learning and innovation," *Administrative Science Quarterly* 35 (1): 128-152.

Cohen, Micahel D., James G. March, and Johan P. Olsen (1972). "A garbage can model of organizational choice," *Administrative Science Quarterly* 17 (1): 1–25.

Cyert, Richard, and James G. March. (1963). "A Behavioral Theory of the Firm." 2nd ed. Prentice Hall, Englewood Cliffs, NJ.

De Loecker, Jan, Jan Eeckhout and Gabriel Unger (2020). "The Rise of Market Power and the Macroeconomic Implications," *The Quarterly Journal of Economics* 135 (2): 561–644.

Dessein, Wouter, Andrea Galeotti, and Tano Santos (2016). "Rational Inattention and Organizational Focus," *American Economic Review* 106:1522-1536.

Einav, Liran, Peter Klenow, Benjamin Klopack, Jonathan Levin, Larry Levin and Wayne Best (2017). "Assessing the Gains from E-Commerce," mimeograph.

Einav, Liran, and Jonathan Levin (2014). "Economics in the age of big data," *Science* 346 (6210): 715-721.

Esselink, Henk, and Lola Hernandez (2017). "The use of cash by households in the euro area," European Central Bank Occasional Paper Series No. 201.

Fabregas, Raissa, Michael Kremer, Jonathan Robinson, Frank Schilbach (2019). "The Value of Local Agricultural Information: Evidence from Kenya," mimeograph.

Fama, Eugene (1970). "Efficient Capital Markets: A Review of Theory and Empirical Work," *Journal of Finance* 25: 383-417.

Farboodi, Maryam, Roxana Mihet, Thomas Philippon, and Laura Veldkamp (2019). "Big Data and Firm Dynamics," *American Economic Review Papers and Proceedings* 109: 38-42.

Forman, Chris, Avi Goldfarb, Shane Greenstein (2012). "The internet and local wages: A puzzle," *American Economic Review* 102 (1): 556-575.

Foster, Lucia, John Haltiwanger, and Chad Syverson (2008). "Reallocation, Firm Turnover, and Efficiency: Selection on Productivity or Profitability?" American Economic Review, 98 (1): 394-425.

Foster, Lucia, John Haltiwanger, and Chad Syverson (2016). "The Slow Growth of New Plants: Learning about Demand?" Economica 83: 91-129.

Galdon-Sanchez, Jose E., and James A. Schmitz (2002). "Competitive Pressure and Labor Productivity: World Iron-Ore Markets in the 1980's," *The American Economic Review* 92 (4): 1222–1235.

Gibbons, Robert and Rebecca Henderson (2013). "What Do Managers Do? Exploring Persistent Performance Differences among Seemingly Similar Enterprises." Chapter 17 in *The Handbook of Organizational Economics*, edited by R. Gibbons and J. Roberts, 680–731. Princeton University Press.

Giorcelli, Michela (2019). "The Long-Term Effects of Management and Technology Transfers," *American Economic Review* 109(1): 121-55.

Goldfarb, Avi, and Catherine Tucker (2019). "Digital Economics." *Journal of Economic Literature* 57(1): 3-43.

Goyal, Aparajita (2010). "Information, Direct Access to Farmers, and Rural Market Performance in Central India." *American Economic Journal: Applied Economics* 2 (3): 22-45.

Grossman, Sanford, and Joseph Stiglitz (1980). "On the Impossibility of Informationally Efficient Markets," *American Economic Review* 70: 393-408.

Hayek, Friedrich (1945). “The Use of Knowledge in Society,” *American Economic Review* 35: 519-30.

Hauswald, Robert, and Robert Marquez (2003). “Information Technology and Financial Services Competition.” *The Review of Financial Studies* 16 (3): 921–948.

Hsieh, Chang-Tai, and Peter Klenow (2009). “Misallocation and Manufacturing TFP in China and India,” *The Quarterly Journal of Economics* 124 (4): 1403–1448.

Hortaçsu, Ali, and Chad Syverson (2004). “Product Differentiation, Search Costs, and Competition in the Mutual Fund Industry: A Case Study of S&P 500 Index Funds.” *The Quarterly Journal of Economics* 119 (2): 403–456.

Jensen, Robert (2007). “The Digital Provide: Information (Technology), Market Performance, and Welfare in the South Indian Fisheries Sector.” *The Quarterly Journal of Economics* 122 (3): 879–924.

Jin, Ginger, and Phillip Leslie (2003). “The Effect of Information on Product Quality: Evidence from Restaurant Hygiene Grade Cards,” *The Quarterly Journal of Economics* 118 (2): 409–451.

Jin, Yizhou, and Zhengyun Sun (2019). “Information Acquisition and the Return to Data.” Harvard University Working Paper.

Jin, Yizhou, and Zhengyun Sun (2020). “Lifting Growth Barriers for New Firms.” Harvard University Working Paper.

Jones, Charles I., and Christopher Tonetti (2020). "Nonrivalry and the Economics of Data." *American Economic Review*, 110 (9): 2819-58.

Kim, Hyunjin (2019). “The Value of Competitor Information: Evidence from a Field Experiment.” Working Paper Harvard Business School.

Levitt, Steven D. (2006). “An Economist Sells Bagels: A Case Study in Profit Maximization.” NBER Working Paper No. 12152.

Li, J., Netessine, S., & Koulayev, S. (2017). Price to compete… with many: how to identify price competition in high-dimensional space. Management Science, 64(9), 4118-4136.

Liberti, Jose, Amit Seru and Vikrant Vig (2019). “Information, Credit and Organization,” Stanford GSB Working Paper.

Lucas, Robert Jr. (1972). “Expectations and the Neutrality of Money,” *Journal of Economic Theory* 4: 103-24.

Matějka, Filip (2016). “Rationally Inattentive Seller: Sales and Discrete Pricing.” *Review of Economic Studies* 83 (3): 1156-1188.

McAfee, Andrew, and Erik Brynjolfsson (2012). “Big Data: The Management Revolution.” Harvard Business Review 26 September.

Milgrom, Paul, and John Roberts (1982). “Limit Pricing and Entry under Incomplete Information: An Equilibrium Analysis.” *Econometrica* 50 (2): 443-459.

Mohnen, Pierre, Michael Polder, and George van Leeuwen (2018). "ICT, R&D and Organizational Innovation: Exploring Complementarities in Investment and Production," NBER Working Paper 25044. http://www.nber.org/papers/w25044.

Muller, Oliver, Maria Fay, and Jan Vom Brocke (2018). "The effect of big data and analytics on firm performance: An econometric analysis considering industry characteristics." Journal of Management Information Systems 35 (2): 488-509.

Nagaraj, Abhishek (2020). "The Private Impact of Public Data: Landsat Satellite Maps and Gold Exploration," Working Paper.

Pettengill, John S. (1979). "Monopolistic Competition and Optimum Product Diversity: Comment." *The American Economic Review* 69 (5): 957–960.

Rothschild, Michael, and Joseph Stiglitz (1976). "Equilibrium in Competitive Insurance Markets: An Essay on the Economics of Imperfect Information," *Quarterly Journal of Economics* 90: 629-649.

Sadun, Raffaella and John Van Reenen (2005). "Information technology and productivity: it ain't what you do it's the way that you do IT." Centre for Economic Performance, London School of Economics and Political Science, London, UK.

Spence, Michael (1973). "Job Market Signaling," *Quarterly Journal of Economics*, 87 (3): 355–374.

Stigler, George (1962). "Information in the Labor Market," *Journal of Political Economy* 70(5): 94-105.

Stigler, George (1964). "A Theory of Oligopoly," *Journal of Political Economy* 72 (1): 44-61.

Syverson, Chad (2004). "Market Structure and Productivity: A Concrete Example." *Journal of Political Economy* 112 (6): 1181-1222.

Syverson, Chad (2011). "What Determines Productivity?" *Journal of Economic Literature* 49 (2): 326–65.

Tirole, Jean (1988). "The Theory of Industrial Organization," MIT PRESS. Cambridge, MA.

Uriz-Uharte, Guillermo (2019). "The Value of Firm Communication for Innovation: Evidence from the Adoption of Communication Technologies," UCL Working Paper.

Figure 1: Number of adopters over time

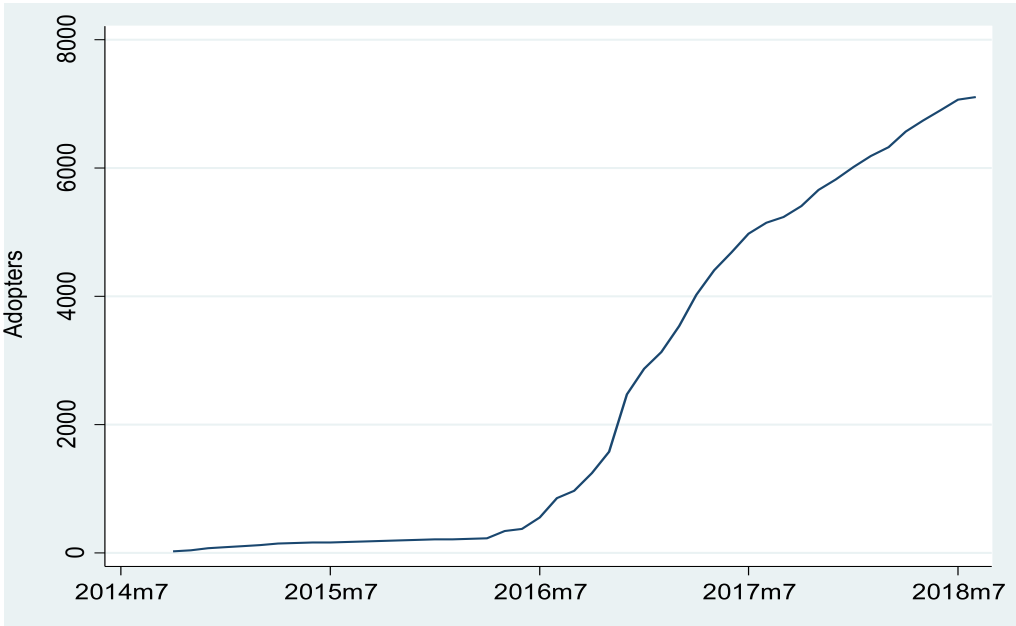

Figure 2: Timeline of Adoption

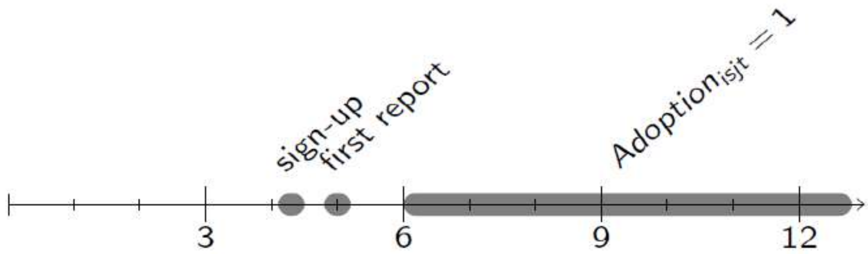

Figure 3: Instrumental variable identification

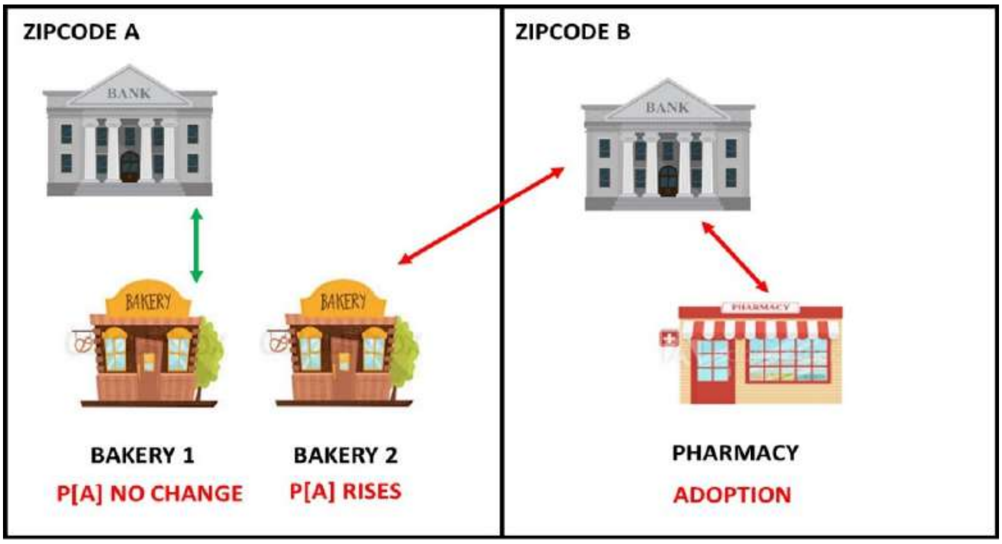

Figure 4: Treatment estimates across sectors, subsectors and regions

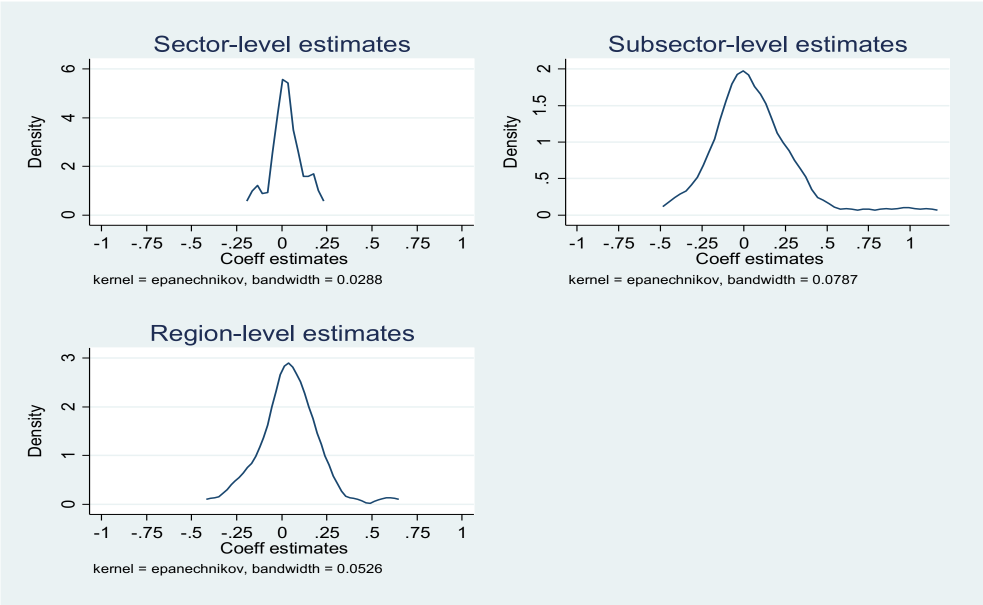

Table 1: Descriptive Statistics

	Observ.	Mean	Std. Dev.	Min	Max
<u>**Full Sample**</u>					
Revenue	4610085	4715	29171	12	7948335
Transactions	4610085	120	710	5	227139
Average Value of Transactions	4610085	64	101	2.4	15000
Customers	4610085	74	338	2	134725
Average Value per Customer	4610085	85	198	1	92066
<u>**Adopters**</u>					
Revenue	63639	6248	18730	15	537791
Transactions	63639	153	462	5	8146
Average Value of Transactions	63639	80	147	3	7006
Customers	63639	92	224	3	5975
Average Value per Customer	63639	102	200	1	10500
Number of competitors	63639	75	96	0	1020
Sophistication	3495	3.53	0.89	1.00	5.00

Notes: Statistics computed from a sample with quarterly level information at the establishment level.

Table 2: Baseline Results

Dependent variable: Δ Log revenue

	OLS (1)	OLS (2)	OLS (3)	1st-stg (4)	2nd-stg (5)	OLS (6)	1st-stg (7)	2nd-stg (8)
Falsification [{Δ Adoption =1}t-1]			0.00978					
			(0.0158)					
Δ Adoption	0.0455***	0.0458***			0.0902**	0.0445***		0.0840**
	(0.0157)	(0.0157)			(0.0386)	(0.0159)		(0.0422)
Lead {Δ Adoption = 1} t+1		0.00263						
		(0.0148)						
Lead {Δ Adoption = 1} t+2		0.00395						
		(0.0161)						
Lead {Δ Adoption = 1} t+3		0.025						
		(0.0164)						
Peers IV				0.00446***			0.00451***	
				(0.00012)			(0.000121)	
Sector-zipcd-quarter FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Dummies first 4 quarters	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Bank-branch time trend						Yes	Yes	Yes
Observations	4,610,085	4,610,085	4,610,085	4,610,085	4,610,085	4,610,085	4,610,085	4,610,085

Notes: *** denotes significance at 1%, ** at 5%, and * at 10%. . Standard errors are clustered at the establishment level and reported in parenthesis.

Table 3: Heterogeneous Effects

Dependent variable: Δ Log revenue

	Sophistication		Size		Competition	
	(1)	(2)	(3)	(4)	(5)	(6)
	OLS	IV	OLS	IV	OLS	IV
Δ Adoption x High	0.0442*	0.0873**	0.0139	0.0725*	0.068***	0.109***
	(0.0232)	(0.0391)	(0.0171)	(0.0378)	(0.0209)	(0.0403)
Δ Adoption x Low	0.0463**	0.0976**	0.0796***	0.146***	0.0206	0.0629
	(0.021)	(0.0458)	(0.0267)	(0.0481)	(0.0216)	(0.0419)
Residual CF		-0.0541		-0.0702		-0.0471
		(0.0454)		(0.0443)		(0.0437)
Sector-zipcd-quarter FE	Yes	Yes	Yes	Yes	Yes	Yes
Dummies first 4 quarters	Yes	Yes	Yes	Yes	Yes	Yes
p-value null equal returns	0.946	0.752	0.0374	0.0226	0.0986	0.106
Observations	4610085	4610085	4610085	4610085	4610085	4610085

Notes: *** denotes significance at 1%, ** at 5%, and * at 10%. Standard errors are clustered at the establishment level and reported in parenthesis.

Table 4: Effects on Other Outcomes

Dep variable: Δ Log number of customers, Δ log revenue per customers, Δ log number of transactions per customer

	Customers		Rev/Cust		Trans/Cust	
	OLS	IV	OLS	IV	OLS	IV
	(1)	(2)	(3)	(4)	(5)	(6)
Δ Adoption	0.0385***	0.119***	0.00701	-0.0293	0.00514	0.0101
	(0.0113)	(0.0301)	(0.00961)	(0.0199)	(0.00325)	(0.00711)
Sector-zipcd-quarter FE	Yes	Yes	Yes	Yes	Yes	Yes
Dummies first 4 quarters	Yes	Yes	Yes	Yes	Yes	Yes
Observations	4610085	4610085	4610085	4610085	4610085	4610085

Notes: *** denotes significance at 1%, ** at 5%, and * at 10%. Standard errors are clustered at the establishment level and reported in parenthesis.

Table 5: Effects on Other Outcomes II

Dependent variable: Δ Log number transaction and Δ log revenue per transaction

	Transactions		Rev/Trans	
	OLS	IV	OLS	IV
	(1)	(2)	(3)	(4)
Δ Adoption	0.0436***	0.130***	0.00187	-0.0394**
	(0.0120)	(0.0316)	(0.00906)	(0.0187)
Sector-zipcd-quarter FE	Yes	Yes	Yes	Yes
Dummies first 4 quarters	Yes	Yes	Yes	Yes
Observations	4610085	4610085	4610085	4610085

Notes: *** denotes significance at 1%, ** at 5%, and * at 10%. Standard errors are clustered at the establishment level and reported in parenthesis.

Table 6: Changes in Composition of Customers

Dependent variable: Δ Share in Prime Customer and Δ Log HHI of Cutomer Types

	Share Prime Customer				Concentration Customer Types			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	OLS	IV	OLS	IV	OLS	IV	OLS	IV
Δ Adoption	0.00168	-0.00578			-0.0249***	-0.0344*		
	(0.00315)	(0.0074)			(0.00715)	(0.0178)		
Δ Adoption x High			-0.0197***	-0.0258***			-0.0576***	-0.0868***
			(0.00474)	(0.0081)			(0.0132)	(0.0232)
Δ Adoption x Low			0.0236***	0.0174**			0.00477	-0.0207
			(0.00401)	(0.00796)			(0.00637)	(0.0174)
Residual CF				0.0069				0.0307
				(0.00836)				(0.0204)
Sector-zipcd-quarter FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Dummies first 4 quarters	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
p-value null equal returns			0.00	0.00			0.00	0.00
Mean dependent variable in levels	0.187	0.187	0.187	0.187	0.288	0.288	0.288	0.288
Observations	4610085	4610085	4610085	4610085	4610085	4610085	4610085	4610085

Notes: *** denotes significance at 1%, ** at 5%, and * at 10%. Standard errors are clustered at the establishment level and reported in parenthesis.

Table 7: Attracting Customers from Other Areas

Dependent variable: Δ Share of revenue from customers from other zipcodes

	OLS (1)	IV (2)	OLS (3)	IV (4)
Δ Adoption	0.00570	0.00929		
	(0.00347)	(0.00583)		
Δ Adoption x Large Share			-0.00467	0.00114
			(0.00334)	(0.00624)
Δ Adoption x Small Share			0.0154***	0.0216***
			(0.00592)	(0.00734)
Residual CF				-0.00676
				(0.00694)
Sector-zipcd-quarter FE	Yes	Yes	Yes	Yes
Dummies first 4 quarters	Yes	Yes	Yes	Yes
p-value null equal returns			0.003	0.002
Mean dependent variable	0.695	0.695	0.695	0.695
Observations	4610085	4610085	4610085	4610085

Notes: *** denotes significance at 1%, ** at 5%, and * at 10%. Standard errors are clustered at the establishment level and reported in parenthesis.

Table 8: Distribution of revenues in peak and off-peak time

Dependent variable: Δ Log revenue in peak and off-peak time of the week

	Peak time		Off-peak time			
	OLS	IV	OLS	IV	OLS	IV
	(1)	(2)	(3)	(4)	(5)	(6)
Δ Adoption	0.0207	0.032	0.0815***	0.170***	0.0382**	0.0815**
	(0.0284)	(0.0651)	(0.0212)	(0.0543)	(0.0156)	(0.0387)
Sector-zipcd-quarter FE	Yes	Yes	Yes	Yes	Yes	Yes
Dummies first 4 quarters	Yes	Yes	Yes	Yes	Yes	Yes
Demand Controls					Yes	Yes
Observations	4610085	4610085	4610085	4610085	4610085	4610085

Notes: *** denotes significance at 1%, ** at 5%, and * at 10%. Standard errors are clustered at the establishment level and reported in parenthesis.

Table 9: Concentration of Revenues over the week

Dependent variable: Δ Log HHI of revenues over the week

	OLS (1)	IV (2)	OLS (3)	IV (4)
Δ Adoption	-0.0152***	-0.0440***	-0.00915*	-0.0312***
	(0.00555)	(0.0127)	(0.00526)	(0.0119)
Sector-zipcd-quarter FE	Yes	Yes	Yes	Yes
Dummies first 4 quarters	Yes	Yes	Yes	Yes
Demand Controls			Yes	Yes
Mean dependent variable in levels	0.419	0.419	0.419	0.419
Observations	4,610,085	4,610,085	4,610,085	4,610,085

Notes: *** denotes significance at 1%, ** at 5%, and * at 10%. Standard errors are clustered at the establishment level and reported in parenthesis.

Table 10: Impact on close competitors

Dependent variable: Δ Log Revenue

	OLS (1)	OLS (2)	IV (3)	OLS (4)	OLS (5)	IV (6)	OLS (7)	OLS (8)
Δ Adoption	0.0430***	0.0397**	0.095466**		0.0418***	0.0966**		0.0397**
	(0.0158)	(0.0160)	(0.0398)		(0.0160)	(0.0398)		(0.0160)
Δ Adoption by competitor	-0.00423*	-0.0146***	-0.0117**	-0.0135***	-0.0114**	-0.0085	-0.0108**	
	(0.002312)	(0.00497)	(0.00536)	(0.00502)	(0.00532)	(0.00571)	(0.00538)	
Δ Adoption by competitor in high competition subsector								-0.0145**
								(0.00597)
Δ Adoption by competitor in low competition subsector								-0.0147*
								(0.00852)
Sector-quarter FE	Yes							
Sector-zipcd-quarter FE		Yes	Yes	Yes	Yes	Yes	Yes	Yes
Sector-zipcd Trends	Yes							
Subsector-zipcd Trends		Yes	Yes	Yes	Yes	Yes	Yes	Yes
Dummies first 4 quarters	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Drop out adopters				Yes			Yes	
Effect only of first adopter					Yes	Yes	Yes	
Observations	4610085	4610085	4610085	4546446	4610085	4610085	4546446	4610085

Notes: *** denotes significance at 1%, ** at 5%, and * at 10%. Standard errors are clustered at the establishment level and reported in parenthesis.

Table 11: Aggregate effect of adoption

Dependent variable: Δ Log Revenue

	OLS	OLS	1st-stage	2nd-stage
	(1)	(2)	(3)	(4)
Δ Adoption	0.0293***	0.0160**		-0.0160
	(0.00823)	(0.00715)		(0.186)
IV			0.0299***	
			(0.00389)	
Sector-quarter FE	Yes	Yes	Yes	Yes
Sector-zipcd FE		Yes		
Observations	75,330	75,330	75,330	75,330

Notes: *** denotes significance at 1%, ** at 5%, and * at 10%. Standard errors are clustered at the sector-zipcode level and reported in parenthesis.

Figure A1: Sample report of monthly information on the adopting establishment

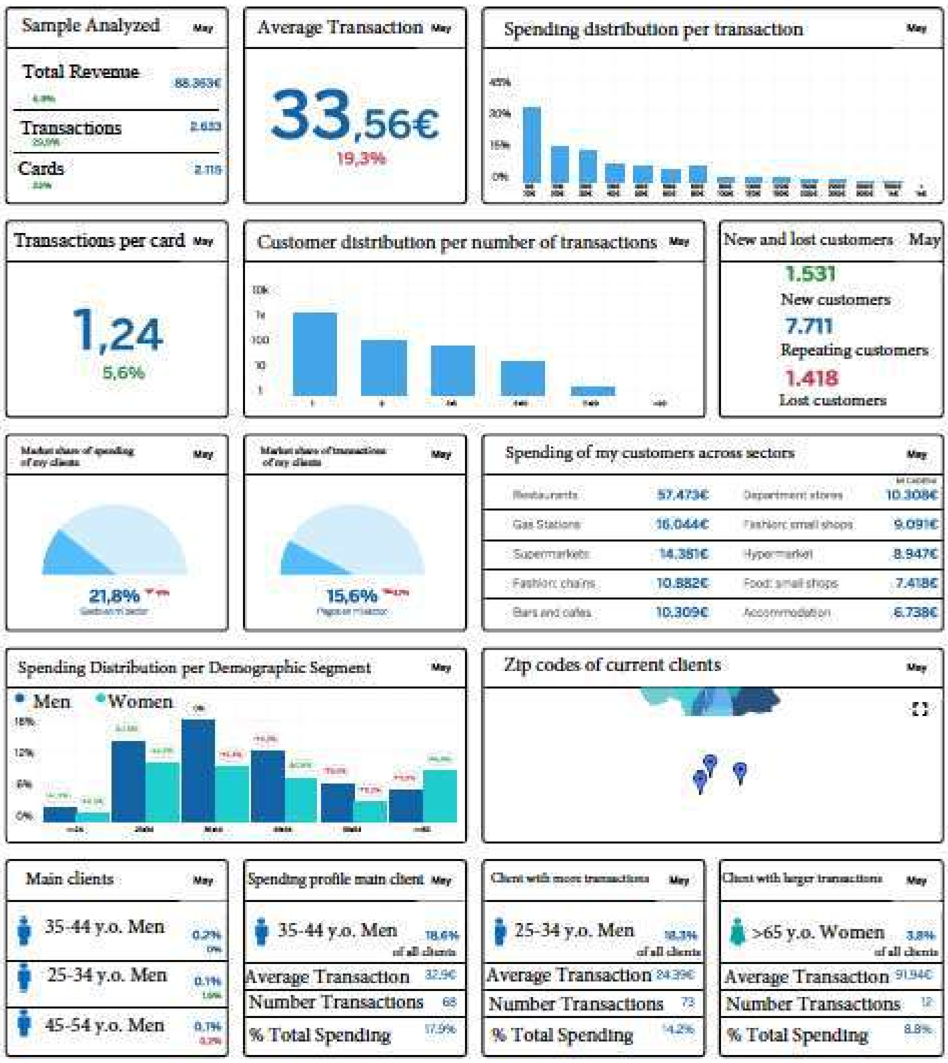

Figure A2: Sample report of monthly information on the competition of the adopting establishment

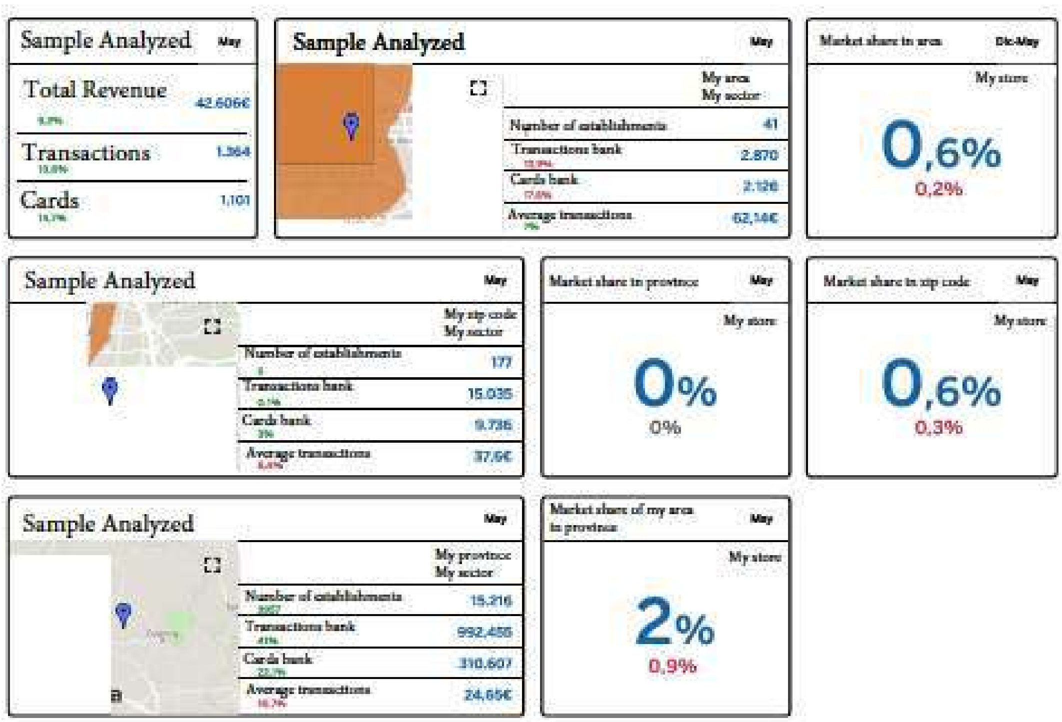

Table A1. Sectors and subsectors

Sector	Subsector
Auto	(1) Carwash; (2) Dealers, garage and spare parts; (3) Vehicle roadworthiness
Bank	(1) Branches; (2) ATM
Bars and restaurants	(1) Fast food restaurants; (2) Restaurants; (3) Pubs and night clubs; (4) Bars and cafes
Books and press	(1) Press; (2) Books
Fashion	(1) Jewelry and watches; (2) Fashion: small shops; (3) Shoes; (4) Fashion: chains; (5) Leather goods
Food	(1) Supermarkets; (2) Food: small shops
Health	(1) Pharmacy; (2) Hospital and medical consultory; (3) Optician's shops
Home	(1) DIY: supermarkets; (2) Gardening and Floristry: chains; (3) Furniture and decoration: small shops; (4) Furniture and decoration: chains; (5) Gardening and Floristry: small shops; (6) DIY: small shops
Accomodation	Accommodation
Hypermarkets	(1) Hypermarket; (2) Department Stores
Leisure	(1) Tickets; (2) Museums and tourist attractions; (3) Shows; (4) Bet
Other services	(1) Public tax; (2) Bazaar; (3) Dry Cleaning & Laundry; (4) Cybercafe; (5) Teaching; (6) Packaging and Storage; (7) Unknown; (8) Veterinary and pets; (9) Funeral; (10) Duty Free; (11) Tobacconist; (12) Donations; (13) Phone Plenish; (14) Insurance; (15) Video Store and pay TV; (16) Other
Real state	Real estate
Sports and toys	(1) Toys and sport articles; (2) Toys: chains; (3) Sportive activities; (4) Sport: chains
Technology	(1) Photography; (2) Musical instruments; (3) Computers and appliances: chains; (4) Computers and appliances: small shops; (5) Telephony: handset sales
Transport	(1) Car rental; (2) Urban transport: bus, metro and train; (3) Airlines; (4) Parkings; (5) Toll; (6) Trains; (7) Taxi; (8) Sea transportation; (9) Boat and aircraft renting; (10) Gas stations; (11) Bus
Travel	(1) Travel agency: local; (2) Travel agency: web
Wellness and beauty	(1) Cosmetics and perfumery: small shops; (2) Cosmetics and perfumery: chains; (3) Wellness and Beauty; (4) Hair and beauty

Table A2: Robustness Results

Dependent variable: Δ Log revenue

	OLS	1st-stg	2nd-stg	OLS	1st-stg	2nd-stg	1st-stg	2nd-stg
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Δ Adoption	0.0380***		0.106**	0.0577***		0.114**		0.0948**
	(0.0163)		(0.0421)	(0.0165)		(0.0447)		(0.0387)
Peers IV		0.00450***			0.00432***			
		(0.00012)			(0.00012)			
Peers IV (no same sector)							0.00448***	
							(0.000119)	
Sector-zipcd-quarter FE	Yes	Yes	Yes				Yes	Yes
Establishment time trend	Yes	Yes	Yes					
Subsector-zipcd-quarter FE				Yes	Yes	Yes		
Dummies first 4 quarters	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	4,610,085	4,610,085	4,610,085	4,610,085	4,610,085	4,610,085	4,610,085	4,610,085

Notes: *** denotes significance at 1%, ** at 5%, and * at 10%. Standard errors are clustered at the establishment level and reported in parenthesis.

Table A3: Event Study

Dependent variable: Δ Log revenue

	OLS (1)	OLS (2)	OLS (3)	OLS (4)	IV (5)	OLS (6)	IV (7)
Δ Adoption	0.0235	0.0250	0.0538**	0.0526**	0.0282	0.0843*	0.0943
	(0.0178)	(0.0181)	(0.0256)	(0.0260)	(0.0500)	(0.0478)	(0.0774)
Quarter FE	Yes						
Sector-quarter FE		Yes		Yes	Yes		
Zipcd-quarter FE			Yes	Yes	Yes		
Sector-zipcd-quarter FE						Yes	Yes
Dummies first 4 quarters	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	63639	63639	63639	63639	63639	63639	63639

Notes: *** denotes significance at 1%, ** at 5%, and * at 10%. Standard errors are clustered at the establishment level and reported in parenthesis. Columns 1-5 present OLS estimates from a regression of log revenue on adoption in which the sample is limited only to adopting establishments. Column 6 instruments the adoption variable.

Table A4: Heterogeneous Effects by Analytical, Marketing and Digital Sophistication

Dependent variable: Δ Log revenue

	Analytical		Marketing		Digital	
	(1) OLS	(2) IV	(3) OLS	(4) IV	(5) OLS	(6) IV
Δ Adoption x High Sophistication	0.00486	0.0682*	0.0337	0.0822**	0.0789***	0.128***
	(0.0233)	(0.0390)	(0.0237)	(0.0391)	(0.0252)	(0.0434)
Δ Adoption x Low Sophistication	0.0728***	0.148***	0.0533**	0.111**	0.0217	0.0693*
	(0.0209)	(0.0459)	(0.0207)	(0.0459)	(0.0199)	(0.0402)
Residual CF		-0.0795*		-0.0610		-0.0543
		(0.0454)		(0.0456)		(0.0436)
Sector-zipcd-quarter FE	Yes	Yes	Yes	Yes	Yes	Yes
Dummies first 4 quarters	Yes	Yes	Yes	Yes	Yes	Yes
p-value null equal returns	0.03	0.014	0.534	0.378	0.075	0.068
Observations	4,610,085	4,610,085	4,610,085	4,610,085	4,610,085	4,610,085

Notes: *** denotes significance at 1%, ** at 5%, and * at 10%. Standard errors are clustered at the establishment level and reported in parenthesis.

Table A5: Heterogeneous Effects by Early vs Late Adopters

Dependent variable: Δ Log revenue

	Total		Pilot		Each Market		Sign up Time	
	(1) OLS	(2) IV	(3) OLS	(4) IV	(5) OLS	(6) IV	(7) OLS	(8) IV
Δ Adoption x Early	0.0503**	0.0956**	-0.0572	-0.0107	0.0635***	0.108***	0.0540**	0.0974**
	(0.0220)	(0.0420)	(0.108)	(0.116)	(0.0171)	(0.0392)	(0.0264)	(0.0424)
Δ Adoption x Late	0.0406*	0.0853**	0.0490***	0.0905**	-0.0263	0.0186	0.0423**	0.0869**
	(0.0223)	(0.0410)	(0.0158)	(0.0384)	(0.0378)	(0.0508)	(0.0191)	(0.0408)
Residual CF		-0.0507		-0.0469		-0.0506		-0.0499
		(0.0437)		(0.0438)		(0.0436)		(0.0438)
Sector-zipcd-quarter FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Dummies first 4 quarters	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
p-value null equal returns	0.758	0.743	0.328	0.353	0.031	0.031	0.72	0.75
Observations	4610085	4610085	4610085	4610085	4610085	4610085	4610085	4610085

Notes: *** denotes significance at 1%, ** at 5%, and * at 10%. Standard errors are clustered at the establishment level and reported in parenthesis. Columns 1 and 2 divide adopters between early adopters (those that are below the median adoption of adoption time - in other words, the 50% of adopters that adopted first) and late adopters (those above the median). Columns 3 and 4 divide adopters between those in the pilot program (early adopters) and those in that adopted once the problem became national (late adopters). There are a total of 221 adopters during the pilot period. Columns 5 and 6 divide adopters between early adopters (the first adopter in each sector-zip code) and late adopters (second or later adopters in a sector-zip code). There are a total of 1,575 late adopters according to this definition. Finally, in Columns 7 and 8, we divide adopters between eaerly and late according to the date in which they sign up to the program. We consider an establishment as an adopter in the quarter subsequent to adoption. In this exercise we exploit variation between those adopting early in the previous quarter or those adopting late in the previous quarter.

Table A6: Changes in Composition of Customers

Dependent variable: Δ Log revenue from sales to prime customer (cols. 1-4), and Δ Log revenue from sales to non-prime customer (cols. 5-6)

	(1) OLS	(2) IV	(3) OLS	(4) IV	(5) OLS	(6) IV
Δ Adoption	0.0514	0.132*				
	(0.0335)	(0.0754)				
Δ Adoption x High share			-0.112***	-0.0216	0.0892***	0.159***
			(0.0403)	(0.0798)	(0.0251)	(0.0515)
Δ Adoption x Low share			0.219***	0.310***	0.00318	0.0745
			(0.0534)	(0.0855)	(0.0254)	(0.0527)
Residual CF				-0.103		-0.0797
				(0.0866)		(0.0553)
Sector-zipcd-quarter FE	Yes	Yes	Yes	Yes	Yes	Yes
Dummies first 4 quarters	Yes	Yes	Yes	Yes	Yes	Yes
p-value null equal returns			0.00	0.00	0.02	0.02
Observations	4610085	4610085	4610085	4610085	4610085	4610085

Notes: *** denotes significance at 1%, ** at 5%, and * at 10%. Standard errors are clustered at the establishment level and reported in parenthesis.

Table A7: Distribution of transactions values in peak and off-peak time

Dependent variable: Δ Log mean transaction value in peak and off-peak time of the week

	Peak time		Off-peak time			
	OLS	IV	OLS	IV	OLS	IV
	(1)	(2)	(3)	(4)	(5)	(6)
Δ Adoption	0.00391	-0.0382	0.0359**	0.0281	0.0193	-0.00221
	(0.0204)	(0.0432)	(0.0142)	(0.0333)	(0.0130)	(0.0298)
Sector-zipcd-quarter FE	Yes	Yes	Yes	Yes	Yes	Yes
Dummies first 4 quarters	Yes	Yes	Yes	Yes	Yes	Yes
Demand Controls					Yes	Yes
Observations	4610085	4610085	4610085	4610085	4610085	4610085

Notes: *** denotes significance at 1%, ** at 5%, and * at 10%. Standard errors are clustered at the establishment level and reported in parenthesis.

Table A8: Credit Card vs Cash Sectors

Dependent variable: Δ Log revenue

	(1) OLS	(2) IV
Δ Adoption x Credit card	0.0511**	0.0924**
	(0.0240)	(0.0401)
Δ Adoption x Cash	0.0412**	0.0866**
	(0.0206)	(0.0436)
Residual CF		-0.0492
		(0.0444)
Sector-zipcd-quarter FE	Yes	Yes
Dummies first 4 quarters	Yes	Yes
p-value null equal returns	0.75	0.86
Observations	4610085	4610085

Notes: *** denotes significance at 1%, ** at 5%, and * at 10%. Standard errors are clustered at the establishment level and reported in parenthesis. Based on information from the Bank of Spain on those sectors with hihg and low share of credit card transactions, we divide sectors in our data as credit card (car dealers, technology, fashion, health care, home, accommodation, and travel) and cash (bars and restaurants, books and press, food, hypermarkets, leisure, other services, sports and toys, transportation and wellness and beauty).