Essays in Economics of Digitization

Guillermo Uriz-Uharte

Thesis submitted in fulfilment of the requirements for the degree of Doctor of Philosophy

Department of Economics
University College London

September 26, 2020

Declaration

I, Guillermo Uriz-Uharte, confirm that the work presented in this thesis is my own. Where information has been derived from other sources, I confirm that this has been indicated in the work.

Chapter 2 was undertaken as joint work with Jose Enrique Galdón Sanchez and Ricard Gil.
Chapter 3 was undertaken as joint work with Jose Enrique Galdón Sanchez, Ricard Gil and Felix Holub.

Signature:

Date: September 25, 2020

Acknowledgements

I would like to express my gratitude to all people and institutions without whom the work on this thesis would have never been possible.

I am grateful to my supervisors, Imran Rasul and Áureo de Paula, for their guidance, challenge, support, and inspiration. I would also like to thank all faculty members and colleges at UCL Economics for providing a stimulating intellectual environment where I have been able to grow both personally and professionally. I acknowledge the financial support and generosity of Fundación La Caixa, ESRC, BBVA Foundation, Bank Sabadell and Club Financiero Español the City of London.

On a personal level, first, I want to thank my parents and my sister, without their unconditional support I would have never been able to walk all this path. Second, I want to thank all friends I have made during this time in London, specially Andrea Pirrone, Gonzalo Paz, Javier Brugues, Jon Piqueras, and Marta Lopes. And, last but not least, I want to thank all my friends from Pamplona who have been next to me during all this time, Felix Holub for all the support and stimulating conversations, and Jorge Alcalde, as none of this would have happened if not for him.

Abstract

While neoclassical economics implicitly assumes that [perfect] information is widely available to firms and decision makers, the crude 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 determinants of numerous economic outcomes. The rise of information technology in the last decades has dramatically lowered the costs of collecting, using, and passing information, originating the eruption of the Information Age. These three essays contribute to the understanding of the impacts of digitization on the development and functioning of economic institutions.

Chapter one studies how a reduction in firm internal communication costs, coming from the adoption of new communication technologies, helps large corporations to achieve higher levels of innovation by overcoming limitations in internal organization.

Chapter two evaluates the impact of an unprecedented Big Data information service, that diffused at zero cost by a large bank, provides information about the competitive environment of the firm. This program presents a unique opportunity to study how access to market information might impact small and medium size firms' performance and strategic decision-making. Results show how adopting establishments are able to increase revenues by *(i)* targeting unexploited business opportunities and diversifying their customer portfolio, and *(ii)* streamlining resource allocation.

Chapter three analyses the implementation of a driving-restriction policy in the city centre of Madrid known as Madrid Central. By restricting access by car to the ban-affected area, Madrid Central achieved its goal of reducing pollution levels in the city centre. However, this can come at the cost of increasing transportation costs for consumers, and discouraging consumption in the area. Results show how information technology, in the form of e-commerce adoption, allowed establishments in the ban-affected area to weather the situation and compensate the decrease in brick and mortar sales with an increase in online sales.

Impact Statement

The analysis and results presented on this thesis are valuable to further our knowledge in Economics through a double channel. First, they contribute to understand the effects of the digital revolution on the economy by providing a rigorous quantitative analysis of the impact of different information and communication technologies (ICTs hereafter) on firm behaviour and market performance. And, second, by analysing the impact of these different information technologies on different economic outcomes, this work helps to gain further insights into complex economic phenomena and market fundamentals that otherwise would be too difficult to study empirically. In a rapidly changing world shaped by the increasing availability of new technologies and data, improving our knowledge on both dimensions is crucial to stay ahead of events and regulate markets accordingly. I hope the work on this thesis helps to shed some light and stimulate future work in these areas.

Chapter one of the thesis studies the impact of ICT adoption in large organizations. By decreasing communication costs, ICT alleviates coordination problems in the internal organization of large corporations. I find this has a significant impact fostering firm innovation and patenting, which in turn brings an increase in firm productivity. These results help to understand the nature of the innovation process as a company-wide endeavor that requires the combination of knowledge inputs from workers with different specializations. Whereas most of the innovation literature has focused on studying firm and market incentives for innovation, neglecting a systematic empirical analysis of firm innovation capacity, this article helps to fill this gap by showing evidence of (i) the importance of firm organizational capacity for innovation activities; and (ii) how technologies can help to overcome limitations in organizational capacity and, as a result, raise innovation. These results should serve to inform the design of more effective government policies promoting innovation and help firms to conceive better innovation strategies. They also provide a possible explanation for the changes in competition and market dynamics observed over the last decades. Improvements in internal organization coming from technology adoption can be one of the factors behind the rise in large firms' market power and reductions in business dynamism documented in other works. These are issues of crucial importance for market regulation and policy design.

Chapter two empirically investigate how getting access to Big Data information would affect small and medium size firms' performance and decision-making. Finding positive returns of adoption allows to confirm that the sparse adoption of Big Data technologies by small and medium firms is not due to lack of returns, but to high costs of adoption or lack of awareness. This serves to open the door for public (and private) interventions intended to correct the scant adoption patterns among these firms and, that way, ameliorate the increasing productivity and market power differences between small and large firms. Moreover, the results show how access to Big Data brings increases in market competition by moving adopters to target new customer segments and diversify their client portfolios. This constitutes further evidence in favour of the introduction of public and private initiatives intended to spread adoption of Big Data technologies and facilitate data sharing initiatives helping to highlight existent market opportunities and favour competition.

Chapter three analyses the costs and benefits of the introduction of a driving restriction policy in the city centre of Madrid. These types of policies have previously been shown to be effective in reducing pollution levels but are usually confronted with high levels of opposition by local commerce arguing the existence of a negative impact on local sales of ban-affected areas that may overturn any positive effect on other dimensions. The findings on this paper provide a detailed and rigorous analysis of the impacts generated by the introduction of Madrid Central that should be taken into account when planning the introduction of any future policy of this type. Results, first, confirm the effectiveness of Madrid Central in reducing traffic and pollution levels. Moreover, they show how potential reductions in brick and mortar sales in the ban-affected areas can be compensated by online sales. These findings highlight that whereas on aggregate terms these policies do not generate relevant distortions, one should be aware of, first, heterogeneous impacts and distributional effects, and, second, the potential important role of new technologies in alleviating distortionary effects.

TABLE OF CONTENTS

List of Figures

List of Tables

Chapter 1

Introduction

Digitization is the process by which technology lowers the costs of storing, sharing, and analyzing data. This process has changed how consumers behave, how firms compete, and how industrial activity is organized. The work on this thesis tries to understand and provide evidence of some important aspects of this transformation process.

Chapter 2 studies the value of firm communication for innovation activities by using information on the adoption of firm communication technologies. Innovation is a company-wide endeavor that requires the combination of knowledge inputs from workers with different specializations. Thus, problems in the internal communication of large corporations hinder their innovation capacity. In this chapter, I show how a reduction in communication costs, coming from the adoption of new communication technologies, enables geographically dispersed firms to achieve higher levels of innovation by lowering the costs of collaboration in R&D activities. I find the increase in innovation is more pronounced for firms with widely dispersed site locations, firms in low competition sectors and firms with low innovation in the past. Moreover, using information on the economic value of innovations, I find better communication is most effective in increasing the number of high-value innovations. Finally, I find a positive response of firm productivity to increases in innovation. I interpret the results as evidence of the impact of new technologies in overcoming limitations in the internal organization of large corporations, potentially increasing their competitive strength as a result.

Chapter 3 provides evidence of the importance of information in competitive markets by analysing the impact of Big Data technologies on the performance and decision-making of small and medium size enterprises. A firm may gain competitive advantage over its rivals through access to market information. Yet evidence suggests only large firms invest in technology that facilitates information provision, potentially contributing to increase their leverage over smaller competitors. This paper aims to empirically investigate how getting access to Big Data information would affect small and medium firms' performance and decision-making. 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 adoption, the bank provided monthly reports with rich information about each firm's clientele portfolio and that of its competitors coming from the analysis of 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 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 underserved customer segments. Not only they increase their number of customers, their new customers also come from underrepresented geographic areas and gender-age groups in their customer portfolio prior to adoption. Our evidence is also consistent with establishments improving their resource allocation efficiency upon technology adoption. These findings suggest that small and medium enterprises obtain substantial returns from information access, and therefore, high adoption costs and lack of awareness are likely to be key barriers preventing these firms from investments in Big Data technology.

Finally, chapter 4, studies the benefits and costs of driving restriction policies by making an analysis of the impact of Madrid Central on congestion, pollution and consumer spending. Driving restrictions in cities aim to reduce congestion and pollution, but they may also unintentionally distort consumer spending decisions. By increasing transportation costs to ban-affected areas, driving restrictions could discourage consumption in stores of those areas. This paper empirically evaluates the effects of a driving restriction regulation in Madrid, Spain, known as Madrid Central. First, using a difference-in-differences identification strategy, we find a decrease of 15% in both congestion and pollution. Second, we rely on a unique dataset on credit card transactions detailing spending for each pair of buyer-seller locations zip codes to analyze how the driving ban changed consumption behavior. Although we find no significant effect on overall consumption spending, our findings show a reduction in brick-and-mortar sales, and a substitution towards online shopping in businesses of the ban-affected area. This implies e-commerce may allow affected establishments to, at least partially, compensate for the reduction in brick-and-mortar sales.

Chapter 2

The Value of Firm Communication for Innovation: Evidence from the Adoption of Communication Technologies

1. Introduction

Small firms have been considered to be responsible for a disproportionate share of significant innovations in the past. Quoting Holmstrom (1989): "The causal evidence suggests the hypothesis that large firms are at comparative disadvantage in managing truly innovative research". However, the patterns of innovation in the last decades reveal a striking increase in patenting concentration by market leaders and large firms (Akcigit and Ates (2019b)). This increase in innovation in large firms can be welfare improving if they were innovating too little in the past. But it can also have profound implications for market competition and business dynamism by increasing large firms' market power and discouraging new entries. To understand the consequences of the increase in patenting of large firms it is necessary, first, to identify the reasons behind it.

In this paper I present an explanation based on an increase in the innovation capacity of large firms. I show how improvements in communication within the firm driven by technology adoption allow large firms to be more innovative. Because the ability to transmit information among workers is central to the innovation process, firms experiencing improvements in communication are able to increase their levels of innovation. To empirically examine this hypothesis I exploit information on the adoption of corporate intranets based on TCP/IP (Transmission Control Protocol / Internet Protocol), i.e. private computer networks based on the same technology as the Internet primarily intended to reduce communication costs and facilitate the access to information in the environment of an organization.[1] Using an original panel dataset of US corporations from 1988 to 2002 with establishment-level information, I find that intranet adoption increases firm patent counts, citation-weighted patents and dollar-weighted patents. The effect of intranet adoption is larger in increasing the number of high economic value patents, and for firms with more impediments to communicate, less incentives to innovate and lower innovation in the past. Moreover, I find that the increase in innovation translates in an increase in firm productivity. I confirm the results are robust to endogeneity concerns in the adoption of intranets by using an instrumental variable identification exploiting variation in historical expertise with network technologies across locations.

[1] At the time of its commercial diffusion, intranets already included a wide variety of applications and functionalities such as videoconferencing, collaboration tools, access to repositories of information located in other parts of the firm, and applications to search for subject matter experts within the organization. A more detailed discussion of the main features and functionalities of intranets is presented in Appendix A.

Innovation requires coordinating efforts and inputs from workers with different specializations. Because of their different fields of expertise, these workers are likely to be spread across different teams and divisions of an organization. Moreover, the idiosyncratic and unpredictable nature of innovation hinders the possibility of coordinating innovation activities by standardizing processes and letting employees stick to some pre-agreed plan, as it is usually done in production and routine tasks.[2] Communication among workers thus becomes crucial to coordinate all relevant knowledge that is dispersed across the firm. Reduction in communication cost will help alleviate coordination problems and will result in higher innovation levels. This is especially relevant for large firms, where it is necessary to coordinate an increasing number of agents with more narrowly-defined tasks and knowledge.[3]

The data on the adoption of intranets comes from the private sector data source Harte-Hanks. This is an establishment-level panel dataset with annual information on the stock of technology, which probably represents the richest source of historical information to study firm's Information and Communication Technology adoption (ICT hereafter).[4] Because firms in my sample are large corporations with many sites, I measure firm's intranet adoption each year as the share of firm sites that have adopted intranet by that moment in time. The point of reference for the commercial diffusion of intranets based on IP standards is the year 1995 (see Forman et al. (2012) and Scott (1998)). However, because of its non-commercial origins, many intranet technologies were quite mature by this time and were quickly adopted by firms. By exploiting variation in the adoption of intranet both across firms and over time, I can estimate its effect on innovation measured as the number of granted patents applied by a firm in a given year.

The main empirical challenge for identification of the causal effect of better communication on innovation is the possible existence of firm-specific transitory shocks to innovation that can be correlated with the adoption decision. To address this concern, first, I perform a series of robustness checks to show that the effects are not driven by simultaneous increases in firm R&D investment, capital investment or other investments in communication technologies, like the internet, that are not directly intended to improve internal communication. Second, I construct an interacted instrument that exploits differences across firms in the costs to adopt IP-based intranets when these became commercially available. This instrument combines two sources of variation: (i) changes in availability of intranet technologies over time; and (ii) cross-sectional variation in familiarity and expertise with network technologies in the regions where firms are located - some regions were more familiar with intranet-type technologies because of the existence of a connection to an early computer network (Bitnet) used by the research community of a local university. Because decisions about the connection to this type of predecessor networks were taken years before the commercialization of

[2] According to March and Simon (1958): ``The type of coordination used in the organization is a function of the extent to which the situation is standardized. (...)We may label coordination based on pre-established schedules coordination by plan, and coordination that involves transmission of new information coordination by feedback. The more stable and predictable the situation, the greater the reliance on coordination by plan."

[3] The saliency of this trade-off between specialization and communication in large organizations was previously studied by Becker and Murphy (1992), Bolton and Dewatripont (1994), Garicano (2000)).

[4] Some influential articles such as Bresnahan et al. (2002), Brynjolfsson and Hitt (2003), and Bloom et al. (2016) have used this dataset to measure hardware utilization. I focus on the adoption of network communication technologies, a much less explored component of these data.

intranets, the instrument should affect the costs and propensity of intranet adoption and should remain exogenous to firm-specific transitory shocks to innovation as argued in Forman et al. (2012).

The heterogeneous effects of intranet adoption provide further evidence on how improvement in communication of a firm impacts its innovation capacity. The increase in innovation is larger for firms with more widely dispersed site locations, firms with low levels of innovation in the past and firms that operate in low competition sectors. This result shows that improvement in communication is more effective in fostering innovation in firms that face greater barriers and incentives to innovate. Using information on the economic value of each patent, I find that intranet adoption is most effective in increasing the generation of high-value patents. By contrast, it does not affect more intensively the generation of innovations of higher scientific quality, as proxied by the number of citations received by a patent. This result is consistent with intranets improving communication not within the research team, but between the research team and other firm departments with different specializations. As a result, the pure scientific quality of innovation is not affected, but the new information flowing between the research department and other parts of the firm is especially valuable to identify profitable innovation ideas and to attune them to consumer preferences and market opportunities.

Finally, I study the effects of intranet adoption and innovation increases on firm productivity. I identify a positive response of firm productivity to increases in innovation. Thus, better communication is indirectly affecting productivity by increasing firm's innovation. This positive impact of innovation on productivity has also been documented in other settings Crepon et al. (1998) and Doraszelski and Jaumandreu (2013)). By contrast, I find no direct contribution of intranet adoption to firm productivity. Production tasks are predictable and can be coordinated ex-ante by standardized processes. As a result, the direct impact of improved communication on productivity is of second-order importance.

This article provides empirical evidence on the importance of internal firm communication for innovation activities and shows how ICTs can foster innovation by alleviating communication problems. Previously, Forman and Zeebroeck (2012) and Agrawal and Goldfarb (2008) showed how better communication between different research teams raises their probability of collaboration in joint projects.[5] My results go one step further than the current literature by showing how better communication can not only raise collaboration among researchers but increase innovation at the organization level. This work also adds to the understanding of the effects of ICT investments on firm productivity. It shows that one of the channels through which ICT investment impacts firm productivity is by increasing firm innovation. This result complements previous works studying the role of ICT as a General Purpose Technology acting as a complement of other factors such as firm decentralization (Bloom et al. (2014)); managerial skills (Bloom et al. (2012)) and human capital (Autor et al. (2003) and Akerman et al. (2015)). Finally, the results in this article contribute to explain the existence of large and persistent differences in productivity levels across businesses (Syverson

[5] There are also a number of articles showing the existence of a positive correlation between better communication and innovation. For instance, Jensen et al. (2007) demonstrate for a sample of Danish firms how those excelling in innovation count with a good system of internal communication and transmission of tacit knowledge. Mansfield and Wagner (1975) shows how a closer integration of the departments of marketing and R&D increases the probabilities of commercialization of technologically successful projects.

(2011)), which are closely connected to other highly relevant phenomena like the rise in large firms' market power and decreases in business dynamism documented for the last decades (De Loecker and Eeckhout (2020) and Akcigit and Ates (2019a)).

The rest of the article is organized as follows. Section 2 describes the data. Section 3 explains the identification strategy. Section 4 presents and discusses the results, and Section 5 concludes.

2. Data

To implement the empirical analysis, I use three different datasets with information on technology adoption, patenting, and accounting data of US firms. First, I obtain information on the adoption of intranets and other ICTs from Harte Hanks Technology database. Second, I use the Kogan et al. (2017) patent dataset which contains information on firm assignee and year of application for each patent in the US. Finally, I complement this with firm productivity estimates and other control variables obtained from Compustat.

2.1. ICT Data

I use the Computer Intelligence Technology Database (CiTDB), a site-level ICT panel produced by the information company Harte-Hanks (hereinafter HH). Since the late 1980s HH has collected IT data in order to sell it to large producers and suppliers of IT products that use it to target their sales efforts. Quoting Bloom et al. (2012) "This exerts strong market discipline on the data quality, as major inaccuracies are likely to be picked up by HH's customers".[6] HH surveys establishments annually on a rolling basis obtaining information of the firm IT stock. The CiTDB contains detailed hardware, equipment, and software information at the establishment-level for US firms with 100 workers or more, which probably represents the richest source of historical information on IT adoption.

In this work I focus on studying the effects of adoption of corporate intranets based on TCP/IP standards.[7] Following previous papers using HH data (Forman (2005), Forman et al. (2002, 2005, 2008, 2012), Forman and van Zeebroeck (2012), and Scott (1998)), I consider the diffusion of TCP/IP intranets began in 1995 and can be considered to be zero before this date.[8]

[6] This dataset also has a number of cleanliness issues when it is used as a panel as there are discontinuities in both the sites that are surveyed and the survey methodology from year to year. With regard to the survey methodology, this is of no concern as the information used in this paper comes from questions that remain unchanged over time. With regard to the establishments covered in the survey, I require an establishment to be covered in all sample years from 1996 to 2002 to be included in the sample. This ensures a consistent measure of adoption at the firm level over time that does not depend on some establishments being dropped or added to the sample.

[7] This information is compiled from telephone surveys in which establishments are asked about the adoption of an intranet based on TCP/IP standards by the moment in which the interview is taking place.

[8] Internal company networks had been around for decades. These networks, though, used their own proprietary software, which was very costly to design, install, operate and evolve. This software used a set of protocols, unique for each company, installed in each computer specifically for the use within this proprietary network. Being very expensive, internal company networks did not provide global accessibility nor did they interconnect all the computers and database resources of corporations. Moreover, being proprietary, they did not provide the market opportunity for the development of cost effective, massively used software. With the booming of the Internet after 1994, the Transport Control Protocol/Internet Protocol (TCP/IP) protocols became standard, leading corporations to design their internal company networks based on them. As a consequence, this transition implied a drastic

This way, my sample includes information from 1988 to 2002, covering the years previous to intranet commercialization, and the years of its diffusion. Furthermore, I do not have information on technology adoption for every year but for every other year.[9] In order to accommodate these two requirements, my sample includes all alternative years from 1988 to 2002, giving a total of 8 firm-year observations. Because HH information is disaggregated at the site-level but patent and accounting information are available only at the firm level, I aggregate intranet adoption at the firm level by calculating the percentage share of firm firm sites reporting to have adopted intranet at a given year. This is my main measure of firm intranet adoption.[10]

My sample is restricted to large corporations because until the late 1990's HH surveyed only firms in the Fortune 1000 list. This is, in general, a limitation to obtain generalized conclusions on the effects of technology adoption on firm outcomes Draca et al. (2006). However, the results are still highly important because (i) firm internal communication problems are greater in large firms than in small ones; and (ii) big corporations are responsible for a very large share of R&D investment and innovation.[11]

2.2. Patent Data

I measure firm innovation by the number of granted patents applied by a firm in a given year. Patents constitute the best available measure of firm innovation outcomes, include interesting technical information, and they are extensively used in the Economic literature. I use a new patent dataset constructed by Kogan et al. (2017). This is an extension of the NBER Patent-citation dataset containing all patents granted by the US Patent and Trademark Office (USTPO) from 1926 to 2010 and all citations received by each patent. It further includes a new measure of the economic value of each patent. This value is estimated by exploiting movements in stock market prices following the day that a patent is issued to a firm.

2.3. Compustat Data

Compustat North America is a database of U.S. and Canadian fundamental and market information on more than 24,000 active and inactive publicly held companies. It is a near census of publicly traded firms, providing thousands of annual reports. First, I use Compustat (Fundamental Annuals) information on sales, number of employees, capital, investment, and

---

reduction in the costs to build and manage an intranet, a huge increase in the availability of applications and a rapid adoption among firms such that any employee with a standard browser and the relevant password was able to access the company's network. Quoting Forman et al (2012) "Dating the rise of the commercial Internet is not an exact science, but a few well-known events provide useful benchmarks for understanding why investment began to boom in 1996 and not before. The first non-beta version of the Netscape browser became available in early 1995, followed by the firm's IPO in August 1995. Bill Gates' internal memo about Microsoft's change in direction ("The Internet Tidal Wave") is dated May 1995." Internet Protocol technologies had not diffused among firms prior to 1995, so it is a common assumption in all previous work using HH data to consider adoption of all TCP/IP based technologies (intranets among them) to be zero prior to this date. See Forman et al. (2002) and Forman (2005) for a detailed analysis of the adoption pattens of TCP/IP technologies by US firms.

[9] This same resource concern is present in Forman and van Zeebroeck (2012).

[10] I show that results do not change by using alternative measures of intranet adoption.

[11] Firms in the sample are responsible for approximately 25% of total R&D investment carried out in the US during the sample years (source: https://www.aaas.org/page/historical-trends-federal-rd).

consumption of intermediate inputs I obtain firm productivity estimates. Appendix D contains a more detailed explanation of the variables used and the estimation procedure. Second, I also use Compustat to retrieve information on the levels of annual R&D investment reported by firms.[12]

2.4. Matching of Datasets

To use all the above information, it was necessary to match firms in the three different datasets. I started by matching patent data to Compustat information. Patent assignees in Kogan et al. (2017) database are identified by CRSP permno number. Using a bridge dataset linking CRSP and Compustat firm identifiers I can directly connect patent assignees to Compustat firms. Matching HH firm information with the two other datasets was a more complicated task. HH does not provide firm identifiers that link firms in HH to other commonly used databases. As a result, I had to resort to string matching algorithms in order to match firms by their names. The details of this matching process are explained in detail in Appendix B.

2.5. Descriptive Statistics

Once firms in the three datasets are matched, I construct a balanced panel of firms including all even years between 1988 and 2002. To do so, I drop firms that are not fully covered by Compustat or HH in all these years. Moreover, I drop firms that are not granted any patent over the sample years as in a Poisson fixed effects regression they do not provide any variation in the estimation.[13] This is a widespread practice in the literature and has no implications for the consistency of the estimates. A more thorough explanation of the reasons for this will be presented in the next section.[14]

My final data is a balanced panel of 348 firms. On average, I have information on intranet adoption for 37 establishments per firm. Table 1 presents some characteristics of these firms in more detail. As mentioned above, these are large firms, which is reflected in the high levels of employment (35,000 employees on average with a firm-year maximum of 1.4 million workers), capital (2,942 million USD on average), and sales (9,485 million USD).[15] The sample covers close to 10% of the total US labor force and around 33% of total sales by US public firms. It is also interesting to highlight the high variation across firms in all these dimensions. This dispersion is even more pronounced in innovation-related variables. Some firms

[12] The US Statement of Financial Accounting Standards No.2 (SFAS 2) requires US firms to disclose in their financial statements material R&D expenditures.

[13] From firms with information in HH for the years 1996 to 2002 I am able to match 764 to Compustat and patent data. Out of this, 387 have at least one patent for one of the sample years. Limiting the sample to a balanced panel by dropping firms with missing information in Compustat for at least one of the sample years (this can be because the firm did not exist at the beginning of the sample or it is missing information in Compustat for one of the sample years), the final sample contains information for 348 firms.

[14] Results in Tables A1 and Table A5 column 1 show baseline results are robust to keeping firms with zero patents over the sample period and using an unbalanced panel.

[15] A referee noted the minimum value of 50 employees is too low given the sample is made of large corporations. I discovered this observation corresponds to a firm that in year 2002 suffered a shock that significantly reduced her size. In appendix Table A10 I present robustness results dropping this firm for all sample years that confirm this does not make any difference in the baseline estimates.

consistently invest heavily in R&D and generate many dozens or even hundreds of patents every year, whereas, in other cases, innovation is a more sporadic and less intense activity.

As mentioned above, I define the level of intranet adoption for a firm in a given year as the percentage share of firm sites that have adopted intranet at that point in time. To identify the effect of intranet adoption on patenting, I exploit both variation over time in intranet penetration and across firms in each year. Figure 1 shows the different levels of intranet penetration for each of the years. Before the year 1996 adoption is 0 for all firms. In 1996, almost 75% of the firms had adopted intranet in at least one site.[16] However, in most of the firms, penetration of intranet is still low with high levels of dispersion (average adoption of 10.2% with standard deviation of 11.5%). This timid diffusion in the beginning was followed by a rapid expansion in subsequent years (average adoption of 27% in 1998 and 44.7% in 2000) that slows down to reach 55% in 2002. This pattern is consistent with the well-known "S" shape observed in the diffusion process of many other innovations (Griliches (1957)). Interestingly, as the level of intranet diffusion increased, so too did the dispersion in penetration across firms (around 20% standard deviation for the years 1998, 2000, and 2002).

A number of reasons can account for the apparently puzzling fact that different establishments of the same firm adopt intranet at different moments in time.[17] The main reason is probably the existence of significant geographical variation in the local availability and costs of Internet and intranet's connectivity during the late 1990s (see Forman et al. (2005)). This was especially important for intranet implementation, as it usually requires high-speed connectivity and not all locations counted with this service over this period. Another reason to mention is that control about IT decisions in firms is often decentralized (see Sambamurthy and Zmud (1999); McElheran (2014)). As a result, IT investment decisions may ignore potential complementarities arising from coordinated investments.

Finally, Figure 2 captures the different trends in patenting over time for those firms adopting intranet more intensively (strong adopters) and those others doing it less intensively (weak adopters). Because treatment in this context is not a binary variable as in many other economic settings, it is not obvious how to define who is treated and non-treated. I consider as strong adopters those firms that are above the median of intranet adoption for all years between 1996 and 2002 (98 firms). Figure 2 shows for each year in the sample the average number of patents generated by a firm in each of the two groups. This graph shows that (1) there exists an increasing trend in patenting over time for all firms in the sample; (2) both strong intranet adopters and weak intranet adopters present parallel trends in patenting

[16] As explained more extensively in Appendix A, because of its non-commercial origins, many intranet technologies were already quite mature by this time and could be rapidly adopted and applied to organizational needs. Moreover, this type of technologies require low co-invention costs to be used successfully (Bresnahan et al. (1996)). As a result, my analysis focuses on short-run changes in innovation and patenting that are made in response to a decline in collaboration costs. This is perfectly in line with other works such as Forman (2005), Forman et al (2005, 2008, 2012), and Forman and van Zeebroeck (2012) that also consider adoption of IP standard technologies to be zero before 1995 and study the impact of adoption during the late 90's.

[17] For a comprehensive analysis of this issue refer to Forman et al. (2012).

growth for years before 1994; and (3) there is a pronounced acceleration in patenting after 1994 for strong intranet adopters coinciding with the diffusion of intranets.[18]

3. Empirical Identification

3.1. Estimation Strategy

Following a literature initiated by Hausman et al. (1984), I consider a firm patent production function in which I include intranet adoption as one of the determinants of patenting. Because the dependent variable, number of granted patent applied by a firm in a given year, is a count variable with overdispersion and many zeros, I use Poisson-based econometric models and estimation methods (see Cameron and Trivedi (2013)). Then, assuming that the patent process follows a Poisson distribution, the expected number of patents for firm *j* at year *t* has the following exponential functional form[19]

$$E[Pat_{jt} | Intranet_{jt}, \boldsymbol{x}_{jt}, \eta_j, \delta_t] = \lambda_{jt} = exp(\alpha\, Intranet_{jt} + \beta\, \boldsymbol{x}_{jt} + \eta_j + \delta_t) \quad (1)$$

where η_j is a firm fixed effect and $\boldsymbol{x}_{jt}$ includes time-varying controls such as R&D investment, number of employees, and year fixed effects. Coefficients of this model should be interpreted as semielasticities. Therefore, α is the semielasticity of patenting with respect to $Intranet_{jt}$. Because I include firm fixed effects, all firm time-invariant characteristics are going to be subsumed and not individually identified. In order to estimate this model by Poisson regression it is necessary to make an assumption of strict exogeneity. Thus, estimates may be inconsistent if regressors are predetermined (e.g., if past shocks to patenting affect R&D investment or intranet adoption in the future).

One big advantage of using Poisson fixed effects regression is that, in contrast with many other non-linear panel data estimators, there is no incidental parameter problem. Even in a short panel (*T* fixed) with $N \to \infty$ one can consistently estimate α given that the conditional mean function is multiplicative in the fixed effect. Similar to the transformation used in the linear fixed effects regression, the fixed effects can be eliminated by using a conditional MLE. Some algebra leads to the simple moment condition

$$\sum_{j=1}^{N} \sum_{t=1}^{T} \boldsymbol{x}_{jt} \left(Pat_{jt} - \frac{\lambda_{jt}}{\lambda_j} \overline{Pat_j} \right) = 0 \quad (2)$$

This moment condition reveals an issue: observations with $Pat_{jt} = 0$ for all T make no contribution to the estimation of α and β. Thus, it is better to drop them out of the sample. As commented on above, this is a common practice in the innovation literature (see for

[18] Figure A1 considers as *strong adopters* those firms that are above the 66 percentile of intranet adoption in all years after 1994 (56 firms in total), and *weak adopters* those firms that are below the 33 percentile of intranet adoption in all years after 1994 (51 firms in total).

[19] The variable $Intranet_{jt}$ constitutes a measure of intranet penetration in firm *j* at year *t*. As argued in Bresnahan et al. (1996), intranets require little adoption or coinvention to be used successfully. Considering also that most patents are applied at an early stage in the innovation process (Pakes (1986)), I focus on short-run changes in innovation driven by intranet adoption.

instance Aghion et al. (2013)) and, as shown by Blundell et al. (2002) it has no implications for the consistency of the estimates.

3.2. Identification Issues

The most basic concern to consider for the identification of the effect of intranet on patenting is a potential selection based on firm characteristics. This would be the case if firms with better or worse innovation capacity were the ones adopting intranets more intensively. Then, the estimated effect of intranet may just be a spurious correlation due to omitted firm characteristics. Using panel fixed effect estimation I control for all firm permanent characteristics and, as a result, concerns about time-invariant endogeneity are solved.

Adopting this approach, the identification strategy exploits variation over time and across firms in the adoption of intranet to estimate its effects on innovation. In other words, given that different firms are adopting intranets with different intensities at different times, I exploit changes in patenting for those firms adopting intranets more intensively at a given moment in time to identify intranet's effect on innovation. This identification strategy hinges on the existence of parallel trends in the propensity to patent for strong intranet adopters and weak adopters in the absence of intranet adoption (after controlling for observable time-varying factors). This may fail if there are firm-specific unobservable transitory shocks that simultaneously affect firm patenting and adoption of intranets. Throughout the results section, I present a series of robustness checks and IV estimations that support the rejection of this hypothesis and favor a causal interpretation of the estimates.

4. Results

This section is structured in three subsections. Subsection 1 reports the baseline results for the effect of intranet adoption on innovation, together with robustness specifications and instrumental variable estimates. Subsections 2 analyzes the heterogeneous effects of intranet adoption for the generation of different types of innovations and for different innovating firms. Subsection 3 presents results about the effects of innovation on firm productivity.

4.1. The Effects of Intranet Adoption on Innovation

Baseline Results

The baseline results are based on a Poisson regression of the number of granted patents applied by firm *j* in year *t* on the level of intranet penetration and different sets of control variables (see equation 1).[20] The first column of Table 2 shows a positive and significant estimate on the effect of intranet when I control only for year fixed effects. Further controlling for firm fixed effects, in column 2 the estimated effect of intranet adoption on patenting is still significant at the 1% level although smaller in magnitude.

[20] Table A1 presents baseline results using an unbalanced panel and keeping firms with zero patents over the sample years.

In column 3 I include as controls the contemporaneous level and the first lag of R&D investment. R&D is the most immediate input for innovation and including it as a regressor will partially capture innovation productivity shocks. Because patents tend to be applied at an early stage of the innovation process (Pakes (1986)), including the contemporaneous level and the first lag of R&D is the best way to capture firm-specific transitory shocks to innovation.[21] Later, I will show how the results do not change by including further lags of R&D or using alternative measures as R&D stock. I also control for firm size by including the number of firm employees as a regressor. Earlier studies including Lanjouw and Lerner (1996) and Lerner (1995) point out that firm size has an effect on innovation due to the presence of economies of scale in the innovation process. It can also be the case that changes in firm size capture variation in unobserved variables affecting innovation incentives, such as firms' perspectives about the future. The inclusion of these controls reduces the size of the effect of intranet by more than one third. The magnitude of the estimated effect implies that a 10 percentage point increase in intranet penetration is associated with a 5.6% increase in patenting once changes in R&D investment and firm size are controlled for. As a result, given the mean of intranet adoption in the year 2002 was 55% (median 57%), intranet is associated with an increase in patenting of 30% approximately. Given the median firm is getting 4 patents in 2002, this implies an increase of slightly above 1 patent.

Following some of the seminal works on the literature of patent production function estimation (Hausman et al. (1984); Hall et al. (1986); etc), in column 4 I include as controls the contemporaneous level and the first four lags of R&D investment. Including these extra lags has minimal consequences in the estimated effect of intranet adoption. Moreover, because of the high autocorrelation in firm R&D investment levels, only the first and last of the years included as controls become significant, which is a common result in the literature.

In column 5 I regress by OLS the log of R&D investment in year *t* on the contemporaneous level of intranet penetration including firm and year fixed effects. Apparently, intranet adoption is not associated with significant increases in R&D investment. However, (i) the coefficient on the effect of intranet on patenting is larger when R&D investments are not controlled for (see columns 2 and 3), and (ii) IV estimates presented in Table 4 will show a positive association between intranet adoption and R&D, so probably one should not directly rule out that intranet adoption causes increases in R&D.[22]

[21] Hall and Ziedonis (2001) and Forman and Zeebroeck (2012) among others adopt a similar approach and control for the contemporaneous level of R&D investment.

[22] Appendix Figures A2 and A3 show heterogeneous treatment estimates of the impact of intranet adoption on patenting across sectors at the 1-digit SIC and 2-digit SIC. These estimates come from a Poisson regression of number of patents on intranet adoption interacted with a dummy variable for each sector, and including as controls year fixed-effects, the contemporaneous level and the first lag of R&D investment, number of workers and sector-specific time trends. We can observe a relative high dispersion in the returns of intranet adoption across firms in different sectors. Moreover, in Table A2 I make a more thorough analysis of the existence of different returns for firms operating in digital sectors (SIC sectors 3570-3579). During the late 90's there was a great expansion and a lot of innovation in digital industries, and, at the same time, firms operating in these sectors intensively adopted intranets and other types of ITs. Therefore, the returns of intranet adoption for firms in digital sectors deserve a more careful analysis. Table A2 shows that (i) there seems to be higher returns of intranet adoption for firms in digital sectors, but (ii) this additional effects disappear once we introduce a specific time trend for firms in digital sectors. This indicates that (i) both innovation and technology adoption were more intense during this time in digital sectors, and (ii) returns of intranet adoption are not higher for digital firms. Finally, Table A3 shows how there are no heterogeneous returns for early vs late adopters of intranets.

Robustness Analysis

It is well known that whereas some patents have a very strong innovative content others represent minor advances. Using simple patent counts can then underestimate or overestimate the importance of innovations. As a result, it could be the case that the higher number of patents firms generate when they adopt intranets represents marginal innovations with a negligible innovative content and the total amount of generated innovation does not change. In order to tackle this problem, the literature has proposed using information on the number of forward citations received by a patent as a proxy for its quality, radicalism, or innovative content (see Hall et al. (2005)). In column 1 of Table 3, I weight patents by an adjusted measure of the number of forward citations received.[23] Doing this, I have to drop five firms that did not receive any citation for any of the patents they obtained during the sample period. The point estimate for the effect of intranet on citation-weighted patents is significant and a bit larger in size than in the baseline regression.

However, it is still possible that some patents represent a breakthrough for a firm despite having low innovative content, or the other way around. Kogan et al. (2017) provide an estimate of the private value of a patent by exploiting movements in stock prices. They further confirm that this new measure is a useful proxy for the value of patents, showing that it is more associated with creative destruction and is more strongly related to firm growth than citation-weighted patent counts. The dependent variable in column 2 is the number of patents generated by the firm in a given year weighted by the adjusted economic value of each patent.[24] The effect of intranet on this adjusted measure of patent counts is positive and significant. It is also interesting to notice that in contrast to simple patent counts and patents weighted by citations, firm changes in R&D investment do not seem to have strong explanatory power in this case. This is an issue probably worthy of further investigation.

Columns 3 to 8, check the robustness of the results to omitted variable bias presenting different specifications with extra added control variables and using different definitions of the variables of interest. In the next section I will resort to instrumental variables estimation to further address this concern.

---

[23] I divide the number of citations received by a patent by the average number of citations of all patents that were issued in the same year. Making this adjustment corrects for the different time spans during which different patents applied for in the same year were able to receive citations. For instance, consider two patents applied for in the year 2000. One of them is granted in 2001 whereas the other is granted in 2005. In my data I have information on citations received until 2010. As a result, patents granted in 2001 were able to receive citations during 9 years whereas patents granted in 2005 could do so only for 5 years. Independently of their qualities, it is likely the case that the patent granted in 2001 received a higher number of citations. Normalizing patent counts by the average number of citations received by patents granted in the same year partially corrects for this problem (see Lerner and Seru (2017) for a monograph on these issues).

[24] I divide again the economic value of a patent by the average of the patent economic values of all the patents that were issued in the same year. This corrects for differences in stock market situation at the time when a patent is granted. Consider again two patents applied for in the year 2002. One of them is issued in 2004, when stock markets were in a good situation and the other in 2008 when stock markets were at a minimum. Independently of the quality of the patents, it is likely that the estimated economic value of the patent granted in 2004 is higher. Dividing by the corresponding average value of patents granted in the same year we partially correct for this problem.

Column 3 shows how results are robust to changes in the way in which firm's intranet adoption is measured. I construct the variable $Weighted\ Intranet_{jt}$ as the share of establishments adopting intranet in which each site is weighted by its number of employees. The effects remain positive and significant at the 5% level.

Column 4 includes a dummy equal to one when at least one establishment of the firm reports having adopted intranet. One source of concern is that when a firm decides to intensify its innovation strategy one of the things they do, together with many other possible initiatives, is to start the implementation of an intranet. As a result, the intranet dummy variable (but not the level of intranet penetration) should be positively biased if it is the decision to initiate the deployment of the intranet what correlates with other firm policies intended to increase innovation. This does not seem to be the case as the dummy variable is not significant and coefficients on the rest of the variables remain similar in magnitude.

Given that intranets can be adopted as part of a wider modernization strategy of the firm one would like to control for other types of investment decisions that can potentially drive the results. More specifically, because the deployment of intranets and Internet was contemporaneous, it could be the case that the effect of the two is being confounded. Moreover, the purpose of each of them is very different; whereas intranets mainly improve internal firm communication, the main effect of the Internet is to facilitate external communication.[25] Therefore, they can both potentially foster innovation but through very different channels that I would like to disentangle. I construct the variable $Internet_{jt}$ as the percentage share of firm establishments reporting to use the Internet for research. Column 5 shows how its effect on patenting is non-significant and the coefficient on intranet becomes somewhat larger in size and significant at the 1% level. Controlling more generally for Internet adoption or for capital installed at the firm, results remain unchanged.

In column 6 I control for an alternative measure of R&D investment. Following Hall and Hayashi (1989) and Klette (1994), I construct a measure of the stock of R&D investment using a permanent inventory method (see Appendix C for a more detailed explanation). The effect of intranet on patenting is perfectly robust to this alternative measure of R&D investment.

Despite the robustness of the results obtained so far, one could still be concerned about the possible existence of firm-specific trends on patenting correlating with the adoption of intranets. To address this concern, I introduce different time trends for different types of firms to confirm that those omitted variables are not driving my results. In column 7 I include time trends for different quartiles of ``firm's innovative capacity'' that I proxy in the following way. Following Blundell et al. (1999), I construct the variable $Patent\ Stock_{jt}$ using a permanent inventory method (assuming a 15% yearly depreciation rate) and use the value of this variable at the beginning of the sample to proxy for firm's innovative capacity. For instance, if it were the case that less innovative firms are increasing their innovation capacity and, for some reason, these are also the firms that are adopting intranets more intensively, by controlling for these trends the endogeneity problem would be solved. Column 7 shows how the effect of intranets seems to be robust to this form of endogeneity. In column 8 I include different time trends for firms that operate in digital sectors. During the '90s there is a high expansion in the digital sectors. At the same time, it is likely that firms operating in digital sectors are

[25] Because the Internet at this time was still at an early stage, those firms reporting to use the Internet for research were basically using it to obtain information about their competitors, markets, or new products.

adopting intranets more intensively. Column 8 shows the effect of intranet adoption on patenting is robust to controlling for these time trends. Controlling for year-specific dummies instead of time trends the results in columns 7 and 8 remain unchanged.[26]

Instrumental Variable Estimation

To further address the concerns of omitted variable bias, in Table 4 I present results of instrumental variable control function estimation (see Wooldridge (2015)) exploiting exogenous variations in the costs of intranet adoption. Under exogeneity of intranet adoption, the following moment condition holds

$$E[e^{v_{jt}} | Intranet_{jt}, \boldsymbol{x}_{jt}, \eta_j, \delta_t] = 1 \quad (3)$$

where v_{jt} is the error term associated with equation (1). This moment condition is not valid if $Intranet_{jt}$ is endogenous. I assume that $Intranet_{jt}$ satisfies the following linear reduced form where z_{jt} is the instrument

$$Intranet_{jt} = \mu_j + \psi_t + \pi\, z_{jt} + \gamma\, \boldsymbol{x}_{jt} + u_{jt} \quad (4)$$

and the following moment condition holds

$$E[u_{jt} | z_{jt}, \boldsymbol{x}_{jt}, \mu_j, \psi_t] = 0 \quad (5)$$

As a result, controlling in (1) for $\rho(u_{jt})$, a non-parametric function of u_{jt}, should be enough to remove the endogeneity bias. I proceed in a two step estimation, estimating first equation (4) and controlling for a polynomial series expansion of the estimated residual in equation (1). In conclusion, I use the following extended moment condition in which a simple test for exogeneity of $Intranet_{jt}$ can be conducted by checking the significance of $\rho(u_{jt})$

$$E[Pat_{jt} | Intranet_{jt}, \boldsymbol{x}_{jt}, \eta_j, \delta_t, u_{jt}] = exp(\alpha\, Intranet_{jt} + \beta\, \boldsymbol{x}_{jt} + \eta_j + \delta_t + \rho(u_{jt})) \quad (6)$$

As I have just shown, after controlling for $\rho(u_{jt})$, $Intranet_{jt}$ becomes exogenous in (6). Thus, I can introduce interactions between $Intranet_{jt}$ and other exogenous variables once $\rho(u_{jt})$ is controlled for. This is exactly the strategy I follow to study the existence of heterogeneous returns to intranet adoption for different types of firms.

The instrument I use is the product between the average number of Bitnet local connections (Bitnet nodes) per firm establishments and the number of Internet Service Providers (ISPs) operating in the US in a given year. The average number of local Bitnet nodes per firm site is

[26] Appendix Table A4 shows results are robust to including as controls a basic measure of internet access (the share of firm sites reporting to have basic internet access), e-commerce (the share of firm sites reporting to use e-commerce), or an alternative measure of R&D stock in which, following Li and Hall (2006), I use sector-specific depreciation rates for R&D investment. Table A4 also shows that using as an explanatory variable intranet adoption at the firm headquarters (a dummy taking the value of one if the firm HQ adopts intranet), or a dummy taking the value of one if at least one firm site has adopted intranet has no impact on patenting.

intended to capture cross-sectional variation in knowledge and expertise on intranet technologies.[27] Bitnet was an early computer network for the research community in US universities. Its first link was created in 1981 between the Universities of CUNY and Yale, and it had become the largest academic network in the US by the end of the 1980s. Bitnet supported interactive transmission of files and email between users based on protocols similar to the TCP/IP later used for the internet and intranets. Because of this, increases in this variable will capture improvements in local expertise with network technologies. Forman (2005) and Forman et al. (2008) argue that availability of local knowledge about IT and network technologies lowered the costs of adoption of Internet and intranets to firms. Following this rationale, other articles such as Forman et al. (2012) and Forman and Zeebroeck (2012) use the number of local nodes of Arpanet (another computer network predecessor of the Internet) to instrument for Internet adoption. Because decisions about the connection to Bitnet are taken at a time before the commercialization of intranets, one should not expect local changes on current economic activity to correlate with them. Figure 2 shows the distribution of Bitnet nodes across US counties. There exists a higher concentration of nodes along the coasts but, in general, one can see a relatively high dispersion across the country.

Because the average number of local connections to Bitnet is going to provide only variation in the cross-section of firms but not over time, I interact average number of Bitnet nodes per firm with the number of ISPs operating in the US at a given point in time. ISPs provide the technology to connect computers to networks and the internet. I obtain data on the number of ISPs from Boardwatch Magazine, a magazine specialized on the Internet that during the 1990s and early 2000s included a directory of ISPs advertisements that constitutes the most comprehensive list of the number of operating ISPs. Changes in the number of ISPs appearing in Boardwatch Magazine have been considered in other works to be a good barometer of the growth of the commercial internet access market (see Stranger and Greenstein (2007)). Before 1995 there were very few ISPs in the country but the number grew exponentially from mid-1995 on with the Internet's commercialization. The number of ISPs went from 1400 in the year 1996 to 7200 in the year 2002. I will consider the number of ISPs to be zero for 1994 and previous years (although probably there was a small number of them, around 30).

In conclusion, my instrument captures variation in local knowledge about network technologies and the availability of these network technologies over time. Thus, increases in the instrument correspond to exogenous reductions in the costs of adoption of intranet technologies that should affect its adoption. By contrast, one should not expect it to affect firm innovation through any other channel. A possible threat to this exclusion restriction would be the existence of different economic trends contemporaneous with the adoption of intranets across regions with more or fewer connections to Bitnet.[28] However, as argued by Forman et al. (2012), because these are historical decisions and there is relative high dispersion of Bitnet nodes across different regions, this variable is unlikely to be correlated with economic activity over our sample period.

Despite the impossibility to empirically check the validity of the exclusion restriction, I can perform a test for the presence of different trends in patenting across firms in regions with

[27] This is calculated in the following way: (i) for each site I take the number of Bitnet connections in the county, and (ii) for each firm I calculate the mean of this number across firm sites.

[28] For analysis of possible endogeneity concerns when using an interacted instrument in panel regression see Christian et al. (2017).

more or less Bitnet nodes in the years before intranet diffusion (1988 to 1994 in my sample). If the exclusion restrictions holds, then being in a region with more Bitnet nodes should not affect patenting but through its effect on intranet adoption. Therefore, I should not be able to find any impact of Bitnet nodes on patenting in the years before intranet commercialization. To do this, I regress the number of firm patents on a time trend and a time trend squared interacted with the variable $Binet\ Nodes_j$ whilst controlling for year and firm FE. The two trend variables are jointly and individually insignificant (p-value of 0.279 for the joint significance). This result is consistent with the validity of the instrumental variable identification.

The first column of Table 4 shows the first-stage regression of intranet adoption on the instrument and controls. The effect of the instrument on intranet adoption is positive and significant (F-statistic of 38.52). Furthermore, this first-stage specification captures 79.9% of the within-firm variation in intranet adoption. Column 2 reports the second stage; a Poisson FE regression of number of patents on intranet, controls, and a polynomial series expansion of the residual from the first stage (the residual and the residual squared). The effect of intranet on patenting is positive and significant at the 5% level. Its size is larger than in the baseline regression, perhaps due to the existence of heterogeneous returns to intranet adoption. If firms located in counties with more Bitnet nodes have higher returns to intranet adoption - because there is more knowledge about the technology and its implementation is going to be more efficient - then the local average treatment effect can be expected to be greater than the average treatment effect. This implies that although the instrument affects innovation only through its impact on intranet adoption, the returns to intranet adoption are largest for those firms whose adoption decision is most strongly affected by the instrument. Despite the coefficient increase, the residual control function is not significant (p-value of 0.138 for the joint significance of the residual and the residual squared) so the null hypothesis of intranet adoption being exogenous cannot be rejected. This increase in the IV estimates is in the same order of magnitude as results in other papers relying on very similar identification strategies for the effects of communication technologies (Forman et al. (2012) and Forman and Zeebroeck (2012)).[29]

In the remaining of columns, I report second-stage estimates for different specifications. I do not present first-stage estimates, but in each column I include the corresponding F-statistics. In column 3 I show the second stage of a regression including all four first lags of R&D investment (the control function version of table 2 column 4). The effect of intranet remains positive and we do not reject the exogeneity of intranet adoption (p-value of 0.164 for the residual control function).

In column 4, instead of using the interacted variable $Binet\ Nodes_j * ISPs_t$ to instrument intranet adoption, I use $Binet\ Nodes_j$ interacted with a different dummy for each of the years 1996, 1998, 2000 and 2002. This allows for more flexible heterogeneous impacts on the effects of the instrument over time. As a result, I have four IVs to instrument intranet

---

[29] In Table 6 column 2 I will show how returns to intranet adoption are quite larger for firms with lower levels of innovation (these firms were getting between 0 and 2 patents per year before intranet adoption). Because small increases in patenting for these firms represent huge increases in percentage terms, the estimated impact of adoption is much larger for them, and in the same order of magnitude as the IV estimates. If these low innovation firms are the ones whose adoption decision is most strongly affected by the instrument, it makes perfect sense the jump in the estimated IV coefficients with respect to the baseline estimates.

adoption. This identification exploits only cross-sectional variation in the costs of intranet adoption for each of the years in which intranet is commercially available, which is the exact same strategy used in Forman and Zeebroeck (2012). Because now I use four IVs, but the four of them are able to explain the same variation in intranet adoption as the original one (using the 4 IVs the R-square in the first stage goes from 79.9 to 80.03), the F-statistic goes down to 11. Apart from this expected change, the rest of the coefficient estimates remain basically unchanged.[30]

Columns 5 and 6 report the second stages for the adjusted versions of patent counts: patents weighted by forward citations and patents weighted by their economic value (the control function version of Table 3 columns 1 and 2). The effect of intranet on patenting is significant in both cases but, for the case of patents weighted by their economic value, exogeneity of intranet adoption is rejected.

Column 7 presents an OLS IV regression of the log of R&D investment on intranet adoption (the IV version of Table 2 column 5). The effect of intranet adoption is significant at the 5% level in this case, and implies that a a 10 percentage point increase in intranet penetration is associated with a 4.5% increase in R&D investment. As a result, given the mean of intranet adoption in the year 2002 was 55%, intranet is associated with an increase in R&D of 25% approximately. Given the median firm is investing 19.9 million in R&D in the year 2002, this implies an increase of almost 5 million. The result is consistent with intranet adoption affecting patenting in two ways. First, it impacts patenting in a direct way by increasing the production of patents once R&D investments are controlled for. Second, it fosters patenting in an indirect way by prompting an increase in R&D investment that in turn will result in more patenting.[31]

In conclusion, the IV results seem to confirm the existence of a positive causal impact of better communication on firm innovation. One further concern to take into account is whether the pathway between intranet adoption and innovation is mediated only by improved communication. It could be the case that as a result of better communication the firm decides to make other complementary investments that also have an impact on innovation capacity. I do not consider this a serious concern for two reasons. First, because I have already controlled for some of the main candidates for complementary investments such as increases in R&D investment, capital investment or investments in other technologies; and none of

---

[30] Appendix Table A6 column 1 shows the first-stage of this IV estimate. The impact of local knowledge on adoption propensity seems to be larger in magnitude and more significant after 1996.

[31] Appendix Table A5 presents robustness analysis of the IV estimates (the corresponding first-stage regressions are in Table A6 columns 2 to 7). Column 1 uses as IV the number of Bitnet nodes per firm ($Binet\ Nodes_j$) interacted with a time trend taking the values 1 to 4 for years 1996 to 2002. Results are very similar in magnitude to the baseline IV estimates. Column 2 uses as IV the number of Bitnet nodes in the headquarters of the firm, instead of the mean of nodes for all firm sites, and interacts this with $ISPs_t$. The estimates are less precise in this case, but similar in magnitude. Column 3 uses the average number of Arpanet nodes per firm site interacted with $ISPs_t$ ($Arpanet\ Nodes_j * ISPs_t$). Results are similar in magnitude and more significant in this case. Column 4 uses a Bartik-type IV in which the instrument is constructed in the following way: (i) for each establishment and year I calculate the share of other establishments in the same county-year that have adopted intranet -excluding the focal establishment-, and (ii) for each firm-year I calculate the variable *Bartik County* as the mean of the previous number across firm sites. The estimated effect of intranet is slightly larger in magnitude but less precise. Column 5 uses another Bartik-type instrument, but this time at the industry level, by calculating for each firm-year the average level of intranet adoption for all other firms in the same industry (at the 2-digit SIC level, or at the 1-digit SIC level if there is no other firm at the 2-digit level). Results are again similar in magnitude to the baseline estimates but less precisely estimated. Finally, column 6 uses the same IV as in the baseline specification, $Bitnet\ Nodes_j * ISPs_t$, but with an unbalanced sample of firms.

them are completely driving the effects. Second, and more importantly, because even if improvements in communication trigger other investments and organizational changes, my estimates would still be capturing the total effect of better communication on innovation. Decomposing this total effect into the different channels that explain it would be a natural next step. In the next sections I am able to provide some evidence about the mechanisms driving the effects, but further research would be necessary to provide a complete answer to this.

4.2. Heterogeneous Effects

This section disentangles some of the mechanisms for the positive effect of better firm communication on innovation by analyzing the heterogeneous effects of intranet adoption on the generation of different types of innovations and for different innovating firms. Results are shown in Tables 5 and 6, respectively.

Heterogeneous Effects by Type of Innovation

I start by analyzing if better firm communication affects the capacity to generate high and low-quality patents in different ways. I will measure quality both in terms of the economic and scientific value of a patent. In column 1 of Table 5 I check the effect of intranet adoption on the number of high economic value patents (patents that are above the median of the economic value of all patents applied in a given year).[32] In column 2 I do the same for the number of patents below the median of economic value. The effect of intranet adoption is highly significant and large in magnitude for the number of high-value patents, whereas it is insignificant for the case of low-value patents.

In column 3 I check the effect of intranet for number of patents generated by a firm that are above the median of forward adjusted citations by application year. These can be considered highly innovative patents. In column 4 I do the same for patents below the median. I find a similar positive effect of intranet adoption for both types of patent qualities. Columns 5 to 8 contain the same specifications as columns 1 to 4 but instrumenting for the variable intranet.[33]

These results are consistent with the company-wide character of innovation. In order to innovate, it is necessary to combine the scientific knowledge of the research team with information about a wide variety of issues such as consumer tastes, market competition and opportunities, firm's production capacity, or firm's business strategy. Whereas the scientific knowledge of the firm is concentrated in the research division, information about the rest of factors is most likely dispersed across different divisions. As a result, an improvement in internal communication between divisions has no effect on the scientific quality of innovations. By contrast, it has a strong impact on the ability to target the most profitable

---

[32] Because in the original sample there are some firms that did not obtain any patent above the median of economic value over the sample period, these firms have to be dropped in the regression.

[33] Appendix Table A7 presents robustness results defining high value patents as those above percentile 66, and low value patens as those below percentile 33. I do the same for high and low innovation patents. Results are qualitatively equivalent to the baseline estimates.

fields for innovation and shape innovation ideas to be in line with consumer tastes, firm capacities, and market opportunities.

Heterogeneous Effects by Firm Type

The effect of intranet adoption fostering innovation should be greater for firms with higher barriers to communication. In this section I show that firms with greater geographical dispersion in their firm's site locations, firms with lower levels of innovation in the past, and firms operating in less competitive markets experience higher increases in innovation as a result of intranet adoption.

Firms with greater geographical dispersion are more likely to have their different sites operating as information silos in the absence of a good system of communication. The lack of interaction between researchers and workers of other departments will handicap innovation capacity. An improvement in communication capacity will bring greater increases in innovation for these firms. To empirically test this hypothesis, I calculate the average distance between each of the firm sites and the firm headquarters (weighting each site by its number of workers). Using this information I construct the variable $Intranet_{jt} * \ High\ Dispersion_j$, where $High\ Dispersion_j$ is a dummy equal to one if firm j is in the top quartile of geographical dispersion. Column 1 of Table 6 reports the estimates on both intranet adoption and the interaction term. The coefficient of $Intranet_{jt}$ remains positive although a little smaller in magnitude. $Intranet_{jt} * \ High\ Dispersion_j$ is also positive and significant at the 5% level. In column 4 I report IV estimations by including a residual control function obtained from a first-stage regression of $Intranet_{jt}$ on $Binet\ Nodes_j * ISPs_t$ and controls. Both terms remain positive and significant and the control function is non-significant (p-value of 0.334).

The next two results show how the effect of intranet is larger for firms and sectors with larger barriers and incentives to innovate. First, I show the effect of intranet is larger for firms that have innovated less in the past. To proxy for the previous level of innovation I use the variable $Patent\ Stock_j$ introduced above and constructed using a perpetual inventory method following Blundell et al. (1999). Then, I create the variable $Intranet_{jt} * \ Low\ Innovation_j$ in which $Low\ Innovation_j$ is a dummy equal to 1 if firm j is in the bottom quartile of $Patent\ Stock_j$ in 1994, just before the diffusion of intranets. Following the same strategy as in the previous case, in column 2 I regress patent counts on $Intranet_{jt}$ and the interacted term. Both terms are positive and significant, confirming the hypothesis that the effect of intranet on patenting is greater for firms with lower initial levels of innovation. I interpret this as evidence that the effect of an improvement in firm communication is larger for firms with stronger barriers to innovation. Actually, the magnitude of the estimated impact of intranet adoption for firms with lower levels of innovation in the past is much larger than for the rest of firms. The sum of the two coefficients gives that the impact for low innovation firms is 0.0234, which implies that a 10 percentage point increase in intranet penetration is associated with a 23.4% increase in patenting, and is not far from the IV estimates (25.6%). Despite how large this estimate may seem, one has to take into account that the levels of innovation for low innovation firms in 1994 are quite low (the median firm is getting 0 patents in 1994, the mean is 0.25 patents, and the maximum 2). Therefore, relatively small increases in the number of patents for these firms would imply large percentage changes in patenting.

In column 3 I show the effect of intranet is also larger for firms operating in less competitive sectors. Less competitive markets have been proved to have lower incentives for innovation (Vives (2008)) and tend to have lower levels of innovation in practice (Geroski et al. (1995); Nickell (1996); and Blundell et al. (1999)).[34] To determine competition level in the sector in which a firm operates I calculate the average of the Herfindahl Index at the two-digit SIC code over the years in my sample. With this information, I construct the variable $Intranet_{jt} * Low\ Competition_j$ in the same way as before. I consider that firm *j* is in a low competition sector if it is in the lowest quartile of the competition variable (which corresponds to being in the highest quartile of the average of the Herfindahl Index over the sample years). In column 3 I report the results of the reduced form regression including both intranet and the interaction term. Column 6 adds the control function. The positive estimates on the effect of both variables confirm that the effect of intranet in stimulating innovation is more intense for firms in sectors with lower incentives for innovation.[35]

4.3. Effects on Productivity

This section examines the effect of intranet adoption on firm productivity. The robust positive impact of intranet adoption on diverse measures of innovation has been documented throughout the article. As a result, and given the extensive literature connecting firm innovation with productivity increases, there is reason to believe that intranet adoption may have affected productivity through a double channel. First, it can do so by improving workers' coordination in production tasks. Second, it can also have an indirect impact on productivity through the increase it generates in firm's innovation.

To empirically study this question, I obtain productivity estimates that I regress on intranet adoption and different measures of firm's innovation stock. First, I use labor productivity measured as the log of sales over number of employees. Second, I employ the method of Levinsohn and Petrin (2003) to calculate firm's TFP. Using this method, I am able to account for the existence of other intermediate inputs such as capital and materials and the

---

[34] Aghion et al. (2005) rationalize this empirical result in a model showing that net profits from innovation are smaller for firms operating in markets with lower competition. They further point out that innovation rents can also decrease in very competitive markets due to the stronger effect of Schumpeterian creative destruction.

[35] Appendix Table A8 shows robustness results in which the heterogeneous effects are time varying. First, in column 1 I create the variable $Intranet_{jt} * Low\ Innovation_{jt}$ in which $Low\ Innovation_{jt}$ is a dummy equal to 1 if firm *j* is in the bottom quartile of $Patent\ Stock_{jt}$ at year *t* (and not in year 1994 as in the baseline specification). Moreover, I include the running variable $Low\ Innovation_{jt}$as a standalone variable (notice this is not necessary in the baseline regression where $Low\ Innovation_j$ is absorbed in the firm fixed-effect). Estimates show that the interacted term is not significant anymore. However, notice that this is likely to happen if firms gradually increase their innovation as a result of intranet adoption and all the effects are not concentrated in the first period after adoption. Imagine, for instance, a low innovation firm adopting intranet in all sites in 1997 and increasing patenting by 50% in 1998 and by an extra 50% in 2000. It is likely the case this firm is not considered anymore as a low innovation firm in year 2000, and, therefore, the estimated impact of the interacted term gets reduced. In column 3 I further include the IV control function. Second, in column 2 I create the variable $Intranet_{jt} * Low\ Competition_{jt}$ in which $Low\ Competition_{jt}$ is a dummy equal to 1 if firm *j* is in the bottom quartile of $Compeition_{jt}$ at year *t* (and not in year 1994 as in the baseline specification). Results are qualitatively equivalent to the baseline estimates. The same applies to specification including the IV control function in column 4. Finally, because firm dispersion does not vary over time (for any firm we use information from establishments covered in all sample years) it is not possible to carry out this robustness analysis in this case.

endogeneity in input levels. One should bear in mind, however, the possible issues pointed by the literature in the use of these structural estimation methods. I obtain firm's TFP estimates using the whole sample of Compustat firms where I estimate production functions at the two-digit SIC level. Appendix D contains a more detailed explanation of the estimation procedure.

In Table 7 column 1, I regress labor productivity on firm intranet adoption and the log of firm patent stock. I further include firm fixed effects, year fixed effects, and industry time-trends (at the two-digit SIC level). The effect of patent stock on labor productivity is positive and significant, whereas there is no significant direct impact of intranet adoption. To confirm the robustness of this finding, in columns 2 and 3 I use alternative measures of innovation stock: the stock of patents weighted by citations and the stock of patents weighted by their estimated economic value. Columns 4 to 6 repeat the same regressions but using as a dependent variable firm's TFP estimates. In all cases, the effect of innovation stock is positive, whereas the estimated direct impact of intranet adoption remains insignificant. Finally, column 7 and 8 show how results do not change by instrumenting for intranet adoption using $Binet\ Nodes_j * ISPs_t$.[36]

In light of these results, one can conclude the non-significant direct contribution of intranet adoption to firm productivity. This is consistent with communication being less critical to ensure the coordination of workers in routine tasks. By contrast, the increase in innovation fostered by intranet adoption has a strong connection with posterior increments in firm productivity.

5. Conclusions

This article studies the role of firm internal communication on innovation and productivity. I provide evidence that problems in the internal communication of large firms can limit their innovation capacity. I also show how the adoption of communication technologies can alleviate this problem, and hence increase the rate of innovation. I find that the effect of technology adoption is larger for firms with higher geographical dispersion and for firms in low competition sectors and low innovation in the past. I also find evidence that the improvement in communication generated by technology adoption is especially effective in increasing the number of high economic value patents. However, it does not affect the scientific quality of innovation. I interpret this as an evidence that a reduction in the costs to transfer knowledge across firm boundaries improves the capacity to identify more profitable innovation ideas and tailor innovations to consumer tastes and market opportunities. Finally, I show how better communication has an indirect impact on productivity through the increase

[36] Appendix Table A9 presents robustness results. Column 1 uses productivity estimates calculated with the method of Olley and Pakes (1996). Column 2 uses productivity estimates calculated with the method of Ackerberg et al. (2015). Column 3 uses productivity estimates calculated with the method of Blundell and Bond (2000). Results are qualitatively equivalent in all the cases. Columns 3 to 6 use productivity estimates calculated with the method of Levinshon and Petrin (2003), but with some variations with respect to the baseline results. Column 4 controls for labour costs when estimating the firm production function by using information from the Compustat variable XLR when available and, otherwise, proxying labour cost by the product of the average total compensation per worker in a given year (using data from the BLS) and the number of workers reported in Compustat. Column 5 controls for advertising expense (XAD variable in Compustat). Finally, column 6 constructs the variable patent stock using the sector-specific depreciation rates suggested in Li and Hall (2006).

in innovation. I find no evidence of a direct effect of better communication on productivity. This is consistent with the view of ICTs as a General Purpose Technology affecting productivity indirectly by inducing other changes in the firm.

Innovation is claimed to be the engine of growth and therefore it is crucial to understand its determinants. However, most of the innovation literature has focused on studying firm and market incentives for innovation, neglecting a systematic empirical analysis of firm innovation capacity. This article helps to fill this gap by showing evidence of (i) the importance of firm organizational capacity for innovation activities; and (ii) how technologies can help to overcome limitations in organizational capacity and, as a result, raise innovation. These results should serve to inform the design of more effective government policies promoting innovation and help firms to conceive better innovation strategies. They also provide a possible explanation for the changes in competition and market dynamics observed over the last decades. Improvements in internal organization coming from technology adoption can be one of the factors behind the rise in large firms' market power and reductions in business dynamism documented in other works. These are issues of crucial importance for market regulation and policy design. I hope this paper helps to shed some light and stimulate future work in these areas.

Bibliography

Ackerberg, Daniel A., Kevin Caves, and Garth Frazer. "Identification properties of recent production function estimators." *Econometrica* 83.6 (2015): 2411-2451.

Aghion, Philippe, et al. "Competition and innovation: An inverted-U relationship." *The quarterly journal of economics* 120.2 (2005): 701-728.

Aghion, Philippe, John Van Reenen, and Luigi Zingales. "Innovation and institutional ownership." *American economic review* 103.1 (2013): 277-304.

Agrawal, Ajay, and Avi Goldfarb. "Restructuring research: Communication costs and the democratization of university innovation." *American Economic Review* 98.4 (2008): 1578-90.

Akicigit, U., and Sina Ates. "Ten facts on declining business dynamism and lessons from endogenous growth theory." *NBER working paper* 25755 (2019).

Akcigit, Ufuk, and S. T. Ates. "What Happened to US Business Dynamism?(No. w25756)." (2019).

Akerman, Anders, Ingvil Gaarder, and Magne Mogstad. "The skill complementarity of broadband internet." *The Quarterly Journal of Economics* 130.4 (2015): 1781-1824.

Autor, David H., Frank Levy, and Richard J. Murnane. "The skill content of recent technological change: An empirical exploration." *The Quarterly journal of economics* 118.4 (2003): 1279-1333.

Becker, Gary S., and Kevin M. Murphy. "The division of labor, coordination costs, and knowledge." *The Quarterly Journal of Economics* 107.4 (1992): 1137-1160.

Bloom, Nicholas, Mirko Draca, and John Van Reenen. "Trade induced technical change? The impact of Chinese imports on innovation, IT and productivity." *The review of economic studies* 83.1 (2016): 87-117.

Bloom, Nicholas, et al. "The distinct effects of information technology and communication technology on firm organization." *Management Science* 60.12 (2014): 2859-2885.

Bloom, Nicholas, Raffaella Sadun, and John Van Reenen. "Americans do IT better: US multinationals and the productivity miracle." *American Economic Review* 102.1 (2012): 167-201.

Blundell, Richard, and Stephen Bond. "GMM estimation with persistent panel data: an application to production functions." *Econometric reviews* 19.3 (2000): 321-340.

Blundell, Richard, Rachel Griffith, and John Van Reenen. "Market share, market value and innovation in a panel of British manufacturing firms." *The review of economic studies* 66.3 (1999): 529-554.

Blundell, Richard, Rachel Griffith, and Frank Windmeijer. "Individual effects and dynamics in count data models." *Journal of econometrics* 108.1 (2002): 113-131.

Boersma, Kees, and Sytze Kingma. "Intranet and organizational learning." *Knowledge management: Concepts, methodologies, tools, and applications*. IGI Global, 2008. 1065-1072.

Bolton, Patrick, and Mathias Dewatripont. "The firm as a communication network." *The Quarterly Journal of Economics* 109.4 (1994): 809-839.

Bresnahan, Timothy F., Erik Brynjolfsson, and Lorin M. Hitt. "Information technology, workplace organization, and the demand for skilled labor: Firm-level evidence." *The quarterly journal of economics* 117.1 (2002): 339-376.

Bresnahan, Timothy, et al. "Technical progress and co-invention in computing and in the uses of computers." *Brookings Papers on Economic Activity. Microeconomics* 1996 (1996): 1-83.

Brynjolfsson, Erik, and Lorin M. Hitt. "Computing productivity: Firm-level evidence." *Review of economics and statistics* 85.4 (2003): 793-808.

Cameron, A. Colin, and Pravin K. Trivedi. *Regression analysis of count data*. Vol. 53. Cambridge university press, 2013.

Christian, Paul, and Christopher B. Barrett. *Revisiting the effect of food aid on conflict: A methodological caution*. The World Bank, 2017.

Crépon, Bruno, Emmanuel Duguet, and Jacques Mairessec. "Research, Innovation And Productivi [Ty: An Econometric Analysis At The Firm Level." *Economics of Innovation and new Technology* 7.2 (1998): 115-158.

De Loecker, Jan, Jan Eeckhout, and Gabriel Unger. "The rise of market power and the macroeconomic implications." *The Quarterly Journal of Economics* 135.2 (2020): 561-644.

Doraszelski, Ulrich, and Jordi Jaumandreu. "R&D and productivity: Estimating endogenous productivity." *Review of Economic Studies* 80.4 (2013): 1338-1383.

Draca, Mirko, Raffaella Sadun, and John Van Reenen. "Productivity and ICT: A Review of the Evidence." (2006).

Forman, Chris. "The corporate digital divide: Determinants of Internet adoption." *Management Science* 51.4 (2005): 641-654.

Forman, Chris, Avi Goldfarb, and Shane Greenstein. *Digital dispersion: An industrial and geographic census of commerical internet use*. No. w9287. National Bureau of Economic Research, 2002.

Forman, Chris, Avi Goldfarb, and Shane Greenstein. "How did location affect adoption of the commercial Internet? Global village vs. urban leadership." *Journal of urban Economics* 58.3 (2005): 389-420.

Forman, Chris, Avi Goldfarb, and Shane Greenstein. "Understanding the inputs into innovation: Do cities substitute for internal firm resources?." *Journal of Economics & Management Strategy* 17.2 (2008): 295-316.

Forman, Chris, Avi Goldfarb, and Shane Greenstein. "The Internet and local wages: A puzzle." *American Economic Review* 102.1 (2012): 556-75.

Forman, Chris, and Nicolas van Zeebroeck. "From wires to partners: How the Internet has fostered R&D collaborations within firms." *Management science* 58.8 (2012): 1549-1568.

Garicano, Luis. "Hierarchies and the Organization of Knowledge in Production." *Journal of political economy* 108.5 (2000): 874-904.

Geroski, Paul A. "Market structure, corporate performance, and innovative activity." *OUP Catalogue* (1995).

Griliches, Zvi. "Hybrid corn: An exploration in the economics of technological change." *Econometrica, Journal of the Econometric Society* (1957): 501-522.

Griliches, Zvi, Bronwyn Hall, and Jerry Hausman. "Patents and R&D: Is there a lag." *International Economic Review* 27.2 (1986): 265-283.

Hall, Bronwyn, and Fumio Hayashi. *Research and Development as an Investment*. No. w2973. National Bureau of Economic Research, 1989.

Hall, Bronwyn H., Adam Jaffe, and Manuel Trajtenberg. "Market value and patent citations." *RAND Journal of economics* (2005): 16-38.

Hall, Bronwyn H., and Rosemarie Ham Ziedonis. "The patent paradox revisited: an empirical study of patenting in the US semiconductor industry, 1979-1995." *rand Journal of Economics* (2001): 101-128.

Hausman, Jerry A., Bronwyn H. Hall, and Zvi Griliches. *Econometric models for count data with an application to the patents-R&D relationship*. No. t0017. national bureau of economic research, 1984.

İmrohoroğlu, Ayşe, and Şelale Tüzel. "Firm-level productivity, risk, and return." *Management Science* 60.8 (2014): 2073-2090.

Jensen, Morten Berg, et al. "Forms of knowledge and modes of innovation." *The learning economy and the economics of hope* 155 (2007).

Klette, Tor Jakob. "R&D, Scope Economies and Company Structure. A" Not-so-Fixed Effect" Model of Plant Performance." (1994).

Kogan, Leonid, et al. "Technological innovation, resource allocation, and growth." *The Quarterly Journal of Economics* 132.2 (2017): 665-712.

Lanjouw, Jean O., and Josh Lerner. *Preliminary injunctive relief: theory and evidence from patent litigation*. No. w5689. National bureau of economic research, 1996.

Lerner, Josh. "Patenting in the Shadow of Competitors." *The Journal of Law and Economics* 38.2 (1995): 463-495.

Lerner, Josh, and Amit Seru. *The use and misuse of patent data: Issues for corporate finance and beyond*. No. w24053. National Bureau of Economic Research, 2017.

Levinsohn, James, and Amil Petrin. "Estimating production functions using inputs to control for unobservables." *The review of economic studies* 70.2 (2003): 317-341.

Li, Wendy CY, and Bronwyn H. Hall. "Depreciation of business R&D capital." *Review of Income and Wealth* 66.1 (2020): 161-180.

Mansfield, Edwin, and Samuel Wagner. "Organizational and strategic factors associated with probabilities of success in industrial R & D." *the Journal of Business* 48.2 (1975): 179-198.

March, James G., and Herbert A. Simon. "Organizations John Wiley & Sons." *New York* (1958).

McElheran, Kristina. "Delegation in multi-establishment firms: Evidence from it purchasing." *Journal of Economics & Management Strategy* 23.2 (2014): 225-258.

Nickell, Stephen J. "Competition and corporate performance." *Journal of political economy* 104.4 (1996): 724-746.

Olley, G. Steven, and Ariel Pakes. *The dynamics of productivity in the telecommunications equipment industry*. No. w3977. National Bureau of Economic Research, 1992.

Pakes, Ariel. "Patents as Options: Some Estimates of the Value of Holding European Patent." *Econometrica* 54.4 (1986): 755-784.

Sambamurthy, Vallabhajosyula, and Robert W. Zmud. "Arrangements for information technology governance: A theory of multiple contingencies." *MIS quarterly* (1999): 261-290.

Scott, Judy E. "Organizational knowledge and the intranet." *Decision Support Systems* 23.1 (1998): 3-17.

Stranger, Greg, and Shane Greenstein. "Pricing at the On-Ramp to the Internet: Price Indexes for ISPs during the 1990s." *Hard-to-Measure Goods and Services: Essays in Honor of Zvi Griliches*. University of Chicago Press, 2007. 197-233.

Syverson, Chad. "What determines productivity?." *Journal of Economic literature* 49.2 (2011): 326-65.

Vives, Xavier. "Innovation and competitive pressure." *The Journal of Industrial Economics* 56.3 (2008): 419-469.

Wooldridge, Jeffrey M. "Control function methods in applied econometrics." *Journal of Human Resources* 50.2 (2015): 420-445.

Appendix A: Intranet Technology

Intranets are formally defined as the application of Internet technology for a prescribed community of users, typically members of an organization. Internet technology standards and protocols are employed but access is restricted by means of passwords, encryption, and firewalls. Internet and intranet technologies were developed for academic and governmental use in the early 1960s. However, the rise of the commercial Internet and intranets based on TCP/IP did not take place until the mid-90's. Because of its non-commercial origins, many intranet technologies were already quite mature by this time and could be applied immediately to organizational needs. At the time of its commercial diffusion, intranets already included a wide variety of applications and functionalities such as videoconferencing, collaboration tools, access to repositories of information located in other parts of the firm, and applications to search for subject matter experts within the organization. As a result, they were fastly adopted by firms (e.g., Forman (2005); Forman et al. (2002)). By contrast, more sophisticated technologies like the ones necessary to conduct Internet transactions required more time to be implemented.

Intranets are built and maintained for information storage and retrieval but mainly for enhancing information flow and communications within an organization. This provided firms with opportunities to create new knowledge and innovate. Numerous case studies highlighting the importance of intranets for fostering innovation can be found in the Business and Information Systems literature (see Scott (1998); Boersma and Kingma (2006)). I bring here three illustrative examples.

> Scientists at the Met Office (the UK's national weather service) *"use an intranet as a discussion forum for ongoing research projects. Staff are able to access each other's webpages to catch up with their colleges research and there are news groups for individual Departments and for special scientific interest groups"*.
>
> A manager in the engineering division of Jaguar (a car manufacturer) claims *"The intranet will allow our engineers to work in what seems to them to be the obvious and natural way. The logic is to ensure that all information is available to the people who need it and that they can access it easily"*.
>
> Olivetti Group (an IT manufacturer) reports to *"use a virtual laboratory to link their main sites and labs worldwide so that researchers access the largest possible amount of current information(…). In an R\&D environment the free exchange of information and ideas is a powerful catalyst of innovation"*. Furthermore, *"if a problem has already been solved by one employee we can find about it intermediately and avoid duplicating efforts. Before the web there was no central repository of information so researchers often spent time looking for information that was already available in-house"*.

Appendix B: Matching of Datasets

In order to conduct the empirical analysis, I needed to match firm information in three different datasets: the patent dataset, Compustat and HH technology dataset. Matching patent assignees to Compustat firms was not specially difficult. Kogan et al. (2017) already contains a link between patent assignees and CRSP firms. This type of work is specially complicated due to inconsistencies in how firm names are listed in patent records. They can contain spelling variations, typographical errors or the name of a subsidiary firm (for example, IBM can be found with more than 100 different names). Then, using a bridge dataset between CRSP and Compustat I am able to connect patent assignees to Compustat firms.

The second step consisted on matching HH firms to Compustat firms. Because of the absence of a common firm identifier in both datasets I had to resort to string matching algorithms. In my matching strategy, I prioritized minimizing the number of false positives over the number of false negatives. As a result, I did a good job preventing wrong matches at the cost of probably losing a number of observations. I started by cleaning firm names getting rid of words like “Corp”, “Incorporated”, “L.S.”, etc. Then, I matched firms in both datasets with the exact same name. The number of false negative cases using only this conservative matching mechanism is too big. For instance, I did not find a match in cases as obvious as ``International Business Machines" and ``Internat Business Machines". To solve this type of problematic cases, I resorted to string matching algorithms. I used different algorithms to reduce as much as possible the number of false positives. Furthermore, I dropped provisional matches that did not coincide in at least one the following pieces of information: firm’s phone number, firm’s web page or firm’s stock market ticker symbol. Finally, I conducted a manual inspection checking in detail any suspicious candidate match. These were a handful of double match cases between firm subsidiaries and parent firms. I did some online research to discover which was the correct match and if this was not clear I dropped the observation.

Appendix C: Variable Definitions

In this section I describe the main variables used in the empirical analysis coming from Harte Hanks Technology Dataset, the patent dataset of Kogan et al. (2017) and Compustat North American Fundamental Annuals. More details about the definition of some of these variables can be found in Imrohoroglu and Tüzel (2014)

- ***Employees***. The number of employees was taken directly from Compustat (Compustat variable EMP). No adjustments were made to this figure.
- ***Investment***. Value of current investment in capital goods calculated as the difference between capital expenditures (Compustat variable CAPX) and funds received for the sale of capital (Compustat variable SPPE) and deflated to prices of 2009.
- ***Capital***. Capital stock is calculated using a perpetual inventory method where the value of capital stock in year t is equal to undepreciated capital in year t-1 plus investment $K_t = K_{t-1}(1-\delta) + I_t$.

For K_0 I use the total book value of capital reported by Compustat (Compustat variable PPENT) deflated to prices of 2009. When available, I use 1983 as the first year to start constructing the series. Otherwise, I use the first available year in Compustat. I assume a yearly depreciation of 0.15.

- ***Output***. Total net sales as reported by Compustat (Compustat variable SALE) deflated to prices of 2009.
- ***Materials***. Materials was calculated by subtracting labor expenses form total expense and deflating this to prices of 2009. I calculate labor expenses as the product of the average total compensation per worker in a given year (using data from the BLS) and the number of workers reported in Compustat. Total expenses was computed as the difference between Operating Income Before Depreciation (Compustat variable OIBDP) and sales (Compustat variable SALE).
- ***R&D investment***. Research and Development expenses as reported in Compustat (variable XRD) deflated to prices of 2009. Investment is measured in thousands of dollars. When I have to take the natural logarithm of R&D investment, as this can be zero for some years, I take the natural logarithm of R&D plus one.
- ***R&D stock***. Constructed applying a perpetual inventory method to R&D investment using a yearly depreciation of 0.15.
- ***Intranet***. Percentage share of firm sites reporting to have adopted intranet in a given year.
- ***Weighted Intranet***. Percentage share of firm sites reporting to have adopted intranet in a given year weighted by their size (number of employees).
- ***Internet***. Percentage share of firm sites reporting to have adopted internet in a given year.
- ***Patents***. Number of granted patents applied by a firm in a given year.
- ***Patent stock***. Constructed using a perpetual inventory method with a yearly depreciation of 0.15. When I have to take the natural logarithm of this variable, as its value can be zero, I take the natural logarithm of Patent stock plus one.

Appendix D: TFP Estimation

I follow the method of Levinsohn and Petrin (2003) to estimate firm-level TFP. This approach tries to address the problem that inputs are choice variables made by the firm to maximize profits, and hence will often depend on unobservable productivity shocks. LP uses assumptions about the information set of the firm at the time of making input decisions.

To see this, consider the following Cobb-Douglass production function

$$y_{it} = \beta_0 + \beta_l l_{it} + \beta_m m_{it} + \beta_k k_{it} + u_{it}$$

Where y_{it} is the logarithm of output of firm *i* at time *t*, and correspondingly, l_{it}, m_{it}, and k_{it} are the firm's (log of) labor, materials and capital inputs. Wlog, assume that $u_{it} = \omega_{it} + \varepsilon_{it}$ where ω_{it} and ε_{it} are unobserved by the econometrician, whereas the firm can observe ω_{it} at time t. The term ε_{it} could be capturing unpredictable shocks, whereas ω_{it} can be interpreted as firm productivity. If ω_{it} is known to the firm, the optimal labor and materials input choice will be a function of ω_{it}, and simple OLS estimation will suffer from simultaneity bias.

LP uses firm's consumption of intermediates inputs (materials) to invert the productivity shock ω_{it} and control for it in the estimation. Assuming the productivity shock follows a markov process, i.e.

$$P(\omega_{it+1}|I_{it}) \ = \ P(\omega_{it+1}|\omega_{it})$$

where I_{it} is firm *i*'s information set at *t* (which includes current and pasts ω_{it}'s). Notice this is not just an assumption on the stochastic process governing ω_{it} but an assumption on firms information set at various points in time. The next important assumption is that labor and materials are fully flexible inputs decided by the firm at *t* once ω_{it} is observed, whereas the level of capital for period *t* has to be made at *t-1* when uncertainty about ω_{it} is not resolved for the firm yet. As a result, the level of intermediate inputs at *t* is a function of k_{it} and ω_{it},

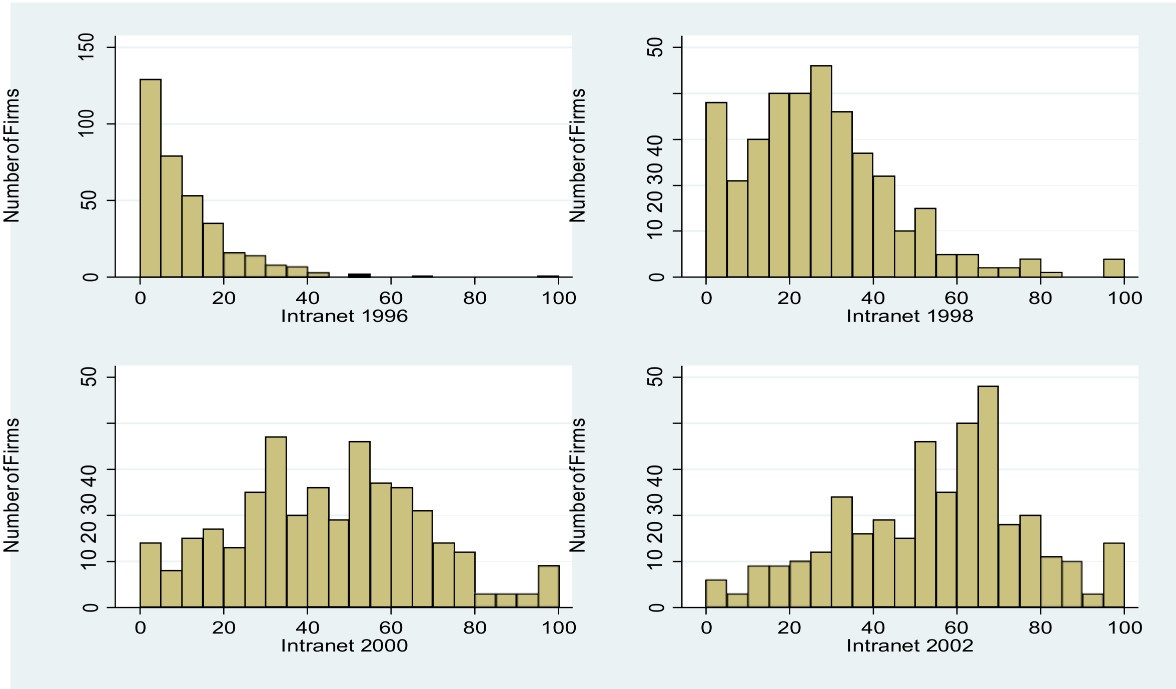
$$m_{it} = h(\omega_{it}\,, k_{it})$$

Under some conditions this expression can be inverted to obtain

$$\omega_{it} = h^{-1}(m_{it}\,, k_{it})$$

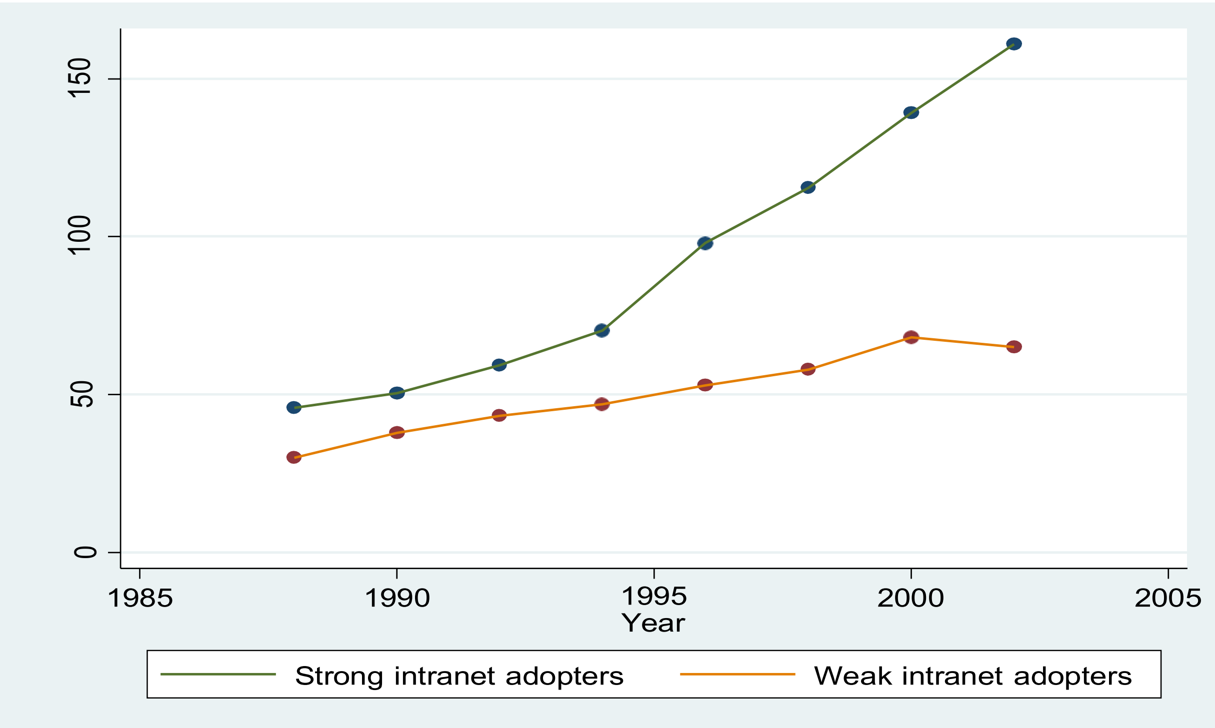
This implies the productivity shock can be written as a function of variables that are observed by the econometric and, therefore solving the endogeneity problem.

Figure 1: Levels of Intranet Penetration



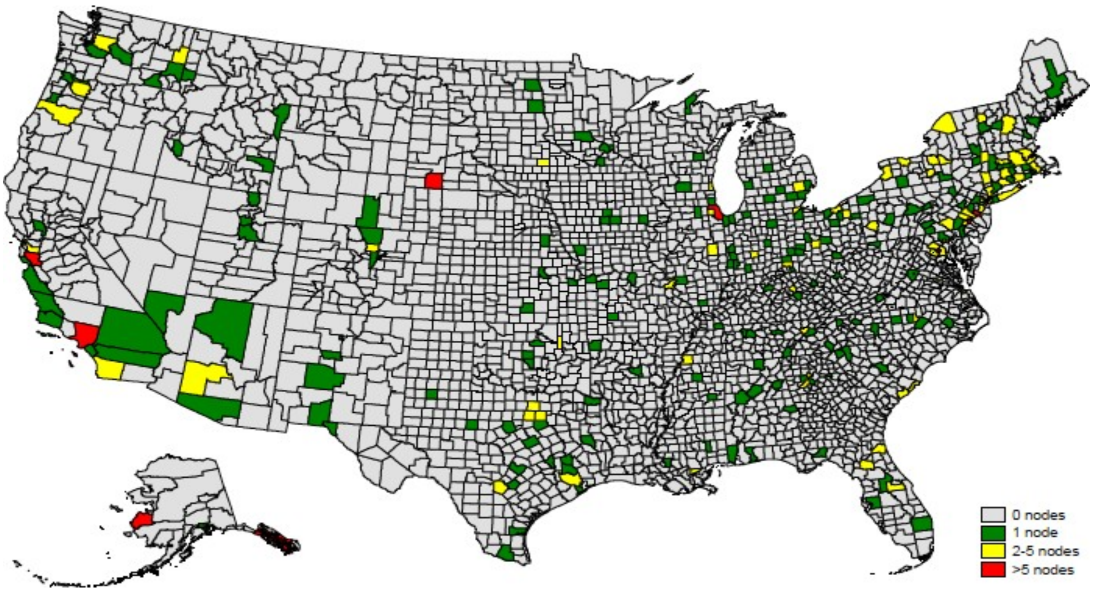
Notes: These graphs show the number of firms in each level of intranet adoption for the years 1996, 1998, 2000, and 2002. Intranet adoption is measured as the percentage share of firm sites having adopted intranet. The horizontal axis shows the different levels of intranet adoption and the vertical axis the number of firms in the corresponding bin of intranet adoption. Source: CiDB Harte Hanks.

Figure 2: Trends in Patenting for Different Levels of Intranet Adoption



Notes: This graph shows the different trends in patenting over time for those firms adopting intranet more intensively (strong adopters) and those doing it less intensively (weak adopters). Strong adopters are those firms that are above the median of intranet adoption in all years after 1994. Intranet adoption starts in 1994 (marked in red). The vertical axis shows the average number of patents per year for firms in each group. Source: CiDB Harte Hanks.

Figure 3 : Distribution of Bitnet nodes across US counties



Notes: This figure shows the number of nodes of the Bitnet network in each county of the US. Source: Agrawal and Goldfarb (2008)

Table 1: Descriptive Statistics

	Observ.	Mean	Overall Std. Dev.	Within-firm Std. Dev.	Median	Min	Max
Employees	2784	35777	83024	28202	11500	50	1400000
Capital (million USD)	2784	2942	7439	2258	710	1	10958
Sales (million USD)	2784	9485	21967	7194	2967	4	28847
R&D (million USD)	2784	228	761	297	15	0	957
Patents	2784	62	226	111	4	0	439

Notes: Data on total firm number of employees, capital, sales, and R&D investment are obtained from Compustat. Capital is calculated using a permanent inventory method. Data on number of patents is obtained from Kogan et al (2017) dataset. The sample comprises 348 firms for all even years between 1988 and 2002.

Table 2: Baseline Results

	(1)	(2)	(3)	(4)	(5)
	Patents	Patents	Patents	Patents	R&D
Intranet	0.0162***	0.00831***	0.00565**	0.00517**	0.00342
	(0.000145)	(0.00304)	(0.00234)	(0.00229)	(0.00439)
Log R&D$_t$			0.0981**	0.101**	
			(0.0461)	(0.0431)	
Log R&D$_{t-1}$			0,05940	-0,00001	
			(0.0612)	(0.0482)	
Log R&D$_{t-2}$				0,00488	
				(0.0539)	
Log R&D$_{t-3}$				-0,02660	
				(0.054)	
Log R&D$_{t-4}$				0.109**	
				(0.0447)	
Employees			0.00317***	0.00310***	0.00512
			(0.00119)	(0.00114)	(0.00380)
Observations	2784	2784	2784	2784	2784
Number of fir	348	348	348	348	348
Year FE	Yes	Yes	Yes	Yes	Yes
Firm FE	No	Yes	Yes	Yes	Yes

Notes: *** denotes significance at 1%, ** at 5%, and * at 10%. The dependent variable Patents is a count of number of granted patents applied by a firm in a given year. Intranet adoption is measured as the percentage share of firm sites having adopted intranet. Coefficients are from Poisson regressions with standard errors clustered by firm (in parentheses).

Table 3: Robustness Analysis

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Citation weight	Dollar weight	Patents	Patents	Patents	Patents	Patents	Patents
Intranet	0.00663**	0.00728**		0.00701**	0.00657***	0.00521**	0.00477**	0.00566**
	(0.00262)	(0.0034)		(0.00281)	(0.00234)	(0.00228)	(0.00206)	(0.0023)
Weighted intranet			0.00296**					
			(0.00124)					
Dummy intranet > 0				-0.307				
				(0.189)				
Internet for Research					-0.00354			
					(0.00332)			
Log R&D$_t$	0.114*	0,0878	0.102**	0.0961**	0.0962**		0.116***	0.101**
	(0.0655)	(0.0839)	(0.0475)	(0.0454)	(0.0465)		(0.0438)	(0.0473)
Log R&D$_{t-1}$	0,0522	0,00567	0,0615	0,0576	0,0592		0,0304	0,0648
	(0.0593)	(0.0366)	(0.0613)	(0.0609)	(0.0603)		(0.0461)	(0.0551)
Log R&D Stock						0.111**		
						(0.0533)		
Employees	0.00248*	0.00332**	0.00320***	0.00308**	0.00319***	0.00308***	0.00298***	0.00324***
	(0.00135)	(0.00163)	(0.00119)	(0.0012)	(0.00119)	(0.00113)	(0.00114)	(0.0012)
Observations	2744	2784	2784	2784	2784	2784	2784	2784
Number of firms	343	348	348	348	348	348	348	348
Firm and Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Trends for innovation groups	No	No	No	No	No	No	Yes	No
Trends for digital sectors	No	No	No	No	No	No	No	Yes

Notes: *** denotes significance at 1%, ** at 5%, and * at 10%.Patents weighted by citations is a count of number of granted patents applied by a firm in a given year weighted by an adjusted measure of number of forward citations received by each patent. Patents weighted by dollars is a count of number of granted patents applied by a firm in a given year weighted by an adjusted measure of the economic value of each patent estimated by Kogan et al (2017). Intranet adoption is measured as the ratio of firm sites having adopted intranet. Weighted intranet is measured as the ratio of firm sites having adopted intranet weighting each site by its number of employees. Dummy Intranet is a dummy taking the value of one if at least one firm site has adopted intranet. Internet adoption is measured as the ratio of firm sites having adopted Internet. R&D stock is constructed using a perpetual inventory method. Coefficients are from Poisson regressions with standard errors clustered by firm (in parentheses).

Table 4: IV Estimation

	(1) Intranet	(2) Patents	(3) Patents	(4) Patents	(5) Citation w.	(6) Dollar w.	(7) R&D
Intranet		0.0256**	0.0239**	0.0272**	0.0265**	0.0262**	0.00454**
		0,0108	0,0107	(0.0109)	0,0118	0,012	(0.00200)
Bitnet nodes$_j$ * ISPs$_t$	0.000792***						
	(0.000128)						
Log R&D$_t$	0,0564	0.0875**	0.0915**	0.0866**	0.102*	0,0699	
	(0.316)	(0.0426)	(0.0403)	(0.0424)	(0.0599)	(0.0743)	
Log R&D$_{t-1}$	0,00448	0,0571	0,0104	0,0569	0,0464	0,00629	
	(0.29)	(0.0599)	(0.0468)	(0.0599)	(0.0565)	(0.0356)	
Log R&D$_{t-2}$			-0,000238				
			(0.0521)				
Log R&D$_{t-3}$			-0,03				
			(0.0538)				
Log R&D$_{t-4}$			0.0965**				
			(0.0423)				
Employees	-0.0259**	0.00394***	0.00383***	0.00400***	0.00325**	0.00412**	0.00628*
	(0.0121)	(0.00131)	(0.00126)	(0.00132)	(0.00146)	(0.00166)	(0.00352)
Observations	2784	2784	2784	2784	2744	2784	2784
Number of firms	348	348	348	348	343	348	348
Firm and Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Interacted IV	Yes	Yes	Yes	No	Yes	Yes	Yes
p-value CF		0,138	0,164	0,126	0,179	0,0043	
F-stat 1st stage	38,52	38,52	37,44	11	38,52	38,52	37,37

Notes: *** denotes significance at 1%, ** at 5%, and * at 10%. Patents is a count of number of granted patents applied by a firm in a given year. Patents weighted by citations is a count of number of granted patents weighted by an adjusted measure of number of forward citations received by each patent. Patents weighted by dollars is a count of number of granted patents weighted by an adjusted measure of the economic value of each patent estimated by Kogan et al (2017). Intranet adoption is measured as the share of firm sites having adopted intranet. Coefficients in column 1 come from an OLS regression with standard errors clustered by firm (in parentheses). In columns 2 to 6, coefficients are from Poisson regressions with standard errors clustered by firm (in parentheses). They include a second-order polynomial series expansion of the estimated error term in the first stage. Coefficients in column 6 come from an OLS regression in which Intranet is instrumented and standard errors are clustered by firm.

Table 5: Heterogeneous Effects I

	Poisson	Poisson	Poisson	Poisson	IV	IV	IV	IV
	(1)	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**	**(7)**	**(8)**
	High Value	Low Value	High Citations	Low Citations	High Value	Low Value	High Citations	Low Citations
Intranet	0.0115***	0,00243	0.00581**	0.00528**	0.0297**	0,025	0.0244**	0.0291**
	(0.00314)	(0.00227)	(0.00256)	(0.00246)	(0.0142)	(0.022)	(0.0104)	(0.0124)
Log R&D$_t$	0.132*	0,127	0.0990**	0.0964*	0.129*	0,0692	0.0884**	0.0849*
	(0.0757)	(0.14)	(0.0487)	(0.052)	(0.0744)	(0.101)	(0.0449)	(0.0488)
Log R&D$_{t-1}$	0,0426	0,181	0,0458	0,0845	0,0356	0,186	0,0428	0,0853
	(0.0678)	(0.158)	(0.0577)	(0.0735)	(0.0661)	(0.146)	(0.0562)	(0.0728)
Employees	0,00203	0,000925	0.00283**	0.00380***	0,00257	0,00283	0.00355***	0.00473***
	(0.00225)	(0.00371)	(0.00121)	(0.00111)	(0.00222)	(0.0043)	(0.00133)	(0.00127)
Observations	2288	1872	2576	2576	2288	1872	2576	2576
Number of firms	286	234	322	322	286	234	322	322
Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Firm FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
p-value CF					0,26	0,0842	0,193	0,0554
F-stat 1st stage					38,52	38,52	38,52	38,52

Notes: *** denotes significance at 1%, ** at 5%, and * at 10%. Patents High Value is a count of number of patents above the median of the economic value of patents generated in a given year. Patent Low Value is the same for patents below the median. Patents High Citations is a count of number of patents above the median of forward citations per year. Patent Low Citations is the same for patents below the median. Intranet adoption is measured as the ratio of firm sites having adopted intranet. Coefficients are from Poisson regressions with standard errors clustered by firm (in parentheses). In the control function regressions (cols 5-8) I control for a second-order polynomial expansion of the residual estimated in a first-stage regression of Intranet on Bitnet nodes$_j$ * ISPs$_t$ and controls. The p-value on the significance of the control function is reported in the table.

Table 6: Heterogeneous Effects II

	Poisson	Poisson	Poisson	IV	IV	IV
	(1) Patents	**(2)** Patents	**(3)** Patents	**(4)** Patents	**(5)** Patents	**(6)** Patents
Intranet	0.00391**	0.00564**	0.00499**	0.0217*	0.0257**	0.0228**
	(0.00178)	(0.00234)	(0.00198)	(0.0127)	(0.0108)	(0.0114)
Intranet * High Dispersion	0.00572**			0.00561**		
	(0.00258)			(0.00269)		
Intranet * Low Innovation		0.0178***			0.0187***	
		(0.00611)			(0.00598)	
Intranet * Low Competition			0.00666**			0.00633**
			(0.00263)			(0.00288)
Log R&D$_t$	0.101**	0.0983**	0.106**	0.0908**	0.0877**	0.0955**
	(0.0483)	(0.0462)	(0.05)	(0.0445)	(0.0427)	(0.0462)
Log R&D$_{t-1}$	0,0732	0,0592	0,0656	0,0716	0,0568	0,0637
	(0.0661)	(0.0612)	(0.0656)	(0.065)	(0.0598)	(0.0645)
Employees	0.00314**	0.00317***	0.00295**	0.00383***	0.00394***	0.00365**
	(0.0013)	(0.00119)	(0.00137)	(0.00142)	(0.00131)	(0.00149)
Observations	2784	2784	2784	2784	2784	2784
Number of firms	348	348	348	348	348	348
Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Firm FE	Yes	Yes	Yes	Yes	Yes	Yes
p-value CF				0,334	0,137	0,274
F-stat 1st stage				38,52	38,52	38,52

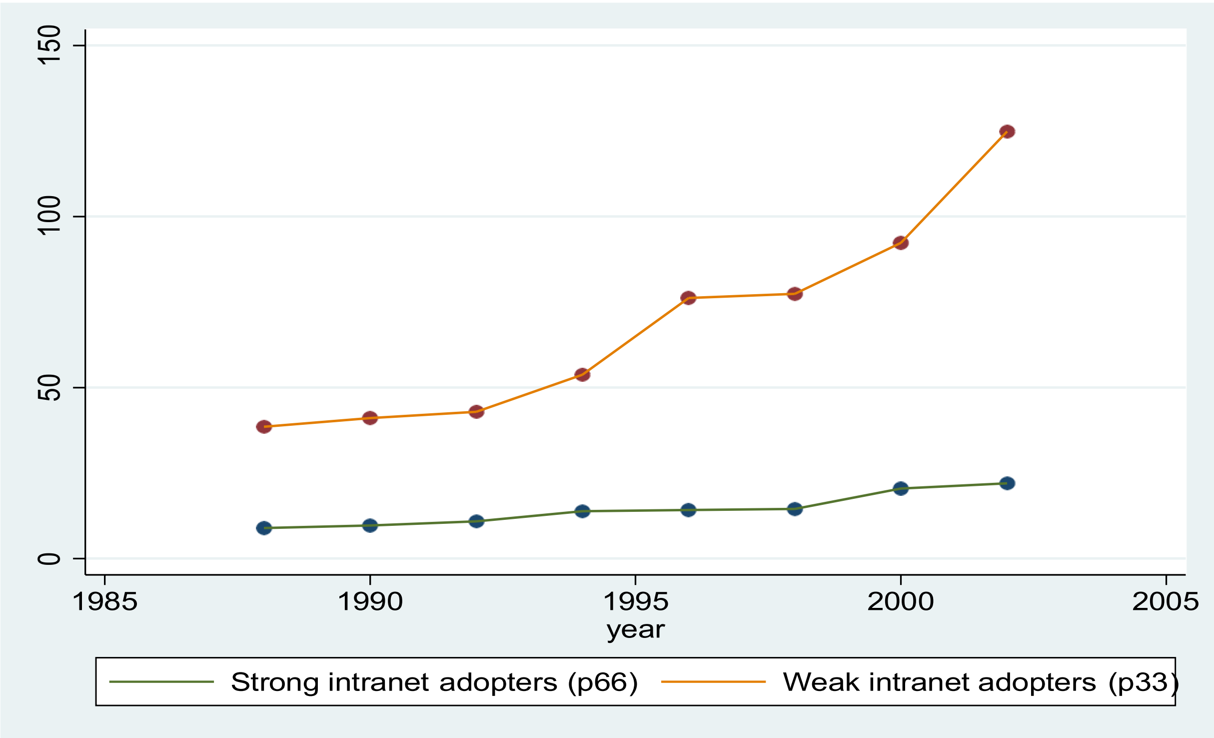
Notes: *** denotes significance at 1%, ** at 5%, and * at 10%.The dependent variable Patents is a count of number of granted patents applied by a firm in a given year. Intranet adoption is measured as the share of firm sites having adopted intranet. High Distance is a dummy equal to one if a firm is in the top quartile of geographical firm dispersion. Low Innovation is a dummy equal to 1 if a firm is in the bottom quartile of patent stock in 1994, just before the diffusion of intranets. Low Competition is dummy equal to 1 if a firm is in the bottom quartile of sector competition. Coefficients are from Poisson regressions with standard errors clustered by firm (in parentheses). In the control function regressions I control for a second-order polynomial expansion of the residual estimated in a first-stage regression of Intranet on Bitnet nodes$_j$ * ISPs$_t$ and controls. The p-value on the significance of the control function is reported inthe table.

Table 7: Effects on Productivity

	OLS	OLS	OLS	OLS	OLS	OLS	IV	IV
	(1)	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**	**(7)**	**(8)**
	Labor Prod	Labor Prod	Labor Prod	TFP	TFP	TFP	Labor Prod	TFP
Intranet	0,000975	0,000941	0,000901	0,000828	0,000786	0,000732	0,000826	-0.000418
	(0.000619)	(0.000614)	(0.000621)	(0.00059)	(0.000586)	(0.000586)	(0.0018)	(0.00236)
Patent Stock	0.0325**			0.0413***			0.0327***	0.0432***
	(0.0163)			(0.0146)			(0.00898)	(0.0153)
Patent Stock Citations		0.0339**			0.0431***			
		(0.0167)			(0.014)			
Patent Stock Dollars			0.0491***			0.0629***		
			(0.0122)			(0.011)		
Observations	2784	2784	2784	2784	2784	2784	2784	2784
Number of firms	348	348	348	348	348	348	348	348
Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Firm FE	Yes	Yes	Yes	Yes	Yes	Yes	No	No
Industry-specific trends	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
F-stat 1st stage							34,6	34,6

Notes: *** denotes significance at 1%, ** at 5%, and * at 10%. The dependent variable in column 1 is labor productivity measured as the log of the ratio of sales over number of employees. Columns 2 to 5 use as a dependent variable an estimate of firm TFP obtained from production function estimation by the methods of Levinshon and Petrin (2003) at the two-digit SIC industry level. Intranet adoption is measured as the share of firm sites having adopted intranet. Patent Stock is a measure of the stock of patents generated by the firm applying a perpetual inventory method. Patent Stock Citations is a measure of the stock of patents applying a perpetual inventory method and weighting each patent by the number of forward citations received. Patent Stock Dollars weights each patent by an estimate of a patent economic value obtained from Kogan et al (2007). Controls for firm and year fixed effects and time-industry trends for two-digit SIC industries are included too. Regression in column 5 instruments Intranet using Bitnet $nodes_j$ * $ISPs_t$. Coefficients are from an OLS regressions with standard errors clustered by firm (in parentheses).

Figure A1: Trends in Patenting for Different Levels of Intranet Adoption (Robustness)



Notes: This graph shows the different trends in patenting over time for those firms adopting intranet more intensively (strong adopters) and those doing it less intensively (weak adopters). Strong adopters are those firms that are above the 66 percentile of intranet adoption in all years after 1994. Weak adopters are those firms that are below the 33 percentile of intranet adoption in all years after 1994. Intranet adoption starts in 1994 (marked in red). The vertical axis shows the average number of patents per year for firms in each group. Source: CiDB Harte Hanks.

Figure A2: Treatment Estimates across Sectors at 1-digit SIC

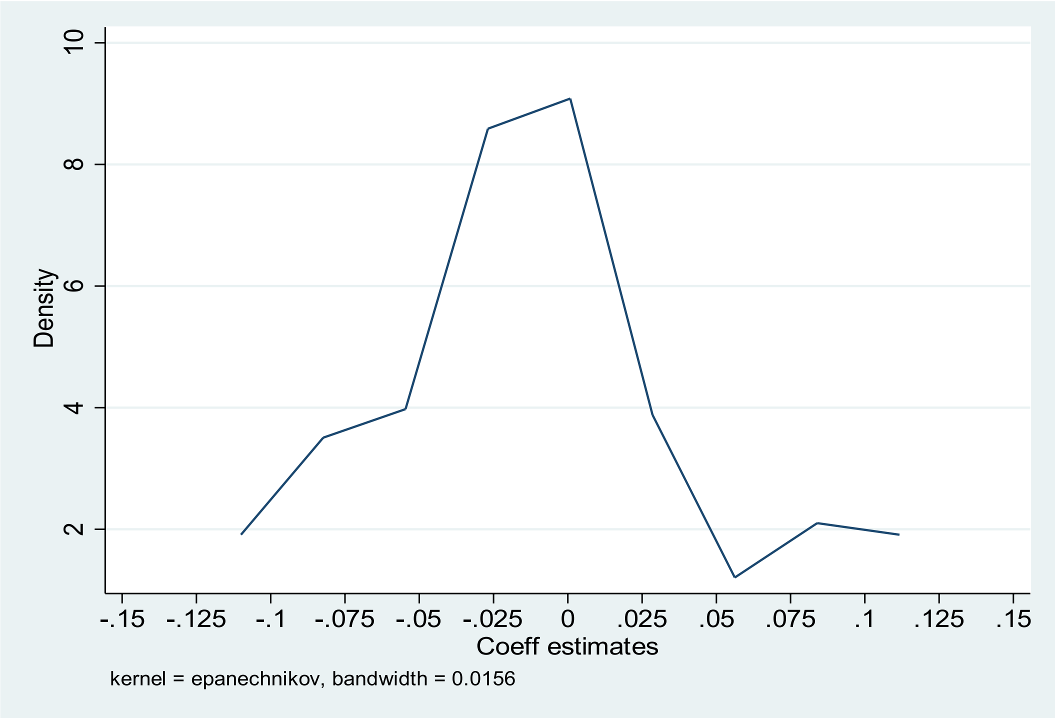


Figure A3: Treatment Estimates across Sectors at 2-digit SIC

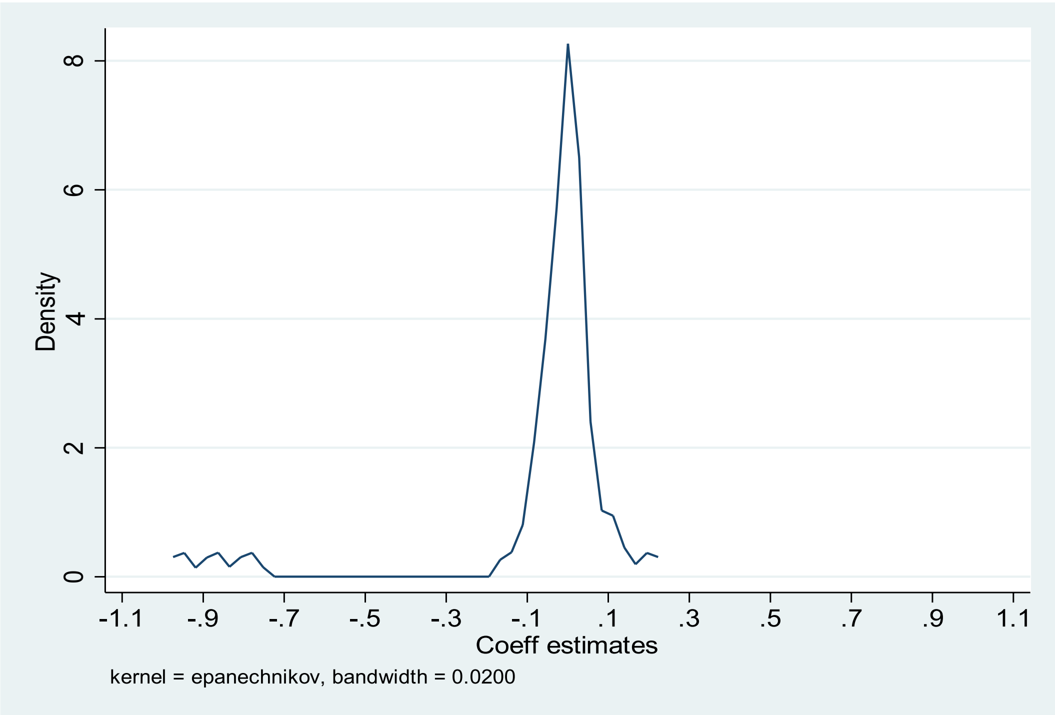


Table A1: Baseline Results with Unbalanced Sample

	(1)	(2)	(3)	(4)
	Patents	Patents	Patents	Patents
Intranet	0.0195***	0.0160***	0.00772***	0.00511**
	(0.000124)	(0.000141)	(0.00298)	(0.00227)
Log R&D$_t$				0.0824*
				(0.0459)
Log R&D$_{t-1}$				0.0800
				(0.0536)
Employees				0.00333***
				(0.00121)
Observations	5687	3030	3030	3030
Number of firms	764	387	387	387
Drop firms with 0 patents	No	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes
Firm FE	No	No	Yes	Yes

Notes: *** denotes significance at 1%, ** at 5%, and * at 10%. The dependent variable Patents is a count of number of granted patents applied by a firm in a given year. Intranet adoption is measured as the percentage share of firm sites having adopted intranet. Column 1 uses an unbalanced panel and keeps firms with zero patents over the sample years. Columns 2 to 4 use an unbalanced panel, dropping firms with zero patens overs the sample period. Coefficients are from Poisson regressions with standard errors clustered by firm (in parentheses).

Table A2: Heterogeneous Effects for Digital Sectors

	(1)	(2)	(3)
	Patents	Patents	Patents
Intranet	0.00494**	0.00537**	0.0209*
	(0.00229)	(0.00225)	(0.0115)
Intranet * Digital Sector	0.0105***	0.00354	0.00304
	(0.00229)	(0.00633)	(0.00677)
Log R&D$_t$	0.0972**	0.103**	0.0939**
	(0.0464)	(0.0472)	(0.0438)
Log R&D$_{t-1}$	0.0566	0.0603	0.0588
	(0.0571)	(0.0560)	(0.0556)
Employees	0.00320***	0.00322***	0.00382***
	(0.00116)	(0.00121)	(0.00132)
Observations	2784	2784	2784
Number of firms	348	348	348
Year FE	Yes	Yes	Yes
Firm FE	No	Yes	Yes
Time trend for Digital Sector	No	Yes	Yes
p-value CF			0.359
F-stat 1st stage			39.96

Notes: *** denotes significance at 1%, ** at 5%, and * at 10%. The dependent variable Patents is a count of number of granted patents applied by a firm in a given year. Intranet adoption is measured as the percentage share of firm sites having adopted intranet and it is also introduced interacted with the dummy Digital Sector that takes the value of one if a firm operates in SIC sectors 3570-3579. Column 1 does not include a specific time trend for firms in digital sectors. Column 2 includes a specific time trend for firms in digital sectors and column 3 presents IV estimates includeing this time trend. Coefficients are from Poisson regressions with standard errors clustered by firm (in parentheses).

Table A3: Heterogeneous Effects for Early Years and Early Adopters

	Poisson	Poisson	IV	IV
	(1)	**(2)**	**(3)**	**(4)**
	Early Years	Early Adopter	Early Years	Early Adopter
Intranet	0.00638**	0.00516**	0.0256**	0.0253**
	(0.00248)	(0.00216)	(0.0107)	(0.0114)
Intranet * Early	-0.00297	0.00119	-0.00153	0.00181
	(0.00405)	(0.00294)	(0.00288)	(0.00296)
Log R&D$_t$	0.0970**	0.0988**	0.0871**	0.0880**
	(0.0456)	(0.0466)	(0.0426)	(0.0428)
Log R&D$_{t-1}$	0.0599	0.0610	0.0574	0.0596
	(0.0611)	(0.0617)	(0.0599)	(0.0607)
Employees	0.00315***	0.00310**	0.00391***	0.00386***
	(0.00119)	(0.00127)	(0.00130)	(0.00138)
Observations	2784	2784	2784	2784
Number of firms	348	348	348	348
Year FE	Yes	Yes	Yes	Yes
Firm FE	Yes	Yes	Yes	Yes
p-value CF			0.146	0.179
F-stat 1st stage			38.52	38.52

Notes: *** denotes significance at 1%, ** at 5%, and * at 10%. The dependent variable Patents is a count of number of granted patents applied by a firm in a given year. Intranet adoption is measured as the percentage share of firm sites having adopted intranet. In columns 1 and 3, Early is a dummy taking the value of 1 for the years 1996 and 1998, the initial years of the deployment of intranets. In columns 2 and 4, Early is a dummy equal to 1 for firms above the median of intranet adoption in the years 1996 and 1998 (notice that this is a dummy equal to 1 not only for the years 1996 and 1998 but for all sample years). Coefficients are from Poisson regressions with standard errors clustered by firm (in parentheses). In the control function regressions I control for a second-order polynomial expansion of the residual estimated in a first-stage regression of Intranet on Bitnet nodesj * ISPst and controls. The p-value on the significance of the control function is reported in the table.

Table A4: Further Robustness Analysis

	(1) Patents	(2) Patents	(3) Patents	(4) Patents	(5) Patents	(6) Patents
Intranet	0.00868**	0.00503*	0.00839**			0.00480**
	(0.00402)	(0.00267)	(0.00405)			(0.00224)
Internet Access	-0.425		-0.472			
	(0.369)		(0.326)			
E-commerce		0.222	0.372			
		(0.290)	(0.347)			
Internet for Research			-0.160			
			(0.344)			
Intranet at HQ				0.0928		
				(0.0797)		
Dummy intranet > 0					-0.113	
					(0.118)	
Log R&D$_t$	0.0993**	0.0967**	0.0965**	0.106**	0.107**	
	(0.0475)	(0.0456)	(0.0468)	(0.0480)	(0.0490)	
Log R&D$_{t-1}$	0.0544	0.0620	0.0581	0.0662	0.0646	
	(0.0578)	(0.0617)	(0.0584)	(0.0631)	(0.0627)	
Log R&D Stock 2						0.0924**
						(0.0376)
Employees	0.00320***	0.00314***	0.00317***	0.00325***	0.00322***	0.00293***
	(0.00120)	(0.00117)	(0.00119)	(0.00118)	(0.00119)	(0.00108)
Observations	2744	2784	2784	2784	2784	2784
Number of firms	343	348	348	348	348	348
Firm FE	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes

Notes: *** denotes significance at 1%, ** at 5%, and * at 10%. The dependent variable Patents is a count of number of granted patents applied by a firm in a given year. Intranet adoption, Internet Acces adoption, E-commerce adoption, and Internet for Research adoption are measured as the percentage share of firm sites having adopted each corresponding technology in a given year. Intranet at HQ is a dummy taking the value of 1 if a firm's headquarters have adopted intranet. Dummy intranet > 0 is a dummy taking the value of 1 if at least one firm site has adopted intranet. Log R&D Stock 2 is measure of firm R&D stock using the depreciation rates estimated in Li and Hall (2016) when available. Coefficients are from Poisson regressions with standard errors clustered by firm (in parentheses).

Table A5: IV Robustness

	Bitnet Trend (1) Patents	Bitnet HQ (2) Patents	Arpanet (3) Patents	Bartik County (4) Patents	Bartik Industry (5) Patents	Unbalanced (6) Patents
Intranet	0.0265**	0.0392*	0.0287***	0.0380*	0.0207*	0.0291**
	(0.0108)	(0.0211)	(0.00985)	(0.0196)	(0.0116)	(0.0119)
Log R&D$_t$	0.0876**	0.0919**	0.0847**	0.0958**	0.0971**	0.0700
	(0.0426)	(0.0443)	(0.0418)	(0.0447)	(0.0468)	(0.0427)
Log R&D$_{t-1}$	0.0563	0.0485	0.0549	0.0493	0.0582	0.0771
	(0.0597)	(0.0566)	(0.0603)	(0.0586)	(0.0600)	(0.0521)
Employees	0.00395***	0.00434***	0.00413***	0.00403***	0.00359***	0.00423***
	(0.00130)	(0.00149)	(0.00128)	(0.00141)	(0.00106)	(0.00135)
Observations	2784	2784	2784	2784	2784	3030
Number of firms	348	348	348	348	348	387
Firm and Year FE	Yes	Yes	Yes	Yes	Yes	Yes
p-value CF	0.119	0.289	0.0565	0.132	0.206	0.105
F-stat 1st stage	38.75	6.805	29.90	13.96	67.39	26.60

Notes: *** denotes significance at 1%, ** at 5%, and * at 10%. The dependent variable Patents is a count of number of granted patents applied by a firm in a given year. Intranet adoption is measured as the percentage share of firm sites having adopted intranet. Column 1 uses as IV the number of Bitnet nodes per firm interacted with a time trend taking the values 1 to 4 for years 1996 to 2002. Column 2 uses as IV the number of Bitnet nodes in the headquarters of the firm, instead of the mean of nodes for all firm sites, and interacts this with ISPs. Column 3 uses the average number of Arpanet nodes per firm site interacted with ISP. Column 4 uses a Bartik-type IV in which the instrument is constructed in the following way: (i) for each establishment and year I calculate the share of other establishments in the same county-year that have adopted intranet -excluding the focal establishment-, and (ii) for each firm-year I calculate the variable Bartik County as the mean of the previous number across firm sites. Column 5 uses another Bartik-type instrument, but this time at the industry level by calculating for each firm-year the average level of intranet adoption for all other firms in the same industry (at the 2-digit SIC level, or at the 1-digit SIC level if there is no other firm at the 2-digit level). Column 6 uses the same IV as in the baseline specification but with an unbalanced sample of firms. All columns include a second-order polynomial series expansion of the estimated error term in the first stage. Coefficients are from Poisson regressions with standard errors clustered by firm (in parentheses).

Table A6: IV First-Stage

	No Trend	Bitnet Trend	Bitnet HQ	Arpanet	Bartik County	Bartik Industry	Unbalanced
	(1)	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**	**(7)**
	Intranet	Intranet	Intranet	Intranet	Intranet	Intranet	Intranet
Bitnet nodes$_j$ * (t=1996)	1.166*						
	(0.698)						
Bitnet nodes$_j$ * (t=1998)	4.346***						
	(1.295)						
Bitnet nodes$_j$ * (t=2000)	5.615***						
	(0.934)						
Bitnet nodes$_j$ * (t=2002)	4.856***						
	(0.885)						
Bitnet nodes$_j$ * Time trend		1.494***					
		(0.240)					
Bitnet nodes HQ$_j$ * ISPs$_t$			0.000316***				
			(0.000121)				
Arpanet nodes$_j$ * ISPs$_t$				0.00556***			
				(0.00102)			
Bartik County					85.03***		
					(22.76)		
Bartik Industry						0.527***	
						(0.0642)	
Binet nodes$_j$ * ISPs$_t$							0.000698***
							(0.000135)
Log R&D$_t$	0.0575	0.0532	0.113	0.0733	-0.00282	-0.0498	0.183
	(0.318)	(0.316)	(0.322)	(0.298)	(0.304)	(0.259)	(0.276)
Log R&D$_{t-1}$	0.00388	0.0107	0.0556	0.0137	0.126	-0.0677	-0.0877
	(0.289)	(0.290)	(0.292)	(0.268)	(0.267)	(0.271)	(0.255)
Employees	-0.0261**	-0.0258**	-0.0289**	-0.0253**	-0.0260**	-0.0223*	-0.0267**
	(0.0121)	(0.0122)	(0.0128)	(0.0128)	(0.0122)	(0.0124)	(0.0117)
Observations	2784	2784	2784	2784	2784	2784	3030
Number of firms	348	348	348	348	348	348	387
Firm and Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
F-stat 1st stage	11.00	38.75	6.805	29.90	13.96	67.39	26.60

Notes: *** denotes significance at 1%, ** at 5%, and * at 10%. Intranet adoption is measured as the percentage share of firm sites having adopted intranet. Column 1 shows first-stage estimates of the seconds-stage IV in Table 4 column 4. Columns 2 to 7 show first-stage estimates corresponding to the second-stage estimates in Table A6. Coefficients come from an OLS regression with standard errors are clustered by firm.

Table A7: Heterogeneous Effects I(Robustness)

	Poisson	Poisson	Poisson	Poisson	IV	IV	IV	IV
	(1) High Value	**(2)** Low Value	**(3)** High Citations	**(4)** Low Citations	**(5)** High Value	**(6)** Low Value	**(7)** High Citations	**(8)** Low Citations
Intranet	0.00858***	0.00287	0.00598**	0.00488*	0.0297**	0.0312	0.0238**	0.0296**
	(0.00291)	(0.00297)	(0.00260)	(0.00252)	(0.0144)	(0.0191)	(0.0105)	(0.0127)
Log R&D$_t$	0.118	0.0961	0.0989**	0.0999*	0.115	0.0116	0.0884*	0.0877
	(0.0731)	(0.125)	(0.0503)	(0.0585)	(0.0723)	(0.0718)	(0.0464)	(0.0549)
Log R&D$_{t-1}$	0.0701	0.463*	0.0462	0.0965	0.0612	0.401	0.0429	0.0983
	(0.0807)	(0.269)	(0.0586)	(0.0795)	(0.0780)	(0.258)	(0.0570)	(0.0788)
Employees	0.00281	0.00301	0.00274**	0.00368***	0.00341	0.00710	0.00343***	0.00465***
	(0.00223)	(0.00522)	(0.00120)	(0.00104)	(0.00219)	(0.00555)	(0.00131)	(0.00121)
Observations	2088	1592	2472	2480	2088	1592	2472	2480
Number of firms	261	199	309	310	261	199	309	310
Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Firm FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
p-value CF					0.142	0.00299	0.214	0.0533
F-stat 1st stage					38.52	38.52	38.52	38.52

Notes: *** denotes significance at 1%, ** at 5%, and * at 10%. Patents High Value is a count of number of patents above the 66 percentile of the economic value of patents generated in a given year. Patent Low Value is the same for patents below the 33 percentile. Patents High Citations is a count of number of patents above the 66 percentile of forward citations per year. Patent Low Citations is the same for patents below the 33 percentile. Intranet adoption is measured as the ratio of firm sites having adopted intranet. Coefficients are from Poisson regressions with standard errors clustered by firm (in parentheses). In the control function regressions (cols 5-8) I control for a second- order polynomial expansion of the residual estimated in a first-stage regression of Intranet on Bitnet nodes$_j$ * ISPs$_t$ and controls. The p-value on the significance of the control function is reported in the table.

Table A8: Heterogeneous Effects II (Robustness)

	Poisson (1) Patents	Poisson (2) Patents	IV (3) Patents	IV (4) Patents
Intranet	0.00554**	0.00484**	0.0249**	0.0222**
	(0.00234)	(0.00198)	(0.0109)	(0.0106)
Intranet * Low Innovation	0.00834		0.00844	
	(0.00534)		(0.00531)	
Intranet * Low Competition		0.00544**		0.00508*
		(0.00269)		(0.00279)
Low Innovation	-2.477***		-2.406***	
	(0.245)		(0.233)	
Low Competition		-0.394***		-0.327**
		(0.146)		(0.144)
Log R&D$_t$	0.104**	0.0970**	0.0936**	0.0886**
	(0.0451)	(0.0470)	(0.0420)	(0.0442)
Log R&D$_{t-1}$	0.0551	0.0544	0.0532	0.0527
	(0.0587)	(0.0592)	(0.0575)	(0.0575)
Employees	0.00312***	0.00294**	0.00387***	0.00361***
	(0.00118)	(0.00119)	(0.00131)	(0.00130)
Observations	2784	2784	2784	2784
Number of firms	348	348	348	348
Year FE	Yes	Yes	Yes	Yes
Firm FE	Yes	Yes	Yes	Yes
p-value CF			0.165	0.258
F-stat 1st stage			38.40	38.31

Notes: *** denotes significance at 1%, ** at 5%, and * at 10%.The dependent variable Patents is a count of number of granted patents applied by a firm in a given year. Intranet adoption is measured as the share of firm sites having adopted intranet. Low Innovation is a dummy equal to 1 if a firm is in the bottom quartile of patent stock in each corresponding year. Low Competition is dummy equal to 1 if a firm is in the bottom quartile of sector competition in each corresponding yeaer. Coefficients are from Poisson regressions with standard errors clustered by firm (in parentheses). In the control function regressions I control for a second-order polynomial expansion of the residual estimated in a first-stage regression of Intranet on Bitnet nodes$_j$ * ISPs$_t$ and controls. The p-value on the significance of the control function is reported inthe table.

Table A9: Effects on Productivity (Robustness)

	OLS	OLS	OLS	OLS	OLS	OLS
	(1)	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**
	TFP OP	TFP ACF	TFP BB	TFP LP	TFP LP	TFP LP
Intranet	0.000425	0.000632	0.000124	-0.000911	-0.000874	0.000851
	(0.000570)	(0.000578)	(0.000584)	(0.000925)	(0.000934)	(0.000590)
Patent Stock	0.0569***	0.0575***	0.0739***	0.0642***	0.0519**	0.0417***
	(0.0154)	(0.0160)	(0.0158)	(0.0215)	(0.0221)	(0.0136)
Observations	2784	2784	2784	2784	2784	2784
Number of firms	348	348	348	348	348	348
Firm and Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Industry-specific trends	Yes	Yes	Yes	Yes	Yes	Yes
Using Labor Costs				Yes		
Control Advertising					Yes	
Alternative Depreciation Rate						Yes

Notes: *** denotes significance at 1%, ** at 5%, and * at 10%. Column 1 uses productivity estimates calculated with the method of Olley and Pakes (1996). Column 2 uses productivity estimates calculated with the method of Ackerberg et al. (2015). Column 3 uses productivity estimates calculated with the method of Blundell and Bond (2000). Columns 3 to 6 use productivity estimates calculated with the method of Levinshon and Petrin (2003), but with some variation with respect to the baseline results. Column 4 controls for labor costs when estimating the firm production function by using information from the Compustat variable XLR when available and, otherwise, proxying labor cost by the product of the average total compensation per worker in a given year (using data from the BLS) and the number of workers reported in Compustat. Column 5 controls for advertising expense (XAD variable in Compustat). Finally, column 6 constructs the variable patent stock using the sector-specific depreciation rates suggested in Li and Hall (2006). Coefficients are from an OLS regressions with standard errors clustered by firm (in parentheses).

Table A10: Main Results Dropping Problematic Firm

	(1) Patents	(2) Citations	(3) Intranet	(4) Patents	(5) Labor Prod	(6) TFP
Intranet	0.00565**	0.00663**		0.0256**	0.000940	0.000835
	(0.00234)	(0.00262)		(0.0108)	(0.000619)	(0.000592)
Patent Stock					0.0331**	0.0405***
					(0.0163)	(0.0146)
Bitnet nodes$_j$ * ISPs$_t$			0.000793***			
			(0.000128)			
Log R&D$_t$	0.0981**	0.114*	0.0820	0.0868**		
	(0.0462)	(0.0655)	(0.326)	(0.0427)		
Log R&D$_{t-1}$	0.0595	0.0522	-0.0140	0.0576		
	(0.0612)	(0.0593)	(0.293)	(0.0599)		
Employees	0.00317***	0.00248*	-0.0258**	0.00394***		
	(0.00119)	(0.00135)	(0.0121)	(0.00131)		
Observations	2776	2776	2776	2776	2776	2776
Number of firms	347	347	347	347	347	347
Firm and Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Industry-specific trends	No	No	No	No	Yes	Yes
p-value CF			0.137			
F-stat 1st stage			38.61			

Notes: *** denotes significance at 1%, ** at 5%, and * at 10%. Patents is a count of number of granted patents applied by a firm in a given year. This table presents a robustness analysis of the baseline results dropping a firm that apparently suffered a shock in year 2002. Column 1 shows estimates of a Poission regression of number of patents on intranet adoption and basic controls. Column 2 instead uses as dependent varible patents adjusted by number of citations. Column 3 and 4 respectively show first-stage and second-stage estimates in which intranet is instrumented with the variable Bitnet nodesj * ISPst. Column 5 and 6 present OLS estimates of the effects of intranet adoption and patent stock on labor productivity and TFP estimates calculated with the method of Levinshon and Petrin (2003).

Chapter 3

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

"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 crude 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 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. In a competitive business landscape information can be as source of competitive advantage in a variety of ways both lowering costs (more efficient resource allocation and improving production processes) and better understanding of business opportunities (product customization and forecasting demand). The rise of information technology in the last few decades 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), originating the eruption of the Big Data revolution and 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."[37]

However, the access and adoption of Big Data technology has concentrated in large corporations and has been anecdotal among small and medium enterprises.[38] Not surprisingly then, the growing literature on information technologies (IT hereafter) adoption, not only Big

[37] https://casestudies.storetrials.com/we-have-data-lets-look-at-data-e8a06e2e3331

[38] 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."

Data, has mainly focused on large firms since they are more likely to adopt. Large firms benefit 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 (McElheran, 2014 and 2015). For this reason, the existing literature presents a gap in understanding the impact of IT and Big Data adoption on DDD in small and medium enterprises (SMEs hereafter).[39]

Identifying what changes are triggered in SMEs' performance and decision-making as a result of Big Data adoption would help answering at least two relevant questions. First, it would allow to disentangle whether SMEs are deterred from adopting Big Data technologies because *(a)* returns from adoption are too low for them, or *(b)* because of a combination of high adoption costs and lack of awareness. Finding positive returns of adoption for SMEs would open the door for public (and private) interventions intended to decrease adoption costs and facilitate data sharing initiatives (Jones and Tonetti, 2020). This would help to correct for the sparse adoption patterns among SMEs that may widen the existent performance differences between firms with "intuition-driven" and "data-driven" decision-making practices, contributing to increase further market concentration with all its consequences on market outcomes such as prices, quality, and innovation. Second, it would allow to understand how access to Big Data information affects firms' competitive strategies and market equilibrium outcomes. Recent years have seen dramatic increases in the availability of data, and the pace is accelerating. Therefore, it is first order importance to understand how this improvement in information 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 this debate by accomplishing two goals. On the one hand, we estimate the distribution of returns to adoption of a Big Data information technology that facilitates the implementation of DDD practices among SMEs. On the other hand, we want to understand the mechanisms behind the effect of Big Data on firm performance, as this will provide evidence on how information affects firms' competitive strategies and market equilibrium outcomes. To do so, we use information on the deployment of a Big Data information-sharing program in Spain from a large European bank among its SMEs customers. Upon voluntarily and freely signing up to this unprecedented and unique Big Data program, SMEs receive a report on their own sales profiles relative to other neighboring establishments in their same sector. Therefore, the program reduces the costs of access of SMEs to Big Data through a double channel. On the one hand, SMEs, most often do not have the capacity to generate large volumes of data about consumer behavior given the limited number of customers they have. On the other hand, it is likely the case that SMEs lack the capacity to analyze large volumes of disaggregated data and extract 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 an unstructured dataset. Despite 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 point-of-sale (hereafter POS). We used comprehensive information on credit and debit card transactions for nearly all POS in the country of study between 2014 and 2018. Our final working data contains quarterly information for 310,610

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

establishments, out of which 7,110 adopted the technology across all provinces in the country, 17 sectors and 70 subsectors.[40]

Our empirical methodology uses OLS regressions of first-differences of quarterly revenue on first-differences of adoption with sector-zip code-quarter fixed effects and establishment-specific time trends as baseline specification. We complement this analysis with an instrumental variable approach where we take advantage of the fact that different establishments within the same sector-zip code dyad are affiliated to different bank branches. Our instrumental variable is then the number of adopters, other than the focal establishment, in the establishment's bank branch. 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 put an effort in the promotion of the program, increasing adoption among its customers. Low affinity branches put no effort in the promotion, and as a result their customers were unlikely to discover the existence of the program and adopt.

We find that adoption is associated with a 4.5% increase in revenue from credit and debit card transactions, and our instrumental variable strategy shows that adoption causally increases establishment revenue by 9% for those establishments whose adoption decisions are 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 high competition environments realize higher returns while establishment sophistication is not a factor driving differences in the returns of adoption. Our evidence also points out that the increase in revenue comes from an increase in both the number of transactions and the number of customers. The average number of transactions per customer did not change after adoption.

We investigate the role of two potential mechanisms behind these findings. On the one hand, adoption may prompt establishments to target existing, yet unexploited, business opportunities. On the other hand, adoption may help establishments improving the efficiency of their internal resource allocation. Our analysis shows direct support for the former mechanism in that adopting establishments increase their sales to underserved customer segments. 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 some evidence consistent with the latter supply-driven mechanism, that is, adopting establishments reshuffle their sales towards idle times of the week while holding constant the demographics of their clientele portfolio.[41]

[40] This represents approximately 1.5% of all potential adopters of the program. We rule out that the low adoption rate of this technology is driven by high adoption costs (adoption was free and did not require high co-invention costs or high sophistication), low returns of adoption (we show returns are positive and sizable) or strategic motives (the amount of information received by an establishment does not depend on the adoption decision of other establishments and there is no reason to strategically delay adoption). Instead we argue this is due to managerial inattention causing lack of awareness, and exploit plausible exogenous variation in the information set of potential adopters for identification purposes.

[41] It would be extremely interesting to provide further evidence on the specific actions that the adopting establishments implement to attract new customers as a result of discovering latent business opportunities. Unfortunately, we cannot say much about it. We find that mean transaction values do not change as a result of adoption, which is consistent with prices remaining stable. However, we have no information about possible

It is important to highlight the fact that adopters not only change their portfolio of customers when they discover new business opportunities, they also 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, that is, more information increases competition among incumbent establishments. While some theories may predict that more information may drive establishments to become more specialized, we find the opposite, that is, establishments with more information start serving more customer types. This finding is important because it has direct consequences for the impact of information on the degree of competition and, ultimately, on consumer surplus and total welfare.[42] If access to more information makes establishments specialize in serving narrow market segments, the degree of competition would go down, prices would increase and welfare could potentially decrease. Instead, our findings suggest a positive association between more information in a market, the degree of competition and total welfare.

Our findings and their implications contribute to three main streams of literature. A first stream 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 follows their approach in that it identifies the adoption of Big Data IT as an input of production that facilitates changes in behavior and strategies, which translates into changes in performance.[43] This way, the sparse adoption patterns of IT and Big Data technologies among small and medium-sized establishments can contribute to exacerbate the productivity gap between large and small firms. Interestingly, our results show how the low adoption rate among SMEs is likely due to high adoption costs and managerial inattention, and not so much to low returns of implementation.

Our second contribution is to the growing literature studying the role of IT in enabling DDD (Brynjolfsson and McElheran, 2016a and 2016b; Brynjolfsson et al., 2011). McAfee and

---

advertising and marketing campaigns, introduction of new products, or other types of strategies implemented by adopting establishments. If anything, given the multidimensional information contained in the reports is likely to highlight very different weaknesses and opportunities for different establishments, one could think the variety of actions implemented can be equally large. In conversations with the bank, we discussed making a survey to adopting establishments to gain a better understanding of these aspects, but unfortunately it could not be implemented. Therefore, our results 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 more detailed evidence on the specific actions implemented to that end.

[42] 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).

[43] 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).

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 skewness of firm size and productivity distribution, and Bajari et al. (2019) show that Big Data allows firms to lower forecasting errors and therefore better decision making. Goldfarb and Tucker (2019) survey the literature on digital economics and IT.

To the best of our knowledge, ours is the first paper to empirically evaluate the gains of adoption of a specific Big Data informational technology among small and medium enterprises in the retail and customer service sectors. 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 organization. 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 technology adopted is homogenous across establishments and it provides information not only about the adopting establishment but about competitor strategies and market opportunities. Third, we are able to provide causal estimates of the impact of adoption on productivity due to our instrumental variable approach.

Our paper also contributes to a third stream of 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).[44] Therefore, understanding whether firms are able to benefit from 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.

This program provides a unique opportunity to study how access to market information might impact a firm's strategic decision-making. On this regard, and to the best of our knowledge, the closest 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.[45] Similar to our findings, she shows that, upon receiving information about their closest rivals, small firms change their strategies to align closer to their competitors' strategies. Moreover,

[44] Relatedly, our findings also have implications for the literature on inattention in organizations to the extent that Big Data 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 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).

[45] Nagaraj (2020) explores the impact of public data infrastructure and shows how better information increases market entry. Fabregas et al (2019) provide empirical evidence of smallholder farmers' valuation for neighboring agricultural information.

she finds suggestive evidence that managerial inattention plays an important role in explaining the firms' lack of awareness. Our paper differs from hers in that we show how firms react to access to rich multidimensional information comparing their own client portfolio to that of their competitors. In our setting, information allows firms to increase their revenues by becoming aware of existing, unexploited business opportunities. Moreover, in line with prior results such as Bloom et al. (2013), Bruhn et al. (2018) or Giorcelli (2019), we find evidence consistent with an improvement in resource allocation of firms upon 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, more sophisticated firms were probably already using a considerable amount of information in their decision-making process prior to adoption. Second, the Big Data technology 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), 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.[46] 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 Big Data IT adoption 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 mechanisms. Finally, section 5 concludes.

2. Institutional Detail and Data Description

2.1. The Bank

Our empirical setting is the market for SMEs in Spain, and our data come from one of the largest European banks with a high market share in the country. Hereafter, we refer to the data provider as "the bank". The bank is a major player in the credit card market both as credit (and debit) card issuer and 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 the country in the fall of 2014 and went national in the spring of 2016. The program aimed to bring Big Data technology to SMEs using the bank's credit card POS.[47] 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 scant adoption of the technology. If anything, bank employees

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

[47] 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."

would offer the adoption of the program as a source of value added to an already existing business relationship with the client.

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 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 for each adopter a monthly report, which became available through the program's online platform. It is necessary to highlight the nature of the information in the report is descriptive and not predictive. This way, 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, this monthly report effectively provided precise market research information on the local market in which each program participant operated.[48]

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, establishment-specific and card-specific identifiers. On the one hand, it is important to note

[48] 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.

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, 17 sectors and 70 subsectors. 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 310,610 establishments, including all 7,100 technology 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.[49] 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 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 used to explore impact heterogeneity for the subsample of 7,100 adopters. 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.[50] Each one of these questions provide Likert

[49] 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.

[50] 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.

scales from 1 to 5. We create a measure of analytical savviness by averaging all three answers of 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.

3. Empirical Methodology and Identification

3.1. Baseline Regressions

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 number of transactions, revenues, 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, 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 regressions specification also includes time-varying controls X_{isjt} such as dummies for the first four quarters an establishment enters our sample 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 technology adoption. 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-

---

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).

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, ours is far from being a stationary context 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 the empirical 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 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 and 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 instrumental variable strategy that exploits the fact that different establishments in the same sector-zip code dyad may hold their corporate bank account in different bank branches located in different zip codes. Hereafter, we call the bank branch where an establishment has its corporate bank account the establishment branch.

Our conversations with bank managers provide a strong foundation for our instrumental variable 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 instrument is the number of adopters per quarter (across sectors and zip codes) other than the focal establishment in the establishment branch. Figure 3 sheds light on the rationale behind our instrumental variable. 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.

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 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.

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 code across different establishments affiliated to different establishment branches would also violate our exclusion restriction. [51]

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

[51] Note that we include bank-branch time trends to control for this possible concern in a robustness specification in Appendix Table A2.

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, note 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. A final and necessary exclusion restriction for the plausibility of our instrumental variable is that sharing the same establishment branch only affects the probability of adoption, but it does not directly affect performance.[52] 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 behind the main results. Third, we conclude this section with a discussion of the results while linking 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. 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 4.6% in revenue. Columns 2 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 in subsequent quarters.[53] 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.

The last two columns of Table 2 implement our IV strategy. Column 4 shows estimates of the first stage results. Column 5 shows the results of running instrumental variables on the baseline specification of Column 1. We find that the effect jumps from 4.6% to 9.0%. We carry

[52] As a robustness check, we produce evidence in Table A2 in the Appendix where the instrumental variable (IV hereafter) does not include peers in the same sector.

[53] We do not observe a reversion to the mean when including further leads up to $t+7$. Results available upon request.

out the Hausman test to check for endogeneity, and we cannot reject the null hypothesis that adoption is exogenous (p-value=0.24).[54]

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 with 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 instrumental variable strategy. In the same manner, if bank branches 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.[55]

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. Answering these questions will grant them access to the monthly reports. We use their answers to compute our measure of sophistication 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

[54] While Table 2 presents our baseline results, Table A1 includes regression specifications with establishment-specific time trends and subsector-zip code-date FE. All our findings are robust to changes in the specification. In addition, appendix Table A3 includes an event-study exercise in which the sample is limited to adopting establishments at some point in time.

[55] 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.

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.[56]

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.[57] 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 instrumental variable 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.[58]

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.[59] 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

[56] 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.

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

[58] Information increases establishment revenues by highlighting untapped business opportunities and streamlining resource allocation. Larger establishments most likely grew as a result of performing well in both extents before adoption. Therefore, the scope for improvement of big establishments in both extents can be expected to be smaller and, accordingly, returns of adoption less significant.

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

markets. These results across more and less competitive sector-zip code dyads are statistically different from each other at the 11% level.[60]

In summary, our heterogeneity results are insightful in depicting scenarios where technology adoption derives in higher returns. Our findings in Table 3 show that those establishments of smaller size and those operating in more competitive markets 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 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. This new information may have two types of direct effects that we define as two distinct mechanisms. On the one hand, the report received may highlight business opportunities that the establishment was not aware of or did not paid 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. On the other hand, the information provided by the report may trigger adopters to reallocate their resources more efficiently towards customer groups and times during the week where their marginal returns to effort are higher. While the former mechanism requires exploiting new business opportunities, the latter implies reallocating existing levels of effort and resources. Hereafter and for simplicity, we call the latter "demand-driven" mechanisms and the former "supply-driven" mechanisms.

4.2.1. Demand-Driven Mechanisms

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 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

[60] Appendix Table A5 shows results on the heterogeneous returns of adoption for early vs late adopters. Interestingly, we find *(i)* that , in general, returns do not differ between early vs late adopters of the program, *(ii)*, however, returns seem to be negligible for establishments adopting during the pilot period, and *(iii)* returns are higher for the first adopter in each sector-zip code than for later adopters.

that new customers purchasing patterns are not statistically different from old customers in two important dimensions as amounts spent and number of transactions.

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 their prices post-adoption.

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 store and the 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 a dummy variable "*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. The dummy variable "*Small Share*" 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 suffers from 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.

In Appendix Table A6, we investigate whether our findings on 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 results 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 group.

Columns 5 to 8 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 a dummy variable "*High Concentration*" that equals 1 if the HHI of the establishment is above the median among all adopters, and 0 otherwise. The dummy variable "*Low Concentration*" 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. Those establishments in the upper half of the concentration distribution decrease concentration by 8.7% upon technology 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 more narrow customer niche.

Finally, Table 7 provides evidence of whether adopters change the spatial composition of their customer base. For simplicity, we compute for each establishment the share of revenue from customers from 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 a dummy variable "*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. The dummy variable "*Small Share*" 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 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.[61]

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. 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

[61] 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 instrumental variable and the share of out-of-zip code customers in our data.

number of customers, but they target new age-gender profiles and look for customers beyond their zip code.[62]

4.2.2. Supply-Driven Mechanisms

Alternatively, we consider a second mechanism for the impact of the newly revealed information contained in the report received by adopters. For simplicity, we call it "supply-driven" mechanisms. 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 to 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 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. Columns 1 to 4 show no change in the sales at peak time and an increase in the sales at off-peak time. This shift of 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 the total sales. These findings suggest that technology adoption triggers changes in business hours not explained by changes in demographics and, therefore, those changes may be due to improvements in supply-side efficiency.

[62] One interesting question is whether the 9% increase in revenue translates into a 9% increase in profits or input consumption (number of employees, open hours, etc.) also increases potentially making profit increase to be zero. We do not have information on these extents, however, and abstracting from any improvement in resource allocation, two things can happen when establishments adopt. On the one hand, if establishments are operating with excess capacity, discovering new opportunities they would be able to increase revenue without increasing input consumption, therefore increasing profits. On the other hand, if establishments are operating at full capacity two things can happen in turn. First, if they are operating at an optimal scale, then they would not be willing to increase production, which is at odds with the increase in revenue we see in the data. Second, if they are not operating at an optimal scale, discovering new business opportunities they would need to increase input consumption to increase revenues, but this would imply an increase in profits. Therefore, we can conclude there must be an increase in profits, although perhaps smaller or larger than the 9% increase in revenue. The improvement in resource allocation would reinforce the profit increase.

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.

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, this section investigates the role of demand-driven and supply-driven mechanisms. 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. We discuss our results in the following section.

4.3. Discussion

We qualify our findings in two ways. First, we place our findings within the existing IT adoption literature. Because of lack of adoption in SMEs (or lack of comprehensive data), the previous literature has focused on large corporations and emphasized the role of IT in improving coordination among employees, departments, and divisions within their organizations (Brynjolfsson and Hitt, 2000; Bresnahan et al., 2002; Brynjolfsson et al., 2011; Bloom et al., 2014). Our paper differs from the existing literature in that diverges attention to SMEs and estimates the returns to Big Data IT to shed light on the puzzle of why are SMEs not investing in these technologies. Our findings show large heterogeneity in the effect of adoption and an average increase in revenues of 9%. Most importantly, our study of mechanisms shows that the increase in revenues is driven by both an improvement of marketing strategies and a more efficient internal organization. Our findings are indicative that SMEs can obtain high returns from adoption of Big Data technologies, and, therefore, the most likely reasons for the low adoption rate are high adoption costs or managerial inattention. Then, intervention may be justified by either providing the technology from government sources or allowing businesses to share information.

Second, we must wonder if the increase in revenues due to technology adoption comes from business stealing (potentially from non-adopters) or is net value generated from better service. For this purpose, 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 include 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 of Column 2 dropping all adopters and so comparing performance of non-adopters before and after adoption in their sector-quarter FE and finding the same exact finding of a drop of revenue of 1.4% upon adoption of a competitor. Finally, 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%).

These findings seem to suggest that some of the gains in revenues following adoption are coming from business stealing effects from competitors. Therefore, we cannot reject with the evidence in Table 10 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 once sector-quarter and sector-zip code fixed effects are introduced 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 instrumental variable. Column 3 shows the instrument 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.[63]

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.[64] Additionally,

[63] Jaravel and Borusyak (2018) point out identification problems 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.

[64] 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).

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.[65]

5. Conclusions

This paper evaluates the impact of the adoption of a Big Data technology in small and medium-sized enterprises providing information about the competitive environment of the firm. 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 about 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 explore whether these effects are heterogeneous and find that smaller adopters and adopters in more competitive markets benefit more from technology adoption. We find no differences across the level of sophistication and digital experience. 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 adoption and economic impact of new technologies. Departing from the existence of PPDs coupled with 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 Big Data IT revolution 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/or competitors' actions). Thus, there is a large scope for small firms to benefit from this new generation of Big Data IT. However, if high adoption costs prevent the adoption of Big Data IT 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 may be socially inefficient, thus opening the door for government intervention or data sharing initiatives to mitigate adoption costs.

[65] Because of the low adoption rate of this technology, we cannot really 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 get no statistically significant causal effect of increases in market adoption rates on sales at the sector and zip code levels.

While our evidence suggests a sizable average return on adoption and heterogeneity across establishments of different sizes, a cautious interpretation of our findings calls for an estimate of the lower bound of the cost of adoption. Consequently, future research should investigate the nature and behavior of these costs under different competitive environments. On the one hand, it is important to understand whether the scarce adoption patterns observed in SMEs can be addressed with mere awareness campaigns, provision of other technologies that exhibit complementarities with the current Big Data technological wave or the provision of sufficient and adequate skilled human capital able 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 technologies and *(ii)* provide further evidence of the specific actions taken by firms as a result of an increase in market information (changes in prices, introduction of new products, marketing campaigns, etc.).

Bibliography

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.

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.

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.

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.

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.

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.

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.

McElheran, Kristina (2014). "Delegation in Multi-Establishment Firms: Adaptation vs. Coordination in I.T. Purchasing Authority," *Journal of Economics and Management Strategy* 23 (2): 225-257.

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

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.

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.

Wu, Lynn, Bowen Lou, and Lorin Hitt (2019). “Data Analytics Supports Decentralized Innovation Communities,” Forthcoming at Management Science.

Figure 1: Number of adopters over time

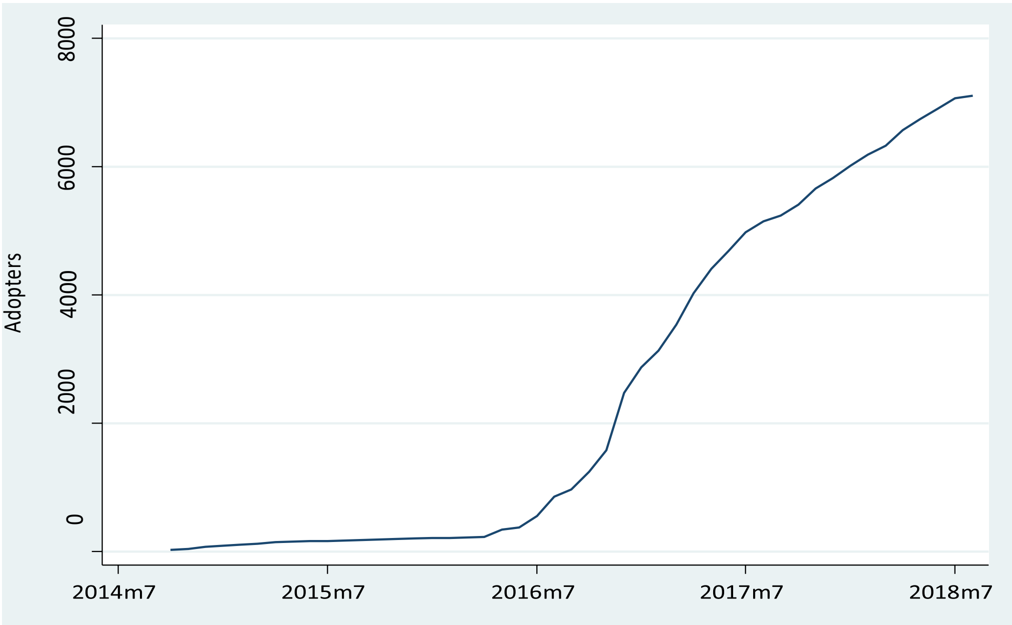


Figure 2: Timeline of Adoption

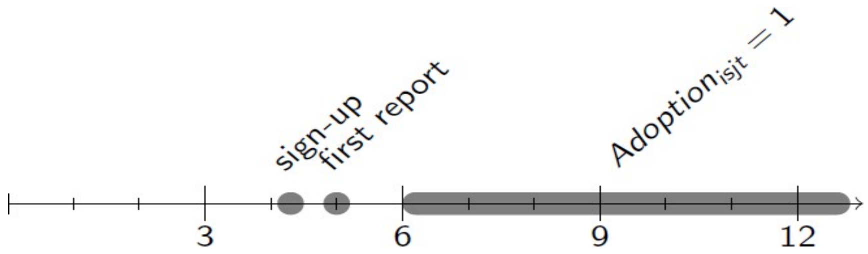


Figure 3: Instrumental variable identification

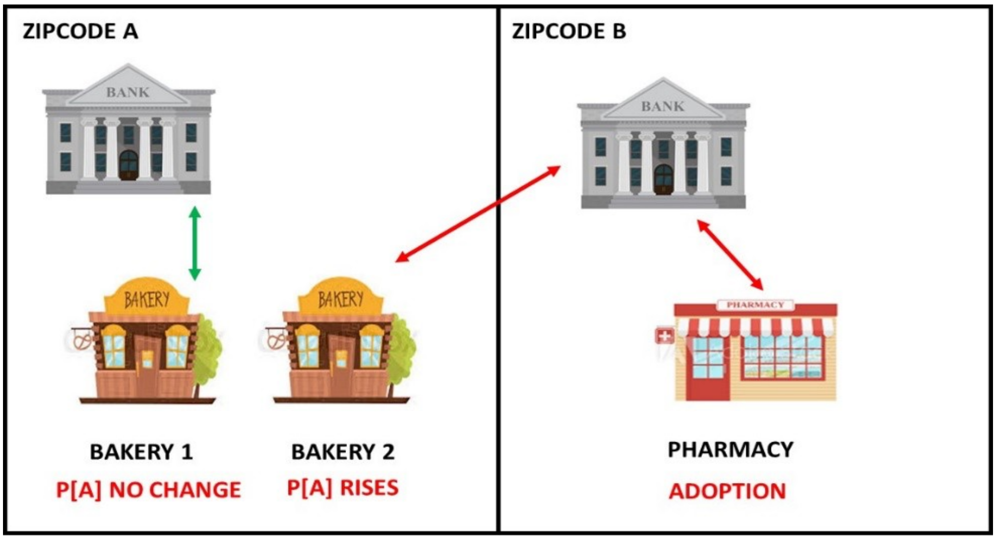


Figure 4: Treatment estimates across sectors, subsectors and regions

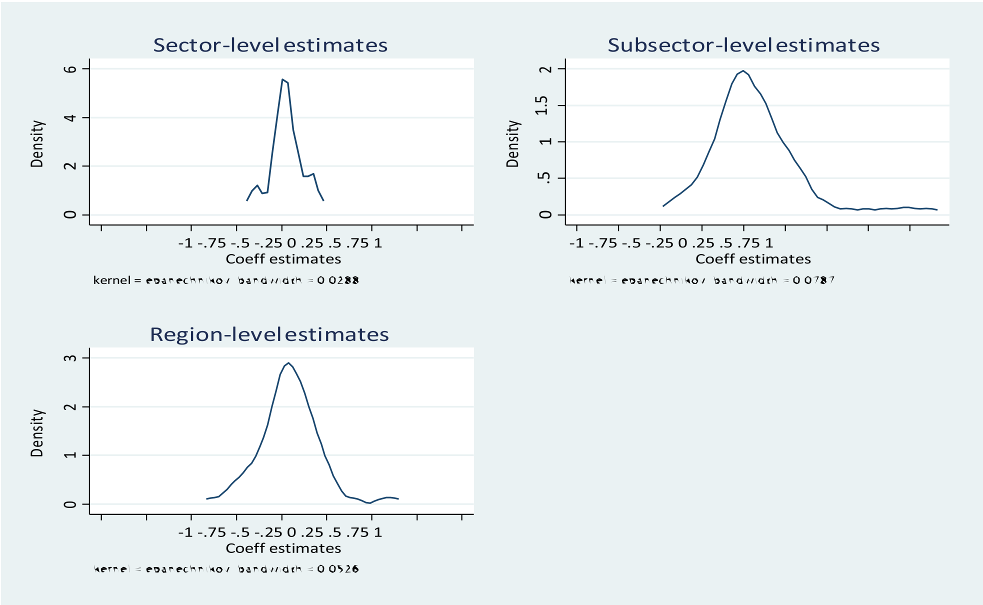


Table 1: Descriptive Statistics

	Observ.	Mean	Overall Std. Dev.	Within-Establishment Std. Dev.	Min	Max
<u>Full Sample</u>						
Revenue	4,610,085	4,715	29,171	11,120	12	7,948,335
Transactions	4,610,086	120	710	293	5	227,139
Average Value of Transactions	4,610,087	64	101	56	2.4	15,000
Customers	4,610,088	74	338	143	2	134,725
Average Value per Customer	4,610,089	85	198	127	1	92,066
<u>Adopters</u>						
Revenue	63,639	6,248	18,730	7,107	15	537,791
Transactions	63,639	153	462	162	5	8,146
Average Value of Transactions	63,639	80	147	94	3	7,006
Customers	63,639	92	224	77	3	5,975
Average Value per Customer	63,639	102	200	128	1	10,500
Number of competitors	63,639	75	96	10	0	1,020
Sophistication	3,495	3.53	3.53		0.89	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	OLS	OLS	1st-stg	2nd-stg
	(1)	(2)	(3)	(4)	(5)
Falsification			0.00978		
			(0.0158)		
Δ Adoption	0.0455***	0.0458***			0.0902**
	(0.0157)	(0.0157)			(0.0386)
Effect t+1		0.00263			
		(0.0148)			
Effect t+2		0.00395			
		(0.0161)			
Effect t+3		0.025			
		(0.0164)			
Peers IV				0.00446***	
				(0.00012)	
Sector-zipcd-quarter FE	Yes	Yes	Yes	Yes	Yes
Dummies first 4 quarters	Yes	Yes	Yes	Yes	Yes
Observations	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) OLS	(2) IV	(3) OLS	(4) IV	(5) OLS	(6) 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	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 4: Effects on Other Outcomes

Dep variable: Δ Log number of customers, Δ log revenue per customers, Δ log average transaction value

	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	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 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	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 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) OLS	(2) IV	(3) OLS	(4) IV	(5) OLS	(6) IV	(7) OLS	(8) 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	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 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	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 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	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 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)
Δ Adoption	0.0430***	0.0397**	0.095466**		0.0418***	0.0966**	
	(0.0158)	(0.0160)	(0.0398)		(0.0160)	(0.0398)	
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)
Sector-quarter FE	Yes						
Sector-zipcd-quarter FE		Yes	Yes	Yes	Yes	Yes	Yes
Sector-zipcd Trends	Yes						
Subsector-zipcd Trends		Yes	Yes	Yes	Yes	Yes	Yes
Dummies first 4 quarters	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Drop out adopters				Yes			Yes
Effect only of first adopter					Yes	Yes	Yes
Observations	4,610,085	4,610,085	4,610,085	4,546,446	4,610,085	4,610,085	4,546,446

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 (1)	OLS (2)	1st-stage (3)	2nd-stage (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

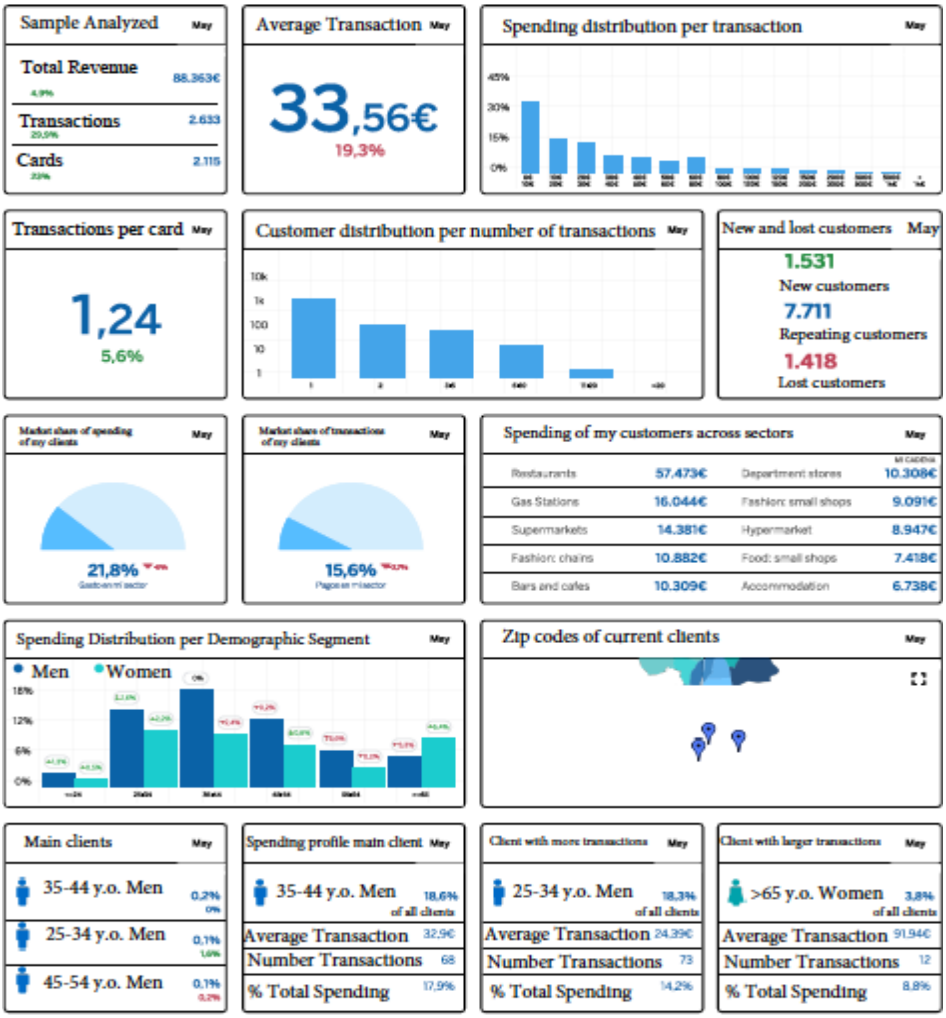


Figure A2: Sample report of monthly information on the competition of the adopting establishment

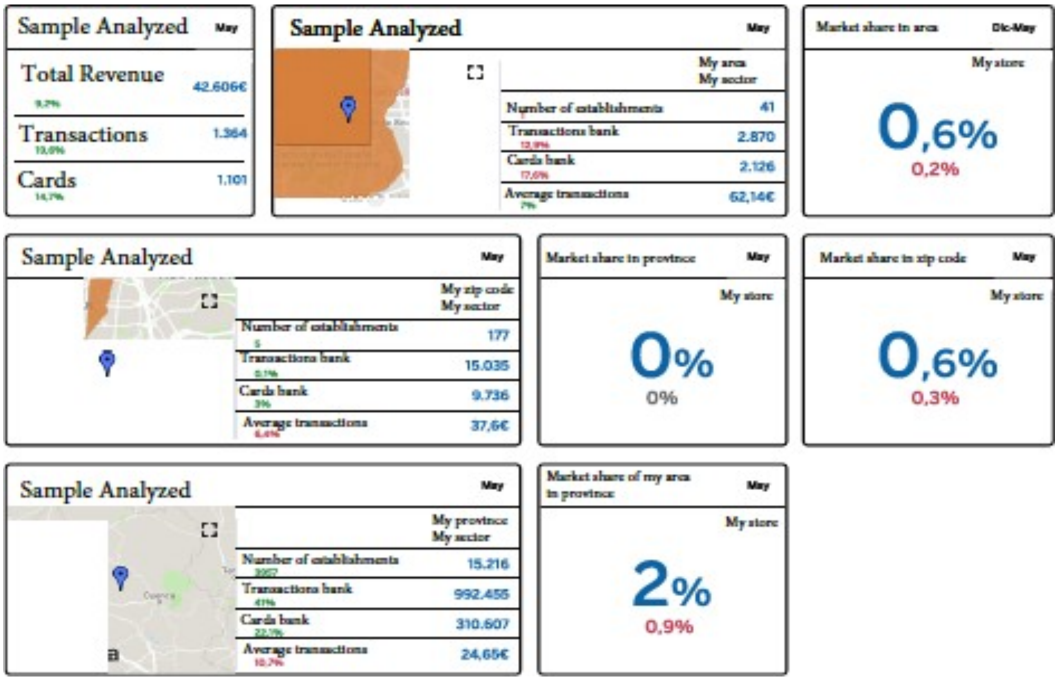


Table A1: Robustness Results

Dependent variable: Δ Log revenue

	OLS (1)	1st-stg (2)	2nd-stg (3)	OLS (4)	1st-stg (5)	2nd-stg (6)
Δ Adoption	0.0380***		0.106**	0.0577***		0.114**
	(0.0163)		(0.0421)	(0.0165)		(0.0447)
Peers IV		0.00450***			0.00432***	
		(0.00012)			(0.00012)	
Sector-zipcd-quarter FE	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
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 A2: IV Robustness Results

Dependent variable: Δ Log revenue

	1st-stg (1)	2nd-stg (2)	OLS (3)	1st-stg (4)	2nd-stg (5)
Δ Adoption		0.0948**	0.0445***		0.0840**
		(0.0387)	(0.0159)		(0.0422)
Peers IV				0.00451***	
				(0.000121)	
Peers IV (no same sector)	0.00448***				
	(0.000119)				
Sector-zipcd-quarter FE	Yes	Yes	Yes	Yes	Yes
Bank-branch time trend			Yes	Yes	Yes
Dummies first 4 quarters	Yes	Yes	Yes	Yes	Yes
Observations	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)	OLS (5)	IV (6)
Δ Adoption	0.0235	0.0250	0.0538**	0.0526**	0.0843*	0.0943
	(0.0178)	(0.0181)	(0.0256)	(0.0260)	(0.0478)	(0.0774)
Quarter FE	Yes					
Sector-quarter FE		Yes		Yes		
Zipcd-quarter FE			Yes	Yes		
Sector-zipcd-quarter FE					Yes	Yes
Dummies first 4 quarters	Yes	Yes	Yes	Yes	Yes	Yes
Observations	133,740	133,740	133,740	133,740	133,740	133,740

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	
	(1) OLS	(2) IV	(3) OLS	(4) IV	(5) OLS	(6) IV
Δ Adoption x Early	0.0503**	0.0956**	-0.0572	-0.0107	0.0635***	0.108***
	(0.0220)	(0.0420)	(0.108)	(0.116)	(0.0171)	(0.0392)
Δ Adoption x Late	0.0406*	0.0853**	0.0490***	0.0905**	-0.0263	0.0186
	(0.0223)	(0.0410)	(0.0158)	(0.0384)	(0.0378)	(0.0508)
Residual CF		-0.0507		-0.0469		-0.0506
		(0.0437)		(0.0438)		(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	0.758	0.743	0.328	0.353	0.031	0.031
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. 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.

Table A6: Changes in Composition of Customers

Dependent variable: Δ Log revenue from sales to prime customer

	(1) OLS	(2) IV	(3) OLS	(4) IV
Δ Adoption	0.0514	0.132*		
	(0.0335)	(0.0754)		
Δ Adoption x High share			-0.112***	-0,0216
			(0.0403)	(0.0798)
Δ Adoption x Low share			0.219***	0.310***
			(0.0534)	(0.0855)
Residual CF				-0.103
				(0.0866)
Sector-zipcd-quarter FE	Yes	Yes	Yes	Yes
Dummies first 4 quarters	Yes	Yes	Yes	Yes
p-value null equal returns			0.00	0.00
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.

Chapter 4

Benefits and Costs of Driving Restriction Policies: The Impact of Madrid Central on Congestion, Pollution and Consumer Spending

1. Introduction

While there is much ongoing research within and outside economics on the causes and consequences of air pollution in the short and long term, there is wide consensus that air pollution is harmful to people's health. There is a large literature establishing the causal link between air pollution and health outcomes (e.g. Chay and Greenstone, 2003; Neidell, 2004; Currie and Neidell, 2005; Deryugina et al., 2019). The consequences go beyond health outcomes, as bad air quality has been associated with less cognitive development (Bharadwaj et al., 2017), lower educational and schooling attainment (Ebenstein et al, 2016), crime (Carillo et al, 2018), or lower productivity (Chang et al., 2016), among others.[66] High air pollution levels across the globe have driven the implementation of a wide array of policies and regulations at different levels of local, regional and national government. Yet, although it is established that these policies are effective in reducing congestion and pollution levels, the question of whether they impose additional costs on society, and the magnitude of these costs, remains open.

Our paper aims to contribute to the policy debate on benefits and costs of pollution reduction policies. While pollution reduction policies, carried through the passing of a traffic restriction regulation or other forms of policy, are effective and reduce outpatient visits (Simeonova et al., 2019), hospitalizations and mortality (He et al., 2018), ambulance calls (Zhong et al., 2017), and pharmaceutical expenditures (Rohlf et al., 2020), policy makers, regulators and researchers know little on the indirect effect of these policies on economic activity. Indeed, a reduction in economic activity may change the perception of these pollution-reducing policies by the public. On the one hand, measuring the costs on economic activity allows regulators and policy makers to measure the net gain of implementing these policies. On the other hand, the implementation of these policies may affect different stakeholders differently by spatially redistributing economic activity and potentially generating a division between winners and losers. In other words, fixing a local pollution hot spot might require measures that impose drastic costs born by few individuals but generate benefits for many others. We contribute to this discussion by evaluating the impact of a driving ban implemented in downtown Madrid, known as Madrid Central, on traffic congestion, air pollution, and economic activity.

Madrid Central (MC hereafter) is a set of rules and regulations, specific to the city of Madrid, aiming to reduce air pollution through a decline in traffic congestion, in order to raise air quality to European Union standards. To achieve these goals, this new regulation restricts the

[66] The list of potentially affected outcomes by air pollution goes beyond those listed here and reaches out to infant mortality and other health outcomes in the developing world (Hammit and Zhou, 2006; Greenstone and Hanna, 2014; Greenstone and Jack, 2015).

entry of cars in the center of the city of Madrid (a zone that we will refer to as “MC area” from now on) to people living elsewhere. This policy raises a clear tradeoff. A direct impact of these regulations will lower pollution emissions in the center of the city. However, by restricting access by car, transportation costs are likely to increase for those consumers living outside the MC area, potentially discouraging consumption and retail sales in businesses within the MC area. Our paper empirically examines and documents this tradeoff between cleaner air and lower retail sales.

We split our empirical investigation and results into two distinct sections. First, using data from both the European Environmental Agency and the city of Madrid on air quality and vehicle traffic, respectively, we assess the direct effect of regulation on traffic congestion and air pollution in downtown Madrid relative to other areas within the city of Madrid and its greater metropolitan area. We use traditional difference-in-difference specifications to estimate the effect of MC on congestion and pollution using the MC area zip codes as treated (by default, we use the non-MC area zip codes as control group) and the period post-MC as treatment period. Our findings show significant decreases in traffic volume and air pollution in the MC area zip codes relative to other areas in Madrid. In particular, in our most conservative specifications, we find that the number of cars per hour in the MC area decreased by 14.7% and the concentration of NO_2 in the MC area decreased by 13.6%. Moreover, we do not find any impact of spillovers to zones adjacent to MC or to the rest of the city of Madrid.

Second, we use credit card transaction level information to evaluate changes in retail spending within the MC area before and after the passing and implementation of MC. In particular, our data on consumer spending spans from the first week of 2015 to the tenth week of 2019, while MC was introduced in the first week of December 2018. The original data set is unique in that it details the date of each transaction, the zip code of residence of the credit-card owner (buyer-zip code) and the zip code of the selling establishment (seller-zip code). Because the raw transaction data specifies all these details for each transaction, we can aggregate information for each *buyer zip code–seller zip code* pair and for each week, pre and post policy introduction. As a result, we can effectively measure “trade flows” between all zip codes within the metropolitan area of Madrid before and after the introduction of MC. In our baseline results, we use a triple differences identification strategy exploiting the fact that MC only has a direct impact on those buyers who, living in zip codes outside of the MC area, make all or part of their purchases in the center of Madrid. Following this strategy, we are able to estimate the impact of MC on consumers effectively traveling to downtown Madrid to do their shopping, (1) relative to the shopping of these same consumers in other areas of the city not directly affected by MC, and (2) relative to the shopping in downtown Madrid of consumers living within the MC area, as they are effectively exempt from the MC regulation. This identification strategy allows us to deal with two potentially important concerns. On the one hand, we are able to control for supply and demand shocks in specific areas of the city that would otherwise compromise identification of the impact of MC. On the other hand, we are also able to control for substitution effects from the MC area towards other areas of the city.

The exceptional granularity of our data allows us to control for these potential confounding effects. By contrast, a traditional difference-in-differences specification not accounting for these concerns is likely to find biased estimates in the evaluation of geographically concentrated policies, and therefore would be an inadequate methodology to identify the real impact of this type of policies. Indeed, we find no impact of MC on consumption spending when using data aggregated at the seller-zip code level. Conversely, when we run gravity-like regressions using data disaggregated at the buyer-zip code and seller-zip code pair level, we find an 8.9% decrease in the value of brick-and-mortar sales and a 12.1% increase in the value of online sales of establishments within the MC area from buyers residing in zip codes outside the MC area. Moreover, it appears that the increase in online sales is, statistically speaking, offsetting the decrease in brick-and-mortar sales. This finding represents our main contribution and opens the possibility of a policy debate linking environmental and e-commerce regulation that favors e-commerce adoption for small and medium-sized enterprises and retail establishments.

These findings contribute to a well-founded literature on the causes and consequences of air pollution, as well as that on the optimization and evaluation of pollution-reduction policies (see reviews and papers by Parry et al., 2007; Graff Zivin and Neidell, 2013; and Currie and Walker, 2011 and 2019). We examine a particular type of policy aiming to reduce traffic congestion and air pollution by limiting the number of cars allowed to circulate in a heavily congested part of a city. We are aware that Madrid is not the first city ever to implement a program of the nature and objectives of those found in MC, and therefore ours is not the first study evaluating the efficiency and efficacy of such types of environmental policies and regulations. Examples of cities that have implemented similar traffic-related policies and their respective studies are Mexico City (Eskeland and Feyzioglu, 1997; Davis, 2008; Salas, 2010; Gallego et al., 2013; Oliva, 2015), Quito (Carillo et al., 2013), Santiago de Chile (Barahona et al., 2019; Gallego et al., 2013; Rivera, 2017), San Jose – Costa Rica (Osakwe, 2010), London (Leape, 2006; Quddus et al., 2007), Bogotá (Zhang et al., 2017), Stockholm (Simeonova et al., 2019), Taipei (Chen and Walley, 2012), Beijing (Chen et al., 2013; Fu and Viard, 2015; Zhong et al., 2017), as well as a number of other Chinese cities (Lin et al., 2011; Li, 2014; Ye, 2017; Li et al., 2019) and German cities (Gehrsitz, 2017; Wolff, 2014).

While the number of papers investigating the health and air quality benefits of different versions of driving bans is extensive, only a few papers study the effects on economic outcomes such as labor supply decisions and local commerce. To the best of our knowledge, Fu and Viard (2015) and Quddus et al. (2007) are the closest papers to ours. While Fu and Viard (2015) show that traffic restriction policies in Beijing reduced the number of hours of labor supplied by affected workers, Quddus et al. (2017) find that a store affected by the traffic ban zone in London, UK, experienced a drop in sales compared to unaffected stores of the same chain. Our paper differs from those in a number of ways. First, our identification relies on a well-defined triple difference strategy where we utilize to our advantage geographical variation on the application of the policy within the city of Madrid. Second, our credit card transaction data allows us to measure economic activity as trade flows between zip codes within Madrid. Third and finally, our data allows us to separate brick-and-mortar from online transactions. Therefore, our findings not only show that the traffic-reducing policy had a negative impact on brick-and-mortar transactions, but we are also able to demonstrate that the diffusion and adoption of e-commerce may dilute part of the potentially negative impact of pollution-reducing policies on retail sales in particular, and on economic activity overall.

We view our findings as novel within the existing literature, and important for policy evaluation and future policy design. On the one hand, our results confirm that pollution-reducing policies aiming at traffic control can be effective. On the other hand, our analysis considers the role of e-commerce attenuating potential backfire of some of these policies on economic activity. An implication of our results is that combining environmental friendly policies with regulation that helps retail and small and medium enterprises transition from brick-and-mortar to e-commerce may be socially beneficial.

The structure of the paper is as follows. Section 2 describes in detail the regulation. In Section 3, we describe the data. Section 4 evaluates whether there was a reduction in traffic congestion and pollution in the MC area. Section 5 examines changes in consumption patterns due to the introduction of MC. Section 6 concludes.

2. Traffic Restriction Policy at Hand: Madrid Central

The city council of Madrid, Spain, enacted a city-specific traffic regulation, known as Madrid Central, on November 30, 2018. This regulation restricted access by car to an area of 472 hectares[67] located in the Madrid city center.[68] Figure 1 shows the extension of the affected area, which is the historic center of Madrid as well as the main commercial and leisure district of the city.

When MC went into effect, local authorities noticeably restricted entry by car to the affected area, so access may only be granted under exceptional circumstances. These exceptions are based on the emission category of vehicles. All vehicles are classified in five different categories according to their emission level (A, B, C, ECO and ZERO in descending order of emissions).[69] Accordingly, the city elaborated a careful list of exceptions that we list as follows:

1. Residents of the MC area can enter the MC area without restrictions. If they were to buy a new car, it would need to belong to category B or cleaner to enter without restrictions. All cars of category B or cleaner can enter the MC area if they park in a public or private garage.[70]
2. Delivery vehicles are subject to time restrictions in their access to the MC area.
3. Commercial and industrial vehicles with an authorization to park in residential areas of the MC area are allowed to access the MC area. Should they require a new parking license, the vehicle would need to be of category B or cleaner to enter without restrictions.
4. People with reduced mobility have no limitations.
5. ZERO emission cars can enter without restriction.
6. ECO emission cars can enter to park for a maximum of two hours.
7. Taxis and ride-hailing vehicles can enter the MC area if they belong to category B or cleaner.

---

[67] The city of Madrid has a total surface of 60,400 hectares.

[68] See Boletin Oficial del Ayuntamiento de Madrid (2018) for more information, https://www.madrid.es/UnidadesDescentralizadas/UDCMovilidadTransportes/AreaCentral/01InfGral/Ac%20Jta%20Gob%2029%20oct%202018_MC.pdf.

[69] Category ZERO refers to electric and hybrid vehicles with a range of more than 40kms. Category ECO refers to hybrid vehicles with a range of less than 40kms and gas vehicles. Category C refers to gasoline vehicles registered after 2006 (EURO 4, 5 and 6) and diesel vehicles registered after 2014 (EURO 6). Category B refers to gasoline vehicles registered after 2000 (EURO 3) and diesel vehicles registered after 2006 (EURO 4 and 5). Lastly, Category A comprises all other vehicles.

[70] Owners and tenants of private garages need a permit. Vehicles getting access to public garages get their plates automatically registered.

8. Public transport vehicles are not subject to restrictions.

As a result, the population most affected by these regulations is the non-residents of the MC area. That segment of the population cannot access the MC area at all with their own vehicles if they belong to category A, and can only access to park in a garage if they belong to category B or cleaner. This implies, for instance, that non-residents are not allowed to park in the street or access the MC area to pick up or drop off passengers if their vehicles are not classified as ECO or ZERO.

The first day of implementation of the MC regulations was November 30, 2018. During its first month, large traffic signals indicated the perimeter of the MC area and the prohibition of entry. Moreover, local police monitored traffic and informed those drivers in violation of the new regulation without imposing any fines. In January 2019, the local authorities introduced an automatic monitoring system based on cameras installed at all access points of the MC area. The system registered license plates and informed violating drivers by postal mail of the infraction, without imposing any fines. From March 16, 2019, violations were fined 90 €. Our data and analysis cover this initial period up to March 16, 2019.[71] Once we have described in this section the regulation and timing of MC, we proceed with our estimation of the impact of MC on traffic congestion, air pollution, and retail sales.

3. Data

To perform our analysis, we combine two different sources of data. First, we use data on traffic, local air pollution and meteorological conditions available from the Government of Spain and the city of Madrid and local governments. Second, we are able to gain access to proprietary data on credit card spending at the transaction level that we can aggregate up to pairs of zip codes for buyer-seller locations within the metropolitan area of Madrid. While the former data allow us to quantify the direct benefits from the driving ban on traffic congestion and air pollution, the latter data will help us quantify indirect costs of the implementation of the driving ban on consumer spending.

We proceed to the aggregation of data not only for the MC area but also for the whole metropolitan area of Madrid. Since there is no legal definition for the metropolitan area of Madrid, we define the metropolitan area of Madrid as the area that includes: (1) all zip codes within the city of Madrid, and (2) all zip codes at least partially inside a buffer of 5 km around the perimeter of Madrid. Note that for our purposes, we divide the city of Madrid into the MC area and the rest of the city. Overall, the full metropolitan area accounts for 126 zip codes (56 within the city of Madrid and 70 outside).[72]

3.1. Traffic and Pollution Data

[71] While it would be interesting to know if compliance increased after March 16, 2019, when fines became enforceable, our findings show a clear and robust reduction in car traffic as a result of the introduction of MC during our period of analysis, confirming the policy also had a disuassive effect during this inital phase.
[72] We check the robustness of our findings in the paper to alternative definitions.

We obtain traffic data from the Madrid Department of Traffic Technology published through the city's open data portal.[73] The majority of data comes from monitors used to control traffic lights, but there are also other data coming from alternative types of sensors. We make use of traffic monitoring data from 2015 to early 2019. The raw data are reported in 15-minute intervals. We drop erroneous observations as well as outliers in the 99.9th percentile. Then, we aggregate each monitor's readings to the daily level if traffic is observed at least 80 times during a given day. Finally, we aggregate all daily monitor data to the weekly level, conditional on observing every day of the week. The resulting dataset is an unbalanced panel of 4,085 traffic monitors across the city of Madrid.[74]

Traffic is measured by the number of vehicles per hour, the share of time (in %) a certain road section is occupied by a vehicle, and the share of road capacity utilized (in %). Summary statistics in Table 1 show that traffic is denser in the city center of Madrid where, on average, 28% of the road capacity is used during the week, compared to 20% outside of the city center. Because highway M-30 is a major ring road that helps intercity traffic bypass the center of Madrid as well as connect commuting traffic to reach the city center, a significant number of traffic monitors are purposely located on this major road, which explains the high number of vehicles observed at monitors outside the city center.

Because EU regulation defines limit values on NO_2 and other pollutants, cities are obliged to install air quality monitoring stations. The European Environmental Agency (EEA) collects measures from all member countries and makes them publicly available. There are 33 stations reporting NO_2 levels across the metropolitan area.[75] Most importantly for our study, one of the 33 stations is located inside the MC area. We use information from this station to estimate treatment effects, considering the rest of the stations as the control group.

The limit value for the mean annual NO_2 concentration specified by the EU regulation is 40 µg/m³.[76] As any reading of a station whose daily average is higher than 40 µg/m³ contributes to the potential violation of this regulation, we create an indicator that takes value one if a station's daily average NO_2 reading exceeds the limit value. The NO_2 data at the monitor level covers the period of time from the first week of 2015 to the tenth week of 2019. For consistency with other data sources, we aggregate all daily NO_2 readings on a weekly basis.[77] Table 1 summarizes weekly and annual mean NO_2 levels and the percentage share of days with NO_2 exceeding 40 µg/m³. One can see that both, at the station inside the MC area and at the stations outside that area, NO_2 levels are very high according to EU standards. The daily average concentration inside the MC area is 47 µg/m³, while it is 38 µg/m³ outside the MC area. We also calculate the share of station-by-year observations that violate the limit value imposed by EU regulation. Table 1 shows that, during the sample period, the station inside

[73] See the following link, https://datos.madrid.es/portal/site/egob/menuitem.c05c1f754a33a9fbe4b2e4b284f1a5a0/?vgnextoid=33cb30c367e78410VgnVCM1000000b205a0aRCRD&vgnextchannel=374512b9ace9f310VgnVCM100000171f5a0aRCRD&vgnextfmt=default

[74] Traffic outside Madrid city is unobserved.

[75] See Figure A1 for the map with locations of all pollution monitoring stations in Madrid. They are represented by pink circles. Blue crosses in the map indicate the location of weather stations.

[76] Directive 2008/50/EU. See https://ec.europa.eu/environment/air/quality/standards.htm

[77] As the treatment is defined on a weekly basis, we could increase the noise due to the aggregation of meteorological control variables and different weekday patterns. Results on air pollution using daily observations including day-of-week fixed effects are consistent with the weekly estimations. Daily results are available from the authors upon request.

the MC area exceeds the limit value every year. Moreover, other stations outside the city center also violate the threshold. This happens in 47% of all observations.

It is worth noting that meteorological conditions can heavily affect air quality. For example, sunlight is a key component in the decomposition of NO_2. Therefore, it is important to control for local weather conditions when studying determinants of air quality (Auffhammer et al., 2013). For this reason, we use data from the European Climate Assessment Dataset (ECAD), which provides daily measures of several meteorological variables across Europe. We match the pollution measurement data collected by each pollution monitoring station in the city to its closest available weather measurements from the ECAD dataset (represented with a blue cross in Figure A1 of the Appendix Section A1). We consider data on daily mean temperature, precipitation, cloud cover, humidity, pressure, wind speed, and wind direction. All these weather variables could influence the complex chemistry of air quality and are commonly used in the literature on air quality. Again, we aggregate all readings to week-level observations. To account for the effect of weather on driving, we repeat this matching procedure for linking weather data to traffic monitors. Table 2 shows summary statistics on key meteorological variables. Due to the matching algorithm of weather conditions to air quality observations, the observation unit in Table 2 is at the pollution monitoring station level.[78] In our data, temperature is measured in degrees Celsius, precipitation in tenths of millimeters, cloud cover in okta,[79] daily sunshine in hours, pressure in hectopascal, humidity in percentage terms, wind speed in tenths of meters per second and wind direction is indicated by eight equally sized bins.

3.2. Consumption Spending Data

Our final database contains data at the credit card transaction-level from a large European bank.[80] The original data set is unique in that it details the date of each transaction, the zip code of residence of the credit-card owner (buyer-zip code) and the zip code of the selling establishment (seller-zip code).[81] Due to our confidentiality agreement with the provider of the data (the bank hereafter), we aggregate transaction information at the buyer-and seller-zip code-week level from the first week of 2015 to the tenth week of 2019. Figure 2 shows all 126 zip codes in Madrid: 6 zip codes belong to the MC area, 50 zip codes to the rest of the city of Madrid, and 70 zip codes are outside the city of Madrid but inside the metropolitan area of Madrid. See that those zip codes (even if partially) inside the MC area appear in black, zip codes outside the MC area and inside the city appear in orange, and violet zip codes are those outside the city of Madrid but inside the greater metropolitan area.

Table 3 presents summary statistics for observations at the seller-buyer-week level. Note that in our data the average value of "trade flows" between zip codes is 2,087 Euros in 54

[78] Descriptive statistics of weather data at the traffic monitor level are reported in Table A1 of the Appendix Section A2.

[79] According to this measure, 0 indicates no clouds and 8 full cloud cover.

[80] For simplicity, hereafter we refer to credit-card transactions, but these include both credit and debit card transactions. The raw data includes all credit-card transactions of consumers living within the metropolitan area of Madrid that are made, either online or offline, in establishments within the metropolitan area of Madrid with a credit-card of the bank provider of the data. Approximately, the data covers 15% of all transactions in the area, and can be considered as a representative sample of the total credit-card purchasing behavior. Galdon-Sanchez et al. (2020) provides a detailed description of the database and its variation.

[81] A zip code in our context is equivalent to a 5-digit zip code in the US.

transactions. Of those, 1,955 Euros from 51 face-to-face transactions (that is, brick-and-mortar transactions, B&M in the tables) and 132 Euros come from 3 online transactions. It is also important to highlight that 23% of observations are zero in brick-and-mortar transactions and 71% of observations are zero in online transactions.

Another unique feature of our data is that we are able to separate transactions into two types: brick-and-mortar transactions (B&M in the tables) and online transactions. This is an important feature because it allows us to test the transportation cost mechanism given that transportation costs increase for brick-and-mortar transactions and they do not for online transactions. Introducing this additional level of heterogeneity enriches the substitution patterns between zip codes within and outside the MC area. On the one hand, when consumers' demand for brick-and-mortar transactions is elastic, higher transportation costs will prompt consumers residing outside the MC area to substitute their former purchases in the MC area for purchases in other areas. On the other hand, those consumers with inelastic demand for products from a specific treated zip code may substitute to online transactions. This second scenario is more likely when the retailer sells a differentiated good and, therefore, it is costly to find a suitable brick-and-mortar transaction substitute outside the MC area.

Table 4 presents basic summary statistics related to selling and purchasing patterns across zip codes in Madrid. The top half of Table 4 details summary statistics of our data at the seller-zip code level. We can see how the share of revenue coming from online sales changes across zip codes in different areas. While zip codes in the MC area produce 85.6% of their revenue from brick-and-mortar sales, the percentage increases for zip codes in the rest of the city of Madrid and outside the city (90% and 95%). Moreover, the mean value of brick-and-mortar and online transactions also changes across zip codes. Finally, the last two rows in the top half show the shares that establishments in the MC area are selling in other areas of Madrid. Not surprisingly, we see that brick-and-mortar sales are tilted towards consumers in their own zip code. Zip codes in the MC area sell, on average, 5.62% of their sales to each of the zip codes in the MC area but only 0.21% to each of the zip codes outside the city of Madrid. Geographical proximity also matters for online sales (as documented by Blum and Goldfarb, 2006). On average, 1.84% of the online sales from zip codes in the MC area go to each of the 6 zip codes in this area; 1.34% go to each of the 50 zip codes in Madrid city and only 0.4% to each of the 70 zip codes outside of the city of Madrid.

The bottom half of Table 4 reports statistics on consumer behavior by buyer-seller-zip code dyad. Consumers living in the MC area make 45.2% of their brick-and-mortar purchases and 20.3% of their online purchases in establishments inside their area. These shares decrease monotonically with the distance to MC. Consumers in other zip codes of the city of Madrid make, on average, 8.9% of their brick-and-mortar purchases and 18.7% of their online purchases in establishments within the MC area. For consumers living outside the city, these numbers decrease as far as 4.3% of brick-and-mortar purchases and 14.8% of online purchases. The last two rows show how much consumers living in different areas of Madrid spend within their local zip code. So far, as the share of brick-and-mortar sales that consumers make in their local zip code is concerned, we see that consumers outside the city tend to spend more (38.7% of their total brick-and-mortar expenditures) than consumers elsewhere. By contrast, we do not see large differences across areas in the propensity to buy online in the local zip code (13.1% for consumers in MC, 15.1% for consumers in the city, and 14.3% for consumers outside the city).

4. Analysis of the Effect of MC on Car Traffic and Air Quality

The main goal of MC is to reduce traffic in the city center of Madrid and thereby lower air pollution. In this section, we study whether the policy achieved that goal. MC focuses on the reduction of NO_2, a pollutant mainly emitted by vehicles, as the city of Madrid repeatedly violated NO_2 limit values defined by the European Union environmental regulation. After defining our empirical strategy, we show our results of the impact of MC on traffic and air pollution.

4.1. Empirical Strategy

We estimate the effect of MC on traffic or NO_2 levels using the following regression equation,

$$Y_{swy} = \beta\ MC_{swy} + \delta X'_{swy} + \mu_{sw} + \tau_{wy} + \varepsilon_{swy} \tag{1}$$

where Y_{swy} stands for the traffic or pollution outcome of interest at the traffic or air quality monitor station *s*, week *w*, and year *y*. It is important to note that the traffic and air quality monitors are not identical. The variable MC_{swy} is a dummy that takes value one if station *s* is inside the MC area in a year-week in which MC is in effect. The vector X'_{swy} includes controls for meteorological conditions at the location of station *s*, week *w*, and year *y*. Therefore, the coefficient δ captures the effect of weather on air pollution levels.[82] For example, these would control for the case that the introduction of MC coincided with the wind blowing from a direction that induces lower pollution levels in the MC area. Moreover, we include station-week fixed-effects μ_{sw} to control for seasonal-specific patterns at each monitoring station. This set of fixed effects would control for instance for the case that during the Christmas season many shoppers go to the city center increasing traffic and pollution levels. The variable τ_{wy} controls non-parametrically for time trends and year-week-specific shocks. This variable would control, for example, for the celebration in Madrid of specific events attracting many visitors to the city and affecting pollution levels. The error term ε_{swy} is potentially serially correlated, so we cluster standard errors at the station level. By using this specification, we aim to consistently estimate the effect of MC on air pollution, captured by β, while controlling for possible confounding factors.

Our estimation strategy requires common trends in treated and untreated stations once we account for all control variables. This could fail, for instance, if people living in the MC area were substituting their old cars for electric vehicles at a faster pace than people in other areas of Madrid were. To account for this, we also allow for unit-specific trends. Our estimates could still be compromised if there were other policies introduced at the same time as MC, affecting traffic or pollution levels in specific areas of the city. If, for instance, a metro line covering the city center opens at the same time as the introduction of MC, we could wrongly attribute the

[82] This includes second order polynomials of temperature, precipitation, cloud cover, humidity, pressure and wind speed, as well as eight wind direction categories interacted with station indicators.

metro's positive effect on air quality to MC. We are not aware of any policy change or intervention of this type during the time span of our data set.[83]

4.2. Results

Table 5 presents the results of estimating equation (1) for the three measures of traffic, in levels and logs, with standard errors clustered at the station level. We find large effects of MC on traffic. The average number of cars dropped by 48.5 (column 1), or 14.7% (column 4). MC reduced the frequency of road segment usage by cars by 1.9 percentage points (column 2), or 17.8% (column 5). A decrease in these two measures implies that roads are used less, in fact, road capacity utilized under MC decreased by 5.7 percentage points (column 3), or 22.8% (column 6). These are not only statistically significant at the 1% level while clustering at monitor level, but also are economically significant magnitudes. Because those that cannot enter the restricted area may park in areas close by, or not drive to the center at all, MC might generate spatial spillover effects (either positive or negative) in traffic levels to nearby areas. In fact, our initial regression specification may be overestimating the decrease in traffic in the restricted area. To account for the spillover, we include a dummy variable in equation (1) that takes value one if station *s* is inside a 1.5 km buffer around the MC area in a year-week in which MC is enforced.[84] Table A2 in Section A3 of the Appendix shows that the net spillovers are positive, i.e. that traffic is also reduced in streets close to the regulated area. As expected, the magnitude of the reduction is smaller than inside the MC area. One can also see that not accounting for positive spillover effects leads to an underestimation of the absolute effect on traffic inside the MC area.

Table 6 presents the results on air quality. We cluster the standard errors in all specifications at the air quality monitor level.[85] In column 1, we use the log of the average weekly level of NO_2 as the dependent variable. Our findings suggest a decrease of 16% in NO_2 in the restricted area due to the introduction of MC. Defining the three closest stations inside the 1.5 km buffer around the MC area (see Figure A1 in Section A1 of the Appendix) as its immedicately adjacent area, the results in column 2 show that (i) the estimated reduction in pollution levels in the MC area remains unchanged, and (ii) there is no evidence of net spillovers to adjacent areas. This result could be due to the fact that the reduction of traffic in the surroundings is not strong enough or that the fleet composition outside MC differs, so that it fails to have a significant effect on air quality. We repeat the same exercise in column 3, considering spillovers to any station within the city of Madrid. The effect on the MC area is now slightly smaller in magnitude, but we find no evidence of spillovers neither towards adjacent areas nor to areas in the rest of the city. These results cannot be compared to the results on traffic, as traffic outside the city of Madrid is unobserved.

In columns 4 to 6, we show results of running the same specification with a different dependent variable, namely, the share of days of a week in which NO_2 levels exceed 40 µg/m³. Our findings here are consistent with those in columns 1 to 3, suggesting a decrease of 12 percentage points in the days of a week in which NO_2 levels exceed 40 µg/m³. This represents

[83] In January 2019 the City Council of Madrid reduced the speed limit on highway M-30 in order to decrease pollution levels. As this route does not cross the MC area, if anything, we would expect the impact of the policy to decrease pollution levels in the control group.
[84] Results are robust to defining alternative buffers around the MC area. These are available upon request.
[85] Results are similar when clustering at monitor and year-week level. These are available upon request.

a 25% reduction relative to the sample mean. These results are confirmed by Pseudo-Maximum Likelihood Poisson regressions in Table A3 of Section A3 in the Appendix, where the outcome is the number of days in a given week in which the limit value was exceeded.

4.3. Robustness checks

Our results appear to remain unchanged both qualitatively and quantitatively when including station-specific trends (Table A4, Section A3 of the Appendix). In a separate specification, we also include the air quality monitoring stations located in Barcelona as a control group and find that results remain mostly unchanged, except for a small reduction in the impact on the share of days exceeding 40 µg/m³ (Table A5, Section A3 of the Appendix).

Finally, as only a single station is treated and the number of clusters is relatively small (33), we also implement a Synthetic Control Method to estimate the impact of MC on air quality in downtown Madrid (Abadie et al., 2010). The station located inside the MC area is matched to a number of monitors outside the MC area based on pre-treatment data of air quality. Each control monitor receives a certain weight, such that the weighted mean of the control monitors' readings predicts air quality at the treated monitor. The algorithm chooses weights to minimize the mean squared error of these predictions. While one could try to find optimal weights by predicting every single observation of air quality at the treated monitor prior to intervention, we only choose a subset of NO_2 readings to be matched. From the beginning of 2015 to mid-2018, i.e. the 25th week of 2018, we only consider air quality from every 20th week to avoid overfitting. After that, we consider all readings until the 47th week of 2018. Treatment begins in the 49th week of 2018. In addition, we also match on the pre-treatment average of NO_2. We do not make use of weather controls as additional matching variables since, by construction, most stations face almost exactly the same weather conditions. Before running the algorithm, we deseasonalize each station's data. The matched stations provide good predictions of pre-treatment NO_2 concentrations at the station inside the MC area with an R-squared of 0.88.

Figure 3 shows the effect on the treated station (in bold). It seems that, at the beginning, MC was not yet effective. However, after some weeks, it decreased NO_2 levels by close to 50%. We cannot calculate standard errors, but repeat the analysis with a placebo treatment for each other monitor (in grey). Comparing the results from these stations, we see that the 50% drop can be interpreted as an unusually large deviation. Abadie et al. (2010) suggest that an effect is significant if the estimated effect of the treated unit is unusually large compared to the distribution of placebo estimates. They propose that one should not simply compare mean squared prediction errors of treated and placebo units in the post-treatment period, but scale these errors by the respective mean squared prediction errors in the pre-treatment period. In our case, we find that the ratio of mean squared prediction errors of the treated air quality station is larger than the ratios of all 32 control stations.

5. The Effect of Madrid Central on Consumer Spending

The results in section 3 appear to indicate that MC achieved its goal of reducing car traffic and pollution levels in the city center of Madrid. However, this may come at the cost of distorting citizens' habits and market outcomes. One of the most salient and controversial dimensions of these distortions is the possible impact that MC may have had on consumption behavior. An increase in the cost of transportation to the MC area can potentially discourage consumption in that area. In this section, we empirically examine whether MC actually affected consumer behavior, and if so, how. Understanding the costs of pollution-reducing policies is as important as evaluating their benefits and, therefore, the results of this section may help policy makers derive conclusions for the introduction of similar policies in the future.

To motivate our theoretical framework and our regression analysis, we start by showing unconditional means of total sales for seller zip codes inside and outside the MC area before and after the MC regulations went into effect. Table 7 shows that sales at the seller-zip code level are higher both inside and outside the MC area after the introduction of MC, but the proportional increase is smaller in zip codes inside the MC area than in zip codes outside the MC area (22% increase in the former vs. 28% increase in the latter). Hence, there is then a relative decrease in sales in MC area after the introduction of MC. However, in order to claim that this decrease in sales can be attributed to the introduction of MC, it is necessary to carry out a more rigorous analysis controlling for potential cofounders.

Following this finding, we develop a theoretical framework that models separately spending by seller and buyer zip code dyads. Our empirical analysis will follow this structure by aggregating the raw credit card transaction data in two meaningful ways. First, we aggregate transaction data at the seller-zip code-week level to follow traditional difference-in-differences estimation methods. By doing so, we study the impact of MC on sales for zip codes within MC area taking as control group all other zip codes in the Metropolitan Area of Madrid outside MC area. Second, we aggregate transactions at the week level for each combination of seller-zip code and buyer-zip code dyad available in the data. The resulting data set contains weekly information on how much consumers of each zip code are buying from sellers of each zip code in Madrid.[86] Our theoretical framework yields predictions of the impact of an increase transportation costs (actual transportation costs or disutility through inconvenience) for consumers living outside MC area when they make purchases of goods and services from businesses within MC area. In our context, the MC policy should not affect: (i) purchases of residents from the MC area in businesses within the MC area; (ii) purchases of residents outside MC in businesses outside the MC area, as the regulation only restricts traffic inside the MC area; and (iii) purchases of residents from the MC area in businesses outside the MC area. In other words, we are able to clearly define "trade flows" affected by the policy (treatment group) and those unaffected (control group). Therefore, the predictions from our theoretical framework and out empirical analysis allow us to identify the impact of the increase in transportation costs for those affected, whilst controlling for demand shocks and supply shocks at different zip codes from the MC area towards other areas of the city.

[86] These data structure is comparable to those data used in the international trade literature for the estimation of gravity equations (Head and Mayer, 2014; Atalay et al., 2019). Analogously to the trade literature, our data allows us to study how "trade flows" between different geographical areas change when transportation costs change exogenously.

5.1. Theoretical framework and identification strategy

We build our identification strategy around a theoretical framework based on seminal work of Anderson (1979), Eaton and Kortum (2002) and Baier and Bergstrand (2007) using a standard gravity model. Assume a city with *N* zip codes, and each zip code has buyers and sellers. For the sake of simplicity, we will consider buyers indexed by their zip code *i* = 1, …, *N* and sellers indexed by their zip code *j* = 1, …, *N*. The sellers in each zip code sells a differentiated item from all other items sold in other zip codes. Buyers may choose to buy items from any zip code, and sellers can sell to buyers from any zip code. While this is effectively a static model, we allow for multiple periods indexed according to their week of the year *w* = 1, …, *W*, and their year *y* = 1, …, *Y*.

Consider then a representative consumer model with a CES demand function in which the buyer residing in zip code *i*, and week *w* of year *y*, has to decide how much to buy from each of the seller zip codes *j* (Q_{ijwy}). There is a seasonal (weekly) taste specific shock θ_{ijw} at the level of the buyer-seller-week. A seller of zip code *j* cannot price discriminate across different buyers and therefore sets a price P_{jwy} common to all buyers. Moreover, buyers have to pay iceberg transportation cost τ_{ijwy}. Because we want to study the impact of the introduction of MC on trade flows between zip codes, we allow transportation costs to vary at the buyer-seller-week level. In our case, we hypothesize that the introduction of MC will affect the purchases in zip codes inside the MC area from buyers in zip codes outside of MC area. Therefore, the objective function U_{iwy} is the following:

$$U_{iwy} = Max \left\{ \left[\int \theta_{ijw} \; Q_{ijwy}^{\frac{\sigma-1}{\sigma}} \; dv_j \right]^{\frac{\sigma}{\sigma-1}} - \int \left(\tau_{ijwy} P_{jwy}\right) Q_{ijwy} \, dv_j \right\}$$

where each consumer maximizes its consumer surplus with respect to Q_{ijwy} taking their preferences, prices and other parameters as given.

Let $\tilde{P}_{iwy} = \left[\int \theta_{ijw} \; \left(\tau_{ijwy} P_{jwy}\right)^{1-\sigma} dv_j \right]^{\frac{1}{1-\sigma}}$ be the price index of buyer *i*, in week *w* of year *y*. Let also $\tilde{Q}_{iwy} = \left[\int \theta_{ijw} \; Q_{ijwy}^{\frac{\sigma-1}{\sigma}} \, dv_j \right]^{\frac{\sigma}{\sigma-1}}$ be the total amount consumed by buyer *i*, in week *w* of year *y*.

Then, the total value of consumption by buyers residing in zip code *i*, in establishments of sellers in zip code *j*, in week *w* of year *y* will be equal to

$$P_{jwy} \, Q_{ijwy} = \left(\tilde{P}_{iwy}^{\sigma} \; \tilde{Q}_{iwy} \right) \left(P_{jwy}^{1-\sigma} \right) \left(\theta_{ijw}^{\sigma-1} \right) \left(\tau_{ijwy}^{-\sigma} \right)$$

Here we can see how an increase in transportation costs τ_{ijwy}, like the one induced by the introduction of MC, will reduce consumption levels.[87] Moreover, this expression can be

[87] The increase in transportation costs induced by the introduction of MC will have a direct impact on the level of purchases from buyer-zip codes outside of MC area in establishments inside the MC area. In turn, if the reduction of consumption in MC area has spillovers in consumption levels in other zip codes, these should be controlled for by the fixed-effects structure. We will not be able to separate this indirect effect of the introduction of MC from aggregate shocks at the buyer-zip code level. Note, however, how this impact should be economically small if the

mapped one-to-one (using logs) to the following equation that we will actually estimate with our data,

$$Y_{ijwy} = \alpha_{iwy} + \gamma_{jwy} + \delta_{ijw} + \beta \, MC_{ijwy} + u_{ijwy} \quad (2)$$

where Y_{ijwy} measures (log) expenditures of residents in zip code *i* in establishments in zip code *j* during week *w* of year *y*. The variable MC_{ijwy} is a dummy variable that takes value one if *i* is a buyer-zip code outside the MC area, *j* is a seller-zip code inside the MC area, and we are in a week-year in which the MC regulations are in effect. Note this dummy is aimed to capture increases in transportation cost between a zip code pair triggered by the introduction of MC, and that β is the coefficient of interest as it measures the effect of MC on purchases of buyers from outside the MC area in establishments inside the MC area once the policy is in effect. Additionally, α_{iwy} is the buyer by week fixed-effect, and γ_{jwy} is the seller by week fixed-effect.[88] The variable δ_{ijw} is the buyer-by-seller fixed-effect specific for each week of the year. We allow this dyad-specific fixed-effect to vary by the week of the year to account for seasonality patterns (e.g. during Christmas time people living in the outskirts of the city may disproportionately increase their shopping in the city center). Finally, u_{ijwy} is the error term.

As a result, through specification (2) we aim to identify the effect of MC on spending levels from buyers living in zip codes outside the MC area in establishments inside the MC area, both relative to the shopping of these same consumers in other areas of the city and relative to the shopping in downtown Madrid of consumers living within MC area. The validity of this identification strategy hinges on the existence of parallel trends across different buyer-zip codes in the share of purchases they make in each seller-zip code. The identification strategy would be compromised if people living in different zip codes change their taste for shopping in the MC area over time in different ways. We will present a number of robustness tests and a falsification exercise to test the validity of the identification strategy.

An interesting departure from specification (2) is one where we split the dummy of interest, MC_{ijwy}, into two dummies: *MC1* for those buyers living in zip codes within the limits of the city of Madrid (and outside the MC area), and *MC2* for those zip codes located outside the city of Madrid and still inside the metropolitan area of Madrid. This alternative specification is aiming to account for plausible large differences in the set of available transportation options to the center of the city of Madrid. Because the policy restricts driving into the city center, consumers now could consider other means of transportation. Those living in the zip codes closer to the MC area are more likely to switch costlessly to move on foot or by public transport. By contrast, consumers living further away from the MC area might find it more difficult to substitute their car for other means of transportation, as they are not able to move

number of zip codes is large enough. We have 126 zip codes, which should make our case comparable to the usual International Trade framework modeling trade across countries.

[88] In standard trade models, the buyer-zip code-week fixed effect α_{iwy} and the seller-zip code-week fixed effect γ_{jwy} would correspond to the importer by period and exporter by period fixed effects, respectively. On the one hand, α_{iwy} controls for changes over time in the average level of expenditures of people living in zip code *i.* For example, if consumers in buyer-zip code *i* often go shopping to seller-zip code *j* that is in the MC area (for instance, because people in zip code *i* predominantly work in the city center) and these consumers are getting richer and spending more money over time, this would be captured by the buyer-week-year fixed effect. On the other hand, γ_{jwy} controls for the possibility that zip codes in the MC area were becoming less attractive over time, as this could bias our estimates.

on foot, and access to public transport is less convenient.[89] Therefore, one would expect the impact of MC on consumers living far away from the city to be more severe than for those living closer to the MC area. To test this hypothesis, we modify equation (2) as follows:

$$Y_{ijwy} = \alpha_{iwy} + \gamma_{jwy} + \delta_{ijw} + \beta_1 MC1_{ijwy} + \beta_2 MC2_{ijwy} + u_{ijwy} \tag{2'}$$

In specification (2'), β_1 and β_2 are the coefficients of interest as they measure the heterogeneity of the effect of MC on purchases of buyers outside the MC area. While the former (β_1) measures the effect of MC on buyer-zip codes residing outside the MC area but inside the city of Madrid, the latter (β_2) measures the effect of MC on buyer-zip codes outside the city of Madrid. All other variables, fixed effects and subscripts are the same as explained above in specification (2). Similarly, u_{ijwy} is the error term.

We must highlight the importance of having access to such geographically disaggregated data, namely sales at the buyer zip code and seller zip code pair level. To do so, we compare our approach above to the bread-and-butter difference-in-difference analysis that we would pursue if our data was aggregated at the seller zip code level. Let us integrate both left-hand and right-hand sides of specification (2) across all buyer zip codes *i* within each seller zip code *j*. The result would be an expression similar to specification (1) in the previous section,

$$Y_{jwy} = \alpha_{wy} + \gamma_{jwj} + \beta\, MC_{jwy} + u_{jwy} \tag{3}$$

where Y_{jwy} measures how much sellers in zip code *j* sell in week *w* of year *y*. The variable MC_{jwy} is a dummy variable that takes value one if seller *j* is in a zip code inside the MC area and in a week-year in which MC is in effect. The parameter β is the coefficient of interest as it measures the effect of MC on sales of establishments within the MC area once the policy is in effect. The parameter α_{wy} is the week-year fixed effect, and γ_{jwj} is the week-year specific seller-zip code fixed effect. Finally, u_{jwy} is the usual error term. Note that γ_{jwj} and MC_{jwy} have the same level of variation and this therefore makes identification of β impossible unless we make certain identification assumptions. To this end, we break γ_{jwj} into a seller zip code fixed effect and a seller zip code specific time trend (γ_j + $\theta_j t_j$) so that identification of the effect occurs out of the zip code specific level and trend.

This specification does not control for potential different demand shocks for different seller zip codes. For instance, if buyers purchasing intensively in seller-zip codes inside the MC area become richer over time and demand more over time, our identification strategy would be compromised. It does not control either for different supply shocks in different areas of the

[89] To check that this is actually the case, we calculate the penalty associated with using a car relative to public transport from each zip code in Madrid to travel to the MC area. For this purpose, we use the Google Maps Distance API and record travel times by car and public transport from the centroid of each zip code to the centroid of MC. We then calculate how much longer it takes to use public transport compared to using the car. In Table A6, Section A4 of the Appendix, we can see the "penalty" of public transport usage is 12 minutes longer for zip codes outside the city of Madrid compared to zip codes inside the city of Madrid. The difference is significant at the 1% level.

city. If it is just the case that the MC area becomes less attractive for consumption over time because many establishments have closed, our identification strategy again would be compromised. These are all concerns that specifications (2) and (2') can easily overcome with the use of time-varying seller and buyer zip code fixed effects and another reason to prefer our gravity approach over the traditional difference-in-difference estimation.

Having gone over our specification, we take our empirical methodology to the data in the next subsection. Our empirical strategy follows three stages. First, we replicate the difference-in-differences approach in Section 3 by examining the results of regressions following specification (3) above with sales at the seller-zip code-week level as dependent variable. Second, we estimate "gravity-like" regressions following specification (2) in which we exploit variation in consumption behavior across buyer-seller-zip code pairs over time. Under this framework, we are able to measure "trade flows" between consumers and businesses located in different zip codes of Madrid. Third and finally, we run specification (2') to test for differences in changes in transportation cost in zip codes inside and outside the city of Madrid.

5.2. Results

Following the empirical strategy described in the previous section, we proceed now by showing estimates of β from the traditional difference-in-differences specification (3) in the previous section, and therefore replicating the methodology used before in Tables 5 and 6 in Section 4. The regression specifications in Table 8 use log revenue and log transactions aggregated at the seller-zip code-week level as dependent variables.

Columns 1 and 2 of Table 8 show a decrease of 9% in revenue and transactions, but this effect is statistically insignificant. Note that this is a rather convenient result for policy makers confronted with opposition by local business in the MC area: MC reduced traffic congestion and air pollution with no statistically significant impact on economic activity. In columns 3 to 6, we continue our analysis by breaking the revenue and number of transactions at the seller-zip code level into brick-and-mortar and online transactions. Interestingly (and counterintuitively), we find that both types of transactions decrease, around 12% for brick-and-mortar transactions and 28% for online transactions. Yet, the effect on brick-and-mortar transactions is statistically insignificant, and only the effect on online revenue and number of transactions appears to be statistically significant at the 5% level. Because we had expected that online transactions are less likely to be affected by the larger transportation costs imposed by MC, these results are inconsistent with our predictions and therefore deserve further investigation.

We proceed next with our "gravity-like" methodology. Because the outcome variables in this section are measured in logs and the trade flows between two zip codes in a given week can be zero, we add the value one to the dependent variable of interest throughout this section.[90] Before running regressions with specification (2), we confirm that "trade flows" across zip codes within the metropolitan area of Madrid actually exhibit gravity.[91] This is important for two distinct and yet equally needed purposes. First, we motivate that transportation costs within a city matter. Second, we motivate the need to control for seller-buyer zip codes pair

[90] Up to 14.1% of the dyad-week flows are zero in our sample. This is substantially lower than in usual setups of country trade flows where there are around 50% of zeros (Helpman et al, 2008).
[91] Table A8 in Section A4 of the Appendix confirms the existence of gravity in our sample.

fixed effects to clear problems of endogeneity in the data. The former is a serious problem because MC is increasing the transportation costs of those buyers located further away from the MC area to a greater extent, and because those buyers located furthest away from the MC area were buying little in the MC area to start with. Therefore, we must include buyer-seller zip codes fixed effects to avoid negative biases in the estimation of the effect of MC on consumption spending.[92]

Once we have established the existence of within-city gravity, we proceed to estimate β from specification (2) and showing results of our triple diff estimation in Table 9.[93] In columns 1 and 2, we use total transaction revenue as the dependent variable and find consistent results with those findings in Table 8. While we find that MC is associated with a decrease of 3 percentage points in revenue, these are all statistically insignificant effects. Columns 3 to 6 examine the impact of MC on brick-and-mortar and online transactions separately.[94] While MC decreases brick-and-mortar transactions between 4.7 and 8.9 percentage points, it increases online transactions between 9.4 and 12.1 percentage points. All four columns offer statistically significant findings. Therefore, these results suggest that, upon the increase in transaction costs due to the implementation of MC, consumers in zip codes outside the MC area switched consumption from brick-and-mortar interactions to online transactions.[95]

Finally, we investigate the incidence of heterogeneity of the impact of MC across buyer and seller zip code pairs. To do so, we estimate regression specification (2') above and show our results in Table 10. Columns 1 and 2 show that MC did not have a statistically significant impact on total revenue or the number of transactions, neither in the city zip codes nor in

[92] Note that Table A9, Section A4 of the Appendix, replicates results of upcoming Table 9 without adding buyer-zip code seller-zip code pair fixed effects and shows the importance of controlling for the underlying variation across zip code pairs that may negatively bias our estimates.

[93] At this point, it is important for us to show that our identifying assumption holds, that is, only the purchases of buyers outside the MC area from establishments inside the MC area are affected by MC. To show that this assumption is supported by the data, we may think of replicating specification (2) explicitly including in the regression equation all possible effects (buyer outside the MC area-seller inside the MC area, buyer and seller outside the MC area, and buyer inside the MC area-seller outside the MC area; while keeping as a baseline the group of buyer inside the MC area-seller inside the MC area). Yet, this exercise presents identification problems due to the original set of fixed effects imposed in our specifications in Table 9. Because our baseline specification is a triple difference-in-differences, all these group-specific fixed effects cannot be identified while keeping buyer-week-year FE and seller-week-year FE. Instead, in addition to buyer-seller-year FE and week-year FE, we can just include buyer-week FE and seller-week FE (notice that actually the buyer-week FE and seller-week FE are subsumed in the buyer-seller-week FE.). Yet, with this new set of fixed effects we are not able to control for changes in a seller zip code over time (for instance, if the MC area becomes more or less attractive over time and there are establishments opening or closing). And we are not controlling either for demand shocks in buyer zip codes (the population of a neighborhood can be changing over time). These are important concerns for identifying the impact of MC that we are able to address in the baseline identification. For that reason, we take a different route to show the validity of our identifying assumptions. We split the sample into two distinct subsamples: purchases from buyers inside and outside the MC area. We run traditional difference-in-difference specifications for each sample and show that there are no changes in revenue sales in those buyers within the MC area while buyers outside the MC area decrease their sales from establishments inside the MC area. Results are available upon request.

[94] Table A13, Section A4 of the Appendix, shows no statistically significant results for the mean transaction value overall, brick-and-mortar transactions and online transactions.

[95] Table A10, Section A4 of the Appendix, regresses share of online revenue, share of online transactions and ratio of transaction values on our treatment dummy and finds consistent results with those in Table 9.

those zip codes outside of Madrid. These findings are consistent with those shown in Tables 9 and 10 when the dependent variable collects all transaction types.

Next, we examine heterogeneity on the impact of MC on brick-and-mortar revenues and number of transactions. The results in columns 3 and 4 show that brick-and-mortar revenues and transactions decrease across the board. The magnitude of the decrease in revenues and transactions is larger for buyer-zip codes outside the city of Madrid, although we cannot statistically reject the hypothesis that they are the same. Interestingly, we find opposite findings regarding online transactions. Columns 5 and 6 show that online revenues and transactions increased across the board. The magnitude of the increase is larger for revenues and transactions in buyer-zip codes outside the city of Madrid, where the increase in transportation cost to the MC area is likely to be largest. Moreover, we can reject the null hypothesis that the effects on online transactions are equal for buyer-zip codes inside and outside the city of Madrid.[96] Again, this indicates a substitution between brick-and-mortar and online consumption due to the increase in transportation costs for residents outside the MC area.[97]

5.3. Robustness checks

A traditional concern when implementing difference-in-differences estimation is the potential existence of different pre-trends across the treatment and control groups. Therefore, different trends in the propensity of different buyer-zip codes to buy in each seller-zip code could invalidate the identification strategy in Table 9. To address this concern, in Table 11, we present results from a falsification test. We consider the introduction of MC taking place approximately one month and a half before it actually did (it was introduced in week 49 of 2018 and we assume it was introduced in week 43).[98] We would expect to capture possible trends, however the results show there is no evidence of different trends in any of the outcome variables. This result also makes us to rule out the existence of potential anticipatory effects that might induce consumers to bring forward consumption in MC area as a result of the inminent increase in transaction costs.

Another potential concern arises from the incidence of zeros in trade flows between some zip code pairs. So far, we have adopted the traditional solution of adding one to the dependent variable of interest to avoid dropping observations once we take logs. Table 12 shows robustness of this result to using Poisson Pseudo Maximum Likelihood. This method accommodates zero trade flows with no transformations of the dependent variables since these are in levels (Silva and Tenreyro, 2006). Column 1 shows very similar results for the

[96] Note that these findings also provide an alternative way of testing the impact of MC on online consumption levels of consumers living outside the city of Madrid by using as a control group online consumption levels of consumers inside the city of Madrid and outside the MC area.

[97] Our heterogeneity results are robust to a number of alternative specifications. Table A11, SectionA4 of the Appendix, shows results of dividing buyer zip codes into two groups, closer and farther than 6 kms from the MC area, as well as results of dropping observations from those zip codes more than 15 kms away from the MC area. Table A12, SectionA4 of the Appendix, includes specifications where buyer zip codes are weighed by their volume of sales. It also includes other specifications with week-year fixed effects interacted with the distance between each zip code and the MC area to control for changes over time in spending behavior correlated with geographic consumer location. Our results are qualitatively robust to these alternative specifications.

[98] Our findings in Table 11 are robust to using other placebo starting dates such as 4 or 8 weeks before week 49 of 2018. These results are available upon request.

effects on brick-and-mortar revenues. Column 2 finds a positive impact of MC on online revenues. Column 3 shows that the impact on total revenues is smaller in magnitude and statistically insignificant. This set of results is consistent with the main findings of the paper in Table 9.

A third set of robustness checks is concerned with the fact that differences in transportation costs may have changed differently across zip codes at the same physical distance. Table 13 presents further evidence consistent with the fact that an increase in transportation costs drives the substitution between brick-and-mortar and online consumption. Using the Google Maps Distance API, we calculate travel times by car and public transport from the centroid of each zip code to the geographic centroid of the MC area. We divide zip codes between those above and below the median of the increase in travel time. We observe that, although the decrease in brick-and-mortar sales is not statistically significant for zip codes with high and low travel time increments, those zip codes with a high increase in travel times drive the increase in online purchases.[99]

6. Conclusions

This paper analyzes the benefits and costs of the introduction of constraints to vehicle circulation in the center of Madrid to improve air quality. By restricting access by car, transportation costs increase for those consumers living outside the area affected by the policy, potentially discouraging consumption in that particular area. We show the regulation had the intended effect of reducing traffic congestion in the affected area, and consequently we observe a significant decrease in air pollution. This first set of results clearly states direct benefits from the implementation of MC in the city of Madrid.

However, our data allow for further investigation on the impact of the policy on economic activity. In particular, we use credit card transaction data from a large bank to examine whether the decrease in traffic congestion in the city center of Madrid caused lower revenues in local establishments due to fewer visitors. The granularity of our data grants the identification of purchases of all possible pairs of buyer zip codes and seller zip codes in the city of Madrid. Our findings show that there was not a statistically significant impact in total revenues and number of transactions due to the policy. Yet, when we separate brick-and-mortar and online transactions, we find that brick-and-mortar revenues and transactions decrease while online revenues and transactions increase. This substitution across transaction types occurs both across seller zip codes within the city and within seller zip codes. We are also able to show that the effect of the policy is larger for those zip codes where buyers face larger transportation constraints and therefore we confirm that the impact of the policy on business volume goes through an apparent increase in transaction costs for some customers and not others.

This paper shows that driving bans come at a cost for consumers and local brick-and-mortar commerce. While air quality improvements are significant and provide large benefits, brick-and-mortar commerce is negatively affected. Our results show that consumers substitute to online purchases, which could compensate the loss in brick-and-mortar commerce. However,

[99] Tables A11 and A12 present further robustness results.

these substitutions are usually made at different types of sellers so that a driving ban might have unintended distributional effects on smaller businesses.

Thus, our paper contributes to the literature in that it provides evidence of the impact of environmental policies on economic activity, more specifically, on revenues and number of transactions of establishments directly affected by the policy. Most importantly, we offer evidence that these effects are not homogenous and vary along different dimensions. The most novel result in our heterogeneity analysis is the important role played by e-commerce to attenuate the impact of environmental regulation, and its implication for policy makers regarding e-commerce and online transactions. Future research on the impact of environmental policies, regardless of the type of pollution regulated, should aim to provide direct evidence of their cost through diminished economic activity. Similarly, understanding the distributional effects of such policies is, to the best of our knowledge, a crucial part of the information necessary for the design of future environmental regulations and their policy implementation. Furthermore, our results speak about the relevant role that e-commerce may play in smoothing the impact of increases in consumer transportation costs generated by other factors than environmental regulations. It would be very interesting to study, for instance, how consumers resorted to online purchasing during the Covid-19 lockdown and how e-commerce adoption allowed establishments to weather the situation.

Bibliography

Abadie, A., Diamond, A. and Hainmueller, J. (2010). Synthetic Control Methods for Comparative Case Studies: Estimating the Effect of California's Tobacco Control Program. *Journal of the American Statistical Association*, 105 (490), 493-505.

Anderson, J. E. (1979): A Theoretical Foundation for the Gravity Equation. *American Economic Review*, 69(1), 106-116.

Atalay, E., Hortaçsu, A., Li, M. J., and Syverson, C. (2019). How wide is the firm border? *The Quarterly Journal of Economics*, *134*(4), 1845-1882.

Auffhammer, M., Hsiang, S. M., Schlenker, W., and Sobel, A. (2013). Using weather data and climate model output in economic analyses of climate change. *Review of Environmental Economics and Policy*, *7*(2), 181-198.

Baier, S. L., and Bergstrand, J. H. (2007). Do free trade agreements actually increase members' international trade? *Journal of International Economics*, *71*(1), 72-95.

Barahona, N., Gallego, F., and Montero, J.-P. (2019). Vintage-specific driving restrictions. Review of Economic Studies, 47.

Bharadwaj, P., Gibson, M., Zivin, J. G., and Neilson, C. (2017). Gray Matters: Fetal Pollution Exposure and Human Capital Formation*. Journal of the Association of Environmental and Resource Economists*, 4(2), 505–542.

Blum, B. S., and Goldfarb, A. (2006). Does the internet defy the law of gravity? *Journal of International Economics*, *70*(2), 384-405.

Boletin Oficial Ayuntamiento de Madrid (2018). Acuerdo de 29 de octubre de 2018 de la Junta de Gobierno de la Ciudad de Madrid por el que se desarrolla el régimen de gestión y funcionamiento de la Zona de Bajas Emisiones "Madrid Central". BOAM numero 8268/2199, 30 de Octubre de 2018.

Carillo, P.E., A. Lopez-Luzuriaga, and A. S. Malik (2018). Pollution or crime: The effect of driving restrictions on criminal activity. *Journal of Public Economics*, Volume 164, 50-69.

Carillo, P.E., A. S. Malik, and Y. Yoo (2013). Driving Restrictions That Work? Quito's Pico y Placa Program. *Canadian Journal of Economics*, 49(4), 1536-1568.

Chang, T., Graff Zivin, J., Gross, T., and Neidell, M. (2016). Particulate Pollution and the Productivity of Pear Packers*. American Economic Journal: Economic Policy*, 8(3), 141–169.

Chay, K. Y. and M. Greenstone (2003). The Impact of Air Pollution on Infant Mortality: Evidence from Geographic Variation in Pollution Shocks Induced by a Recession. *The Quarterly Journal of Economics*, 118, 1121 – 1167.

Chen, Y., G. Z. Jin, N. Kumar, and G. Shi (2013). The Promise of Beijing: Evaluating the Impact of the 2008 Olympic Games on Air Quality. *Journal of Environmental Economics and Management*, 66, 424 – 443.

Chen, Y. and A. Whalley (2012). Green Infrastructure: The Effects of Urban Rail Transit on Air Quality. *American Economic Journal: Economic Policy*, 4 58 – 97.

Currie, J., and Neidell, M. (2005). Air Pollution and Infant Health: What Can We Learn from California's Recent Experience? *The Quarterly Journal of Economics*, 120(3), 1003–1030.

Currie, J. and R. Walker (2011). Traffic Congestion and Infant Health: Evidence form E-ZPass. *American Economic Journal: Applied Economics*, 3, 65 – 90

Currie, J., and Walker, R. (2019). What Do Economists Have to Say about the Clean Air Act 50 Years after the Establishment of the Environmental Protection Agency? *Journal of Economic Perspectives*, 33(4), 3–26.

Davis, L. W. (2008). The Effect of Driving Restrictions on Air Quality in Mexico City. *Journal of Political Economy*, 116(1), 38–81.

Deryugina, T., Heutel, G., Miller, N. H., Molitor, D., and Reif, J. (2019). The Mortality and Medical Costs of Air Pollution: Evidence from Changes in Wind Direction*. American Economic Review*, 109(12), 4178–4219.

Eaton, J., and Kortum, S. (2002). Technology, geography, and trade. *Econometrica*, *70*(5), 1741-1779.

Ebenstein, A., Lavy, V., and Roth, S. (2016). The Long-Run Economic Consequences of High-Stakes Examinations: Evidence from Transitory Variation in Pollution. *American Economic Journal: Applied Economics*, 8(4), 36–65.

Eskeland, G. S. and T. Feyzioglu (1997). Rationing Can Backfire: The 'Day without a Car' in Mexico City. *The World Bank Economic Review*, 11, 383 – 408.

Fu, S., and Viard, B. (2015). The Effect of Beijing's Driving Restrictions on Pollution and Economic Activity. *Journal of Public Economics*, Vol. 125, 98 – 115.

Galdon-Sanchez, J.E., Gil, R., and Uriz-Uharte, G. (2020). The Value of Informaton in Competitive Markets: The Impact of IT and Big Data on Small and Medium Enterprises. *Working paper.*

Gallego, F., Montero, J.-P., and Salas, C. (2013). The effect of transport policies on car use: Evidence from Latin American cities. *Journal of Public Economics*, 107, 47–62.

Gehrsitz, M. (2017). The effect of low emission zones on air pollution and infant health. *Journal of Environmental Economics and Management*, 83(Supplement C), 121–144.

Graff Zivin, J., and Neidell, M. (2013). Environment, Health, and Human Capital. *Journal of Economic Literature*, 51(3), 689–730.

Greenstone, M. and R. Hanna (2014). Environmental Regulations, Air and Water Pollution, and Infant Mortality in India. *American Economic Review*, 104, 3038 – 3072.

Greenstone, M. and B. K. Jack (2015). Envirodevonomics: A Research Agenda for an Emerging Field. *Journal of Economic Literature*, 53(1), 5-42.

Hammitt, J. K. and Y. Zhou (2006). The Economic Value of Air-Pollution-Related Health Risks in China: A Contingent Valuation Study. *Environmental & Resource Economics*, 33, 399 – 423.

He, J., Gouveia, N., and Salvo, A. (2018). External Effects of Diesel Trucks Circulating Inside the São Paulo Megacity. *Journal of the European Economic Association*, 17(3), 947-989.

Head, K., and Mayer, T. (2014). Gravity equations: Workhorse, toolkit, and cookbook. In *Handbook of International Economics* (Vol. 4, pp. 131-195). Elsevier.

Helpman, E., Melitz, M., and Rubinstein, Y. (2008). Estimating trade flows: Trading partners and trading volumes. *The Quarterly Journal of Economics*, *123*(2), 441-487.

Leape, J. (2006). The London Congestion Charge. *Journal of Economic Perspectives*, 20(4), 157–176.

Li, S. (2014). Better Lucky than Rich? Welfare Analysis of Automobile License Allocations in Beijing and Shanghai. Working paper.

Li, S., Liu, Y., Purevjave, A., and Yang, L. (2019). Does subway expansion improve air quality? *Journal of Environmental Economics and Management*, Volume 96, 213-235.

Lin, C.-Y., W. Zhang, and V. I. Umanskaya (2011). The Effects of Driving Restrictions on Air Quality: São Paulo, Beijing, and Tianjin. Available athttp://purl.umn.edu/103381.

Neidell, M. J. (2004). Air Pollution, Health, and Socio-Economic Status: The Effect of Outdoor Air Quality on Childhood Asthma. *Journal of Health Economics*, 23, 1209 – 1236.

Oliva, P. (2015). Environmental Regulations and Corruption: Automobile Emissions in Mexico City. *Journal of Political Economy*, 123(3), 686–724.

Osakwe, R. (2010). An Analysis of the Driving Restriction Implemented in San José, Costa Rica. Environment for Development Policy Brief.

Quddus, M. A., Carmel, A., and Bell, M. G. H. (2007). The Impact of the Congestion Charge on Retail: The London Experience. *Journal of Transport Economics and Policy*, 41(1), 113–133.

Parry, I. W. H., M. Walls, and Harrington, W. (2007). Automobile Externalities and Policies. *Journal of Economic Literature*, 45, 373 – 399.

Rivera, N. (2017). The Effectiveness of Temporary Driving Restrictions: Evidence from Air Pollution, Vehicle Flows, and Mass-Transit Users in Santiago. 2017 Annual Meeting, July 30-August 1, Chicago, Illinois 259182, Agricultural and Applied Economics Association.

Rohlf, A., Holub, F., Koch, N., and Ritter, N. (2020). The effect of clean air on pharmaceutical expenditures. *Economics Letters*, 192, 109221.

Salas, C. (2010). Evaluating Public Policies with High Frequency Data: Evidence for Driving Restrictions in Mexico City Revisited. Catholic University of Chile Working Paper No. 374.

Silva, J. S., and Tenreyro, S. (2006). The log of gravity. *The Review of Economics and Statistics*, *88*(4), 641-658.

Simeonova, E., Currie, J., Nilsson, P., and Walker, R. (2019). Congestion Pricing, Air Pollution, and Children's Health. *Journal of Human Resources*.

Wolff, H. (2014). Keep Your Clunker in the Suburb: Low Emission Zones and Adoption of Green Vehicles. *Economic Journal*, 124, F481 – F512.

Ye, J. (2017). Better safe than sorry? Evidence from Lanzhou's driving restriction policy. *China Economic Review*, Volume 45, 1-21.

Zhang, W., Lin Lawell, C.-Y. C., and Umanskaya, V. I. (2017). The effects of license plate-based driving restrictions on air quality: Theory and empirical evidence. *Journal of Environmental Economics and Management,* 82, 181–220.

Zhong, N., Cao, J., and Wang, Y. (2017). Traffic Congestion, Ambient Air Pollution, and Health: Evidence from Driving Restrictions in Beijing. *Journal of the Association of Environmental and Resource Economists*, 4(3), 821–856.

Figure 1: Map of Madrid Central within the City of Madrid

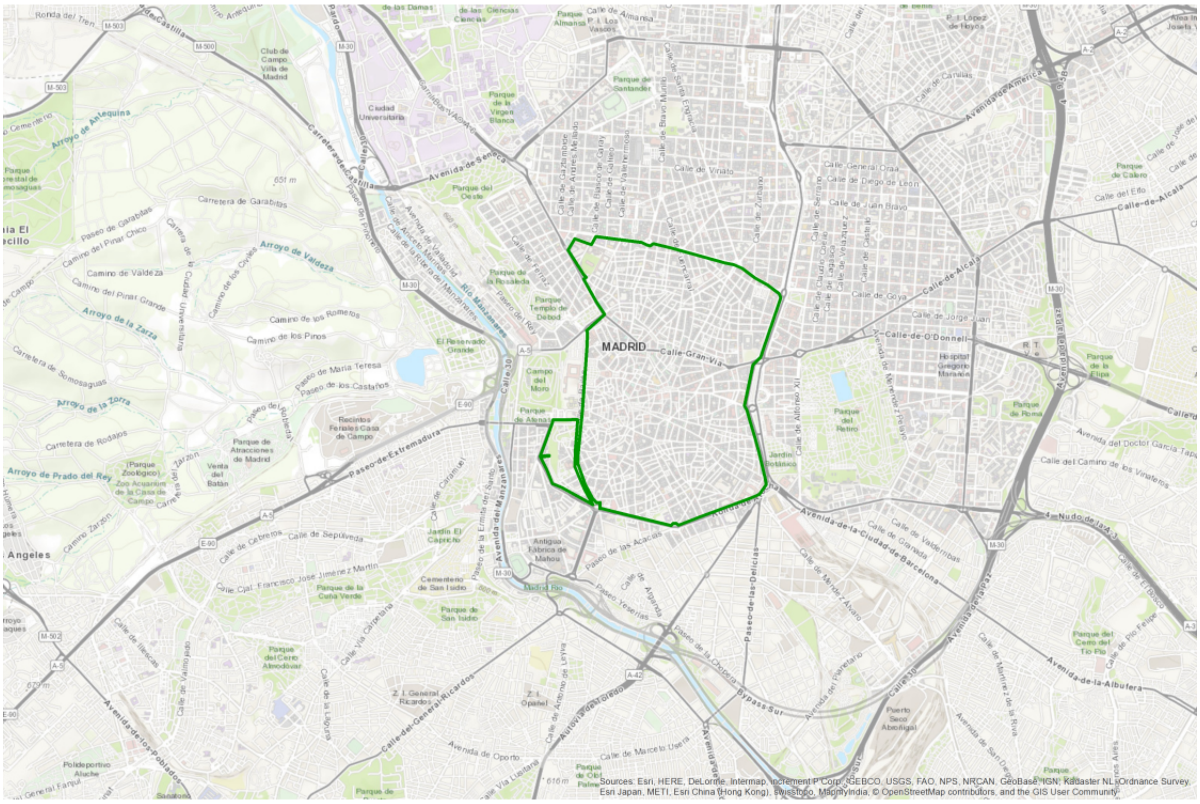


Figure 2: Map of zones and postal codes

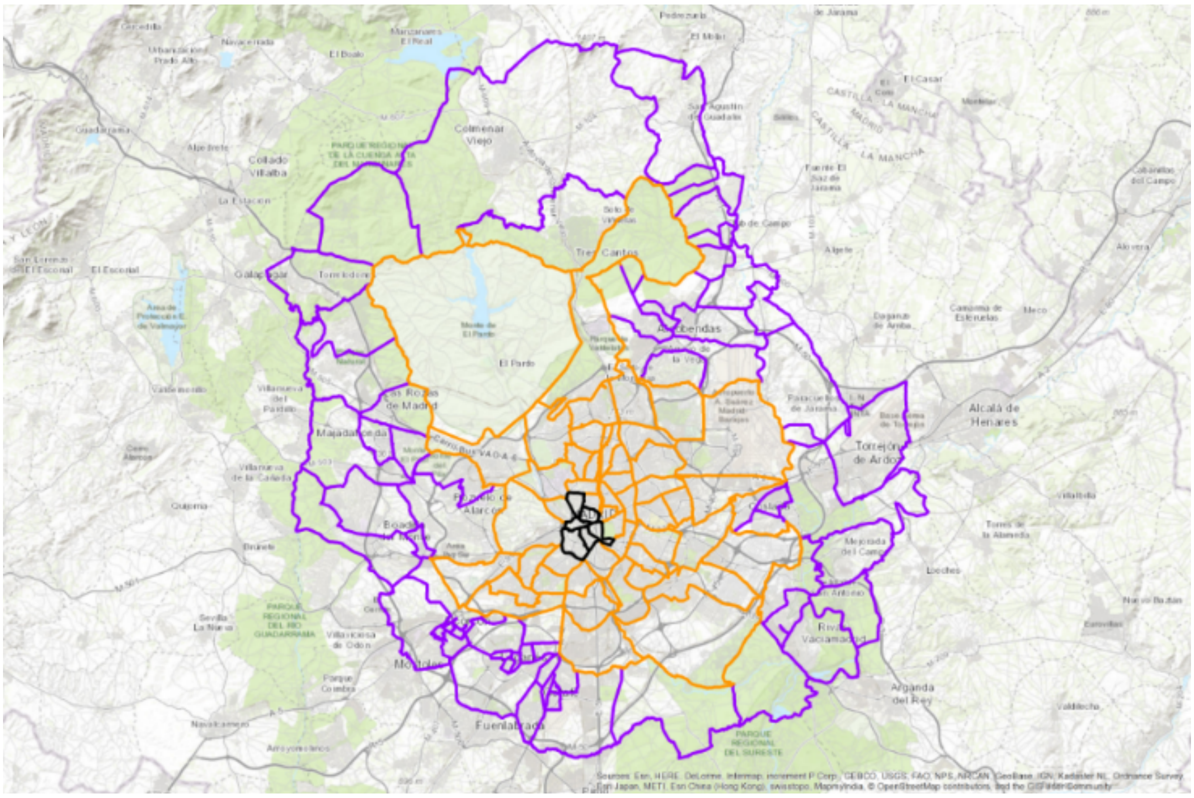


Figure 3: Synthetic control group for pollution levels

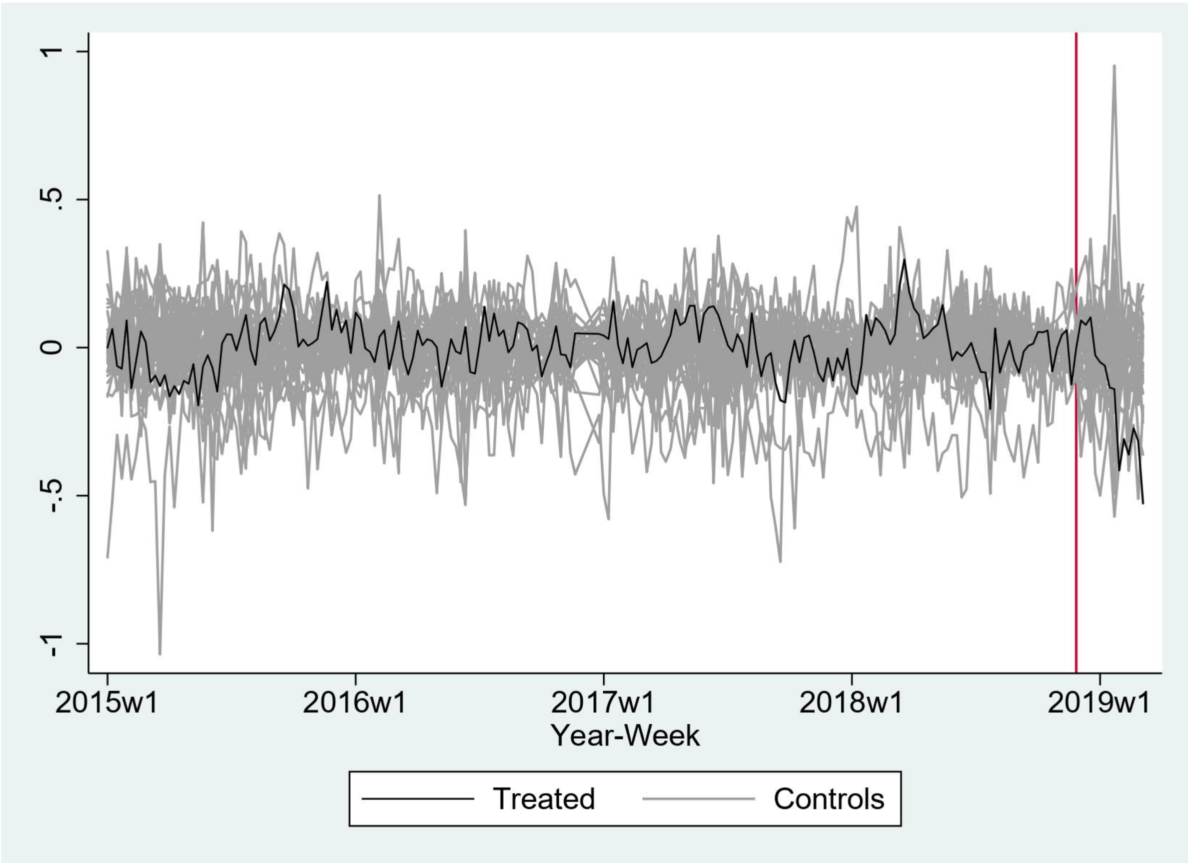


Table 1: Descriptive statistics on traffic and pollution levels

	Mean (1)	SD (2)	Min (3)	Max (4)	Obs (5)
Inside Madrid Central area					
Traffic					
Vehicles per hour	334.76	291.8	0	1,715.28	15,548
Time occupied [%]	10.65	9.73	0	98.51	15,544
Utilized capacity [%]	27.66	10.8	3.51	61.16	15,548
Pollution					
NO_2	47.43	12.13	26.96	95.69	216
$NO_2 > 40$	0.65	0.3	0	1	216
Yearly NO_2	47.84	2.46	44.39	49.99	5
Yearly $NO_2 > 40$	1	0	1	1	5
Outside Madrid Central area					
Traffic					
Vehicles per hour	454	509.41	0	4,354.98	604,81
Time occupied [%]	6.51	7.28	0	98.33	604,56
Utilized Capacity [%]	19.92	10.84	0	99.56	603,92
Pollution					
NO_2	38	16.96	3.82	133.44	6,971
$NO_2 > 40$	0.4	0.35	0	1	6,971
Yearly NO_2	40.05	10.95	14.93	76.14	160
Yearly $NO_2 > 40$	0.47	0.5	0	1	160

Table 2: Descriptive statistics on weather conditions

	Mean (1)	SD (2)	Min (3)	Max (4)	Obs (5)
Temperature [°C]	15.68	7.69	1.23	30.87	7,187
Precipitation [0.1mm]	10.69	18.21	0	158.29	7,187
Cloud cover [okta]	3.42	1.77	0	7.71	7,187
Sunshine [h]	8.19	2.97	0.97	13.91	7,187
Pressure [hPa]	1,017.38	6.2	994.9	1,035.49	7,187
Humidity [%]	57.96	15.62	22.29	92.86	7,187
Wind speed [0.1 m/s]	22.56	10.36	1.71	80.14	7,187
0° ≤ Wind direction < 45°	0.22	0.22	0	1	7,187
45° ≤ Wind direction < 90°	0.14	0.18	0	0.86	7,187
90° ≤ Wind direction < 135°	0.09	0.13	0	0.86	7,187
135° ≤ Wind direction < 180°	0.05	0.1	0	0.57	7,187
180° ≤ Wind direction < 225°	0.12	0.16	0	1	7,187
225° ≤ Wind direction < 270°	0.2	0.2	0	1	7,187
270° ≤ Wind direction < 315°	0.11	0.15	0	1	7,187
315° ≤ Wind direction < 360°	0.07	0.12	0	1	7,187

Table 3: Descriptive statistics on sales

	Mean (1)	SD (2)	Min (3)	Max (4)
Total value	2,087.29	9,644.52	0	606,469
B&M value	1,954.69	9,488.51	0	606,469
Online value	132.59	633.51	0	126,363
Share online value	0.08	0.18	0	1
Total transactions	53.87	268.20	0	16,696
B&M transactions	51.13	264.92	0	16,605
Online transactions	2.73	14.08	0	701
Share online transactions	0.06	0.14	0	1
Total transaction value	28.73	39.44	0.09	9,227.47
B&M transaction value	27.59	38.74	0.09	9,227.47
Online transaction value	16.18	72.46	0.12	31,539.1
Share of obs with 0 total value	0.22			
Share of obs with 0 B&M value	0.23			
Share of obs with 0 online value	0.71			

Note: These are statistics from observations at the seller-buyer-week level.

Table 4: Descriptive statistics on consumption

	Madrid Central	Madrid City	Outside Madrid City
	(1)	(2)	(3)
Number of zip codes	6	50	70
Seller-zip code statistics			
Share of revenue coming from B&M sales	85.6%	90%	95%
Mean value of B&M sales	38.28	36.75	41.1
Mean value of online sales	63.2	44.58	54.79
Mean share of B&M sales by zip codes in Madrid Central to each of the zip codes in MC, Madrid City, or Outside Madrid City	5.62%	1.12%	0.21%
Mean share of online sales by zip codes in Madrid Central to each of the zip codes in MC, Madrid City, or Outside Madrid City	1.84%	1.34%	0.40%
Buyer-zip code statistics			
Share of B&M purchases in MC	45.2%	8.9%	4.3%
Share of online purchases in MC	20.3%	18.70%	14.80%
Share of B&M purchases in local zip code	27%	28.50%	38.70%
Share of online purchases in local zip code	13.10%	15.10%	14.3%

Table 5: Effects on traffic levels

	Vehicles per hour	Time occupied [%]	Utilized capacity [%]	Log Vehicles per hour	Log Time occupied [%]	Log Utilized capacity [%]
	(1)	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**
Madrid Central	-48.50***	-1.936**	-5.679***	-0.147***	-0.178***	-0.228***
	(12.13)	(0.85)	(0.776)	(0.0247)	(0.0495)	(0.0378)
Location-Week FE	Yes	Yes	Yes	Yes	Yes	Yes
Year-Week FE	Yes	Yes	Yes	Yes	Yes	Yes
Mean dep. var.	456.3	6.571	20.19	5.648	1.493	2.905
NxT	597,221	596,895	596,328	597,031	592,874	571,518
N	3,948	3,948	3,948	3,948	3,927	3,823

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

Table 6: Baseline regression on NO_2 levels

	Log NO_2	Log NO_2	Log NO_2	NO_2 >40	NO_2 >40	NO_2 >40
	(1)	(2)	(3)	(4)	(5)	(6)
Madrid Central	-0.162***	-0.163***	-0.136***	-0.121***	-0.122***	-0.111***
	(0.0165)	(0.0182)	(0.0375)	(0.0145)	(0.0159)	(0.0306)
Madrid Central Surroundings*Post		-0.008	0.018		-0.015	-0.004
		(0.032)	(0.046)		(0.029)	(0.039)
City of Madrid*Post			0.0357			0.015
			(0.0421)			(0.0337)
Station-Week FE	Yes	Yes	Yes	Yes	Yes	Yes
Year-Week FE	Yes	Yes	Yes	Yes	Yes	Yes
Mean dep. var.	3.545	3.545	3.545	0.408	0.408	0.408
NxT	7,187	7,187	7,187	7,187	7,187	7,187
N	33	33	33	33	33	33

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

Table 7: Unconditional means of sales

	Pre (1)	Post (2)
Inside the MC area	546,489.0	665,388.5
Outside the MC area	244,061.1	312,626.0

Note: These are means of sales values pre and post adoption for zip codes inside and outside the MC area. These means are calculated from observations at the seller-week level.

Table 8: Seller-zip code sales

Dependent variables: Log revenue and log transactions for all transactions, and B&M and online transactions separately

	Total		B&M		Online	
	Rev	**Trans**	**Rev**	**Trans**	**Rev**	**Trans**
	(1)	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**
Treatment	-0.0987	-0.094	-0.123	-0.119	-0.284**	-0.282**
	(0.148)	(0.148)	(0.16)	(0.158)	(0.118)	(0.119)
Week-year FE	Yes	Yes	Yes	Yes	Yes	Yes
Seller FE	Yes	Yes	Yes	Yes	Yes	Yes
Seller trends	Yes	Yes	Yes	Yes	Yes	Yes
Observations	27,594	27,594	27,594	27,594	27,594	27,594

Notes: *** denotes significance at 1%, ** at 5%, and * at 10%. Standard errors are clustered at the seller-zip code level. The dependent variable is the log of revenue for total transactions, and brick-and-mortar and online transactions at the seller-zip code level in a given week. The variable Treatment takes value 1 when a seller-zip code is within the MC area and the MC regulations are in place, and 0 otherwise. In all columns, we control for week-year specific FE, and seller zip code FE, and include seller-zip code specific trends and trends squared.

Table 9: Baseline Results

Dependent variable: Log of revenue and log of number of transactions for all transactions, B&M and online transactions

	Total		B&M		Online	
	Rev	Trans	Rev	Trans	Rev	Trans
	(1)	(2)	(3)	(4)	(5)	(6)
Treatment	-0.0385	-0.0291	-0.0895**	-0.0472*	0.121**	0.0943**
	(0.0380)	(0.0243)	(0.0404)	(0.0253)	(0.0606)	(0.0444)
Buyer-week-year FE	Yes	Yes	Yes	Yes	Yes	Yes
Seller-week-year FE	Yes	Yes	Yes	Yes	Yes	Yes
Buyer-seller-week FE	Yes	Yes	Yes	Yes	Yes	Yes
Observations	3,460,968	3,460,968	3,460,968	3,460,968	3,460,968	3,460,968

Notes: *** denotes significance at 1%, ** at 5%, and * at 10%. Standard errors are clustered at the origin-destination zip code pair level. The dependent variable is log of revenue and log of number of transactions for all transactions, brick-and-mortar and online transactions at the seller-zip code by buyer-zip code level in a given week. The variable Treatment takes the value one when (1) a seller-zip code is within the MC area and the MC regulations are in place, and (2) the buyer-zip code is outside the MC area, and 0 otherwise. In all columns, we control for buyer-week-year specific FE, seller-week-year specific FE and buyer-zip code by seller-zip code by week of the year FE.

Table 10: Heterogeneous effects

Dependent variable: Log of revenue and log of number of transactions for all transactions, B&M and online transactions.

	Total		B&M		Online	
	Rev	**Trans**	**Rev**	**Trans**	**Rev**	**Trans**
	(1)	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**
Zip codes City	-0.0376	-0.0299	-0.0804*	-0.0399	0.049	0.0763*
	(0.0394)	(0.0251)	(0.0418)	(0.0261)	(0.0629)	(0.0454)
Zip codes Out of city	-0.0391	-0.0285	-0.0959**	-0.0524**	0.172***	0.107**
	(0.041)	(0.0254)	(0.0441)	(0.0265)	(0.064)	(0.0451)
Buyer-week-year FE	Yes	Yes	Yes	Yes	Yes	Yes
Seller-week-year FE	Yes	Yes	Yes	Yes	Yes	Yes
Buyer-seller-week FE	Yes	Yes	Yes	Yes	Yes	Yes
p-val equal effects	0.956	0.921	0.581	0.4	0.00155	0.0699
Observations	3,460,968	3,460,968	3,460,968	3,460,968	3,460,968	3,460,968

Notes: *** denotes significance at 1%, ** at 5%, and * at 10%. Standard errors are clustered at the origin-destination zip code pair level. The dependent variable is log of revenue, and log of number of transactions for the full sample, brick-and-mortar and online transactions at the seller-zip code by buyer-zip code level in a given week. The variable Zip codes City takes value one when (1) the seller-zip code is within the MC area and the MC regulations are in place and (2) the buyer-zip code is outside the MC area but within the city of Madrid, and 0 otherwise. The variable Zip codes Out of city takes value one when (1) the seller-zip code is within the MC area and the MC regulations are in place and (2) the buyer-zip code is outside the city of Madrid, and 0 otherwise. In all columns, we control for buyer-week-year specific FE, seller-week-year specific FE and buyer by seller by week of the year FE.

Table 11: Falsification test

Dependent variable: Log of revenue

	B&M	Online	Total
	(1)	(2)	(3)
Treatment	-0.0895**	0.121**	-0.0385
	(0.0409)	(0.0606)	(0.038)
Falsification	0.00509	0.000127	0.0437
	(0.048)	(0.102)	(0.0482)
Buyer-week-year FE	Yes	Yes	Yes
Seller-week-year FE	Yes	Yes	Yes
Buyer-seller-week FE	Yes	Yes	Yes
Observations	3,460,968	3,460,968	3,460,968

Notes: *** denotes significance at 1%, ** at 5%, and * at 10%. Standard errors are clustered at the origin-destination zip code pair level.

Table 12: Poisson Pseudo Maximum Likelihood

Dependent variable: Revenue

	B&M	Online	Total
	(1)	(2)	(3)
Treatment	-0.0797**	0.0668***	-0.0286
	(0.031)	(0.0246)	(0.0299)
Buyer-week-year FE	Yes	Yes	Yes
Seller-week-year FE	Yes	Yes	Yes
Buyer-seller-week FE	Yes	Yes	Yes
Observations	3,460,968	3,460,968	3,460,968

Notes: *** denotes significance at 1%, ** at 5%, and * at 10%. Standard errors are clustered at the origin-destination zip code pair level. The dependent variable is level of revenue for brick-and-mortar, online and total transactions at the seller-zip code by buyer-zip code level in a given week. The variable Treatment takes value 1 when (1) a seller-zip code is within the MC area and the MC regulations are in place, and (2) the buyer-zip code is outside the MC area, and 0 otherwise. In all columns, we control for buyer-week-year specific FE, seller-week-year specific FE and buyer by seller by week of the year FE. Estimation is done by Poisson Maximum Likelihood.

Table 13: Increases in Travel Time

Dependent variable: Log of revenue

	B&M	Online	Total
	(1)	(2)	(3)
Zip codes Low Increase	-0.0915**	0.0906	-0.0546
	(0.0418)	(0.0635)	(0.0387)
Zip codes High Increase	-0.0876*	0.148**	-0.0239
	(0.0448)	(0.0641)	(0.042)
Buyer-week-year FE	Yes	Yes	Yes
Seller-week-year FE	Yes	Yes	Yes
Buyer-seller-week FE	Yes	Yes	Yes
p-val equal effects	0.893	0.151	0.276
Observations	3,460,968	3,460,968	3,460,968

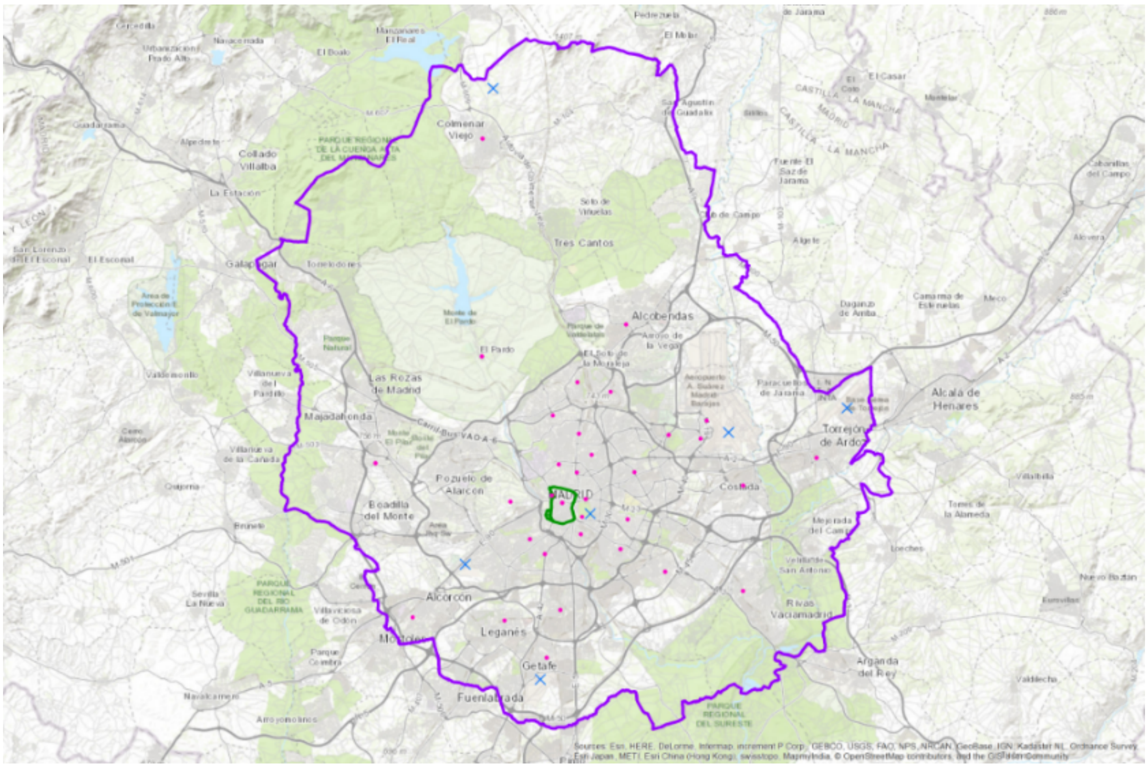
Notes: *** denotes significance at 1%, ** at 5%, and * at 10%. Standard errors are clustered at the origin-destination zip code pair level.

Appendix

Section A1. Maps

See the location of the 33 air quality (pollution) monitoring stations within the city of Madrid in Figure A1. In the map, the pollution monitoring stations appear with pink circles in the map whereas weather stations appear in blue crosses.

Figure A1: Map of stations



Section A2. Summary statistics

Table A1 aims to provide descriptive statistics at the traffic monitor level for the same variables reported in Table 2.

Table A1: Descriptive statistics on weather conditions

	Mean (1)	SD (2)	Min (3)	Max (4)	Obs (5)
Temperature [°C]	15.63	7.74	3.39	30.87	616,297
Precipitation [0.1mm]	10.43	17.78	0	131.43	616,297
Cloud cover [okta]	3.45	1.76	0	7.57	616,297
Sunshine [h]	8.2	2.92	1.13	13.41	616,297
Pressure [hPa]	1017.25	6.09	998.21	1035.54	616,297
Humidity [%]	58.29	14.54	26.29	90.29	616,297
Wind speed [0.1 m/s]	19.73	8.46	0.43	75.43	616,297
0° ≤Wind direction < 45°	0.19	0.19	0	1	616,297
45° ≤ Wind direction < 90°	0.17	0.18	0	0.71	616,297
90° ≤ Wind direction < 135°	0.1	0.14	0	0.86	616,297
135° ≤ Wind direction < 180°	0.05	0.1	0	0.57	616,297
180° ≤ Wind direction < 225°	0.1	0.14	0	0.86	616,297
225° ≤ Wind direction < 270°	0.22	0.21	0	1	616,297
270° ≤ Wind direction < 315°	0.12	0.15	0	0.86	616,297
315° ≤ Wind direction < 360°	0.06	0.1	0	0.86	616,297

Section A3. Alternative specifications of congestion and pollution analysis

Table A2 replicates the specifications in Table 5 including a dummy for whether the traffic monitoring station is located in the surroundings of the MC area.

Table A2: Effects on traffic levels: Spillovers

	Vehicles per hour	Time occupied [%]	Utilized capacity [%]	Log Vehicles per hour	Log Time occupied [%]	Log Utilized capacity [%]
	(1)	(2)	(3)	(4)	(5)	(6)
Madrid Central	-52.63***	-1.987**	-6.043***	-0.153***	-0.179***	-0.241***
	(12.14)	(0.85)	(0.777)	(0.0247)	(0.0496)	(0.0379)
Madrid Central Surroundings*Post	-25.51***	-0.318*	-2.248***	-0.0375***	-0.00897	-0.0899***
	-3.741	(0.171)	(0.199)	(0.00623)	(0.0171)	(0.0109)
Station-Week FE	Yes	Yes	Yes	Yes	Yes	Yes
Year-Week FE	Yes	Yes	Yes	Yes	Yes	Yes
Mean dep. var.	456.3	6.571	20.19	5.648	1.493	2.905
NxT	597,221	596,895	596,328	597,031	592,874	571,518
N	3,948	3,948	3,948	3,948	3,927	3,823

Notes: *** denotes significance at 1%, ** at 5%, and * at 10%. Standard errors are clustered at the station level and p-values are reported in parentheses.

Table A3 replicates the exercise in Table 6 with Poisson regressions.

Table A3: Effects on Air Pollution Levels: Poisson Regressions

	NO_2>40 (1)	NO_2>40 (2)	NO_2>40 (3)
Madrid Central	-0.178***	-0.182***	-0.158***
	(0.0223)	(0.0248)	(0.0571)
Madrid Central Surroundings*Post		-0.0255	-0.00185
		(0.0461)	(0.0683)
City Madrid* Post			0.0308
			(0.06)
Station-Week FE	Yes	Yes	Yes
Year-Week FE	Yes	Yes	Yes
Mean dep. var.	0.46	0.46	0.46
NxT	6,381	6,381	6,381
N	33	33	33

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

Table A4 replicates the exercise in Table 6 with station-specific trends.

Table A4: Regression of NO_2 levels with station-specific trends

Dependent variable: Log of NO_2 (cols 1-3), and a dummy equal to one if NO_2>40 (cols. 4-6)

	Log NO_2	Log NO_2	Log NO_2	$N0_2$>40	$N0_2$>40	$N0_2$>40
	(1)	(2)	(3)	(4)	(5)	(6)
Madrid Central	-0.158***	-0.157***	-0.146***	-0.135***	-0.136***	-0.130***
	(0.0157)	(0.0173)	(0.0398)	(0.0165)	(0.0168)	(0.0301)
Madrid Central Surroundings*Post		0.0149	0.0254		-0.00735	-0.00172
		(0.0299)	(0.0467)		(0.0618)	(0.066)
City of Madrid*Post			0.0146			0.0078
			(0.044)			(0.0318)
Station-Week FE	Yes	Yes	Yes	Yes	Yes	Yes
Year-Week FE	Yes	Yes	Yes	Yes	Yes	Yes
Mean dep. var.	3.545	3.545	3.545	0.408	0.408	0.408
NxT	7,187	7,187	7,187	7,187	7,187	7,187
N	33	33	33	33	33	33

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

Table A5 introduces the city of Barcelona as control group.

Table A5: Regression of NO_2 levels with Barcelona in the control group

Dependent variable: Log of NO_2 (cols 1-3), and a dummy equal to one if NO_2>40 (cols. 4-6)

	Log NO_2	Log NO_2	Log NO_2	$N0_2$>40	$N0_2$>40	$N0_2$>40
	(1)	(2)	(3)	(4)	(5)	(6)
Madrid Central	-0.157***	-0.147***	-0.167***	-0.103***	-0.0847***	-0.0723***
	(0.0154)	(0.023)	(0.0185)	(0.0127)	(0.0185)	(0.0181)
Madrid Central Surroundings* Post		0.0182	-0.00232		0.0339	0.0464*
		(0.0315)	(0.0278)		(0.0256)	(0.0264)
City of Madrid*Post			-0.0391			0.0239
			(0.0471)			(0.0398)
Station-Week FE	Yes	Yes	Yes	Yes	Yes	Yes
Year-Week FE	Yes	Yes	Yes	Yes	Yes	Yes
Mean dep. var.	3,563	3,563	3,563	0,418	0,418	0,418
NxT	8,708	8,708	8,708	8,708	8,708	8,708
N	40	40	40	40	40	40

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

Section A4. Alternative specifications of the consumption spending analysis

Using the Google Maps Distance API, we calculate travel times by car and public transport from the centroid of each zip code to the centroid of MC. We then calculate how much longer it takes to use public transport compared to using the car. As mentioned in the text, Table A6 shows the "penalty" of public transport usage is 12 minutes larger for zip codes outside the city of Madrid compared to zip codes inside the city of Madrid. The difference is significant at the 1% level.

Table A6: Changes in Travel Time

Dependent variable: Changes in travel time to the MC area (in minutes)

	(1)
Zip codes out of city	12.40***
	(2.46)
Observations	126

Notes: *** denotes significance at 1%, ** at 5%, and * at 10%. We regress the difference between travel time to MC by public transportation and car on a dummy if a zip code is out of city.

Table A7 shows results of alternative difference-in-differences specification from those in Table 8 with sales aggregated at the seller-zip code level without time trends and time trends squared for each seller-zip code. The original results in Table 8 are qualitatively robust.

Table A7: Seller-zip code sales with no trends

Dep var: Log revenue and log transactions for all, B&M and online transactions

	Total		B&M		Online	
	Rev	**Trans**	**Rev**	**Trans**	**Rev**	**Trans**
	(1)	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**
Treatment	-0.0987	0.00298	-0.123	0.0302	-0.284**	-0.419**
	(0.148)	(0.0604)	(0.160)	(0.0606)	(0.118)	(0.0786)
Week-year FE	Yes	Yes	Yes	Yes	Yes	Yes
Seller FE	Yes	Yes	Yes	Yes	Yes	Yes
Seller trends	No	No	No	No	No	No
Observations	27,594	27,594	27,594	27,594	27,594	27,594

Notes: *** denotes significance at 1%, ** at 5%, and * at 10%. Standard errors are clustered at the seller-zip code level. The dependent variable is the log of revenue for brick-and-mortar, online and total transactions at the seller-zip code level in a given week. The variable Treatment takes value 1 when a seller-zip code is within the MC area and the traffic restriction is in place. In all columns we control for week specific FE, and seller zip code FE. In columns 2, 4, and 6 we further include seller-zip code specific trends and trends squared.

Table A8 shows results of regressing our measure of weekly revenues at the buyer-zip code and seller-zip code dyad level on the distance between centroids of the buyer and seller zip codes. Column 1 of Table A8 confirms that trade flows decrease with distance between zip codes. In columns 2 and 3 we separate brick-and-mortar from online transactions, and we show that both transaction types exhibit gravity. Intuitively, brick-and-mortar sales are far more sensitive to distance between zip codes than online sales. Finally, because we worry about unobserved cross-sectional differences in the dyads of trade flows, we regress the share of online revenue over total sales on distance between zip codes. Consistently, with our gravity results in columns 1 to 3, we find that the share of online revenue between two zip codes increases with distance between zip codes.

Table A8: Gravity

Dependent variable: Log of revenue for all transactions, B&M and online transactions separately and percentage share of online revenue.

	Total revenue	B&M revenue	Online revenue	Share online revenue
	(1)	(2)	(3)	(4)
Log distance between seller-buyer zip codes	-1.722***	-1.767***	-0.608***	2.080***
	(0.0196)	(0.0196)	(0.0157)	(0.0792)
Buyer-week-year FE	Yes	Yes	Yes	Yes
Seller-week-year FE	Yes	Yes	Yes	Yes
Observations	3,460,968	3,460,968	3,460,968	3,460,968

Notes: *** denotes significance at 1%, ** at 5%, and * at 10%. Standard errors are clustered at the origin-destination zip code pair level. The dependent variable is the log of revenue for all transactions, brick-and-mortar and online transactions at the seller-zip code by buyer-zip code level in a given week (cols. 1-3), and the percentage share of online revenue for that seller-buyer-week triple (col. 4). The independent variable is the log of distance measured in km between the centroid of the seller-zip code and the centroid of the buyer-zip code. In all columns, we control for seller-week-year specific FE and buyer-week-year specific FE.

Table A9 replicates results of Table 9 without adding buyer-zip code seller-zip code pair fixed effects and shows the importance of controlling for the underlying variation across zip code pairs that may negatively bias our estimates.

Table A9: No buyer-seller pair specific FE

Dependent variable: Log of revenue and log of number of transactions for all transactions, B&M and online transactions

	Total		B&M		Online	
	Rev (1)	Trans (2)	Rev (3)	Trans (4)	Rev (5)	Trans (6)
Treatment	-1.835***	-1.909***	-2.007***	-1.994***	-1.360***	-0.779***
	(0.158)	(0.161)	(0.163)	(0.163)	(0.133)	(0.0932)
Buyer-week-year FE	Yes	Yes	Yes	Yes	Yes	Yes
Seller-week-year FE	Yes	Yes	Yes	Yes	Yes	Yes
Buyer-seller-week FE	No	No	No	No	No	No
Observations	3,460,968	3,460,968	3,460,968	3,460,968	3,460,968	3,460,968

Notes: *** denotes significance at 1%, ** at 5%, and * at 10%. Standard errors are clustered at the origin-destination zip code pair level. The dependent variable is log of revenue and log of number of transactions for all transactions, brick-and-mortar and online transactions at the seller-zip code by buyer-zip code level in a given week. The variable Treatment takes value 1 when (1) a seller-zip code is within the MC area and the traffic restriction is in place, and (2) the buyer-zip code is outside the MC area, and 0 otherwise. In all columns we control for buyer-week-year specific FE, and seller-week-year specific FE.

Table A10 replicates the specification in Table 9 with different dependent variables, the share of online revenue, share of online transactions and ratio of transaction values. The results are showing that decreases in brick-and-mortar transactions and increases in online transactions are taking place within buyer zip code and seller zip code pairs.

Table A10: Online Shares

Dependent variable: Percentage share of online revenue, online number of transactions and ratios between online and brick-and-mortar transaction values

	Share online revenue	Share online transactions	Ratio transaction values
	(1)	(2)	(3)
Treatment	3.380***	1.434***	-0.0141
	(0.675)	(0.495)	(0.143)
Buyer-week-year FE	Yes	Yes	Yes
Seller-week-year FE	Yes	Yes	Yes
Buyer-seller-week FE	Yes	Yes	Yes
Observations	3,460,968	3,460,968	3,460,968

Notes: *** denotes significance at 1%, ** at 5%, and * at 10%. Standard errors are clustered at the origin-destination zip code pair level.

Table A11 presents robustness checks in which we change the sample definition and the definition of strongly affected and mildly affected zip codes. In columns 1 to 3, we consider that those zip codes mildly affected by MC are zip codes within six km (3.7 miles approximately) of the MC area. We consider a zip code is within six kms of the MC area if the centroid of the zip code is within six kms of *Puerta del Sol* (a square that represents the centroid of the MC area). We use the same criterion to determine which zip codes are within 15 km of the MC area. The results are qualitatively unaffected by this change in definition. In columns 4 to 6, we reduce the sample to only those zip codes within 15 km of the MC area, and consider those within six km to be mildly affected. Again, our heterogeneity results are qualitatively similar to those in Table 10.

Table A11: Robustness Results I

Dependent variable: Log of revenue

	All Sample			Zipcodes within 15km		
	B&M	**Online**	**Total**	**B&M**	**Online**	**Total**
	(1)	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**
Zipcodes < 6 km	-0.0719*	0.0185	-0.0463	-0.0612	0.04	-0.0312
	(0.0436)	(0.0668)	(0.0407)	(0.0423)	(0.071)	(0.0398)
Zipcodes > 6 km	-0.0948**	0.152**	-0.0361	-0.0939**	0.147**	-0.0387
	(0.0422)	(0.062)	(0.0393)	(0.0416)	(0.0698)	(0.0395)
Buyear-week-year FE	Yes	Yes	Yes	Yes	Yes	Yes
Seller-week-year FE	Yes	Yes	Yes	Yes	Yes	Yes
Buyer-seller-week FE	Yes	Yes	Yes	Yes	Yes	Yes
p-val equal effects	0.413	0.0015	0.707	0.259	0.0291	0.792
Observations	3,460,968	3,460,968	3,460,968	1,575,050	1,575,050	1,575,050

Notes: *** denotes significance at 1%, ** at 5%, and * at 10%. Standard errors are clustered at the origin-destination zipcode pair level. The dependent variable is log of revenue for brick-and-mortar, online and total transactions at the seller-zipcode by buyer-zipcode level in a given week. The variable Zipcodes < 6 km takes the value of one when the seller-zipcode is within Madrid Central Area and the traffic restricion is in place and (2) the buyer-zipcode is outside Madrid Central Area but within 6 km of Madrid Central. The variable Zipcodes > 6 km takes the value of one when the seller-zipcode is within Madrid Central Area and the traffic restricion is in place and (2) the buyer-zipcode is further than 6 km from Madrid Central Area. In all columns we control for buyer-week-year specific FE, seller-week-year specific FE and buyer by seller by week of the year FE.

Table A12 presents further robustness checks. First, we are concerned with the fact that the decrease in brick-and-mortar purchases or the increase in online purchases originates in smaller buyer-zip codes so that the results would not reflect the average impact of MC. To address this issue, in columns 1 to 3 of Table A12, we weight information from each of the buyer-zip codes by the average volume of total sales of the given zip code during the year 2015. The estimates remain similar in magnitude. The main difference is that now the decrease in total revenue from buyer-zip codes outside the city of Madrid becomes significant.

We also worry about unobserved changes over time in the propensity of consumers to buy in zip codes located further away. For this reason, columns 4 to 6 in Table A12 introduce a week-year fixed-effect interacted with the distance between the buyer-zip code and the seller-zip code. This set of interactions controls for a potential increase in propensity to buy in zip codes that are further away from the local zip code. We can see how results are robust to this concern and, if anything, they are larger in magnitude and more significant.

Table A12: Robustness Results II

Dependent variable: Log of revenue

	Weights			Extra Controls		
	B&M	Online	Total	B&M	Online	Total
	(1)	(2)	(3)	(4)	(5)	(6)
Zipcodes City	-0.0788**	0.0413	-0.049	-0.0831**	0.105*	-0.0442
	(0.0338)	(0.0647)	(0.0322)	(0.0419)	(0.0628)	(0.0395)
Zipcodes Out of city	-0.138***	0.186***	-0.0847**	-0.101**	0.284***	-0.0525
	(0.0357)	(0.0694)	(0.0337)	(0.0444)	(0.0639)	(0.0414)
Buyer-week-year FE	Yes	Yes	Yes	Yes	Yes	Yes
Seller-week-year FE	Yes	Yes	Yes	Yes	Yes	Yes
Buyer-seller-week FE	Yes	Yes	Yes	Yes	Yes	Yes
week-year FE * zipcode distance	No	No	No	Yes	Yes	Yes
p-val equal effects	0.0124	0.000946	0.115	0.519	0.00000444	0.766
Observations	3,460,968	3,460,968	3,460,968	3,460,968	3,460,968	3,460,968

Notes: *** denotes significance at 1%, ** at 5%, and * at 10%. Standard errors are clustered at the origin-destination zipcode pair level.

Table A13 presents results of the impact of MC on mean transactaion values. Using specification (2), now we use as the dependent variable the mean value of transactions for all transactions in column one, for brick-and-mortar transactions in column 2, and for online transactions in column 3. There are no significant changes in mean transaction values for any of the three categories.

Table A13: Effects on Transaction Values

Dependent variable: Log of mean transaction value for all transactions, B&M and online transactions

	Total	B&M	Online
	(1)	(3)	(5)
Treatment	-0.00942	-0.0423	0.0265
	(0.0267)	(0.0288)	(0.0461)
Buyer-week-year FE	Yes	Yes	Yes
Seller-week-year FE	Yes	Yes	Yes
Buyer-seller-week FE	Yes	Yes	Yes
Observations	3,460,968	3,460,968	3,460,968

Notes: *** denotes significance at 1%, ** at 5%, and * at 10%. Standard errors are clustered at the origin-destination zip code pair level. The dependent variable is log of the mean transaction value (calculated as the ratio between total revenue over number of transactions) for all transactions, brick-and-mortar and online transactions at the seller-zip code by buyer-zip code level in a given week.