UNIVERSITA' COMMERCIALE "LUIGI BOCCONI"

PhD SCHOOL

PhD program in Economics and Finance

Cycle: 32^{nd}

Disciplinary Field (code): SECS-P/01

Essays in Applied Microeconomics and Finance

Advisor: Francesco DECAROLIS

PhD Thesis by

Angelo D'Andrea

ID number: 1656479

Year 2022

Acknowledgements

I would first like to thank my supervisor, Francesco Decarolis, and my co-author, Nicola Limodio, whose expertise was invaluable in formulating the research questions and methodology. Your insightful feedback pushed me to sharpen my thinking and brought my work to a higher level.

I would like to acknowledge the faculty at Bocconi University, whose collaboration was key to identifying and nurturing my research interests. In particular, I want to thank Donato Masciandaro for always being there during these years of intense work.

I would also like to acknowledge faculty outside of Bocconi, whose feedback has been invaluable in defining my agenda. A special thank goes to Marco Di Maggio, whose support has had substantial implications on my current (and future) research.

I would like to thank Enrico Sette and Marco Pelosi for having joined me in the job market project. Our zoom calls and days spent coding and interpreting results, are reflected in the high quality of the paper.

I would like to thank my colleagues at the Ph.D. for the wonderful time spent together. I would particularly like to single out my friends Sergio and Francesca, who provided stimulating discussions as well as happy distractions to rest my mind outside of research.

I would like to thank my friends. Those in Milan, who followed me throughout this long path. My friends in Rome and those in San Giorgio. A special thank goes to the "Gruppo Storico" and my old friends, Antonio, Michele, and Pierluigi, for always being a major source of support.

Finally, I would like to thank my family. I have no words to describe how important you are to me. If I have reached this prestigious goal, I owe it to your patience and your supporting role throughout the roller-coasting of my emotions.

A special thank goes to my wife. You are the brightest ray on sunny days and the sunlight on dark nights.

Thesis Abstract

This dissertation consists of three essays that examine how new technologies help to alleviate financial frictions, and how monetary incentives can improve the performance of local politicians. In two chapters, I study the effects of high-speed internet on banking and credit. In another chapter, I estimate the causal effect of higher mayoral wages on public procurement outcomes.

The first chapter studies the impact of access to broadband internet on bank credit supply to non-financial firms. We rely on granular micro-data from Italy and instrument the staggered arrival of broadband through the historical presence of telephone infrastructures. We find that banks with branches in municipalities reached by fast internet increase loan supply, both at the extensive and the intensive margin, and reduce credit price. We document that the expansion of credit goes through three main channels: connected banks become more productive; they expand their geographical scope; and competition in municipalities reached by fast internet intensifies. Broadband internet also leads banks to make internal credit reallocation by moving credit away from small municipalities towards bigger municipalities. The latter effect is nevertheless associated with local economic growth.

The second chapter exploits the staggered arrival of submarine cables in Africa and shows that high-speed internet lifted financial markets. We document a novel mechanism through which fast internet promotes banking and credit supply. Plummeting telecommunication costs induce banks to adopt new financial technologies, the real-time gross settlement system (RTGS), which lower transaction costs and promote credit. We find that upon connecting to high-speed internet, banks adopt the RTGS, reduce liquidity hoarding and increase interbank transactions and lending. We also observe that fast internet particularly strengthens firms in countries with weak pre-existing interbank markets.

The third chapter documents the causal effect of mayoral wages on procurement outcomes in Italy, using a regression discontinuity analysis. To identify the effect of mayoral wages, I use a unique characteristic of the Italian legislation, namely that mayoral remuneration varies at predetermined population thresholds. The main results are as follows. First, I show that higher mayoral wages are not related to differences in aggregate measures of procurement. Second, I show that specific procurement outcomes are affected by higher mayoral wages: the number of admitted offers increases, as do the final rebates on the reserve price; the probability that the same firm is awarded a contract repeatedly decreases; and for a limited sample, cost overruns go down. Finally, I provide evidence that re-election incentives play a role in explaining the effect of mayoral wages.

Contents

Chapter 1: Broadband and Bank Intermediation **1**

1 Introduction **2**

2 Institutional Background **7**

2.1 Broadband Internet 7

2.2 The Italian Bank Credit Market 8

3 Data **10**

4 Empirical strategy **12**

5 Results **15**

5.1 Motivating Evidence 15

5.2 The Effects of Broadband on Bank Credit 16

5.3 The Supply Channel 17

5.4 Robustness 18

6 Mechanisms **20**

6.1 Lending Efficiency 20

6.2 Banks' Geographical Reach 21

6.3 Competition 22

7 Extensions: Digital Highways, Credit, and the Real Effects of Broadband **23**

7.1 Credit Reallocation 23

7.2 The Spatial Distribution of the Effects: Does Broadband Internet Create Banking Deserts? 24

8 Conclusion **27**

Figures **29**

Tables **33**

Chapter 2: High-speed Internet, Financial Technology and Banking **49**

1 Introduction **50**

2 Empirical framework and data **54**
2.1 Fiber-optic submarine cables and banking . . . 54
2.2 Markets of liquidity in Africa . . . 57
2.2.1 Real-time Gross Settlement Systems . . . 58
2.3 Data and descriptive statistics . . . 59
2.4 Empirical Methodology . . . 62

3 Results **64**
3.1 Banking . . . 64
3.2 The role of interbank markets . . . 67
3.2.1 Interbank Markets and Weak Users . . . 69
3.3 Firms . . . 70

4 Robustness **72**

5 Conclusion **74**

Figures **76**

Tables **84**

Chapter 3: Mayoral Wage and Public Procurement **94**

1 Introduction **95**

2 Institutional Background **100**

3 Data and Sample **103**
3.1 Descriptive Statistics . . . 105

4 Regression Discontinuity Design **107**
4.1 Estimation . . . 108

5 Empirical Analysis **108**
5.1 Testing the continuity assumption . . . 108
5.2 Procurement outcomes: Graphical analysis . . . 110
5.3 Procurement outcomes: Nonparametric analysis . . . 111
5.3.1 Mechanisms . . . 116

6 Robustness checks **119**

7 Conclusion **123**

Figures **125**

Tables **130**

References **144**

Appendices **158**

Appendix: Broadband and Bank Intermediation 159

Appendix: High-speed Internet, Financial Technology and Banking 174

Appendix: Mayoral Wage and Public Procurement 195

Chapter 1

Broadband and Bank Intermediation *

Angelo D'Andrea[†] Marco Pelosi[‡] Enrico Sette[§]

Abstract

This paper studies the effects of access to broadband internet on bank credit to non-financial firms. We rely on granular micro-data from Italy and an IV empirical strategy to address the endogeneity of broadband diffusion. We find that banks with branches in municipalities reached by fast internet increase loan supply, both at the extensive and the intensive margin, and reduce credit price. We document that the expansion of credit goes through three main channels: banks increase productivity; they expand their geographical scope; and competition in municipalities reached by fast internet intensifies. Broadband internet also leads banks to make internal credit reallocation by moving credit away from smaller municipalities.

*We are grateful to Francesco Sobbrio for sharing the data on the instrument that we use in our empirical analysis. We also thank Francesco Decarolis, Marco di Maggio, Nicola Gennaioli, Nicola Limodio, Daniel Paravisini, Nicolas Serrano-Velarde, Fabiano Schivardi, Emanuele Tarantino and seminar participants at Bocconi University, the Bank of Italy and the Norges Bank. The views expressed in this paper are those of the authors and should not be attributed to the Bank of Italy or the Eurosystem.

[†]Bocconi University. E-mail: angelo.dandrea@unibocconi.it

[‡]Bank of Italy, Structural Economic Analysis Directorate, and LSE. E-mail: marco.pelosi@bancaditalia.it

[§]Bank of Italy, Structural Economic Analysis Directorate. E-mail: enrico.sette@bancaditalia.it

1 Introduction

The arrival of fast internet has been one of the most disruptive innovations in history, and as such, it had a substantial impact on economic activity. The availability of a massive amount of information and the ability to communicate them quickly transformed many industries' size and operations. As an information-intensive business, banking was particularly prone to this transformation. When information flows are limited, banks face higher information asymmetries, communication costs, and more severe agency problems (Leland and Pyle, 1977). Innovations in information and communication technology, from hardware such as computers to software such as credit scoring and client profitability programs, up to the recent diffusion of internet technologies, mitigate these frictions and can play a crucial role in shaping banking activity (Mishkin and Strahan, 1999). As proof of this, banks have long relied on cutting-edge technologies to deliver innovative products, streamline loan-making processes and improve their back-office efficiency (Frame et al., 2018).

Despite its relevance and the importance of the banking industry for the smooth functioning of the economy (Levine, 1997), the evidence on the effects of fast internet on bank activities, in particular lending, is scant[1] [2]. A key reason for this is the lack of high-quality administrative data and of an identification strategy to deal with the endogeneity of the introduction of the internet[3].

This paper studies the effect of access to broadband internet on bank credit to non-financial firms, and it sheds light on the mechanisms behind this effect. We rely on granular micro-data from Italy and instrument the staggered arrival of broadband through the historical presence of telephone infrastructures, controlling for prior local characteristics. We find that fast internet favors credit expansion and is associated with a reduction in the average price of credit. Then, we separately identify the component of the total effect which is attributable to supply and show that it accounts for half of the total effect[4].

[1] D'Andrea and Limodio (2019) is an exception. They focus on the effects of high-speed internet in Africa and show that fast internet favored new financial technologies in the interbank market, thus alleviating banks' liquidity risk and promoting lending to the private sector.

[2] By contrast, a large literature studies the effects on the banking industry of regulatory reforms (Bertrand et al., 2007), removals of barriers to entry (Cetorelli and Strahan, 2006), shocks of various nature (financial, real, natural disasters), institutional quality, and even the role of culture and ethnicity (Caprio et al., 2007; Grosjean, 2011; Calomiris and Carlson, 2016; Pascali, 2016; Fisman et al., 2017; D'Acunto et al., 2019).

[3] The latter refers to the fact that local economic conditions can influence the rolling of internet infrastructures.

[4] Indeed, the effect of fast internet on lending is an equilibrium effect determined by supply and demand factors. On the one hand, broadband access may favor credit expansion through a more efficient acquisition and process of information. On the other hand, it may boost credit demand through an increase in firms' productivity and profitability.

When dealing with the mechanisms behind our results, we document that broadband internet affects the organizational design of banks, with effects on banks' productivity, market geography, and local banking competition.

We resort to a unique dataset from Italy that includes detailed information on the dates of broadband internet arrival and the geographical location of the necessary infrastructure for broadband. This information is matched with loan-level data from the comprehensive Italian credit register and other administrative details on the location of bank branches and banks' assets, liabilities, and employees. We observe these data between 1998 and 2008, which mark the expansion of broadband internet in Italy, and perform an econometric analysis based on instrumental variables (IV).

Measuring the impact of access to broadband on credit is challenging. Since high-speed internet is not randomly assigned to municipalities, bank credit could be affected by hidden factors other than (but related to) broadband connection. To deal with this source of endogeneity, we use an instrumental variable strategy and leverage the position of the municipality in the pre-existing voice telecommunications (telephone) infrastructure to instrument for broadband availability (Campante et al., 2018). As fast internet services in Italy could only be offered in municipalities connected to high-order telecommunication exchanges via optic fiber, we use the distance between the municipality centroid and these exchange infrastructures (a proxy for the required investment to connect the municipality with the fiber) as a source of variation for the availability of high-speed internet (Ciapanna and Sabbatini, 2008). Because the pre-existing telecommunication network was not randomly distributed, our instrumental variable relies on the interaction between the abovementioned distance and a dummy variable for the period after broadband internet became available. Our identification assumption is that whatever correlation existed between the distance and relevant municipality characteristics, this did not change at the time of introduction of broadband for reasons other than broadband itself. Results from our 2SLS estimates shed light on the causal effect of fast internet on bank credit.

The main findings of the paper can be summarized as follows. We find a positive and statistically significant effect of broadband internet on the extensive and intensive margin of the credit relationship. Going from zero to high broadband coverage is associated with an increase in the number of loans issued by banks of 12% (0.08 of a standard deviation, s.d.), and an increase in the amount of credit granted of 28% (0.13 of a s.d.)[5]. Then, we find a negative and statistically significant effect of broadband on the price of bank credit. Moving from zero to high coverage of high-speed internet is associated with a decrease in the average interest rate of 30 b.p. (0.18 of a s.d.).

[5] "High" broadband coverage means at least 75% of the population in the municipality connected to fast internet.

These results represent equilibrium outcomes that take into account both demand and supply. The effect of fast internet on credit demand has been only indirectly documented in the literature by focusing on firm's productivity (Akerman et al., 2015). On the other hand, the evidence on the effect of broadband on bank's internal activity and its organizational design remains scant. However, there is consensus that information technologies have revolutionized the way lending is conducted by traditional banks (He et al., 2021). For that reason, in this paper, we apply a particular lens on how broadband internet affects bank credit supply.

The granularity and the structure of our data help us isolate credit supply from other confounding factors. Similar to Khwaja and Mian (2008), in the most demanding specification, we exploit the panel structure of the data and the diffusion of multiple bank relationships in Italy (Gobbi and Sette, 2014) to compare the amounts of credit extended to firms that have relationships with banks in municipalities served differently by broadband. In this way, we can isolate credit supply from demand, and we can exclude that total credit variation is driven by firm-specific needs (which may be affected by the availability of fast internet). We use high dimensional fixed-effect regressions that control for time-varying firms' loan demand and for features specific to the firm-bank-municipality relationship. Finally, we add bank-year fixed effects to control for time-varying credit variation that originates from bank-specific policies.

When restricting the analysis on credit supply by controlling for firm-level demand for credit, we find that branches in municipalities reached by fast internet expand their amount of credit 19% more than other branches (0.14 of a s.d.).

To qualify the effect of broadband internet on credit supply, we study how banks' productivity, geographical scope, and local competition are affected by fast internet. As far as the first is concerned, we show that the lending efficiency of banks, measured by banks' labor productivity and credit quality, increases as a consequence of broadband availability. Internet helps increase credit extended per employee by 24%. Standard models suggest that a surge in loan supply may lead to a worse quality distribution (Berger and Udell, 2004; Foos et al., 2010). If anything, in our setting, we find that credit quality improves as the share of non-performing loans (NPLs) per bank decreases. These findings are in line with Petersen and Rajan (2002) and Berger (2003), who argue that richer hard information and more efficient back-office technology helps both ex-ante screening and ex-post monitoring.

Moreover, the geographical reach of banks widens. Banks operating in municipalities reached by fast internet expand their markets beyond standard geographical borders, which typically coincide with provinces. We find that they are more likely to originate loans outside the province where they are located. In addition, the physical distance between banks' municipalities and borrowers increases, too. These results are consistent with the view that improved screening and monitoring, together with a reduction in communication costs, allow banks to reduce the dis-economies of distance and increase "proximity" (Berger, 2003; Felici and Pagnini, 2008).

Finally, municipalities reached by fast internet experience a rise in banking competition. This is confirmed by the increase in the number of available bank brands in the municipality and the dynamics of standard proxies of competition. The concentration ratio of the top 5 and top 3 banks decreases, as well as the Herfindahl–Hirschman Index (HHI) of deposits. Consistently, we find that increased competition pushes down loan prices (in line with Hauswald and Marquez (2003); Vives and Ye (2021)).

The arrival of fast internet has additional effects on banks and borrowers. To this extent, we find that banks tend to implement internal credit reallocation across municipalities, with new loans managed by larger, more distant branches. In the test using granular data, we show that broadband internet does not have any effect on credit supply for branches located in small municipalities[6]. At the same time, banks tend to open new branches in places reached by fast internet, but not if these places are small. Access to high-speed internet creates digital highways that carry bank credit from connected peripheries to the center, (i.e., from smaller municipalities connected to broadband towards bigger municipalities). Yet, this local credit desertification is not accompanied by slower economic growth. Broadband access boosts GDP per capita both in bigger and smaller municipalities, showing off the virtues of the credit flows that broadband contributes to creating.

Our results are robust to several robustness checks, most notably, different measures of broadband coverage and the inclusion of several control variables at the municipality level, aiming to control municipality time trends. We also run placebo IV specifications for the years before fast internet was available and simulate internet availability as if we were in the post broadband period. Reassuringly, we find no impact of broadband internet on bank credit.

This paper builds and extends on different strands of the literature. It contributes to the broad literature on the effects of new telecommunication infrastructures on the economy (Roller and Waverman, 2001; Forman et al., 2009; Czernich et al., 2011; Kolko, 2012; Akerman et al., 2015; Pascali, 2017; DeStefano et al., 2018; Donaldson, 2018; Steinwender, 2018; Hjort and Poulsen, 2019). In this respect, it is one of the few that concentrate on the role of ICTs in banking. D'Andrea and Limodio (2019) exploit the staggered arrival of fiber-optic submarine cables in Africa and show that high-speed internet lifts banking. They highlight one possible mechanism behind this effect that is related to more efficient interbank markets.

[6] We define a municipality as small if its population is below the in-sample median of 4,639 inhabitants.

Lin et al. (2021) study China in the late 19th century and show that the telegraph significantly expanded banks' branch networks. This paper adds to the existing literature by focusing on a specific technology, broadband internet, and a specific instrument, enhanced bank credit to firms, and by showing the channels through which it operates.

The paper also contributes to the literature on information technology and banking. Hauswald and Marquez (2003) provide a theoretical framework on the effects of new information technologies on loan prices and competition. They show that the advent of new information technologies generates ambiguous effects depending on whether these technologies are readily available to all competitors rather than exclusive use to some of them. Using similar arguments, Vives and Ye (2021) show that IT progress involves an increase in competition intensity when it weakens the influence of bank-borrower distance on monitoring costs. Petersen and Rajan (2002) and Berger (2003) provide intuitions and empirical evidence on the effects of new technologies on the distance between lenders and borrowers. New technologies allow financial intermediaries to substitute soft information with hard information, thus increasing the distance between borrowers and lenders. Related to this topic, Felici and Pagnini (2008) find that the geographical reach of entry decisions increases for those banks that resort more to information and communication technologies and that the latter has important pro-competitive effects. Degryse and Ongena (2005) and Keil and Ongena (2020) show the effects of new technologies on bank de-branching, while in a recent paper, Ahnert et al. (2021) show that job creation by young firms in the US is stronger in counties that are more exposed to IT-intensive banks. Our paper adds to this literature by showing the effects of broadband internet on bank lending and the organizational design of banks. The paper is also close to the literature on internet banking and bank performance (DeYoung, 2005; Ciciretti et al., 2009) and that of information technology and financial stability (Pierri and Timmer, 2020). Finally, the paper builds a bridge between the traditional literature on technology and banking and the fast-growing literature on FinTech (De Roure et al., 2016; Buchak et al., 2018; Tang, 2019; Braggion et al., 2020; Di Maggio and Yao, 2020), which documents the economic effects of state-of-the-art financial technologies.

To conclude, our paper contributes to the large literature on information in financial intermediation (Leland and Pyle, 1977; Campbell and Kracaw, 1980). Stiglitz and Weiss (1992) show that despite the richer strategy space available to lenders, market equilibria can be characterized by credit rationing if information asymmetries are relevant. New technologies such as credit scoring (Einav et al., 2013), fax machines, and the internet can help reduce these information asymmetries and improve bank lending (Liberti et al., 2016; Liberti and Petersen, 2019).

The rest of the paper is organized as follows. Section 2 presents the institutional background in

Italy. Section 3 describes our data. Section 4 presents the empirical specifications and the identification strategy. Section 5 shows the main results with robustness checks. Section 6 investigates the mechanisms behind our findings. Section 7 discusses the effects of new digital highways on bank credit. Finally, Section 8 concludes.

2 Institutional Background

Italy represents an ideal laboratory for our analysis. First, the long history of human settlement in the country allowed for the existence of several relatively small municipalities located at short distance from one another, often separated by geographical barriers (rivers, lakes, mountains). This creates large variation in the distribution of the infrastructure needed to bring broadband to different municipalities, which often have a very similar level of economic activity and development and are just a few miles away. Second, Italy in our sample period did not experience an especially fast growth in credit, nor it experienced a housing bubble, contrary for example to the US, UK, Ireland or Spain. Third, Italy is a developed economy, mostly bank dependent, with an economic structure similar to that of other major countries. Finally, Italy has very granular administrative micro-data that are crucial to implement our identification strategy.

2.1 Broadband Internet

Broadband internet connection in Italy has been traditionally provided through asymmetric digital subscriber lines (ADSL). ADSL technology is a data communications technology that enables faster data transmission than a conventional voiceband modem, and it was introduced by the Italian telecommunications incumbent operator, Telecom Italia, in 1999. The development of the ADSL infrastructure was relatively slow in the first years. By the end of 2000, only 117 out of 8,100 Italian municipalities had access to the new technology. Instead, it sped up sensibly during the subsequent years. By the end of 2005, about half of all municipalities owned an ADSL line, accounting for approximately 86% of the population. Figure 1 reports the time series of broadband adoption in terms of the number of municipalities with ADSL access, between 2000 and 2008. Given the low access and penetration rates until 2001, we consider this as the last "pre-broadband" year throughout our analysis.

ADSL technology relies on information transmission over conventional copper phone wires. Henceforth, ADSL access depends crucially on the user's position in the pre-existing voice telecommunications infrastructure. Technically, the voice telecommunications infrastructure consists of three levels: the Line

Stage (LS), the Urban Group Stage (UGS), and the Transit Group Stage (TGS). The LS is the last structure where all the providers connect with their equipment, after which the famous last mile that reaches the end-users begins. In Italy, the 10,500 LSs are linked to one of the 628 UGS, which are connected to one of the 65 TGS. To complete the physical architecture of the network, some TGSs are tied to the three international gateways (Milan, Rome, and Palermo), which allow for international communications.

Two parameters are of specific importance for ADSL deployment and performance. The first is the distance between the end user's premises and the closest telecommunication exchange (the LS), known as the "local loop". If the length of the local loop is above a certain threshold, the ADSL connection cannot be implemented through traditional copper wires, but it needs fiber optic cables. The second is the distance between the LS and the closest higher-order telecommunication exchange (the UGS). Independently from the length of the local loop, for ADSL to be available, the connection between the LS and the UGS must be through fiber optic cables. In Italy, the length of the local loop has not constituted a limiting factor for the development of the broadband infrastructure. Since the local loop was a key element in the voice telecommunications network, its length was generally short and distribution capillary. Instead, the distance between the LS and the UGS, which was irrelevant for voice communication purposes, has become the primary determinant of the investment needed to provide ADSL to a given area and, consequently, of the timing of ADSL adoption (Ciapanna and Sabbatini, 2008). The latter is behind the choice of our instrumental variable.

To build our instrument, we exploit the fact that the 628 UGSs were inherited from the pre-existing voice telecommunication system, so their location was determined long before the advent of the internet[7]. As a consequence, the position of the telecommunication infrastructures was not influenced by the ADSL technology (Impiglia et al., 2004; AGCOM, 2011). Our IV builds on the assumption that ceteris paribus, the closer a municipality happened to be to a UGS when the ADSL became available, the more likely that municipality had access to high-speed internet earlier in the ADSL diffusion process.

2.2 The Italian Bank Credit Market

During the twenty years between 1980 and 2000, the Italian financial industry has changed substantially, modernizing its operations and performance. Following the implementation of the Second Banking Coordination Directive, Italy enacted a comprehensive banking law ("Testo Unico Bancario") in 1993, which drastically reduced government ownership of banks. The share of assets in the hands of banks

[7] The network of physical infrastructures needed to provide voice telephony services to the Italian citizens was built in the post World War II period, between 1945 and 1960.

owned by central and local government or foundations accounted for 12%, from the 18% in 1998 and the 58% in 1990, (ABI, 2001). Under the joint effect of deregulation and technological changes, the system became much more "market-oriented", and substantial advances occurred in terms of the quantity, productivity and prices of banking services and the diversification, depth and efficiency of the markets (Bank of Italy, 2003; Angelini and Cetorelli, 2003).

In the same period, the Italian banking system has undergone substantial restructuring. At the end of 2000, there were more than 800 banks, one-third of which were part of a banking group. The reorganization of the banking system took place mainly by ways of mergers and acquisitions, increasing the overall degree of concentration. The five largest groups in Italy held more than 50% of total banking assets, up from 35% in 1996. In this regard, Italy lined up with other European economies. However, the extension of individual banks' branch networks and numerous competitors in the same markets heightened competition. A series of standard indicators confirms that the increasing concentration of the Italian credit system has come within a framework of intensifying competition.

The branch network at the beginning of the 2000s was also very dense. At the end of 2000, there was one branch for every 2,100 inhabitants (ABI, 2001), which means that about four-fifths of the population could choose between (at least) three banks in their town of residence. Similarly, the ATM network was very capillary, with 31,750 ATMs (more than one every 2,000 inhabitants) and POS terminals widespread, about 570,800 (more than one every 110 inhabitants).

The Italian financial system is mostly bank-dependent. In 2000, deposits and money market fund shares were equivalent to 87% of GDP. Loans accounted for about one-third of non-financial companies' outstanding liabilities, and this share was fast growing. Considering the average annual flows over the period 1998-2000, loans accounted for 55% of the total increase in firms' liabilities. The majority of these loans were granted by resident banks (Bank of Italy, 2003).

Throughout the 2000s, Italian banks have supported the increasing demand for credit by non-financial firms through loosened supply conditions. Firm leverage has increased conspicuously between 2000 and 2007, moving from 34% to 39%. The available evidence (see Bugamelli et al. (2018) for a detailed review) suggests that the Italian banking system has sustained productivity growth before the great financial crisis by supporting firm-level innovation and exporting and by improving the allocation of capital across firms. These dynamics have been similar in other major European countries.

3 Data

Our final dataset combines information from several sources. It includes details on ADSL coverage and the infrastructural characteristics of the Italian municipalities between 2004 and 2008. It pairs these information with matched firm-bank data related to the period 1998-2008. We have information on the amount of credit granted by bank b to firm f, and the specific features of the credit relationship (loan type, presence of collateral, interest rate). Then, we also use data on the location and the opening and closing of bank branches during the period of analysis, together with data on bank employees and bank deposits. Finally, we gather information on the balance sheets of Italian non-financial incorporated firms.

Data on ADSL coverage are from the "Osservatorio Banda Larga", backed by the Italian Ministry of Telecommunications (Campante et al., 2018). The data include information on the percentage of households with access to ADSL-based services, for each municipality and year between 2005 and 2008, on an asymmetric six-point scale: 0%, 1%–50%, 51%–75%, 76%–85%, 86%–95%, and above 95%. No data are available for years prior to 2005. We view 2002 as the first year the ADSL technology became available (as discussed in the previous section). Hence, for years prior to 2002, we consider the percentage of households with access to ADSL equal to zero. We then use information from the Annual reports of the Italian Communications Authority (AGCOM) to retrieve data for 2004, whereas we treat 2002 and 2003 as missing.

Throughout the analysis, we use the asymmetric six-point-scale variable as our baseline measure of broadband internet access. However, in robustness checks we also experiment with alternative measures. First, we create dummy variables for *good access* (which takes value 1 if broadband access is above 50%) and *some access* (1 if broadband access is above 0%) to broadband. These measures facilitate the interpretation of the coefficients, as they do not rely on the asymmetric scale. Second, we define a proxy for good internet as the number of years since at least 50% of households in a municipality have had access to the ADSL. The latter has the advantage to provide a dynamic to the broadband effect but comes with the disadvantage of considering 2004 as the first year of ADSL adoption, introducing some noise in the first years of the sample.

Figure 2 reports the distribution of access to broadband across Italian municipalities in 2004, the first year of data availability, and 2008, the last year considered in our sample, with darker colors indicating high or full access. The figure documents the rapid diffusion of high-speed internet throughout the country.

Data on ADSL coverage are complemented with those on the infrastructural features of the internet technology. In particular, we collect information on the number and geographical location of LSs and UGSs. Then, we compute the geodesic distance between the centroid of each municipality and the closest UGS and use this variable, interacted with a dummy post-2001, to instrument access to the broadband[8].

Data on matched firm-bank relationships are from the Italian Credit Register (CR) held by the Bank of Italy. The CR contains information on the universe of loans and guarantees banks and financial companies issued to their customers above 30,000 euros (75,000 euros before 2008). For each credit relationship, we observe data on the bank and the firm involved, the total amount granted, the amount utilized, the composition of the credit (three loan types are distinguished: credit lines, term loans, loans backed by receivables), its status (performing or not) and the timing of the relationship. Moreover, we also observe the municipality of the branch that the borrower selects as the reference for the management of the credit relationship. This feature is essential as it allows to observe the bank's location with which the borrower interacts. It also allows us to match data on the loans issued to a firm by banks in different municipalities, with information on internet access in each municipality.

Data on interest rates are from *Taxia*, which is part of the CR. While a subset of the CR, it provides detailed information on interest rates covering more than 80% of total bank lending (Rodano et al., 2013). Such data include the rates charged on outstanding loans (distinguished into credit lines, term loans, and loans backed by receivables) and newly issued term loans.

Data on bank deposits and bank employees are from the Supervisory Reports that banks submit to the Bank of Italy, the banking supervisor of the country, during our sample period.

Data on bank branches are from the Bank of Italy "Lista succursali". For each bank branch, we observe its name, bank identifier, group to which it belongs (when relevant), location, and period of activity (initial and closing date).

We match data at the bank-level using the unique bank identifier.

Data on firm balance sheets are from the firm register collected by CERVED Group. These data provide balance sheets and income statements for the universe of incorporated firms in Italy from 1998-2008. Firms not covered are mainly small firms (sole proprietorship or small household producers). Throughout our analysis, our sample includes all the firms covered by the CERVED database. We match these data with the credit data using the unique firm tax identifier.

Finally, we collect information on the local economic activity and the social background of different areas by using publicly available data from the Bank of Italy, the Italian national statistical institute

[8] Data on the location of LSs and UGSs have been kindly provided by Francesco Sobbrio and are used in Campante et al. (2018).

(ISTAT), and the Ministry of Economy and Finance (MEF).

Table 1 reports summary statistics associated with our final sample. Panel A refers to data at the municipality level and shows municipalities geographical distribution (*North*, *Center*, *South*), as well as statistics on access to broadband (*Internet*) and the ADSL underlying infrastructure (the number of LSs and their average distance from the municipality; the number of UGSs and their average distance from the municipality). Panel B refers to data at the bank-municipality level that we use as our baseline setting throughout the analysis. It shows the number of loans issued by a bank in a given municipality, the amount of credit granted, and the average interest rate charged. Finally, panel C refers to loan-level data and reports statistics on (granted) loan amounts.

4 Empirical strategy

We test the effects of banks' broadband availability on several outcomes, at the bank-level, at the loan (bank-firm relationship) level, and the municipality level.

The existing literature provides some guidance on what could be the effect of broadband internet on bank credit. First, we expect credit outcomes, both the extensive (loans issued) and the intensive (amount of credit granted) margin, to be positively related to broadband (Petersen and Rajan, 2002; Berger, 2003). The effect of high-speed internet on interest rates, instead, is a priori ambiguous. Following the framework by Hauswald and Marquez (2003), the effect of new information technologies on credit price is negative (an increase in interest rates) when the informational advantage of the intermediary that gathers information leads to less competition in the credit market. On the contrary, the effect is positive (a decrease in interest rates) when access to information makes data much more widely and readily available to all competitors. In this case, an improvement in IT generates spillover effects that erode informational advantages and serve to level the playing field, with a consequent reduction of interest rates.

To estimate the intention-to-treat (ITT) effect of the increased access to broadband internet, we rely on the following specification:

$$Y_{(r)bmt} = \nu + \beta Broadband_{mt} + \gamma X_{(r)bmt} + \text{fixed effects} + \varepsilon_{(r)bmt} \quad (1)$$

where subscripts r, b, m, and t indicate, respectively, relationship, bank, municipality and year (*r* or *b* depending on whether the specification refers to credit relationship characteristics or aggregate bank characteristics); *Y* is the outcome variable; *Broadband* represents the percentage of households that have access to the ADSL in the municipality of the branch, based on the asymmetric six-point scale described in the previous section; X includes time-varying controls at the relationship or bank level; fixed effects are different sets of dummies depending on the outcome variable. The latter may include time (year), branch municipality, and bank-municipality of the branch fixed effects. Furthermore, in the regressions at the loan level, we saturate the model with the inclusion of bank-year, firm-year, and firm-bank-municipality of the branch fixed effects[9], in order to isolate the effect of broadband internet on credit supply. We estimate equation (1) with standard errors clustered at the municipality level.

Our main outcome variables are the following. In the regressions aggregated at the bank level: i. the (log) number of loans issued by bank b, in municipality m, at time t, which measures the effect of broadband internet on the extensive margin of bank credit; ii. the total (log) amount of credit granted by bank b, in municipality m, at time t, which measures the intensive and the extensive margin combined; iii. the average interest rate charged by bank b in municipality m, which proxies for the price of credit[10]. In the regressions at the bank-firm relationship level, when we turn on credit supply: i. the (log) amount of credit commitments by bank b to firm f, at time t, which evaluates the effect of broadband on the intensive margin of the credit relationship.

Next, we further explore the mechanisms through which broadband internet affects lending (section 6). We keep the same econometric specification at the bank level and focus on different outcome variables. We use indicators of bank productivity and credit quality to test the effect of broadband internet on the lending efficiency of the bank. We also study measures of distance and location of the borrowers to test for the geography of the loans once broadband internet is available. Then, we look at outcomes at the municipality level to test for local competition. In that regard, we concentrate on the number of local competitors and standard measures of market concentration. Finally, we use a mix of the above-mentioned specifications to provide evidence on the reallocation effects of broadband internet and the associated real effects.

A major concern with the estimates of a model in which we regress bank credit features on access to ADSL is that broadband availability is unlikely to be randomly allocated across municipalities, potentially generating a bias in the estimates of model parameters (Comin and Hobijn, 2004). We use an instrumental variable approach that exploits exogenous variation in broadband adoption across differ-

[9] Notice that this is finer than firm-bank fixed effects, since it focuses on the relationship between the firm and the bank in a specific municipality.

[10] In the main analysis, we use a weighted average of bank's interest rates, where credit is not distinguished between different loan types (term loans, loans backed by receivables, credit lines) and weights depend on the amount of firms' utilized credit for each loan type.

ent geographical areas to deal with this concern.

We take advantage of the geographical distribution of ADSL physical infrastructures and leverage differences across Italian municipalities in the distance between a municipality and the closest UGS, where the latter represents a key determinant of the cost of supplying ADSL services. The underlying assumption is that the distance to the closest UGS affects the pattern of ADSL roll-out, with municipalities located farther away from UGSs getting access to broadband internet later on, ceteris paribus.

Even though the presence and the location of the UGSs precedes the development (and even the existence) of broadband in Italy, the spatial distribution of UGSs is itself non-random. To address this issue, we exploit the panel structure of the data and add bank-municipality (or more granular, i.e., firm-bank-municipality) fixed effects to our specifications. Then, to account for potential time-varying confoundings, we interact the distance of a municipality to the closest UGS with a dummy for the post-2001 period (i.e., after the introduction of high-speed internet). The latter constitutes our instrument to ADSL coverage.

The main identifying assumption behind our IV is that whatever correlation existed between the distance to the closest UGS and relevant municipality characteristics, this did not change at the time of the introduction of the ADSL in the municipality. Indeed, we identify the effect of the change in the impact of distance on the outcomes of interest, under the assumption that any change in that impact occurs solely because of the availability of broadband internet (Campante et al., 2018)[11].

Hence, our econometric model relies on the following two-stage specification:

$$Broadband_{mt} = \rho + \delta DistanceUGS_m \times Post2001 + \omega X_{(r)bmt} + \text{fixed effects} + \xi_t + \epsilon_{(r)bmt} \tag{2}$$

$$Y_{(r)bmt} = \nu + \beta \widehat{Broadband}_{mt} + \gamma X_{(r)bmt} + \text{fixed effects} + \phi_t + \varepsilon_{(r)bmt} \tag{3}$$

where *Distance UGS* is the (time invariant) distance of the bank's municipality centroid to the closest high-order telecommunication exchange (UGS), and we interact this variable with a dummy *Post2001*, that takes value 1 for the years after 2001, and zero otherwise[12]. We estimate equation (3) via two-stages least squares (2SLS) regressions.

Table 2, reports first stage estimates as presented in equation (2). Column 1 refers to our baseline measure of ADSL coverage. Columns 2 and 3 refer to the dummies *good access* and *some access*, which

[11] In the appendix, table B1, we provide a balance table that compares mean values of geographical and socioeconomic indicators for municipalities below and above the median of distance from the closest UGS.

[12] This approach is similar to that in Paravisini et al. (2015), Campante et al. (2018), Manacorda and Tesei (2020) and Guriev et al. (2021).

are equal to 1 if at least 50% or more than 0% of households have access to the internet, respectively. Finally, column 4 focuses on the *years since good internet* (the number of years since at least 50% of the population has ADSL access). As we can see from the table, coefficients are negative and statistically significant, in line with our underlying hypothesis. Moreover, F-statistics are generally high and well above the rule of thumb thresholds.

5 Results

This section shows the main results on the effects of broadband internet on bank credit in Italy. We first present motivating evidence using a difference-in-differences (DiD) setting. Next, we implement our preferred empirical strategy and show estimates from the instrumental variable (IV) analysis.

5.1 Motivating Evidence

To gain intuition, before implementing the two-stage specification as in equations (2) and (3), we run a standard DiD event study on the (log) number of loans and the (log) amount of credit granted by Italian banks.

We simulate a DiD setting and divide the sample into two groups. Treated banks are those in municipalities where at least 50% of the households were connected to the ADSL in 2006. Control banks are those in municipalities where ADSL was unavailable in 2006 or solely to a restricted share of the households. Then, we consider 2001 as the baseline year (in line with the main analysis) and show the heterogeneous effects of broadband internet at the extensive and intensive margin of bank credit.

Figure 3 displays estimated event study coefficients, together with the corresponding 95% confidence interval. The top panel focuses on the number of loans. We consider the semi-dynamic regression proposed by Borusyak and Jaravel (2017), and drop the farthest (negative) relative year from the event, in addition to the baseline category [13]. First, we do not find evidence of pre-trends, meaning that the two groups of banks are on a parallel trend before the arrival of high-speed internet. Second, the treatment dynamics show the positive and statistically significant effect of broadband on bank credit. The effect of fast internet takes one year to materialize and then monotonically increases, with a long-term effect amounting to roughly a 9% increase in the number of loans issued. The bottom panel focuses on the amount of credit extended. While standard errors are somewhat larger, the event study points to similar results. Treated and control banks are on a parallel trend before the treatment, and fast internet is

[13] Borusyak and Jaravel (2017) show that this is the minimum number of restrictions for point identification.

associated with more credit.

Overall, these preliminary findings indicate a positive relationship between access to broadband internet and credit granted by banks.

5.2 The Effects of Broadband on Bank Credit

We now implement the baseline two-stage specification on our main dependent variables, following equations (2) and (3).

We first aggregate data at the bank-municipality-year level and focus on the extensive margin (the number of loans issued by banks to non-financial firms) and the intensive margin (the amount of credit issued) of bank credit. Results are shown in Table 3. Columns 1 and 2 report the basic OLS estimates. The coefficients associated with broadband access are positive and statistically significant, indicating that the advent of high-speed internet is associated with more credit granted by banks.

Interestingly, OLS coefficients are lower than the average treatment effect from the event study in section 5.1. As regards the extensive margin, the coefficient *broadband* is 0.007 (meaning that moving from zero to full broadband increases the number of loans granted by 3.9%), while the corresponding mean effect implied by the event study is around 6.7%[14].

This difference may signal the short-term bias of the two-way fixed effects estimator when treatment effects take time to materialize[15]. In order to test this hypothesis, we follow De Chaisemartin and d'Haultfoeuille (2020) and compute the weights associated with the OLS estimates. Results are reported in table 4. In line with our hypothesis, more than 60% of the OLS weights are associated with the first years after the treatment, and weights are strongly decreasing over time[16].

Next, we instrument broadband access with the interaction between distance to the closest UGS and the post-2001 dummy and consider this as our preferred specification[17]. Columns 3 and 4 show the

[14] Computed on the semi-dynamic model that also excludes 2002 and 2003, for which broadband data are missing in our dataset. Notice that estimates from a standard event study considering the first year of ADSL access provide a mean effect of 7%.

[15] Borusyak and Jaravel (2017) show that the OLS coefficient from two-way fixed effects staggered DiD is a weighted average of the OLS coefficients. When treatment effects are heterogeneous over time, those weights are higher for low values of the relative time from the event and may even become negative in the long run.

[16] The downward bias of the OLS coefficient could also have different sources. First, coverage might be related to unobservable municipality characteristics associated with higher credit. Second, it could be related to the coarseness of the measure of broadband access, especially at the bottom, where going from 1% to 49% access would be entirely missed and yet have the strongest impact (if diminishing returns to broadband are present).

[17] Importantly, when we compute the weights for the reduced form regression of Y on our instrument, we find equal weights.

main results. We find a positive and statistically significant effect of broadband internet on bank credit. The effect is larger than in the basic OLS specification: moving from zero to high broadband coverage is associated with an increase in the number of loans issued of 12% (0.08 of a s.d.) and an increase in the amount of credit granted of 28% (0.13 of a s.d.)[18]. This result is qualitatively in line with Liberti et al. (2016) that show an improvement in the allocation of credit when more and better information is available.

Then, we test the effects of having access to broadband internet on the price of bank credit. We aggregate data at the bank-municipality-year level and run a specification similar to equations (1) and (3). Results, shown in table 5, indicate that the average effect of high-speed internet on interest rates is negative and statistically significant. Passing from zero to high ADSL coverage is associated with a decrease in the average interest rate of 30 b.p. (0.18 of a s.d.)[19]. The latter is in line with the theoretical argument by Hauswald and Marquez (2003) that, being broadband internet a general purpose information technology whose benefits are widely and easily available among bank competitors, the spillover effects from information dissemination dominate the negative effects from informational rents[20].

5.3 The Supply Channel

The findings in the previous section represent equilibrium outcomes. Indeed, the effect of broadband internet on bank credit results from two forces acting simultaneously: credit demand (firms that ask for credit) and credit supply (banks that offer credit).

The effect of fast internet on credit demand has been indirectly investigated in the literature by studying firm productivity (Akerman et al., 2015; DeStefano et al., 2018). Although there is consensus on the fact that banks use cutting-edge technologies to deliver innovative products, streamline loan-making processes, and improve back-office efficiency (He et al., 2021), evidence on the direct effect of broadband on banks' productivity is relatively scant[21]. Moreover, the fact that lending increases while interest

Results are reported in table 4.

[18] "High" broadband coverage means a value of *internet* equal to 3 (at least 75% of the population connected). The latter is also the interquartile range of ADSL coverage for municipalities included in our sample during the period of broadband availability.

[19] This accounts for a reduction of 5% in the average rate, that is in line with Brown and Goolsbee (2002) which find that the growth of the internet has reduced term life prices by 8–15 percent.

[20] Schenone (2010) finds evidence of these spillover effects in initial public offerings (IPOs), using the Securities Data Company (SDC) dataset by Thomson Reuters.

[21] In the appendix, figures 5 and 6, we report descriptive statistics on the % growth of the use of web technologies within bank branches, during the period 2001-2007. These statistics, obtained from the Economic Analysis of the Italian Banking Association

rates go down is difficult to reconcile with a pure demand-side story.

In this paper, we further contribute to the existing literature by isolating the component of the total effect of broadband on bank credit due to credit supply.

We perform this exercise in table 6, where we exploit the granularity of our dataset to further characterize the effect of broadband on the intensive margin of the credit relationship. We use loan-level data and leverage variation within firm-bank (Ali Choudhary and Limodio, 2021)[22]. Columns 1 and 3 report OLS and 2SLS estimates related to this specification. Passing from zero to high broadband coverage is associated with an increase in the amount of credit granted by bank b to firm f of 41% (0.28 of a s.d). Next, following Khwaja and Mian (2008), we saturate the model with the inclusion of firm by year fixed effects that capture the component of the total effect that is related to demand. Importantly, we also add bank-time fixed effects to attenuate the concern that the (exogenous) arrival of broadband internet simultaneously affects both the supply and the demand for credit. Finally, this specification also includes firm-bank-municipality of the branch fixed effects to capture time-invariant characteristics of the credit relationship and possible confounders related to the specialization of different bank branches[23].

Estimates are reported in columns 2 and 4 and refer to the effect of broadband internet on credit supply. As we can see from the table, almost half of the total effect is a pure supply-side effect[24]. Going from zero to high broadband coverage is associated with an increase in bank credit supply of 19% (0.14 of a s.d.).

5.4 Robustness

We subject our main results to several robustness checks.

First, we verify that our instrument passes the Angrist and Imbens (1995) instrument's monotonicity test. By instrumenting our endogenous variable (D_{mt}, Z=[0,1]) with the distance to the closest UGS interacted with the dummy post-2001 (Z_{mt}), we are implicitly assuming that the effect of distance on access to broadband is monotone, that is, either $D_{mt}^{high} \geq D_{mt}^{low}$ or $D_{mt}^{high} \leq D_{mt}^{low}$, $\forall mt$. The assumption is not verifiable, but has testable implications on the CDFs of internet for municipalities close or far from

(ABI), suggest a sensible increase in the use of web technologies in the back office activities of banks during the examined period.

[22] These specifications at the firm-bank-municipality-year level always include controls for the loan (one-year lagged) share of revolving loans of the firm, and the loan share of extended credit of the issuing bank

[23] Notice that this specification is more demanding than that used in Paravisini et al. (2015) as we observe the municipality of the specific branch that manages the loan.

[24] These results always have to be interpreted with the caveat that the variation used in these specifications is that of firms that borrow from multiple banks.

the UGS, that is, they should never cross. In fact, if $D_{mt}^{high} \geq D_{mt}^{low}$ with probability 1, then $Pr(D^{high} \geq j) \geq Pr(D^{low} \geq j), \forall j \in supp\ D$ (Decarolis and Rovigatti). Figure 7 plots the CDFs of internet for banks close to the UGS (blue solid line) and far from the UGS (red solid line)[25]. Since the two CDFs never cross, the instrument passes the test.

Second, our basic identification assumption would be violated if there are underlying trends affecting the outcomes of interest that correlate with our instrument. To control for these confounding factors, we augment our specifications in equations (2) and (3) with several economic and socio-demographic municipal characteristics available from the 2001 Census, interacted with a second-order polynomial-time trend. In this way, we control flexibly for the possibility of differential time trends. The baseline group of controls includes the natural logarithm of the total population, elderly population, the number of private firms operating in the municipality, the number of employees, the distance from the provincial capital[26]. Table B2 reports the estimates related to the extensive and the intensive margin of bank credit. All the coefficients keep the same sign as in the baseline specification, and they remain statistically significant at standard levels. Furthermore, the magnitudes are not sensibly affected by the inclusion of these control variables.

Next, we test the robustness of the main results by using different measures of ADSL coverage. Table B3 focuses on the extensive margin of the credit relationship. In column 1, we report our baseline IV estimate for reference. In column 2, we use *Years Since Good Internet* and find a similarly positive effect. One more year of good internet is associated with an increase of 6.3% in the number of loans. Considering an average of four years of good internet, the total effect of broadband on credit is an increase of 25%. In this case, OLS and IV estimates get much closer to one another, showing the importance of considering the treatment effect dynamic in our analysis. In columns 3 and 4, we consider two alternative dummy variables: *good access* and *some access*. The coefficients are close in magnitude to our baseline *Broadband* measure, although they slightly vary between each other.

In table B4, we address the possibility that our results are picking up some underlying trend in credit that happens to be correlated with the diffusion of broadband. We run placebo IV specifications for the years from 1998 to 2003, assuming that the level of ADSL access in 2006 was already present in 2001 (and the following years). Reassuringly, we see no impact of this fictitious introduction of broadband internet, supporting the view that pre-existing trends do not drive our findings.

[25] Notice that, for this exercise, we use a dummy variable below/above the median of UGS distance to proxy for our instrument (that is a continuous variable).

[26] *Population* is the only variable for which a yearly time series is available.

6 Mechanisms

Results so far show that the arrival of broadband internet leads to an expansion of bank credit. This materializes on the extensive margin (number of loans granted), on the intensive margin, and on the price of credit issued by banks (interest rates go down). Moreover, part of the total effect is due to factors independent of credit demand, as we have seen that loan amounts are affected by access to broadband, which also controls for firm-specific time-varying unobservables. In what follows, we further explore the channels through which all these effects take place.

6.1 Lending Efficiency

It is often argued that IT advances play a substantial role in boosting productivity (Draca et al., 2007). It is thus essential to test whether bank lending efficiency increases as a consequence of broadband internet availability.

We measure bank lending efficiency using two different indicators. First, loans per bank employee, which is a measure of the bank's labor productivity. Second, the share of non-performing loans (NPLs), which proxies banks' credit quality.

Petersen and Rajan (2002) suggest that new technologies allow banks to collect more hard information about borrowers, enabling them to change the nature of lending from an emphasis on strict ex-ante screening and costly ex-post monitoring, to fine-tuned screening and frequent ex-post monitoring. Similarly, Berger (2003) documents the increase in profit productivity due to improvements in ”back-office” technologies, as well as consumer benefits from new ”front-office” technologies. Since high-speed internet enhances screening and monitoring, we expect the effect of broadband on productivity to be positive and significant. On the other hand, the effect of broadband on credit quality is a priori ambiguous. Since marginal borrowers are generally worse than incumbent customers, credit quality could worsen as a consequence of credit expansion. However, hard information on borrowers becomes richer and timely once new technologies are available. Improved screening and monitoring activities can thus offset the negative effect of credit growth.

Table 7 shows the results. Columns 1 and 3 show that bank's access to broadband internet has a positive effect on its labor productivity, measured as the amount of extended credit per bank employee. Going from zero to high broadband coverage is associated with an increase in credit per employee of 24%. On the other hand, columns 2 and 4 show that high-speed internet is associated with a slight decrease in the share of NPLs, meaning that credit quality, on average, improves with the expansion of credit[27].

Taken together, these two findings support the thesis that banks' overall lending efficiency sensibly increases after the introduction of broadband internet[28].

6.2 Banks' Geographical Reach

Lending is traditionally a "local" business, and the distance between lenders and borrowers is a crucial factor shaping credit relationships, especially those that involve SMEs (Degryse and Ongena, 2005). However, Petersen and Rajan (2002) suggest that technology helps break the "tyranny of distance". By improving screening and monitoring activities of banks, new technologies allow for increasing capital intensity of lending and thus lending to more distant borrowers. Along the same lines, Berger (2003) shows that technological progress facilitates the geographic expansion of banking organizations by reducing distance-related dis-economies. New services created by technological progress with higher value added, traditional banking services delivered more efficiently, bank monitoring and the control of risk exposures at longer distances and lower costs, and reduced managerial diseconomies of distance all contribute to ease the way banks find and finance new clients.

On the other hand, Wilhelm (2001) argues that advances in communication technology and increased capacity for information do not imply greater exchange of information inside the bank. This is due to the limited incentives for loan officers to transfer information on which they hold monopoly power. Similarly, advances in communication technology may not lead to more exchange of information between firms and banks (Bhattacharya and Chiesa, 1995) and between different banks (Padilla and Pagano, 1997). In this regard, technological developments may have no effects on the distance between lenders and borrowers[29].

To address these critical issues, we look at the effects of broadband internet on the geography of the credit relationship by focusing on new loans originated by Italian banks during the period of our analysis[30]. We define a dummy variable for the loan being originated outside the province of the branch

[27] The latter is in line with Pierri and Timmer (2020), which study the implications of IT in banking for financial stability. The authors find that pre-crisis IT adoption that was higher by one standard deviation led to 10% fewer NPLs during the 2007–2008 financial crisis.

[28] Casolaro and Gobbi (2007) find that banks adopting IT capital intensive techniques are more efficient and interpret the latter as evidence of a catching-up effect consistent with the usual pattern of diffusion of new technologies.

[29] See also Degryse and Ongena (2005) for empirical evidence on the static nature of the relationship between firms and banks in Belgium, between 1973 and 1997.

[30] To substantiate our hypothesis that broadband reduces communication costs and increases "proximity", in table B5 of the

(*Diff. Province*) to measure the effect of broadband internet on the geographical borders (the market) within which the bank operates[31]. Then, we create a direct measure of the distance between lenders and borrowers by computing the geodesic distance between the centroid of the municipality of the branch that manages the loan and the exact location of the firm.

Estimates from model (3), aggregated at the bank-municipality-year, are reported in table 8. Columns 1 and 3 refer to the share of new loans originated out of the province of the branch. Columns 2 and 4 refer to the inverse hyperbolic sine of the geodesic distance between the municipality of the branch and the firm. The table shows that access to high-speed internet increases the probability that the bank extends credit outside its province. At the same time, broadband internet is associated with firm-bank relationships exhibiting longer distances[32]. Findings are in line with the literature and document the shrinking effects of new information technologies on the distance between lenders and borrowers (Petersen and Rajan, 2002; Berger, 2003; Felici and Pagnini, 2008).

These results also suggest that broadband internet can trigger a re-definition of local credit markets, with all the consequences for agents involved in the credit relationship and for regulatory and supervisory authorities.

6.3 Competition

The expansion of credit following the arrival of broadband internet may be driven by tougher local competition. Hauswald and Marquez (2003) find that when the information gap between banks becomes smaller, because of ICT diffusion, there is a softening of the winner's curse that leads to an increase in the intensity of bank competition. Similarly, Vives and Ye (2021) find that when IT progress involves a weakening in the influence of bank-borrower distance on monitoring costs, then banks' competition intensity increases. Finally, Felici and Pagnini (2008) show that new communication and information technologies have significant pro-competitive effects in local banking markets. By increasing the geographical reach of bank entry decisions, these new technologies augment local credit market contestability.

The effects of broadband internet, and more efficient information technologies in general, on banking competition, is an interesting question in itself. We explore this question in what follows, focusing

---

appendix, we preliminary show that broadband access is positively associated with the share of loans granted to firms connected themselves to broadband.

[31] In Italy, before the advent of fast internet, provinces defined the borders of bank credit markets (Crawford et al., 2018).

[32] In table B6 of the appendix we propose the same exercise at the firm level, where we can control for firm fixed effects (that capture time-invariant firm characteristics). As we can see, results are qualitatively in line with those just presented.

on two measures of competition: the number of (physical) bank competitors in the municipality; and measures of concentration of the local credit market.

Vesala (2000) shows that loan mark-ups were decreasing in Finland, in lock-step with the rapid development of the internet. On the other hand, Gropp et al. (2009) find only a small increase in contestability in the European loan markets despite the impressive technological advances experienced in many countries.

Our results on the effects of broadband on local competition are shown in tables 9 and 10. In table 9, the main dependent variable is the (log) number of bank competitors in a municipality. Estimates from table 9 show a significant increase in local competition[33]. Indeed, the number of banks competing in the market increases when the municipality is reached by fast internet.

In table 10, the main dependent variables are standard indicators of concentration of the local (municipality) credit market: the concentration ratio of the top 5 and 3 banks, and the Herfindahl–Hirschman Index (HHI), computed using data on bank deposits[34]. As we can see from the table, all the coefficients are negative and statistically significant at the 99% level.

Overall, high-speed internet is associated with an increase in competition in the local credit markets. Results are in line with Frame et al. (2018) on the effects of new technologies on banking.

7 Extensions: Digital Highways, Credit, and the Real Effects of Broadband

In this section, we present additional results associated with the arrival of broadband internet on the activity of banks, with a focus on credit allocation and the spatial distribution of the effects.

7.1 Credit Reallocation

An extension that helps characterize our main results relates to the effects of access to broadband internet on credit allocation. The expansion of credit documented in the previous sections may coincide with an enlargement of the set of borrowers served by the banking industry. Otherwise, it can result from an increased amount of credit granted to borrowers that already benefited from banking services. The

[33] Competition within the municipality.

[34] As robustness, we replicate the same estimates by computing standard indicators of competition using data on extended credit. Results are similar to those in table 10.

latter is particularly likely to occur in the case of credit flows from branches of the same bank, where transaction and operating costs are relatively lower (Cetorelli and Goldberg, 2012a,b).

We analyze these possibilities in table 11. Columns 1 and 3 test whether new loans originated by banks reached by fast internet are issued to firms already having a credit relationship. Columns 2 and 4 focus on new loans originated towards firms that already have a credit relationship with the same bank (in a different municipality). Results show that banks connected to high-speed internet have a higher probability of granting a loan to firms out of the credit market before. Moreover, they also have a higher probability of serving firms with a relationship with the same bank.

These findings provide evidence on the complementary effects of broadband on bank credit allocation. First, access to broadband internet determines an expansion in the set of borrowers served. Second, it allows banks to implement internal capital reallocation.

7.2 The Spatial Distribution of the Effects: Does Broadband Internet Create Banking Deserts?

The previous sections show that fast internet leads banks to expand their geographical reach (outside of their province) and reallocate part of their credit within their internal organization. These dynamics may lead to a movement away from small municipalities, typically in the countryside, towards larger municipalities.

Physical infrastructures as highways and railroads, for example, are known to have affected the spatial allocation of economic activities. The direction of their effects has been heterogeneous between rural and urban areas, the periphery and the center, places crossed by the facility and adjacent areas (Rephann and Isserman, 1994; Chandra and Thompson, 2000; Baum-Snow, 2007; Michaels, 2008; Atack et al., 2010; Banerjee et al., 2012; Duranton and Turner, 2012; Donaldson, 2018). Digital infrastructures, sharing some of the underlying features, may have similar effects.

To explore this issue, in this section, we focus on the heterogeneity that distinguishes between small and bigger municipalities and show the existence of potential winners and losers from the process of broadband diffusion (Akerman et al., 2015). This heterogeneity is of particular interest because it assesses whether technological progress can determine credit desertification and, eventually, local economic stagnation. Furthermore, it is relevant in the political economy literature as the internet has been identified as a potential driver of polarization of the political spectrum and behind the recent rise of populist parties.

Table 12 replicates the analysis in table 6 for small municipalities. We exploit the granularity of our dataset and test the effect of broadband internet on the intensive margin of the credit relationship. We use loan-level data and leverage variation within firm-bank. Columns 1 and 3 report OLS and 2SLS estimates related to this specification. Then, we follow Khwaja and Mian (2008) and concentrate on the supply of loans by saturating the model with the inclusion of firm-year fixed effects, which capture demand-side confounders. The coefficients associated with this specification are reported in columns 2 and 4 and isolate the effect of broadband internet on credit supply. As we can see from the table, there seems to be no effect of fast internet on the amount of credit granted to non-financial firms in small municipalities[35].

In the appendix, tables B8 to B12, we also replicate the analysis on the mechanisms behind our results and show that for banks in small municipalities: lending efficiency goes down (in particular, productivity goes down), competition increases[36], and banks do not react in terms of loans geographical expansion and credit reallocation.

The test in table 12 highlights how broadband does not have any positive effect on credit supply for branches located in small municipalities. In this section, we argue that a reshaping can explain part of this null effect in the markets served by banks in different municipalities.

As we have argued in the introduction, broadband internet is a multi-dimensional information technology that reduces information asymmetries, communication costs, and agency problems in the banking industry. To work properly and activate a cheaper and timely communication channel between lenders and borrowers, both agents need to be connected. One possible explanation behind the null effect of broadband on credit in small municipalities is that firms in such places get access to larger credit markets. Firms in small municipalities reached by fast internet may become the "easiest target" for banks in bigger municipalities connected to broadband. As a result, they can exploit this situation to borrow a larger amount of money from outside the municipality and rip better credit prices.

We test this hypothesis in table 13. Columns 1 and 2 report OLS and IV estimates where the dependent variable is a dummy equal to one when the loan is from a bank in a bigger municipality to a firm in a small municipality, out of the province of the bank, and connected to fast internet.

The results from table 13 provide supporting evidence that improved lender-borrower communica-

[35] In table B7 of the appendix, we report the results based on the specification on interest rates. Our results show that access to broadband internet is associated with an increase in the price of credit (contrary to the specification that refers to the entire sample). The increase in the average rate in small municipalities connected to high-speed internet can be explained, in part, by the nature of the credit relationships that remain anchored in those municipalities. These credit relationships usually rely on soft information and are less sensitive to credit price (Ioannidou and Ongena, 2010).

[36] Even if some of the IV coefficients are not statistically significant at standard levels.

tion increases the probability of credit relationship formation. Firms in small municipalities reached by fast internet face new (and broader) investment opportunities and reallocate part of their borrowing out of the municipality towards banks that operate in larger cities (which offer higher amounts of credit, at a lower cost). Consequently, both borrowers in small municipalities and banks in bigger municipalities benefit from the arrival of broadband internet. The former, as they face higher credit supply at a lower price. The latter can reach borrowers that were previously out of their market and can expand their customer base.

This phenomenon of credit centralization has the potential to create local banking deserts. In the same direction, it is interesting to analyze whether banks react to the new technological advancements with a change in their geographical location. We investigate this aspect by looking at the evolution of bank branches.

Italy was in a phase of sensible expansion of the number of branches since the beginning of the 90s. Figure 4 provides evidence in this direction. Furthermore, although we are used to associate informatization with de-branching, following the idea that automatic lending diminishes the value of geographical proximity and so the relevance of local branch presence (Kroszner and Strahan, 1999; Berger and DeYoung, 2001; Petersen and Rajan, 2002), the substitution of a brick-and-mortar model with a click-and-mortar model of banking is a very recent phenomenon[37] [38]. Therefore, during the period of our analysis, the choice that Italian banks faced was not whether to open or close branches but rather whether and where to open new branches.

In table 14 we test the (heterogeneous) effects of having broadband internet on the number of branches in a municipality. Columns 1 and 3 refer to the whole sample, and columns 2 and 4 focus on the sample of small municipalities. Results from the table show that banks tend to increase their branch presence in places reached by broadband, where they can exploit the potential of fast internet. However, this does not happen in small municipalities, in which the effect of broadband is null. Indeed, we consider the latter as another sign of credit centralization from the perspective of the bank.

To conclude our analysis, we check whether these developments, credit flows towards bigger municipalities that lead to local credit desertification, are accompanied by economic underdevelopment in

[37] The latter is probably due to the limited diffusion of smartphones and software allowing for sufficient cyber-security, until recent times.

[38] Different studies have attempted to measure the impact of technology on branching empirically. Degryse and Ongena (2004) argue that new technologies may have only a limited impact on branch presence because of the importance of bank branch proximity for customers. Keil and Ongena (2020) show that broadband and mobile internet access explain well the recent de-branching of banks at the country level, but not that at the US county or bank branch level.

small areas.

Tables 15 and 16 replicate the analysis in equations (2) and (3), where the dependent variables are: the natural logarithm of population, the natural logarithm of income, and the natural logarithm of income per capita in the municipality[39]. Table 15 accounts for the entire sample. Table 16 focuses on the restricted sample of small municipalities. As we can see from the tables, the effect of broadband internet on the real economy is generally positive and more pronounced in small municipalities. Even if not conclusive, these findings suggest that new credit flows allowed by high-speed internet are the expression of broader investment opportunities and alleviated financial frictions, rather than bank desertification followed by local economic underdevelopment.

8 Conclusion

In this research, we provide empirical evidence on the effects of broadband internet on bank credit to non-financial firms. To address this point, we combine data on access to the ADSL technology in Italy with firm-bank matched data from the Bank of Italy. We follow 901 banks in 5271 municipalities, during the period 1998-2008, and show the effects of broadband at the extensive and at the intensive margin of the credit relationship and on credit price.

Our quasi-experimental design relies on the staggered adoption of the ADSL technology across Italian municipalities and an instrumental variable strategy that exploits the municipality's position in the pre-existing voice telecommunications infrastructure.

To explore our research question, we implement two-stages least squares analysis and focus on the effects of broadband on credit by isolating the effects on credit supply, on interest rates, and the underlying mechanisms that elucidate our main results.

Our findings highlight that high-speed internet fosters bank credit towards non-financial firms. The total amount of credit increases with broadband availability, while the average interest rate goes down. Many channels contribute to this aggregate effect. The internal efficiency of banks goes up as a consequence of broadband access. Banks reached by fast internet expand their markets and increase the distance towards their borrowers. At the same time, local competition increases, as reflected by the growth in the number of physical branches and competitors and by standard indicators of competition. Finally, banks connected to broadband internet tend to reach new borrowers and implement internal

[39] Data on income are from the publicly available dataset of the Italian Ministry of Economy and Finance (MEF). The time series starts in 2000, meaning that we have no information for the period 1998-1999.

credit reallocation.

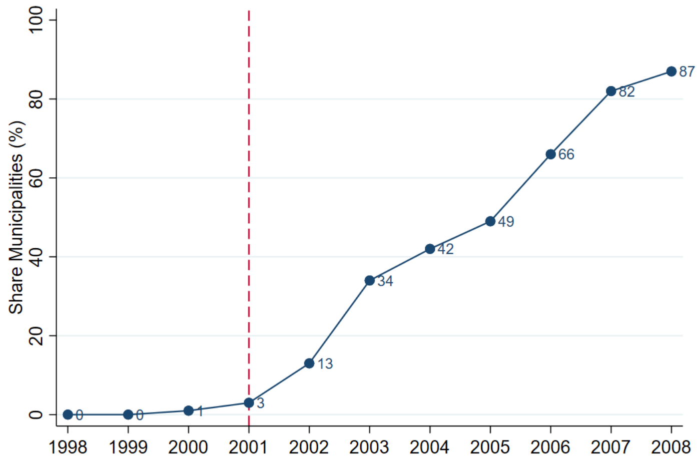
The effect of broadband, however, is heterogeneous. Access to high-speed internet creates digital highways that carry bank credit from the periphery to the center (i.e., from small municipalities to bigger municipalities). Nevertheless, credit desertification in small municipalities does not lead to economic underdevelopment, showing off the virtues of the credit flows generated by internet technologies.

Overall, our results are consistent with high-speed internet promoting bank credit and creating new credit opportunities for non-financial firms.

To conclude, our paper directly speaks to policymakers as we document the multifaceted effects of investments in new hi-tech infrastructures. The latter can serve as a guide for the introduction of future technologies, as the ultra-high-speed internet and the 5G mobile technology. Moreover, the issue of the effects of technological innovation on the operativity and the structure of the banking system is of utmost importance for central banks, both as monetary policy authorities (since it involves banks' lending activity) and as micro and macroprudential supervisors (since it involves banks' risk profiling).

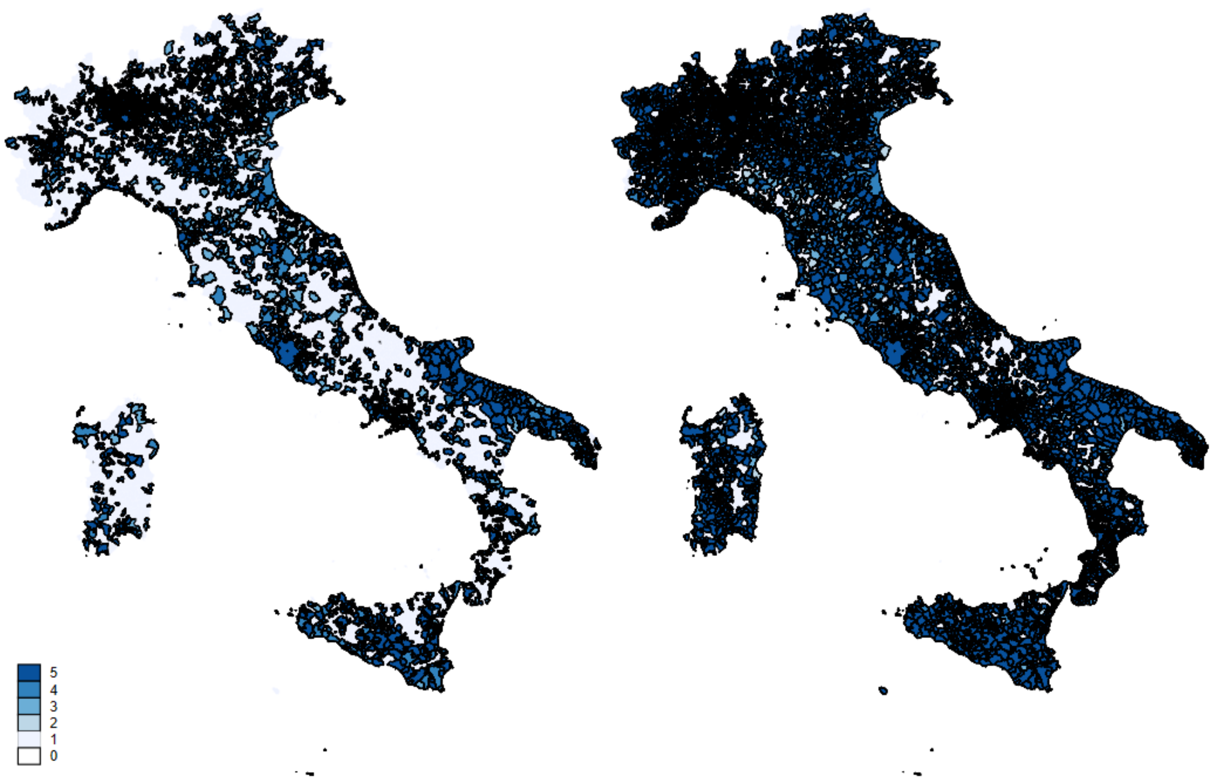
Figures

Figure 1: Broadband internet in Italy



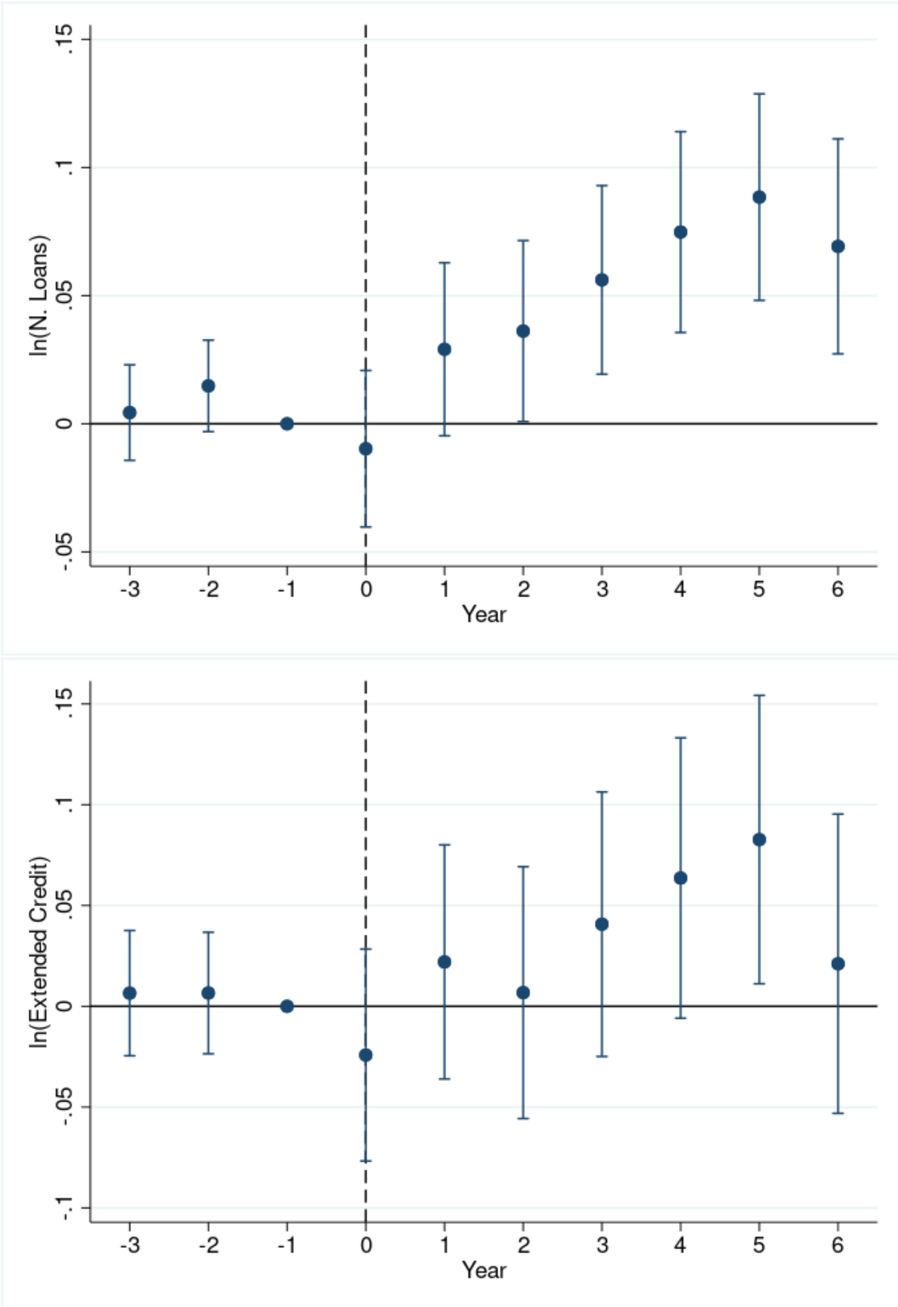
Notes: Broadband diffusion in Italy between 2000 and 2008. On the y-axis, we report the share of municipality with access to the ADSL technology. On the x-axis, the years. The dashed vertical line indicates the separation between the pre-broadband and the post-broadband period, that we make coincide with 2001.

Figure 2: Geographical distribution of Broadband internet



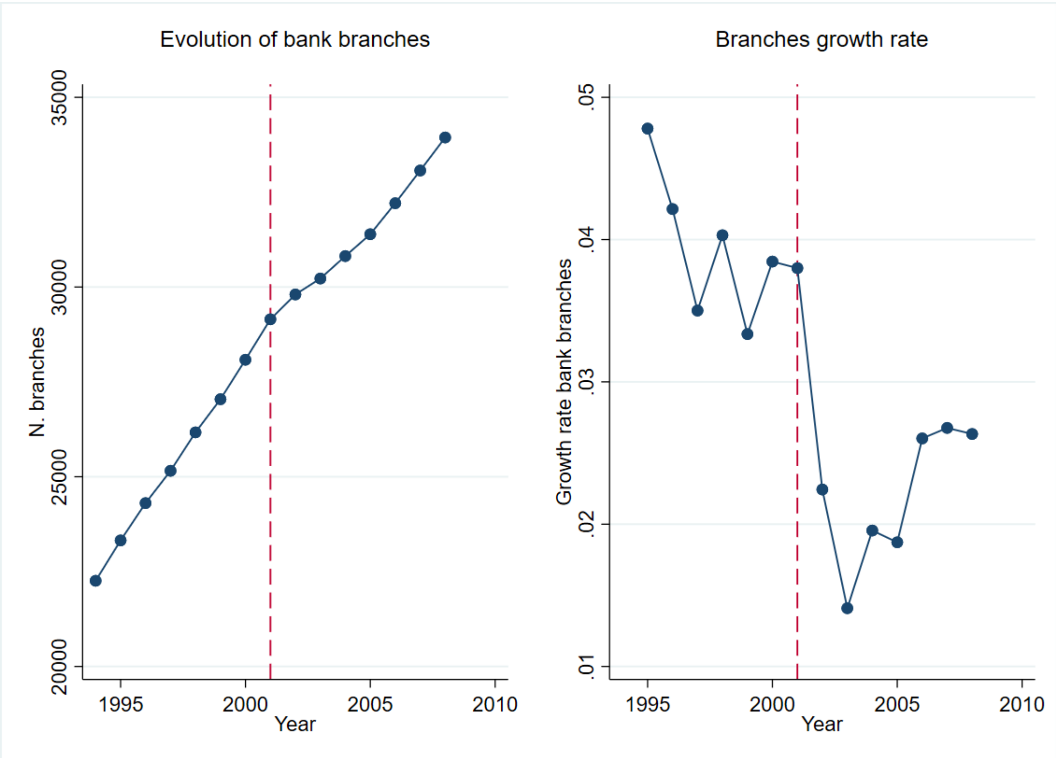
Notes: Geographical distribution of ADSL access in Italian municipalities. The left panel refers to 2004, the first year for which data on ADSL are available. The right panel refers to 2008, the last year in our sample. The measure of broadband internet is the baseline six-point asymmetric scale of ADSL coverage. Lighter colors indicate no or low access. Darker colors indicate high or full access.

Figure 3: DiD Event study: number of loans and credit amount



Notes: DiD setting. The treatment group is made by banks in municipalities with access to ADSL in 2006 (early adopters). The control group is made by banks in municipalities with no access to ADSL in 2006 (late adopters). Year 0 corresponds to 2002, the first year in which broadband internet is available. In the top panel, on the y-axis is $ln(N.\ loans)$, the natural logarithm of the number of loans issued by each bank. In the bottom panel, on the y-axis is $ln(Ext.\ credit)$, the natural logarithm of the total amount of credit granted by each bank. Both panels follow the indications of Borusyak and Jaravel (2017) and drop 1998, in addition to the baseline year (2001), from the computations.

Figure 4: Bank Branches in Italy, 1995-2010



Notes: This figure plots the evolution of bank branches in Italy during our sample period. On the left is reported the time series of the total number of branches. On the right is the growth rate of branches by year.

Tables

Table 1: Summary Statistics

	Mean	sd	p50	N
	Panel A: Municipality			
Municipalities				5,271
Years				11
North	0.61	0.49	1.00	51,290
Center	0.15	0.36	0.00	51,290
South	0.24	0.43	0.00	51,290
Internet	2.04	2.35	0.00	42,058
Number SLs	1.79	4.04	1.00	51,290
Distance SL	0.40	1.23	0.00	51,290
Number UGSs	0.13	1.10	0.00	51,290
Distance UGS	12.49	8.87	11.07	51,290
Distance prov. capital	21.96	12.93	20.00	50,859
	Panel B: Bank-municipality			
Number of loans	28.23	147.37	8	153,120
Extended credit	29,086.22	282,980.90	3,584.40	153,120
Average interest rate	6.10	1.70	5.98	86,382
	Panel C: Loan			
Extended credit	1,028.48	8,159.02	299.32	4,330,369

Notes: This table reports summary statistics for our final dataset. Panel A refers to data at the municipality level. We provide information on the municipality geographical distribution, as well as on access to broadband and the ADSL underlying infrastructure. Panel B refers to data at the bank-municipality level, that we use quite intensively throughout our analysis. We provide information on the number of loans issued by a bank in a given municipality, the amount of credit granted (in thousands of euros), and the average interest rate charged. Finally, panel C refers to loan level data and reports the credit amount.

Table 2: First stage regressions

	0-5 (Internet)	Dummy (Good access)	Dummy (Some access)	Years since good internet
distance UGS	-0.053***	-0.010***	-0.009***	-0.035***
$\times$ post 2001	(0.007)	(0.001)	(0.001)	(0.005)
Mun FE	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes
Method	OLS	OLS	OLS	OLS
F-statistic	52.6	46.5	46.6	47.2
Mean	2.041	0.437	0.451	1.206
R-squared	0.760	0.750	0.763	0.818
N	41932	41932	41932	41932

Notes: This table reports estimates from OLS as presented in equation (2). The dataset is at the municipality-year level. The dependent variables are: *Internet*, the percentage of households with access to ADSL-based services, in municipality m and year t, on an asymmetric six-point scale; *Good access*, a dummy variable that takes value 1 if broadband access is above 50%, and zero otherwise; *Some access*, a dummy variable that takes value 1 if broadband access is above 0%, and zero otherwise; and *Years since good internet*, a variable that counts the number of years since the percentage of households with access to the ADSL was above 50%. The main predictor is our instrument: the interaction between *Distance from UGS* and a dummy variable *post2001*. *Method* reports the used estimator; *F-statistic* reports the F-statistic from the regression; *Mean* refers to the mean of the dependent variable; *R-squared* is the adjusted R^2; and *N.* refers to the number of observations. Fixed effects are at the municipality and year level. Standard errors, in parentheses, are clustered at the municipality level. ***, ** and * indicate significance at the 1%, 5% and 10% level, respectively.

Table 3: Regressions of Internet on Banks' Number of Loans and Extended Credit

	(1)	(2)	(3)	(4)
	Ln	Ln	Ln	Ln
	(N. loans)	(Ext. credit)	(N. loans)	(Ext. credit)
Internet	0.007***	0.009***	0.039**	0.081***
	(0.002)	(0.004)	(0.016)	(0.024)
Bank-Mun FE	X	X	X	X
Year FE	X	X	X	X
Method	OLS	OLS	IV	IV
F-statistic			134.9	133.6
Mean	28.79	30094.456	28.79	30094.456
R-squared	0.901	0.860		
N	124243	123762	124243	123762

Notes: This table reports estimates from OLS and 2SLS as presented in equations (1) and (3). The dataset is at the bank-municipality-year level. The dependent variables are: *Ln(N. loans)*, the natural logarithm of the number of loans issued by bank b in year t; and *Ln(Ext. Credit)*, the natural logarithm of the amount of loans granted by bank b in year t. The main predictor is *Internet*, a measure of ADSL coverage in the municipality, based on a six-point asymmetric scale. Columns 1 and 2 refer to basic OLS estimates. Columns 3 and 4 refer to 2SLS estimates, where the variable *Internet* is instrumented by the interaction between *Distance from UGS* and a dummy variable *post2001*. *Method* reports the used estimator; *F-statistic* reports the Sanderson-Windmeijer multivariate F-statistic, when the 2SLS methodology is adopted; *Mean* refers to the mean of the dependent variable; *R-squared* is the adjusted R^2; and *N.* refers to the number of observations. Fixed effects are at the bank-municipality and year level. Standard errors, in parentheses, are clustered at the municipality level. ***, ** and * indicate significance at the 1%, 5% and 10% level, respectively.

Table 4: Weights OLS estimates on the extensive margin of the credit relationship

Internet		Instrument	
Year	weight	Year	weight
2004	0.34	2004	0.20
2005	0.27	2005	0.20
2006	0.19	2006	0.20
2007	0.11	2007	0.20
2008	0.08	2008	0.20

Notes: This table reports the weights associated with OLS estimates from equation (1). Estimates are reported for years from 2004 to 2008 only (post broadband). The left panel accounts for the regression of the number of loans (extensive margin of the credit relationship) on *Internet*. The right panel reports weights for the reduced form regression of the number of loans on the instrument, *Distance from UGS* $\times$ *post2001*. The weights associated with the coefficients of *Internet* are decreasing over time. Those associated with the coefficients of the instrument are, instead, constant.

Table 5: Regressions of Internet on Average Interest Rates

	(1) Average Rate	(2) Average Rate
Internet	-0.018***	-0.107**
	(0.007)	(0.045)
Controls	X	X
Bank-Mun FE	X	X
Year FE	X	X
Method	OLS	IV
F-statistic		318.4
Mean	6.81	6.81
R-squared	0.678	
N	112834	112834

Notes: This table reports estimates from OLS and 2SLS as presented in equations (1) and (3). The dataset is at the bank-municipality-year level. The dependent variable is *Average Rate*, the (weighted) average interest rate on loans issued by bank b in year t. The main predictor is *Internet*, a measure of ADSL coverage in the municipality, based on a six-point asymmetric scale. Column 1 refers to the basic OLS estimate. Column 2 and refers to the 2SLS estimate, where the variable *Internet* is instrumented by the interaction between *Distance from UGS* and a dummy variable *post2001*. *Method* reports the used estimator; *F-statistic* reports the Sanderson-Windmeijer multivariate F-statistic, when the 2SLS methodology is adopted; *Mean* refers to the mean of the dependent variable; *R-squared* is the adjusted R^2; and *N.* refers to the number of observations. Fixed effects are at the bank-municipality and year level. Standard errors, in parentheses, are clustered at the municipality level. ***, ** and * indicate significance at the 1%, 5% and 10% level, respectively.

Table 6: Regressions of Internet on Firms' Extended Credit

	(1) Ln (Ext. Credit)	(2) Ln (Ext. Credit)	(3) Ln (Ext. Credit)	(4) Ln (Ext. Credit)
Internet	0.007***	0.004***	0.114***	0.058***
	(0.002)	(0.001)	(0.017)	(0.009)
Controls	X	X	X	X
Bank-Year FE	X	X	X	X
Bank-Mun FE	X		X	
Firm-Year FE		X		X
Firm-Branch FE		X		X
Method	OLS	OLS	IV	IV
F-statistic			259.8	335.9
Mean	1057.814	1180.751	1057.814	1180.751
R-squared	0.153	0.948		
N	2115962	1643157	2115962	1643157

Notes: This table reports estimates from OLS and 2SLS as presented in equations (1) and (3). The dataset is at the firm-bank-municipality-year level. The dependent variable is *Ln(Ext. Credit)*, the natural logarithm of the amount of loans granted by bank b to firm f, in year t. The main predictor is *Internet*, a measure of ADSL coverage in the municipality, based on a six-point asymmetric scale. Columns 1 and 2 refer to basic OLS estimates. Columns 3 and 4 refer to 2SLS estimates, where the variable *Internet* is instrumented by the interaction between *Distance from UGS* and a dummy variable *post2001*. *Controls* refer to the loan (one-year lagged) share of revolving loans of the firm, and the loan share of extended credit of the issuing bank. *Method* reports the used estimator; *F-statistic* reports the Sanderson-Windmeijer multivariate F-statistic, when the 2SLS methodology is adopted; *Mean* refers to the mean of the dependent variable; *R-squared* is the adjusted R^2; and *N.* refers to the number of observations. Fixed effects are at the bank-municipality and bank-year level, in columns 1 and 3. The model is saturated with firm-bank-municipality fixed effects and firm-year fixed effects in columns 2 and 4. The latter aims at isolating the supply effect. Standard errors, in parentheses, are clustered at the municipality level. ***, ** and * indicate significance at the 1%, 5% and 10% level, respectively.

Table 7: Regressions of Internet on Banks' Lending efficiency (productivity and quality)

	(1)	(2)	(3)	(4)
	Ln	Asinh	Ln	Asinh
	(Ext./Empl.)	(NPLs/N. loans)	(Ext./Empl.)	(NPLs/N. loans)
Internet	0.010**	-0.000	0.072***	-0.002**
	(0.004)	(0.000)	(0.023)	(0.001)
Bank-Mun FE	X	X	X	X
Year FE	X	X	X	X
Method	OLS	OLS	IV	IV
F-statistic			123.4	134.9
Mean	1487.241	0.020	1487.241	0.020
R-squared	0.759	0.597		
N	116743	124243	116743	124243

Notes: This table reports estimates from OLS and 2SLS as presented in equations (1) and (3). The dataset is at the bank-municipality-year level. The dependent variables are: *Ln(Ext./Empl.)*, the natural logarithm of the amount of credit issued by bank employee; and *Asinh(NPLs/N. loans)*, the inverse hyperbolic sine of the share of non performing loans on total loans. The main predictor is *Internet*, a measure of ADSL coverage in the municipality, based on a six-point asymmetric scale. Columns 1 and 2 refer to basic OLS estimates. Columns 3 and 4 refer to 2SLS estimates, where the variable *Internet* is instrumented by the interaction between *Distance from UGS* and a dummy variable *post2001*. *Method* reports the used estimator; *F-statistic* reports the Sanderson-Windmeijer multivariate F-statistic, when the 2SLS methodology is adopted; *Mean* refers to the mean of the dependent variable; *R-squared* is the adjusted R^2; and *N.* refers to the number of observations. Fixed effects are at the bank-municipality and year level. Standard errors, in parentheses, are clustered at the municipality level. ***, ** and * indicate significance at the 1%, 5% and 10% level, respectively.

Table 8: Regressions of Internet on Loan Geography

	(1) Share (≠ Prov.)	(2) Asinh (Avg. Distance)	(3) Share (≠ Prov.)	(4) Asinh (Avg. Distance)
Internet	0.001	0.003	0.021***	0.062*
	(0.001)	(0.006)	(0.006)	(0.032)
Bank-Mun FE	X	X	X	X
Year FE	X	X	X	X
Method	OLS	OLS	IV	IV
F-statistic			94.0	133.5
Mean			0.16	18.17
R-squared	0.292	0.324		
N	81851	79425	81851	79425

Notes: This table reports estimates from OLS and 2SLS as presented in equations (1) and (3). The dataset is at the bank-municipality-year level, and focuses on new loans. The dependent variables are: *Share(Diff. Province)*, the share of the loans originated outside the province of the bank; and *Asinh(Avg. Distance)*, the inverse hyperbolic sine of the average geodesic distance between the centroid of the municipality of the bank and the location of the firm. The main predictor is *Internet*, a measure of ADSL coverage in the municipality, based on a six-point asymmetric scale. Columns 1 and 2 refer to basic OLS estimates. Columns 3 and 4 refer to 2SLS estimates, where the variable *Internet* is instrumented by the interaction between *Distance from UGS* and a dummy variable *post2001*. *Method* reports the used estimator; *F-statistic* reports the Sanderson-Windmeijer multivariate F-statistic, when the 2SLS methodology is adopted; *Mean* refers to the mean of the dependent variable; *R-squared* is the adjusted R^2; and *N.* refers to the number of observations. Fixed effects are at the bank-municipality and year level. Standard errors, in parentheses, are clustered at the municipality level. ***, ** and * indicate significance at the 1%, 5% and 10% level, respectively.

Table 9: Regressions of Internet on Bank Competitors in the municipality

	(1) Ln (Competitors)	(2) Ln (Competitors)
Internet	0.010***	0.071***
	(0.001)	(0.009)
Mun FE	X	X
Year FE	X	X
Method	OLS	IV
F-statistic		52.3
Mean	3.28	3.28
R-squared	0.962	
N	41858	41858

Notes: This table reports estimates from OLS and 2SLS as presented in equations (1) and (3). The dataset is at the municipality-year level. The dependent variables is *Ln(competitors)*, the natural logarithm of the number of bank (physical) competitors in municipality m, in year t. The main predictor is *Internet*, a measure of ADSL coverage in the municipality, based on a six-point asymmetric scale. Column 1 refers to basic OLS estimates. Column 2 refers to 2SLS estimates, where the variable *Internet* is instrumented by the interaction between *Distance from UGS* and a dummy variable *post2001*. *Method* reports the used estimator; *F-statistic* reports the Sanderson-Windmeijer multivariate F-statistic, when the 2SLS methodology is adopted; *Mean* refers to the mean of the dependent variable; *R-squared* is the adjusted R^2; and *N.* refers to the number of observations. Fixed effects are at the municipality and year level. Standard errors, in parentheses, are clustered at the municipality level. ***, ** and * indicate significance at the 1%, 5% and 10% level, respectively.

Table 10: Regressions of Internet on Competition (Deposits)

	(1) HHI	(2) Top 3 Share	(3) Top 5 Share	(4) HHI	(5) Top 3 Share	(6) Top 5 Share
Internet	-0.004***	-0.005***	-0.003***	-0.025***	-0.026***	-0.022***
	(0.000)	(0.000)	(0.000)	(0.003)	(0.002)	(0.002)
Mun FE	X	X	X	X	X	X
Year FE	X	X	X	X	X	X
Method	OLS	OLS	OLS	IV	IV	IV
F-statistic				88.4	88.4	88.4
Mean	0.68	0.96	0.99	0.68	0.96	0.99
R-squared	0.924	0.651	0.326			
N	49566	49566	49566	49566	49566	49566

Notes: This table reports estimates from OLS and 2SLS as presented in equations (1) and (3). The dataset is at the municipality-year level. The dependent variables are: *HHI*, the Herfindahl–Hirschman Index of bank deposits in municipality m and year t; *Top 3 share*, the share of deposits owned by top 3 banks in the municipality; and *Top 5 share*, the share of deposits owned by top 5 banks in the municipality. The main predictor is *Internet*, a measure of ADSL coverage in the municipality, based on a six-point asymmetric scale. Columns 1 to 3 refer to basic OLS estimates. Columns 4 to 6 refer to 2SLS estimates, where the variable *Internet* is instrumented by the interaction between *Distance from UGS* and a dummy variable *post2001*. *Method* reports the used estimator; *F-statistic* reports the Sanderson-Windmeijer multivariate F-statistic, when the 2SLS methodology is adopted; *Mean* refers to the mean of the dependent variable; *R-squared* is the adjusted R^2; and *N.* refers to the number of observations. Fixed effects are at the municipality and year level. Standard errors, in parentheses, are clustered at the municipality level. ***, ** and * indicate significance at the 1%, 5% and 10% level, respectively.

Table 11: Regressions of Internet on Existing Relationships

	(1) Dummy (Multiple)	(2) Dummy (Multiple Bank)	(3) Dummy (Multiple)	(4) Dummy (Multiple Bank)
Internet	0.001	0.001	-0.031***	0.034**
	(0.001)	(0.002)	(0.009)	(0.014)
Bank-Year FE	X	X	X	X
Bank-Mun FE	X	X	X	X
Firm FE	X	X	X	X
Method	OLS	OLS	IV	IV
F-statistic			191.5	191.5
Mean	0.91	0.18	0.91	0.18
R-squared	0.650	0.525		
N	633732	633732	633732	633732

Notes: This table reports estimates from OLS and 2SLS as presented in equations (1) and (3). The dataset is at the firm-bank-municipality-year level, and focuses on new loans. The dependent variables are: *Dummy(Multiple)*, a dummy variable for the loan issued to a firm already having a credit relationship; and *Dummy(Multiple Bank)*, a dummy variable for the loan issued to a firm already having a credit relationship with the same bank (in a different place). The main predictor is *Internet*, a measure of ADSL coverage in the municipality, based on a six-point asymmetric scale. Columns 1 and 2 refer to basic OLS estimates. Columns 3 and 4 refer to 2SLS estimates, where the variable *Internet* is instrumented by the interaction between *Distance from UGS* and a dummy variable *post2001*. *Method* reports the used estimator; *F-statistic* reports the Sanderson-Windmeijer multivariate F-statistic, when the 2SLS methodology is adopted; *Mean* refers to the mean of the dependent variable; *R-squared* is the adjusted R^2; and *N.* refers to the number of observations. Fixed effects are at the bank-municipality, firm and bank-year level. Standard errors, in parentheses, are clustered at the municipality level. ***, ** and * indicate significance at the 1%, 5% and 10% level, respectively.

Table 12: Regressions of Internet on Firms' Ext. Credit - Small Municipalities

	(1) Ln (Ext. Credit)	(2) Ln (Ext. Credit)	(3) Ln (Ext. Credit)	(4) Ln (Ext. Credit)
Internet	0.001	-0.001	-0.031	-0.005
	(0.003)	(0.004)	(0.029)	(0.046)
Controls	X	X	X	X
Bank-Year FE	X	X	X	X
Bank-Mun FE	X		X	
Firm-Year FE		X		X
Firm-Branch FE		X		X
Method	OLS	OLS	IV	IV
F-statistic			22.4	18.9
Mean	665.857	746.809	665.857	746.809
R-squared	0.243	0.967		
N	130647	47709	130647	47709

Notes: This table reports estimates from OLS and 2SLS as presented in equations (1) and (3). The dataset is at the firm-bank-municipality-year level, and includes information on small municipalities (below the median of population) only. The dependent variable is *Ln(Ext. Credit)*, the natural logarithm of the amount of loans granted by bank b to firm f, in year t. The main predictor is *Internet*, a measure of ADSL coverage in the municipality, based on a six-point asymmetric scale. Columns 1 and 2 refer to basic OLS estimates. Columns 3 and 4 refer to 2SLS estimates, where the variable *Internet* is instrumented by the interaction between *Distance from UGS* and a dummy variable *post2001*. *Controls* refer to the loan (one-year lagged) share of revolving loans of the firm, and the loan share of extended credit of the issuing bank. *Method* reports the used estimator; *F-statistic* reports the Sanderson-Windmeijer multivariate F-statistic, when the 2SLS methodology is adopted; *Mean* refers to the mean of the dependent variable; *R-squared* is the adjusted R^2; and *N.* refers to the number of observations. Fixed effects are at the bank-municipality and bank-year level, in columns 1 and 3. The model is saturated with firm-bank-municipality fixed effects and firm-year fixed effects in columns 2 and 4. The latter aims at isolating the supply effect. Standard errors, in parentheses, are clustered at the municipality level. ***, ** and * indicate significance at the 1%, 5% and 10% level, respectively.

Table 13: Regressions of Internet on the Easiest Target

	Dummy (Firm in small muni,out prov, with internet)	Dummy (Firm in small muni,out prov, with internet)
Internet	0.007***	0.054***
	(0.002)	(0.018)
Bank-Year FE	X	X
Bank-Mun FE	X	X
Firm FE	X	X
Method	OLS	IV
F-statistic		86.2
Mean	0.04	0.04
R-squared	0.598	
N	550197	550197

Notes: This table reports estimates from OLS and 2SLS as presented in equations (1) and (3). The dataset is at the firm-bank-municipality-year level, and focuses on new loans. The dependent variable is a dummy that identifies loans to firms in small municipalities connected to fast internet, out of the province of the bank, granted by banks in municipalities with at least 10,000 inhabitants. The main predictor is *Internet*, a measure of ADSL coverage in the municipality, based on a six-point asymmetric scale. Column 1 refers to basic OLS estimates. Column 2 refers to 2SLS estimates, where the variable *Internet* is instrumented by the interaction between *Distance from UGS* and a dummy variable *post2001*. *Method* reports the used estimator; *F-statistic* reports the Sanderson-Windmeijer multivariate F-statistic, when the 2SLS methodology is adopted; *Mean* refers to the mean of the dependent variable; *R-squared* is the adjusted R^2; and *N.* refers to the number of observations. Fixed effects are at the bank-municipality, firm and bank-year level. Standard errors, in parentheses, are clustered at the municipality level. ***, ** and * indicate significance at the 1%, 5% and 10% level, respectively.

Table 14: Internet on Bank Branches in the municipality

	(1)	(2)	(3)	(4)
	Ln	Ln	Ln	Ln
	(Branches)	(Branches)	(Branches)	(Branches)
		Small		Small
Internet	0.004***	0.000	0.054***	0.003
	(0.001)	(0.000)	(0.006)	(0.002)
Bank-Mun FE	X	X	X	X
Year FE	X	X	X	X
Method	OLS	OLS	IV	IV
F-statistic			111.9	166.2
Mean	1.88	1.04	1.88	1.04
R-squared	0.950	0.894		
N	137691	45837	137691	45837

Notes: This table reports estimates from OLS and 2SLS as presented in equations (1) and (3). The dataset is at the bank-municipality-year level. The dependent variables is *Ln(branches)*, the natural logarithm of the number of branches of bank b in municipality m, in year t. The main predictor is *Internet*, a measure of ADSL coverage in the municipality, based on a six-point asymmetric scale. Columns 1 and 2 refer to basic OLS estimates. Columns 3 and 4 refer to 2SLS estimates, where the variable *Internet* is instrumented by the interaction between *Distance from UGS* and a dummy variable *post2001*. Columns 1 and 3 refer to the all sample. Columns 2 and 4 refer to the sample of small municipalities. *Method* reports the used estimator; *F-statistic* reports the Sanderson-Windmeijer multivariate F-statistic, when the 2SLS methodology is adopted; *Mean* refers to the mean of the dependent variable; *R-squared* is the adjusted R^2; and *N.* refers to the number of observations. Fixed effects are at the bank-municipality (or municipality) and year level. Standard errors, in parentheses, are clustered at the municipality level. ***, ** and * indicate significance at the 1%, 5% and 10% level, respectively.

Table 15: Regressions of Internet on Income and Population

	(1) ln (Income)	(2) ln (Pop.)	(3) ln (Income p.c.)	(4) ln (Income)	(5) ln (Pop.)	(6) ln (Income p.c.)
Internet	0.003***	0.002***	0.002***	0.040***	0.028***	0.017***
	(0.000)	(0.000)	(0.000)	(0.003)	(0.002)	(0.002)
Mun FE	X	X	X	X	X	X
Year FE	X	X	X	X	X	X
Method	OLS	OLS	OLS	IV	IV	IV
F-statistic				52.6	53.7	53.7
Mean	123.914	11318	10.062	123.914	11318	10.062
R-squared	0.998	0.999	0.986			
N	33268	41932	33268	33268	41932	33268

Notes: This table reports estimates from OLS and 2SLS as presented in equations (1) and (3). The dataset is at the municipality-year level. The dependent variables are: *Ln(Income)*, the natural logarithm of income; *Ln(Pop.)*, the natural logarithm of population; and *Ln(Income p.c.)*, the natural logarithm of income per capita. The main predictor is *Internet*, a measure of ADSL coverage in the municipality, based on a six-point asymmetric scale. Columns 1 to 3 refer to basic OLS estimates. Columns 4 to 6 refer to 2SLS estimates, where the variable *Internet* is instrumented by the interaction between *Distance from UGS* and a dummy variable *post2001*. *Method* reports the used estimator; *F-statistic* reports the Sanderson-Windmeijer multivariate F-statistic, when the 2SLS methodology is adopted; *Mean* refers to the mean of the dependent variable; *R-squared* is the adjusted R^2; and *N.* refers to the number of observations. Fixed effects are at the municipality and year level. Standard errors, in parentheses, are clustered at the municipality level. ***, ** and * indicate significance at the 1%, 5% and 10% level, respectively.

Table 16: Regressions of Internet on Income and Population - Small municipalities

	(1) ln (Income)	(2) ln (Pop.)	(3) ln (Income p.c.)	(4) ln (Income)	(5) ln (Pop.)	(6) ln (Income p.c.)
Internet	0.004***	0.002***	0.002***	0.064***	0.053***	0.021***
	(0.000)	(0.000)	(0.000)	(0.007)	(0.005)	(0.004)
Mun FE	X	X	X	X	X	X
Year FE	X	X	X	X	X	X
Method	OLS	OLS	OLS	IV	IV	IV
F-statistic				122.8	126.4	122.8
Mean	25.125	2517	9.988	25.125	2517	9.988
R-squared	0.992	0.995	0.979			
N	16630	20913	16630	16630	20913	16630

Notes: This table reports estimates from OLS and 2SLS as presented in equations (1) and (3). The dataset is at the municipality-year level, and includes information on small municipalities (below the median of population) only. The dependent variables are: *Ln(Income)*, the natural logarithm of income; *Ln(Pop.)*, the natural logarithm of population; and *Ln(Income p.c.)*, the natural logarithm of income per capita. The main predictor is *Internet*, a measure of ADSL coverage in the municipality, based on a six-point asymmetric scale. Columns 1 to 3 refer to basic OLS estimates. Columns 4 to 6 refer to 2SLS estimates, where the variable *Internet* is instrumented by the interaction between *Distance from UGS* and a dummy variable *post2001*. *Method* reports the used estimator; *F-statistic* reports the Sanderson-Windmeijer multivariate F-statistic, when the 2SLS methodology is adopted; *Mean* refers to the mean of the dependent variable; *R-squared* is the adjusted R^2; and *N.* refers to the number of observations. Fixed effects are at the municipality and year level. Standard errors, in parentheses, are clustered at the municipality level. ***, ** and * indicate significance at the 1%, 5% and 10% level, respectively.

Chapter 2

High-speed Internet, Financial Technology and Banking *

Angelo D'Andrea ** and Nicola Limodio ***

Abstract

Exploiting the staggered arrival of submarine cables, we show that high-speed internet lifted financial markets in Africa. We document a novel mechanism through which high-speed internet promotes the role of banks and credit supply. Plummeting telecommunication costs induce banks to adopt critical financial technologies, like the real-time gross settlement system (RTGS), which lower transaction costs and promote credit. We find that upon connecting to high-speed internet, banks adopt the RTGS more extensively, reduce liquidity hoarding and increase interbank transactions and lending. We also observe that high-speed internet particularly strengthens firms in countries with weak pre-existing interbank markets.

JEL: G2, G21, O16, O12

Keywords: Fintech, Banking, Investment, Financial Development

* This paper has benefited from discussions with and the suggestions of Francesco Amodio, Giorgia Barboni, Matteo Benetton, Laura Blattner, Martin Brown, Emanuele Colonnelli, Francesco Decarolis, Ralph De Haas, Filippo De Marco, Taryn Dinkelman, Jon Frost, Nicola Gennaioli, Itay Goldstein, Andrew Hertzberg, Sean Higgins, Eliana La Ferrara, Joseph Kaboski, Christian Meyer, Tommaso Oliviero, Marco Pagano, Marco Pagnozzi, Lorenzo Pandolfi, Daniel Paravisini, Jacopo Ponticelli, Manju Puri, Ricardo Reis, Claudia Robles-Garcia, Farzad Saidi, Antoinette Schoar, Enrico Sette, Amir Sufi, Chad Syverson, Pietro Veronesi, Luigi Zingales, and seminar participants at Bank of Italy, BIS-BoE-CEPR Conference on Financial Innovation, Bocconi University, Copenhagen Business School, Econometric Society 2020 Virtual World Congress, European Finance Association 2020, Federal Reserve Bank of New York - 2nd FinTech Conference, FIMAD Webinar Series, NBER Summer Institute 2020 Corporate Finance, University of Naples - CSEF, Humboldt University, University of Sussex Business School, and other seminars and conferences. Daniele Goffi, Federico Lenzi and Giovanni Scotti Bentivoglio provided superb research assistance. We thank for financial support the BAFFI - CAREFIN and IGIER research center at Bocconi University. All errors are our own.

** angelo.dandrea@unibocconi.it, Bocconi University, Department of Economics, Via Roentgen 1, 20136 Milan, Italy.

*** nicola.limodio@unibocconi.it, www.nicolalimodio.com, Corresponding Author, Bocconi University, Depart- ment of Finance, BAFFI CAREFIN and IGIER, Via Sarfatti 25, 20136 Milan, Italy.

1 Introduction

High-speed internet induced a political and economic revolution in Africa by emancipating participation (Manacorda and Tesei, 2020; Guriev et al., 2021), promoting mobile money (Jack and Suri, 2014; Beck et al., 2018), and stimulating employment opportunities (Hjort and Poulsen, 2019; Bircan and De Haas, 2020). At the same time, local financial markets experienced profound and unprecedented changes due to this technological upgrade. Such disruption was particularly relevant for banks, as the largest financial intermediary in the continent (Levine, 2005; Beck and Levine, 2018). In fact, as McKinsey and Company (2018) reports, African banks reacted to the availability of fast internet by restructuring their business model toward novel financial technologies, FinTech, which offered an opportunity to reduce financial frictions (Buera et al., 2011; Kaboski and Townsend, 2012) and information asymmetries (Hertzberg et al., 2010, 2011, 2018).

In this research, we explore the direct relationship between high-speed internet, financial technology and banking. To do so, we combine extensive datasets with the staggered installation of fiber-optic submarine cables, which introduced extensive and long-lasting declines in telecommunication costs. Our evidence indicates that high-speed internet lifted the African banking system with sizeable and persistent gains in both lending and deposits. Results from the existing literature (Hjort and Poulsen, 2019; Bircan and De Haas, 2020) indicate that high-speed internet promoted credit demand, by boosting employment and business opportunities. Our work adds to this established dimension a new perspective: high-speed internet can lead to a stronger credit supply by encouraging the bank adoption of novel financial technologies.

In particular, we document a mechanism through which high-speed internet affects banks and credit supply in line with the work of Bolton et al. (2011), Bolton et al. (2021) and Denbee et al. (2021). As faster and more reliable connections become available, banks find convenient to adopt previously unavailable financial technologies, like the RTGS, which lower transaction costs in interbank exchanges and enhance integration in liquidity markets (Townsend, 1978; Zilibotti, 1994; Guerrieri and Lorenzoni, 2009). As a result, banks change their liquidity management by reducing inside liquidity (hoarding cash and short-term assets) and increasing outside liquidity (use of interbank assets and liabilities). This improvement in liquidity management strengthens lending, resulting in more extensive firm credit and performance. These findings complement a growing empirical literature analysing the mechanisms behind the role of banks and long-term credit in emerging markets (Breza and Liberman, 2017; Breza et al., 2018; Breza and Kinnan, 2018; Bustos et al., 2020).

We complement the staggered installation of fiber-optic submarine cables with four comprehensive datasets. First, we track the exact geography and timing of submarine cable arrival in Africa from the mid-1990s onward, following the work of Hjort and Poulsen (2019). Second, we build a bank-level dataset with balance sheet information following 629 African banks between 1997 and 2018. Third, we introduce a hand-collected dataset containing information on the adoption of a financial technology central for intermediation: the real-time gross settlement system (RTGS), for all African countries and banks in our sample. Finally, we construct a dataset following 32,761 firms in Africa from the World Bank Enterprise Surveys, and observe some indicators of firm credit (access to finance, credit amount, loan maturities) and performance (sales, employment).

To study how high-speed internet affects banking and FinTech adoption, we face a particularly challenging empirical constraint: the power of the test. Both faster internet connections and novel financial technologies tend to be gradually adopted over time, which makes statistical power a problem. The installation of fiber-optic submarine cables in Africa offers an ideal source of variation for two reasons. First, the magnitude of the effect is large, as this novel technology generated a radical improvement in reliability and a 98% decline in the cost of operating the RTGS compared to satellite technology (Detecon, 2013), commonly used in the local financial sector (African Development Fund, 2002). Second, African banking systems present considerable frictions in interbank markets, as clarified by Figure 1. Its left panel shows that while poorer countries display underdeveloped interbank markets, this is particularly the case in Africa (indicated with a square). The right panel reveals that African banks hoard vast amounts of liquid assets, between 45 and 55 per cent of their liabilities, in line with presenting dysfunctional interbank markets.

We begin our empirical analysis by studying how the arrival of high-speed internet affects some key banking variables: lending, government bonds, deposits, and equity. In this respect, we explore an event-study design with a five-year window and a staggered difference-in-difference, in which a bank is treated when its country of operation receives a submarine cable connection. Our results show that we cannot reject the existence of parallel trends before the treatment. At the same time, after the cable arrival, treated banks present sizeable gains in lending and deposits, which increase by 43% and 30% respectively, without differences in the holding of government bonds and equity over liabilities.

Given the current interpretation of these results as due to changes in credit demand, we add a novel dimension to studying this problem by investigating the supply factors behind this banking expansion. For this reason, we exploit a source of bank connection to high-speed internet that is unrelated to local demand factors. A bank operating in country c may gain a fast internet connection because its headquarter, located in country g, is connected. As a result, this measure of connectedness does not depend on the submarine cable in country c, which induces changes in demand. We verify that both the country and headquarter connection imply similarly-sized and positive coefficients, which may be identifying the presence of both demand and supply effects. Our key specification focuses on the supply effect by including country by year fixed effects, which allows us to compare banks operating in the same country and year, which only differ in the timing of high-speed internet access due to the heterogeneous headquarter connection. These estimates imply the presence of positive and significant supply effects of fast internet.

After presenting evidence on the effects of high-speed internet on credit supply, we explore a specific mechanism through which this technological progress may lower the marginal cost of bank funding and increase credit supply. In particular, we show that cheaper and more reliable internet connections promote the adoption of novel financial technologies, which lower transaction costs in liquidity markets and stimulate interbank activities. Through a set of event studies and staggered difference-in-difference specifications, we show two main facts. First, once the fiber-optic submarine cables arrive in a country, the probability of adopting the RTGS, an essential technology to operate interbank transfers, increases by 16% for the country and by 5.5% at the bank level, while it lies on parallel trends before the cable connection. Second, local interbank markets grow significantly upon the arrival of the cable and RTGS adoption.Interbank loans and deposits grow by 30% and 70% respectively, and liquidity hoarding declines by 14 percentage points. For a limited subset of banks, we also explore a sharper specification looking at the maturity of interbank transactions and observe a notable increase in short-term maturities. This fact offers additional support to the hypothesis that high-speed internet promotes a more extensive and frequent bank use of the interbank market.

To further investigate our mechanism, we analyse a crucial cross-sectional heterogeneity which allows us to characterize the effect of fast internet in greater detail. We test the assumption that high-speed internet mainly benefits banks with higher ex-ante transaction costs to operate in the interbank market. Such cost difference is established through an interbank "weak user" dummy, taking unit value for banks that were below the median usage of the interbank market in a country, before the submarine connection. Our results suggest that most of the findings in the previous analysis are driven by these banks becoming more active. The latter remains true even by focusing exclusively on the cross-sectional variation within a country-year unit of observation, hence comparing weak versus strong interbank users by including country by year fixed effects and factoring out the arrival of the cable itself.

Finally, we employ firm level data to investigate the relationship between high-speed internet and firms. We structure this exercise in two tests. First, we verify that firms in countries experiencing the arrival of the submarine cable exhibit an increase in some of their financial variables (access to finance, loan maturities), without significant differences in sales and workforce. Second, we note that such an impact is quantitatively relevant for countries presenting weak interbank markets before the arrival of the cable. We observe that the arrival of high-speed internet in countries with weak pre-existing interbank markets is associated with a 25% increase in access to finance, a 7% higher likelihood of receiving a bank loan, a doubling in loan maturities, a sizeable expansion in yearly sales and workforce.[1]

To verify the robustness of our results, we devote an extended section and appendix to explore alternative specifications and explanations. First, we provide empirical evidence on the lack of correlation between the arrival of the cable and economic and financial variables. Second, we complement the analysis of banking variables through several additional tests. We report an even-study design relying on the year in which the headquarter receives high-speed internet and find results in line with those that use the arrival of the cable in the country of operation. We offer a deeper analysis on weak interbank users by presenting a balance table, to understand which observables characterize these banks, and a scatter plot on their loan/deposits composition. In methodological terms, we check the sensitivity of our graphical results to the choice of different event windows and control for time-varying variables reflecting the institutional quality of the country, particularly focusing on financial regulation and contract enforcement. We include different layers of fixed effects, and also provide estimates on different samples. Finally, we offer additional evidence using firm-level data and show that inputs, as well as outputs, increase with the arrival of the cable.

The paper contributes to the literature on financial technology and banking. We highlight how new technologies help to innovate traditional bank functions, as liquidity management,[2]and go beyond the competition-enhancing effects of FinTech (Goldstein et al., 2019).[3]for investment management; Benetton

[1] The results on loan maturities are particularly interesting and in line with a liquidity story. As banks face lower costs for interbank transactions, this reduces their liquidity risk and makes long-term loans cheaper and more extensive, as (Choudhary and Limodio, 2017) find in Pakistan.

[2] This paper also brings to the data the predictions of theoretical papers on credit and growth (Townsend, 1978; Diamond and Dybvig, 1983; Bencivenga and Smith, 1991; Saint-Paul, 1992; Zilibotti, 1994; Acemoglu and Zilibotti, 1997; Goldstein and Pauzner, 2005; Guerrieri and Lorenzoni, 2009), in line with a growing empirical literature on liquidity risk and credit (Choudhary and Limodio, 2017; Limodio and Strobbe, 2017). On the specific institution of the interbank market, the findings of this paper are in line with the recent work of Heider et al. (2015) on equilibrium liquidity hoarding, Allen et al. (2018) on the heterogeneity in interbank access, Craig and Ma (2018) on the structure of interbank networks, Coen and Coen (2019) on risk propagation and Denbee et al. (2021) on the liquidity multiplier induced by the interbank network.

[3] Part of this literature highlights that the rise of FinTech changed the competitive landscape in many sectors, for example: Buchak et al. (2018) and Fuster et al. (2019) for lending; Bartlett et al. (2018) and Tang (2019) for consumer and peer-to-peer

et al. (2019) for crypotcurrencies and Beck et al. (2018) for mobile payment systems. Recently, D'Acunto et al. (2020) observe that novel financial technologies, like Robo-Advising, lower cultural biases and reduce discrimination in financial decisions. The closest papers on this topic focus on the effects of payment systems and technology on individual and firm behavior. Higgins (2019) exploits a natural experiment in debit card adoption and shows that financial technology changes individual transaction costs and affects consumption. On a similar ground, Crouzet et al. (2019) verify that the temporary shock of the Indian demonetization induced the persistent adoption of novel payment technologies, in line with a dynamic adoption model with network complementarities. Finally, Lin et al. (2021) study the effect of the telegraph introduction in 19th century China and find a significantly positive effect on banks branch network and geographic scope. These results are in line with our mechanism, but involving banks and liquidity management, instead of firms and individuals. At the same time, two papers show that high-speed internet promotes the adoption of information and communication technologies (ICT) (Augereau and Greenstein, 2001), and that this fosters the country-level adoption of the RTGS and other novel practices in central banking (Bech and Hobijn, 2006).

Our results shed also new light on the mechanisms through which high-speed internet can affect the economy. Eichengreen et al. (2016) study the effect of submarine cables on the foreign exchange market, and show results compatible with ours: as submarine cables arrive, local banks respond with their forex trades. Hjort and Poulsen (2019) show that submarine cables improve business opportunities and productivity, especially for high-skill sectors and workers. Finally, Bircan and De Haas (2020) show that historical credit access affects innovation and technology adoption across Russian firms and generate economic growth.

The rest of the paper is organized as follows. Section 2 introduces the empirical framework and presents the data. In section 3, we show the results from the empirical analysis. Section 4 presents some robustness checks. Finally, we offer concluding remarks in section 5.

2 Empirical framework and data

2.1 Fiber-optic submarine cables and banking

Submarine cables in the oceans have a long history, since in 1842 Samuel Morse demonstrated the feasibility of transmitting telegraphic signals over long distances. In 1850-1851, the first telegraphic cable under the sea connected England and France, while in 1866, the first long-term successful transatlantic ca-

lending; Berg et al. (2018) for credit scoring and Hertzberg et al. (2018) for screening; D'Acunto et al. (2019) and Abis (2017).

ble was laid between Canada and Ireland. Despite their rapid diffusion, early submarine cables suffered from reliability and capacity problems. In the absence of repeater amplifiers, they required high voltages to transmit signals over long distances, creating distortion, limiting carrying capacity and heightening the risk of short-circuiting.

At the turn of the 19th and the 20th century, the new science of transmitting higher frequencies was established, through the introduction of coaxial cables.[4] Nevertheless, commercialization was delayed by the two World Wars and the Great depression. The first modern submarine cable, TAT-1 (Transatlantic No. 1), a coaxial cable insulated using polyethylene and utilizing vacuum tubes as repeaters, was only laid in 1955.

In the 1980s, coaxial cables were replaced by modern fiber-optic cables– glass fibers conveying signals by light rather than electric current. The advantages of fiber-optic cables were several: from greater reliability, to higher capacity and faster speed of transmission. The first submarine fiber-optic cable was laid in 1986 between England and Belgium, while the first transatlantic cable connected France, the United Kingdom and the United States in 1988. At that time, the Internet was beginning to take shape, and the development of the global fiber-optic network and the Internet proceeded simultaneously.

Over the last 30 years, more than 1 million kilometres of fiber-optic submarine cables have been constructed. This extraordinary network provides the communication infrastructure at the base of modern Internet (Carter, 2010). The path of construction has been fairly irregular. After a great burst during the period 2000-2002, in conjunction with the dot-com bubble, the cable-laying industry has contracted severely, eventually coming back to the previous growth rates after the 2008 and the great financial crisis.

Nowadays, it is estimated that more than 95% of Information and Communications Technology data are carried on low cost modern fiber-optic submarine cables.[5] Transmission of data through cables has several advantages: it increases the reliability of connection, it increases the speed of the signal and the overall capacity, and reduces transmission costs. Figure 2 shows the average unit cost per Mb/s transported capacity, based on 2008 prices. The price was about 740,000 US$ for satellite transmission, compared with 14,500 US$ for submarine fiber-optic transmission (Detecon, 2013).

The submarine cables network is a core infrastructure of the modern financial system. Each day,

[4] In coaxial cables, the copper or copper-plated steel wire is surrounded by an insulating layer which is in turn enclosed by a metallic shield.

[5] These data refer to the testimony of D. Burnett before the Senate Foreign Relations Committee on the United Nations Law of the Sea Convention, 4 October 2007.

the Society for Worldwide Interbank Financial Telecommunications (SWIFT) transmits more than 15 million messages over cables to over 8,300 banking organizations, securities institutions and corporate customers in 208 countries all over the world (Burnett et al., 2013). Referring to the fiber-optic submarine network, the previous Staff Director for Management of the Federal Reserve, Steve Malphrus, observed: *"when the communication networks go down, the financial sector does not grind to a halt, it snaps to a halt"*(Burnett et al., 2013).[6]

High-speed connections improve the functioning of the banking system in several directions. First, banks benefit from better screening and scoring algorithms, as well as from enhanced information processes and human resource management, which can result in greater ability to interact with firms, households, and other banks. Second, the internal liquidity management of the banks is highly affected by faster connectivity, mostly through the access and use of the interbank market, where sizeable monetary transactions take place with intense frequencies. Being connected to the submarine cables network determines whether a bank can operate in real-time with a long list of counterparties. As a consequence, the network of the bank, its interbank operations and the speed of transactions can vary dramatically depending on whether the bank has access to the fiber technology.

The staggered arrival of fiber-optic submarine cables in Africa at the turn of the 20th and the 21st century, is key to understand the development of the banking sector in the continent. In this paper, we provide evidence on the fact that high-speed internet improves banking and facilitates access to credit. In particular, we show that the advent of fast internet in Africa has been paralleled by the development of domestic interbank markets, with a sensible increase in interbank liquidity. We propose this channel to explain the main effects of high-speed internet on banking.

In this process, we take the arrival of fiber-optic submarine cables as an exogenous technological shock, since it was mostly due to the need to increase connectivity between America, Europe and Asia. Our empirical analysis capitalizes on three main facts. First, before the arrival of submarine cables, the interconnection of national banking networks in Africa was mostly based on satellites (African Development Fund, 2002). In that regard, the fiber-optic technology represents a major shock to banks, as they transition from a more expensive to a cheaper technology and also from a less reliable to an extremely

[6] On December 26 of 2006, the Hengchun earthquakes occurred off the southwest coast of Taiwan, in a zone which connects the South China Sea with the Philippine Sea. The earthquakes not only caused casualties and building damage, but several submarine communications cables were cut, disrupting telecommunication services in various parts of Asia. The earthquakes catastrophically disrupted Internet services in Asia, affecting many Asian countries as China and Hong Kong. Consequences on financial transactions were important as well. In particular, the foreign exchange market was seriously affected. It was only the capillary presence of other submarine cables that avoided the instant halt of foreign exchange transactions.

reliable system. Second, geography matters for the timing of arrival. Distance from Europe was crucial for earlier cable receivers, whereas being on the route between America, Europe and Asia was key for more recent connections. Third, endogeneity issues, coming from the fact that submarine cables had the explicit purpose to improve the banking sector, are of second order. Indeed, the first submarine cable laid for the express purpose to foster electronic trading in financial markets was tested in September 2015 (Eichengreen et al., 2016), only toward the end of our sample, and it is not included in our analysis.[7] Before that time, fiber-optic cables had the broader aim to accommodate general telecommunication needs, namely long-distance telegraphic communication, telephone calls, fax and internet transmission.

2.2 Markets of liquidity in Africa

Africa constitutes the best laboratory to explore our research question. First, banks in Africa experience substantial liquidity risk, due to imperfect risk sharing and high volatility of deposits. Second, African countries are characterized by a limited functioning of local liquidity markets, which exacerbates deposit shocks. Third, the staggered arrival of submarine cables provides the ideal setting for econometric investigation. We think that statistical power in our analysis is high and that, among the reasons behind the positive effects of high-speed internet on banking, there is a critical impact caused by the reduction of interbank transaction costs.

Banks in Africa are severely impaired in their access to international capital markets, because of local regulation or, even more importantly, because of low international reputation. At the same time, most of the central banks in Africa are either legally unable or *de facto* unwilling to provide liquidity on a predictable basis. Figure 3 shows data on the status of discount window facilities for all countries in Africa, as described by Choudhary and Limodio (2017). Based on documentation either local or published by the International Monetary Fund and the World Bank, this figure confirms that more than 50% of African central banks are not actively engaged in discount window operations.

Moreover, local interbank markets are generally very small or non-existent, forcing African banks to rely on the hoarding of reserves and liquid assets to cushion themselves from liquidity shocks. Hence, a major reduction in the cost of interbank transactions, as the one caused by the arrival of fast internet, can dramatically affect the banking system and generate cascade effects on credit supply and firms activity.

Recent orientation of policymakers also acknowledges that lack of credit is mostly a supply problem, where liquidity risk and banks play a major role. For example, World Bank (2015) presents a survey of

[7] Our analysis explicitly focuses on the first fiber-optic cable landed in a country. The last cable in our sample approached Africa in December 2012 (see appendix A).

financial development among financial sector practitioners (bankers, central bankers, regulators, academics), from which two important messages emerge: 1) access to finance is a supply problem (75% of respondents agree); 2) domestic banks are core institutions determining how firms and households have access to finance (61% of respondents agree).

Our paper contributes to this literature by showing that, improving the functioning of liquidity markets, new technologies can generate positive effects on banks risk sharing, efficiency, credit supply, and economic growth.

2.2.1 Real-time Gross Settlement Systems

The arrival of high-speed internet in Africa, through fiber-optic submarine cables, led to a profound transformation of the banking system. Among the critical changes that modernized local banking networks, a prominent role has been played by the real-time gross settlement systems (RTGSs).

Real-time gross settlement systems are special interbank transfer systems where the transfer of money and securities takes place on a "real-time" and "gross" basis. RTGSs are typically used for low-volume, high-value transactions and their purpose is to reduce credit risks due to settlement lags. These systems exhibit an inherent trade-off. On the one hand, they reduce information asymmetries by giving an accurate picture of an institution's account at any point in time, thus lowering settlement risk. On the other hand, they require large amounts of liquidity in the system to work properly, which is a major shortcoming in a market that deals with short-term liquidity shocks.

In Africa, settlement risk represents a core friction for the expansion of interbank markets. As a result, RTGSs are seen in favourable light by practitioners and their adoption is related to more advanced and efficient interbank markets.[8][9] This technology has been often promoted by transnationals economic unions, like the Southern African Development Community in Southern Africa (SADC), the West African Economic and Monetary Union in West Africa (WAEMU), and the Economic Community of Central African States in Central Africa (CEMAC). At the same time, individual countries and banks

[8] It is important to notice that the 24 principles for financial markets infrastructures published in April 2012 by the Committee on Payment and Settlement System (CPSS) and the Technical Committee of the International Organisation of Securities Commission (IOSCO) emphasise final settlement in central bank money, in real-time, as the new global standard.

[9] RTGS systems have become crucial infrastructures in the modern financial system. In 2014, during the Second Libyan civil war, the Government of National Accord, through the Central Bank based in Tripoli, disconnected its two eastern branches from its automated clearing system, the Real-Time Gross Settlement (RTGS). The eastern branches were under the control of the competing faction, the Libyan National Army of general Khalīfa Belqāsim Haftar, and this move had the specific aim to prevent east-based authorities from accessing government accounts and funds and to limit their access to finance.

have invested in RTGS with the explicit purpose to reduce credit risk and deepen their access to liquidity markets.

In section 3, while exploring different mechanisms behind the effect of fast internet on banking, we show that connection to high-speed internet is associated to more liquid interbank markets. The increase in interbank liquidity, indeed, is partly due to the diffusion of real-time gross settlement systems fostered by the availability of the fiber-optic technology (see section 3.2).

2.3 Data and descriptive statistics

In the first part of our analysis, we focus on the effects of high-speed internet on banking. The main data sources are Bankscope and BankFocus from Bureau van Dijk. These databases contain financial variables and finance reports for about 30,000 public and private banks across the globe. We employ data from Bankscope until 2013, its last year of operations, integrated with data from BankFocus which allow to extend the sample up to 2018. We use Bankscope and BankFocus to construct the main dependent variables and some of the control variables that we use in the robustness section.[10]

We complement bank-level data with hand collected country level data for fiber-optic submarine cables. Our main source is TeleGeography maps, a Telecommunications market research and consulting firm providing data on the telecom industry since 1989. TeleGeography provides general information about the fiber cables – name, total length, owners (generally a consortium of public and private companies), list of landing points (country and landing location), and year from which the cable is ready to serve (RTS). Moreover, it supplies the shapefiles of the worldwide submarine cable network, that we use to generate our maps.

To construct the main predictor at the headquarter level, we extract information from the CvH database on bank ownership (Claessens and Van Horen, 2015) and complement it with information from individual bank's report. The CvH database contains full ownership information for the period 1995-2013 for most of the commercial banks included in our sample. We extend the panel until 2018 by hand-collecting the remaining information.

We finally combine our main dataset with two ancillary sources from the World Bank– the World Bank Global Financial Development Database (WB GFDD) and the World Bank Worldwide Governance Indicators (WB WGI). The WB GFDD is an extensive dataset of financial system characteristics for 214 economies, capturing various aspects of financial institutions and markets. The WB WGI contains aggregate and individual governance indicators for over 200 countries and territories over the period 1996-

[10] Our data have been homogenized following the routine by Duprey and Lé (2016).

2018, for six dimensions of governance. We use both the datasets to retrieve control variables at the country level.

In the second part of our analysis, we focus on a specific mechanism that can partially explain the effect of high-speed internet on banking and credit supply: the development of interbank markets. In addition to the datasets already mentioned, and to show the positive relationship between fast internet and RTGSs adoption at the country and bank level, we extract information from central banks websites and individual bank reports. While the year of adoption of RTGS at the country level is often public information, released by central banks and telecommunication authorities, it is generally difficult to retrieve information on RTGS at the bank level. For this reason, we examined individual bank's annual reports and digitized the year of adoption of RTGSs when available. This provided us with information on a restricted sample of banks that we integrated with data on RTGS at the country level to form our final panel dataset.

In the third part of our analysis, we focus on the real effects of connection to high-speed internet. We use part of the data collected for the banking analysis, combined with information on firm's characteristics, business activity, and funding, from the World Bank Enterprise Survey (WB ES). The WB ES offers an array of economic data for 164,000 firms in 144 countries gathered through different surveys. For the purpose of this paper, we focus on surveys conducted in African coastal countries, during the period 2002-2018.

Our final dataset includes 629 banks, located in more than 90 cities, distributed among 37 coastal countries in Africa, during the period 1997-2018.[11] Our firm dataset includes 32,761 firms in 31 African countries. Countries are all coastal and the amount of firms participating into the surveys is well-spread across them.[12]

For each country, we use the arrival of the first fiber-optic submarine cable to proxy for the positive technological shock on the access to high-speed internet. Moreover, we narrow the scope of our investigation to interbank markets, interpreting the connection to high-speed internet as a shock that reduces transaction costs for interbank transactions. We assume that once the cable lands in a country, banks in the sample are automatically connected. This assumption is motivated by two facts. First, banks in our sample are typically located in capital cities, which are usually the places receiving high-speed internet

[11] We focus on coastal countries because for those that are landlocked is not clear whether (and when) terrestrial connections have made available the access to the fiber-optic technology.

[12] There are only a few exceptions as Egypt-2013 and Nigeria-2014. Together, those two surveys account for 20% of our observations.

first. Second, among companies, banks are likely to be early-adopters since new technologies are generally associated with sizeable profits, as discussed by Hannan and McDowell (1984) and Frame et al. (2018).

The arrival of fiber-optic submarine cables in Africa has been staggered over time.[13] The years of arrival span from 1994 to 2013, and a graphical representation is provided in figure 4.

In this paper, we exploit the staggered arrival of submarine cables to identify the effect of high-speed internet on banking and then, on the real economy. In order to evaluate the effects of the new technology on banking, we focus on the following dependent variables: *Loans*, that proxies for credit to the private sector; *Government*, how much is invested in government securities; *Deposits*, that indicates deposits and short-term funding; and *Equity/Deposits and short-term funding*,[14] that shows the share of equity over deposits. To highlight that more efficient interbank markets can potentially explain the positive effect of fast internet on banking, we study: *Loans to banks* and *Deposits from banks*, that proxy for the liquidity of the interbank market; and *Liquid Assets/Deposits and short-term funding*, that proxies for the hoarding of liquid assets.[15] Finally, to assess the effects of the new technology on firms credit and business activity, we define: *Access to Finance*, a dummy variable that shows whether the firm considers access to finance to be an issue; *Loans from Banks*, whether the firm has issued at least one loan with a commercial bank in the last fiscal year; *Loans Maturity*, the term, in months, of loans from banks; *Sales*, the amount of total annual sales; and *Workforce*, the number of permanent and temporary full-time employees.[16]

As the main predictor in our analyses, we use a dummy, $Submarine$, that is a binary variable for the arrival of the fiber-optic submarine cable in the country (in the country of the headquarter when we attempt to isolate the supply effect). This dummy takes value 0 before the arrival of the cable, and 1 from the time of the arrival on. In some of our specifications, we also refine this variable to check for bank's and country's heterogeneity. In the analysis at the bank level, we concentrate on $Submarine \times Weak\ user$, namely the interaction between $Submarine$ and a dummy variable that specifies whether the bank was below the median of interbank volumes, in the country, before the arrival of the cable. In the analysis of the firms outcomes, we concentrate on $Submarine \times Weak\ Interbank$, where the latter

[13] Appendix A offers a table in which, for all the countries in Africa, we have the name of the first submarine cable landed, and the month and year from which that cable was ready to serve.

[14] This variable is trimmed to remove the first and the last percentile. The latter is made necessary by the presence of abnormal values that represent unreasonable outliers.

[15] Before, we also study how high-speed internet affects the adoption of RTGSs at country and bank level.

[16] *Sales* and *Workforce* are trimmed to remove the first and the last percentile. The latter is to avoid that outliers drive the results. Moreover, *Workforce* only considers firms with at least 5 employees.

represents the interaction between $Submarine$ and a dummy that specifies whether the country was below the median of interbank volumes before the arrival of the cable.

Table 1 provides summary statistics for both dependent (bank's and firm's) and independent variables. Column 1 refers to the number of observations. Columns 2 and 3 refer to mean and standard deviation. Finally, columns 4 to 6 show the 50th, 5th and 95th percentile.

2.4 Empirical Methodology

Our empirical strategy relies on four different methodologies. First, we develop an event study design meant to test for pre-trends and to investigate the dynamics of the treatment effect. Second, we implement a staggered difference-in-difference specification using two-way fixed effects regressions. The staggered difference-in-difference provides compact estimates of the average treatment effect under the assumptions of no pre-trends and constant treatment. Third, we offer a specific test to identify the existence of a distinct supply effect, by focusing on multinational banks. Fourth, we refine our analysis with the inclusion of a bank specific (or country specific, in the case of firm analysis) heterogeneity, being a (pre) weak user of the interbank market. In the last cases, we also augment our specifications with the inclusion of country by year fixed effects, which factor out the confounding effects associated to domestic unobservable variables. The next paragraphs provide a detailed description of each of the aforementioned methodologies.

The first specification that we propose is an event study based on the year of arrival of the fiber-optic submarine cable. The event study allows to check for pre-trends and, in a lesser extent, to provide evidence on the dynamics of the treatment effect. The empirical specification is as follows:

$$Y_{ict} = \alpha_i + \beta_t + \gamma_{-5} I\left\{K_{ct} \leq -5\right\} + \sum_{k=-4}^{4} \gamma_K I\left\{K_{ct} = k\right\} + \gamma_{5+} I\left\{K_{ct} \geq 5\right\} + \varepsilon_{ict} \qquad (1)$$

where: Y_{ict} represents the dependent variable,[17] for bank i, in country c, at time t; α_i and β_t are bank and year fixed effects; K_{ct} is the relative year from the activation of the cable (ACT_c), $K_{ct} = t - ACT_c$; γ_{-5} is the single coefficient for far leads; and γ_{5+} is the single coefficient for longer-run effects.

[17] Dependent variables are: *Loans* (natural logarithm of net loans (in million of US dollars)); *Government* (natural logarithm of government securities (in million of US dollars)); *Deposits* (natural logarithm of deposits and short-term funding (in million of US dollars)); and *Equity* (share of total equity over deposits and short-term funding), for the first part of the analysis. They are: *Loans to Banks* (natural logarithm of loans to banks (in million of US dollars)); *Deposits from Banks* (natural logarithm of loans to banks (in million of US dollars)); and *Liquid Assets over DST* (share of liquid assets over deposits and short-term funding), for the second part of the analysis.

We use the event study only for banking variables. The observation window is 1997–2018, while we restrict the event window to be the interval [-5;+5] from the year of arrival of the cable.[18] We assign value 1 to the dummies that are at the extremes of the event window, where $-5 \geq K_{ct} \geq 5$, and set the year before the arrival of the submarine cable as the baseline category, as standard in the literature.

The second specification that we propose is a canonical (two-way fixed effects) staggered difference-in-difference regression. We use this specification to study the effects of fast internet on banking and firm outcomes. Compared to the dynamic specification, it imposes no pre-trends and constant treatment effects. Hence, the staggered difference-in-difference provides a compact measure of the average causal effect of fiber-optic technology on our dependent variables. The empirical specification is as follows:

$$Y_{ict} = \alpha_{i(c)} + \beta_t + \gamma D_{ct} + \varepsilon_{ict} \tag{2}$$

where: Y_{ict} represents the dependent variable, for bank (firm) i, in country c, at time t; $\alpha_{i(c)}$ and β_t are bank (country) and year fixed effects; and D_{ct} is a dummy variable that equals one after the arrival of the first submarine cable in country c, and zero before.

The third specification that we propose is meant to isolate the effect of supply from that of demand (where the former is mostly induced by lower marginal costs in the interbank market). To identify the supply effect, we focus on multinational banks and define high-speed internet connectedness at the level of the bank headquarter, moving from D_{ct} to D_{gt}. The advantage of this specification is that banks can be connected to fast internet because the group to which they belong is connected, independently from their country being reached by the fiber-optic cable. The empirical specification is as follows:

$$Y_{igct} = \alpha_i + \beta_t + \sigma_{ct} + \gamma D_{gt} + \varepsilon_{igct} \tag{3}$$

where: Y_{igct} represents the dependent variable, for bank i, belonging to group g, in country c, at time t; α_i and β_t are bank and year fixed effects; and D_{gt} is a dummy variable equal to one after the arrival of the first submarine cable in the country of the headquarter. In this setting, γ captures the average effect of fast internet on bank i, when the headquarter of group g, regardless of the country, receives high-speed internet. To specifically account for country time-varying unobservables, we also include country by year fixed effects, σ_{ct}.

Finally, the fourth specification that we propose is a modified version of the staggered difference-in-difference regression in equation (2), that allows for the inclusion of a specific heterogeneity. We test

[18] We try with different specifications of the event window. Results are particularly stable (see appendix F.1 for the 3 years window).

the basic idea that the magnitude of the effect of the technology shock depends on the bank's relative decline of transaction costs in interbank transactions. In particular, banks that had higher transaction costs before the arrival of high-speed internet were the ones most exposed to the shock. With that purpose in mind, we define an indicator of (pre) weak user that takes value 1 if the bank was below the median of interbank volumes, in the country, before the arrival of high-speed internet, and zero otherwise.[19] Similarly, when dealing with the firm analysis, we define an indicator of weak interbank market that takes value 1 if the country amount of interbank transactions before the arrival of high-speed internet was below the median, and zero otherwise. Then, we implement the following empirical specification:

$$Y_{ict} = \alpha_{i(c)} + \beta_t + \gamma_1 D_{ct} \times X_{i(c)} + \gamma_2 D_{ct} + \varepsilon_{ict} \tag{4}$$

where: Y_{ict} represents the dependent variable, for bank (firm) i, in country c, at time t; $\alpha_{i(c)}$ and β_t are bank (country) and year fixed effects; D_{ct} is a dummy variable equal to one after the arrival of the first submarine cable in country c; and $X_{i(c)}$ is the bank (country) specific heterogeneity. Notice that the presence of the dummy D_{ct} and its interaction with $X_{i(c)}$ is not coupled by the inclusion of $X_{i(c)}$ alone, as the latter is absorbed by bank (or country) fixed effects. We also strengthen our findings for banking outcomes by augmenting equation (4) with the inclusion of country by year fixed effects, σ_{ct}.

3 Results

Following the structure of the paper, we divide this section into three subsections. In the first subsection, we study the effect of high-speed internet on banking, and isolate its supply-driven component. The second subsection identifies a specific mechanism whereby the effect of fast internet on banking is partly due to more liquid interbank markets. In the third and final subsection, we study the real effects on firms outcomes.

3.1 Banking

We exploit the staggered arrival of fiber-optic submarine cables on the African coast to show the effect of high-speed internet on four dependent variables: *Loans*, the natural logarithm of net loans, that proxies

[19] Following this definition, the subsample of banks considered in equation (4) excludes those for which data previous to the cable arrival are not available. Additionally, our measure of weak user, only considers data in the five-years before the cable arrival.

for bank credit to the private sector; *Government*, the natural logarithm of government securities, that quantifies the investments of banks in government bonds; *Deposits*, the natural logarithm of bank deposits and short-term funding; and *Equity/Deposits and short-term funding*, that measures the amount of shareholders' equity over bank deposits.

The first exercise that we propose is the event study as defined in equation (1). Results are reported in figure 5. The top-left panel refers to bank loans and shows no pre-trends, i.e., before the arrival of the cable, point estimates are close to zero and none of them is statistically significant. However, the coefficients become positive and statistically significant after the cable arrival. In particular, we observe a small jump at year zero, followed by a sizeable and gradual increase after year one. The top-right panel refers to the investments in government securities. Similarly to before, no pre-trends can be detected and coefficients are positive and stable after year one, even though none of them is statistically different from zero. The magnitudes of these coefficients is informative of the fact that much of the new credit granted by banks is directed to profitable (more risky) investments, rather than to safer government bonds. The bottom-left panel refers to deposits and short-term funding. The pattern is very similar to that of loans– no pre-trends and a gradual increase after the arrival of the fiber-optic cable, but with lower magnitudes. Finally, the bottom-right panel refers to the share of equity over deposits. Interestingly, the increase in deposits is perfectly paralleled by an equivalent increase in total equity.

Table 2 reports the estimates from the staggered difference-in-difference specification as defined in equation (2). This two-way fixed effects regression provides a compact measure of the average causal effect of high-speed internet on our four banking outcomes. It imposes no pre-trends[20] and assumes constant treatment effects. Results from table 2 confirm those from the event studies. Connection to the fiber-optic cable is associated with a significant increase in loans and deposits, while the coefficients associated with government securities and the share of equity over deposits are not statistically different from zero. The estimates for loans and deposits are also large in magnitude: having access to high-speed internet increases loans by 43% and deposits and short-term funding by 30%.

Results from figure 5 and table 2 represent core findings of the paper. They show the effect of high-speed internet on banking in Africa, emphasizing the increase in credit associated with the technological shock. However, these specifications do not help to distinguish between the demand-driven component and the supply-driven component of the aggregate effect. High-speed internet directly benefits firms and contributes to the rise of their productivity and production efficiency (Akerman et al., 2015; Hjort

[20] This justifies the order of our results. We present the event study before the staggered diff-in-diff to graphically show the lack of pre-trends.

and Poulsen, 2019), thus boosting their demand for credit. To make a step forward, in this paper we present a tailored exercise that aims to isolate the supply channel.

We exploit information on the multinational bank-group composition of banks in our sample to construct a proxy of the cable connection at the level of the headquarter.For banks belonging to multinational groups, high-speed internet may be directly available for the headquarter but not for the subsidiary located in a different country. In this way, we make use of a source of exogenous variation, the arrival of fast internet in the country of the headquarter (different from the country where the bank operates), immune from domestic demand factors. Using bank headquarters to identify effects on foreign branches is theoretically in line with the work of Bloom et al. (2012), who show how management practices tend to be persistent within multinational groups. This stylized fact applies also to banks as discussed by Detragiache et al. (2008), which we employ in our setting.

Figure D1 and table D1, in appendix D, report the results from the event study and the staggered diff-in-diff as in equation (3). Here, connection of the headquarter, $Submarine_{gt}$, is used as the only predictor. This variable is a dummy that takes unit value when the country of the headquarter is connected, and zero otherwise.[21]Estimates are in line with those in figure 5 and table 2.

In this setting, demand-side effects are alleviated but not eliminated. To fully control for credit demand, we proceed in two steps. First, we replicate the specification in equation (3) but using both $Submarine_{ct}$ and $Submarine_{gt}$ as predictors. Table 3 reports the main results and it is suggestive of a separate effect coming from the supply side.[22]

Second, we augment the specification in equation (3) with the inclusion of country by year fixed effects, which directly absorb time-varying demand factors. Results are reported in table 4, and provide robust evidence on the existence of a separate supply channel behind the overall effect of high-speed internet on banking. Comparing table D1 and table 4, we see that 66% of the effect of fast internet on bank loans and 70% of the increase in deposits are not driven by domestic demand factors. Rather, credit supply plays a major role. This finding has important implications for the policy debate over government investment in broadband infrastructure. Our results suggest that investing in new technologies stimulates banks efficiency and promotes the supply of credit to the private sector.

The supply-driven effect of high-speed internet on credit in Africa represents the main finding of

[21] Notice that, when the bank does not belong to an international group, the headquarter corresponds to the bank itself.

[22] In case the two variables were perfectly collinear, the coefficients associated to $Submarine_{ct}$ and $Submarine_{gt}$ would tend to abnormal numbers. In case only demand mattered, the coefficients associated to $Submarine_{gt}$ would not be statistically different from zero.

the paper. In the next section, we provide supportive evidence on the existence of a specific mechanism behind this effect: the development of domestic interbank markets.

3.2 The role of interbank markets

The previous section shows how connection to fast internet affects banking outcomes. In the last paragraphs, we documented the existence of a separate supply channel and quantified its magnitudes. Here, we offer a foundation behind this channel, and provide empirical evidence in favour of a specific mechanism: the increase in liquidity of interbank markets.

We propose three pieces of evidence to corroborate our hypothesis. First, we show that connection to fiber-optic submarine cables has a positive effect on the probability that countries adopt real-time gross settlement systems, and that banks within countries take part to these systems. Second, we show that high-speed internet positively affects the amount of loans to banks and deposits from banks, simultaneously reducing the dependence of African banks to the hoarding of liquid assets. Finally, we show that these effects are larger for banks that were weak users of the interbank market before the arrival of the cable (see section 3.2.1).

We start considering the real-time gross settlement system. The left panel of figure 6 plots the event study for RTGS adoption, where the x-axis refers to the relative year of arrival of the cable, and the y-axis is the probability that the country adopts the RTGS. The right panel of figure 6 plots the same event study but at the bank level, distinguishing between banks that participate to the real-time system, and those that remain out of it. Two things are worth noting. First, coefficients in the left panel are positive and monotonically increasing after the cable arrival, underlining a positive relationship between connection to high-speed internet and country adoption of RTGS. This is in line with Augereau and Greenstein (2001); Bech and Hobijn (2006), and our hypothesis that internet speed and connection reliability foster the take-up of financial technologies. Second, this positive relationship is confirmed for individual banks. The fact that the coefficients at the country level are larger in magnitude with respect to the coefficients at the bank level clarifies that not necessarily all the banks adopt RTGS when it is available. The latter is of particular interest, and in line with what we know about the US interbank market and Fedwire adoption.

Table 5 presents compact estimates from the staggered diff-in-diff specification as defined in equation (2). Column 1, refers to the RTGS adoption at the country level. Column 2, shows the results for individual banks. Consistently with the event studies, coefficients in table 5 show a positive relationship between high-speed internet and the adoption of RTGS at the country level. Column (1) of Table 5 indicates that upon the arrival of the submarine cable, countries increase their probability of adopting the RTGS by 16%, which is particularly high given that the endline mean dependent variable is only 45%. Column (2) complements this result, by showing that the high-speed internet connection also induces a higher probability that the single banks join the real-time gross settlement system once the country adopt it. The magnitude in this case is smaller, 5.4%, but very precise and sizeable, as it implies a 10% increase relative to the baseline probability that a bank adopts the RTGS.

We continue by showing the direct effects of high-speed internet on interbank transactions. The event study in figure 7 provides evidence in that direction. The bottom-left panel refers to bank's loans to other banks, and shows two phases: a) no pre-trends before the arrival of the submarine cable, with point estimates that are close to zero and non-statistically significant; b) an upward trend with statistically significant coefficients after the arrival of the fiber-optic cable, with a jump at year zero and a gradual increase in loans to banks in the following years. The bottom-right panel refers to bank's deposits from other banks. Similarly to before, the pattern is almost flat previous to the arrival of the cable and increasing from then on– we observe a jump at year zero and a gradual increase in the next five years. The magnitude of the effect is even larger for deposits from banks.[23] Finally, the top panel refers to bank's liquid assets as a share of deposits and short-term funding. The pattern for the share of liquid assets is almost flat before the arrival of the submarine cable– with none of the estimates statistically significant– and then it sharply declines at year zero, remaining negative and stable over time.

Table 6 reports the estimates from the staggered diff-in-diff specification as defined in equation (2). Evidence from table 6 is in line with that presented in the event studies. Access to the fibre-optic connection determines an increase in loans to banks and deposits from banks, and a correspondent decrease in the share of liquid assets over deposits and short-term funding. All the coefficients are statistically significant and they are large in magnitude. The introduction of high-speed internet increases by 30% the amount of loans that banks in the sample provide to other banks and by 70% the amount of deposits from banks. Considering a hypothetical bank that has median values of both loans to banks (46 million of US$) and deposits from banks (19 million of US$), the access to the new technology causes an increase in loans to banks by 13.8 million of US$ and an increase in deposits from banks by 13.3 million of US$. The coefficient associated to the share of liquid assets is negative and significant. Having access

[23] This finding seems to suggest that the new technology, reducing transaction costs for interbank operations and lending risk within the interbank market, changes the structure of the banking network. Small and marginal banks benefit from the reduced interbank lending risk and lend their excess liquidity to big and core banks. Big and core banks, that mostly compose our sample, mostly act as liquidity hubs. This occurrence is even more plausible if lending risk outside of the interbank market does not decrease.

to the fiber-optic technology causes a decrease in the share of liquid assets over deposits and short-term funding of about 14 percentage points (that is a huge number if we consider that the average share in our sample is 46%). Our interpretation of these findings is that connection to high-speed internet reduces transaction costs for interbank operations, allows interbank markets to be effective in smoothing for liquidity shocks, and leads banks to substitute unprofitable hoarding of liquid assets with real-time interbank transactions.[24]

To corroborate our interpretation, we implement an additional exercise looking at interbank maturities. Following our hypothesis, once interbank transactions become an effective tool for smoothing liquidity shocks, banks substitute hoarding of liquid assets with shorter-term interbank securities. In that regard, we expect a relative increase in short-term interbank transactions with respect to longer-term transactions once the fiber-optic technology is available. To show that, we repeat our staggered diff-in-diff regression as in equation (2), but using as dependent variables a set of dummies identifying different interbank maturities. Results are reported in figure 8, and show that the coefficient associated with short-term interbank maturities (less than three months) is positive and statistically significant, while those associated with longer-term maturities are lower in magnitude and indistinguishable from zero. While this exercise offers valuable insights, we acknowledge that the limited availability of data poses serious concerns on its accuracy. Hence, we interpret the corresponding results with caution, as suggestive evidence of the effects of fast internet on interbank maturities.

3.2.1 Interbank Markets and Weak Users

Given the previous results, we show that the effect of high-speed internet on banking is heterogeneous among banks. Our proposed mechanism suggests that connection to fast internet (mainly) affects bank credit supply through the reduction of transaction costs in the interbank market. Following this reasoning, we expect banks that were minor players in the interbank market before the arrival of the cable, to be mostly affected by access to fast internet. To test this hypothesis, we augment the event study and the staggered diff-in-diff specification in equations (1) and (2) with the inclusion of our heterogeneity as defined in equation (4). We define a dummy variable for weak interbank user that, for each country, takes value 1 if the bank was below the median of loans to banks and deposits from banks before the arrival of the fiber-optic cable, and zero otherwise. In appendix E, table E1, we provide a balance table for this variable. *Weak users* tend to be smaller than other banks and have more assets over deposits. In

[24] This chain of events eventually unblocks new investment opportunities and determines the expansion in private sector credit that we have seen in the previous section.

appendix E, figure E1, we also plot a scatter of *weak user* vs *non-weak user* banks, and show that the two groups do not differ in their ratio of net loans over deposits and short-term funding. Then, we interact this predetermined variable with the dummy that identifies the presence of high-speed internet, D_{ct}, and focus on the specification in equation (4).[25]

Results from this specification are reported in figure 9 and table 7. In line with our hypothesis, most of the findings from figure 7 and table 6 come from the changing behaviour of (pre) weak users. The increase in deposits from banks and the reduction in the share of liquid assets over deposits and short-term funding, common to all banks, are sensibly higher for weak users. On the other hand, the increase in loans to banks has to be exclusively attributed to weak users.

In table 7, beyond providing compact estimates of the heterogeneous effect of fast internet on interbank outcomes, we also add a column for net loans. The table shows a differential increase in loans to banks by 50%, 0.397 log points, and in deposits from banks by 80%, 0.586 log points, for weak users. The reduction of liquid assets over deposits and short-term funding is 7.7 p.p. on average, with weak users reducing this ratio by further 11 p.p.. The last column shows that loans to the private sector increase for all banks, with weak users further increasing loans by 39%, 0.274 log points. The results from this table confirm our hypothesis: connection to high-speed internet lowers transaction costs in interbank exchanges, allowing constrained banks to change their liquidity management by reducing inside liquidity (hoarding short-term assets) and increasing outside liquidity (use of interbank assets and liabilities), thus strengthening lending to the private sector.

To conclude, we modify equation (4) by including country by year fixed effects. Results are presented in table 8, and show that point estimates remain basically unchanged with respect to table 7 (if something they increase), while standard errors decrease. We interpret this finding as a powerful test supporting our main results.

3.3 Firms

This section studies the effects of high-speed internet on firms, and constitutes the last link in the chain that connects fast internet to the real economy through the banking sector. We use data from the WB ES and focus on survey waves for African coastal countries, from 2002 to 2018.

We use similar empirical methodologies as the ones implemented in the previous sections, and study the effects of fast internet on the following dependent variables: *Access to Finance*, a dummy variable that

[25] Notice that the variable *weak user* is available for a restricted sample of banks, those that have data before the arrival of the submarine cable. In appendix E, table E2, we show that our main results remain unchanged for this restricted sample.

indicates whether managers in the firm consider access to finance a minor problem; *Loans from Banks*, a dummy variable that indicates whether the firm took at least one loan with a commercial bank to finance its activity during the last fiscal year; *Loans Maturity*, the duration, in months, of loans maturities; *Sales*, the amount of total annual sales; and *Workforce*, the number of permanent and temporary full-time employees.

In table 9, we present the results from the staggered diff-in-diff specification as defined in equation (2). As usual, this two-way fixed effects regression provides a compact measure of the average causal effect of high-speed internet on firms access to finance, their ability to borrow from banks, loans maturities, their total annual sales, and workforce. The baseline assumptions are absence of pre-trends and constant treatment effect among groups and through time. Results from table 9 suggest a positive relationship between high-speed internet and firms activity. Being connected to the fiber-optic cable is associated with an easier access to finance and an increase in loans maturities. To a lesser extent, we also see an increase in the probability of receiving a loan from commercial banks, and total annual sales, while the coefficient on workforce is close to zero and non-significant.[26]

In table 10, we present the results from the staggered diff-in-diff specification as defined in equation (4). Mirroring the analysis in section 3.2.1, we test the hypothesis that the real effects associated with access to fast internet are larger for those countries where the interbank market was underdeveloped before the arrival of the cable. To test this hypothesis, we define an indicator of weak interbank market that takes value 1 if the amount of interbank volumes in the country was below the median before the arrival of high-speed internet, and zero otherwise. Then, we interact this predetermined variable with the dummy that identifies connection to the submarine cable, D_{ct}. Results from table 10 corroborate our hypothesis. The effect of high-speed internet on corporate finance, loans maturity, sales and workforce is especially marked (and positive) for firms in countries where the interbank market was relatively underdeveloped. We observe that the arrival of fast internet in these countries is associated with a 25% increase in access to finance, a 7% higher likelihood of receiving a bank loan, a doubling in loan maturities, a sizeable expansion in yearly sales, and an increase in workforce by 4%.

Although this part of the analysis mainly offers suggestive evidence on the real effects of fast internet, it complements in a coherent way the findings on banking. Consistently with our story, the arrival of high-speed internet in Africa reduced transaction costs in the interbank markets, encouraging countries to invest in financial infrastructures and banks to reduce their hoarding of liquid assets. The latter,

[26] The coefficient associated to total sales is particularly high in magnitude. We take this coefficient with caution and look for new and comparable data in order to corroborate our findings.

induced banks to reallocate funds towards the private sector, thus, promoting local business activities and economic growth.

4 Robustness

In the previous sections we presented the main results of the paper. First, we documented a positive effect of fast internet on banking, isolating its demand-driven component. Then, we investigated a specific channel, the development of liquid interbank markets, to explain the effect of high-speed internet on credit supply. Finally, we reported a corresponding beneficial effect on firms. In this section, we provide additional checks to test the robustness of our results.

In table A1, appendix A, we present the list of countries included in our sample. For each country, we report their first fiber-optic submarine cable and the year from which it is ready to serve. Anecdotal evidence indicates that submarine cables approached Africa in their route from Europe and America to Asia. Furthermore, the first cables aimed to accommodate general telecommunication needs, namely long-distance telegraphic communication, telephone calls, fax and internet transmission. Connecting African banks was not the primary purpose behind the cable installation, though, concerns about the exogeneity of cables arrival remain. In tables C1 and C2, appendix C, we provide empirical support to the fact that the timing of connection is exogenous to the economic and financial characteristics of the countries being connected. Table C1 shows cross-country regressions where the year of arrival of the cable is regressed on three indicators that proxy for the economic outlook of the country and its banking sector before the cable arrival.[27] Pre-levels of economic and banking activity do not predict the timing of connection. Table C2 reports panel regressions where submarine connectedness is regressed over the current values of the economic and banking indicators, and their 10–years average. Again, the evidence suggests that none of the variables is correlated with fiber-optic submarine connection.

In the main analysis we exploit connection of the headquarter to disentangle the effect of supply from that of demand. There, we presented a table with both connection at the country level and at the headquarter level to show the existence of a separate supply channel. Then, we introduced country by year fixed effects to isolate the effect of supply. In figure D1 and table D1, appendix D, we substantiate these findings by providing the event study and the diff-in-diff regression in which the connection of the headquarter is used as the main predictor. Both the difference-in-difference estimates and the dynamics of the effects resemble those in figure 5 and table 2.

[27] These indicators come from the WB GFD database.

In the section dedicated to the interbank market, a central role is played by the bank heterogeneity on weak and non-weak (pre) users. Through it, we highlighted the heterogeneous effect of high-speed internet on banking associated with differences in transaction costs in the interbank market. Table E1, appendix E, refers to the balance table of the variable *Weak user*. The first two rows show the differences in loans to banks and deposits from banks that define the variable. Then, weak users differ from non-weak users also in terms of size, being them generally smaller. Instead, the two groups are substantially similar concerning assets over deposits ratio, loans over assets, government securities over assets, and liquid assets over total assets. In figure E1, we present a scatter plot where weak users and non-weak users are compared on the basis of their loans/deposits composition. Importantly, the two categories lie on the same line and the cloud of points is mostly overlapping. To complete our analysis, we also replicate the exercise in table 2 on the restricted sample of banks for which the variable *Weak user* is non-missing. Results are reported in table E2, and show that the basic estimates are unaffected.

Figures F1a and F1b, in appendix F.1, perform the same event studies as in the main analysis but focusing on a 3-years event window. Graphical analysis indicates that our results are not sensitive to the choice of the event window.

Throughout the panel, there are cases in which banks change the consolidation procedure used to draw up their balance sheets. To control for the confounding effect of a change in the consolidation procedure, in tables F2a and F2b, appendix F.2, we add bank by consolidation code fixed effects. The results are mostly unchanged.

Allen et al. (2018) show that existing differences in interbank market usage can be explained by the trust of the market participants in the stability of the country's banking sector and counterparties. To check for this alternative hypothesis, in tables F3a and F3b, appendix F.3, we include different control variables from the World Bank WGI database. *Regulatory quality*, captures perceptions of the ability of the government to formulate and implement sound policies and regulations that permit and promote private sector development; *Rule of law*, captures perceptions of the extent to which agents have confidence in and abide by the rules of society, and in particular the quality of contract enforcement, property rights, and the courts; *Control of corruption*, captures perceptions of the extent to which public power is exercised for private gain, including both petty and grand forms of corruption, as well as capture of the state by elites and private interests. Results illustrate that our estimates are highly resilient to the inclusion of these controls.

In our dataset, some of the main outcome variables present missing values. To overcome this limit, in tables F4a and F4b, appendix F.4, we replicate our analysis on an extended sample, where missing values are imputed using multivariate normal regressions.[28] Estimates from these specifications validate those in the main analysis and show that our results are robust to the inclusion of imputed data.

The increase in loans to the private sector, documented in the main analysis, may be associated with an increase in the riskiness of banking assets. In figure F5, appendix F.5, we investigate this aspect focusing on loan loss reserves. Our results suggest that the level of risk associated with loans remains basically unchanged after the arrival of fast internet.

Finally, we provide some robustness at the firm level. Tables F6a and F6b, appendix F.6, focus on expenses for electricity and raw materials. As the tables reflect, high-speed internet is associated with higher investments in firms inputs, especially in those countries that had weaker interbank markets before the cable arrival. The increase in expenses in firms inputs is in line with the evidence on access to finance and sales that we document in the main analysis.

5 Conclusion

In this research, we offer empirical evidence on the impact of high-speed internet on banking in Africa. To address this question, we combine data on African banks with those on the arrival of fiber-optic submarine cables, and country and bank reports with information on domestic financial infrastructures. We follow 629 banks in 37 African coastal countries, during the period 1997-2018, and show the effects of connection to fiber-optic cables in a quasi-experimental design. The latter relies on the staggered arrival of submarine cables, and the fact that African countries were primarily connected to increase the connectivity between America, Europe and Asia. To explore our research question, we offer a variety of econometric methods and focus on credit supply. We use the event studies to show the dynamics of banking outcomes around the year of cable arrival, while the difference-in-difference specifications provide compact estimates. Moreover, we exploit a key source of bank heterogeneity based whether a bank was a weak interbank users prior to the cable installation, to help characterize the main mechanism. Our findings highlight that high-speed internet fosters countries adoption of a financial technology central for bank intermediation, the real-time gross settlement system (RTGS), which contributes to a systematic increase in interbank transactions and promotes lending to the private sector.

In a supplementary exercise, we also follow 32,761 African firms and show the real effects of high-speed internet. Results in this second part indicate that connection to fast internet increases firms access

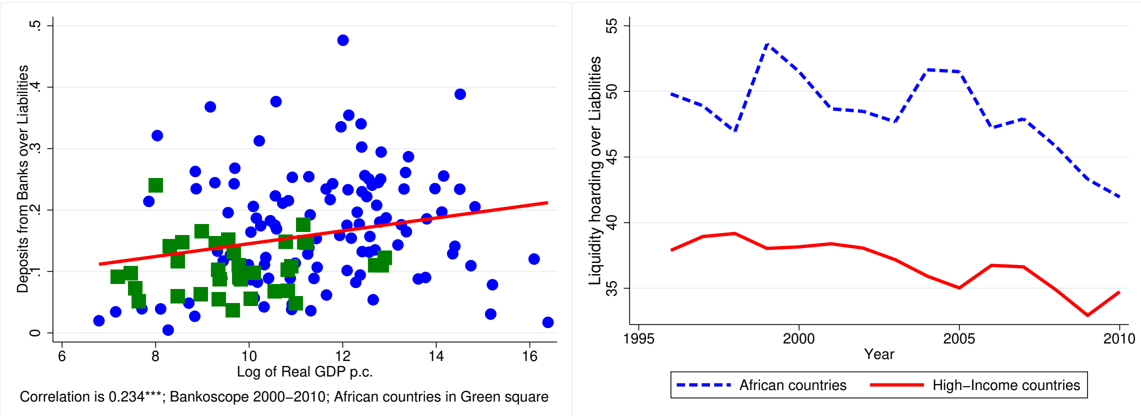
[28] We use *mi impute mvn*, from STATA, to fill in missing values using multivariate normal regression. Our procedure accommodates arbitrary missing-value patterns and uses an iterative Markov chain Monte Carlo (MCMC) method to impute missing values.

to finance and loans maturity. This effect, together with an increase in sales and full-time employment, is particularly relevant in countries that had an underdeveloped interbank market before the arrival of the fiber-optic cable.

Overall, we believe that our results are consistent with high-speed internet promoting financial technology adoption, liquidity and credit. In particular, this paper sheds light on two critical elements for future research. First, the adoption of innovative financial technologies can shape both the business outside the bank and its inside functioning, as the liquidity management. Second, promoting the size and the speed of interbank markets can improve financial integration, risk-sharing and ultimately credit and development. To conclude, our paper directly speaks to policymakers, as it suggests that investments in new technologies can favour capital markets integration, and help the convergence of underdeveloped countries.

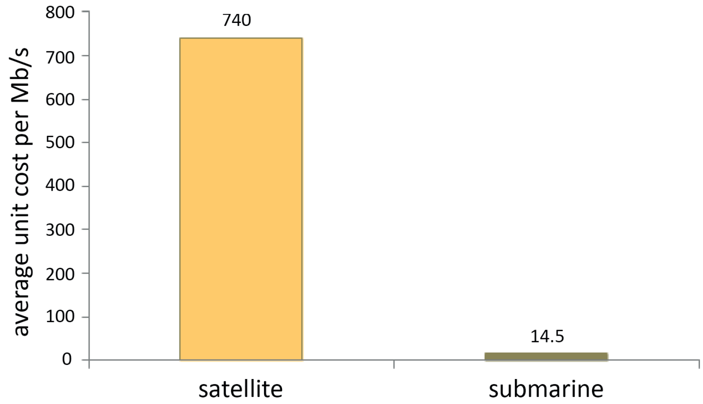
Figures

Figure 1: Interbank Markets and Liquidity Hoarding



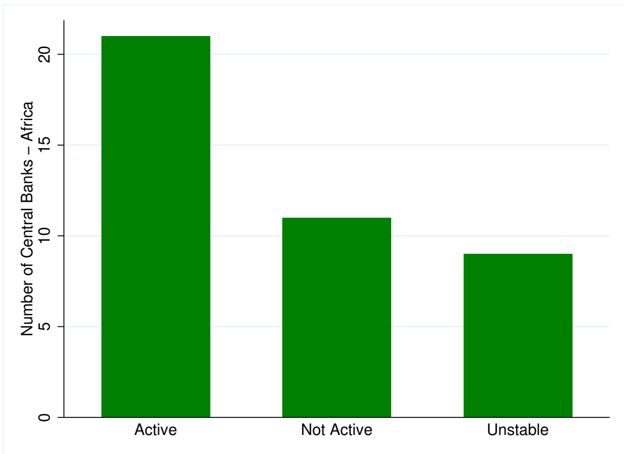
Notes: This figure reports information on interbank transactions and hoarding of liquid assets by banks in Africa. The left panel uses data from Bankscope on deposits from banks and total liabilities, while data on GDP per capita are extracted from the World Development Indicators published by the World Bank. Each dot represents a country, while green squares indicate banks operating in Sub-Saharan Africa. The right panel uses data from the World Bank Global Financial Development Database. The blue dashed line indicates liquid assets over liabilities for countries belonging to the Sub-Saharan region of Africa. The red line refers to the group of high-income countries according to the World Bank classification.

Figure 2: Internet Cost per Mbit/s Transported



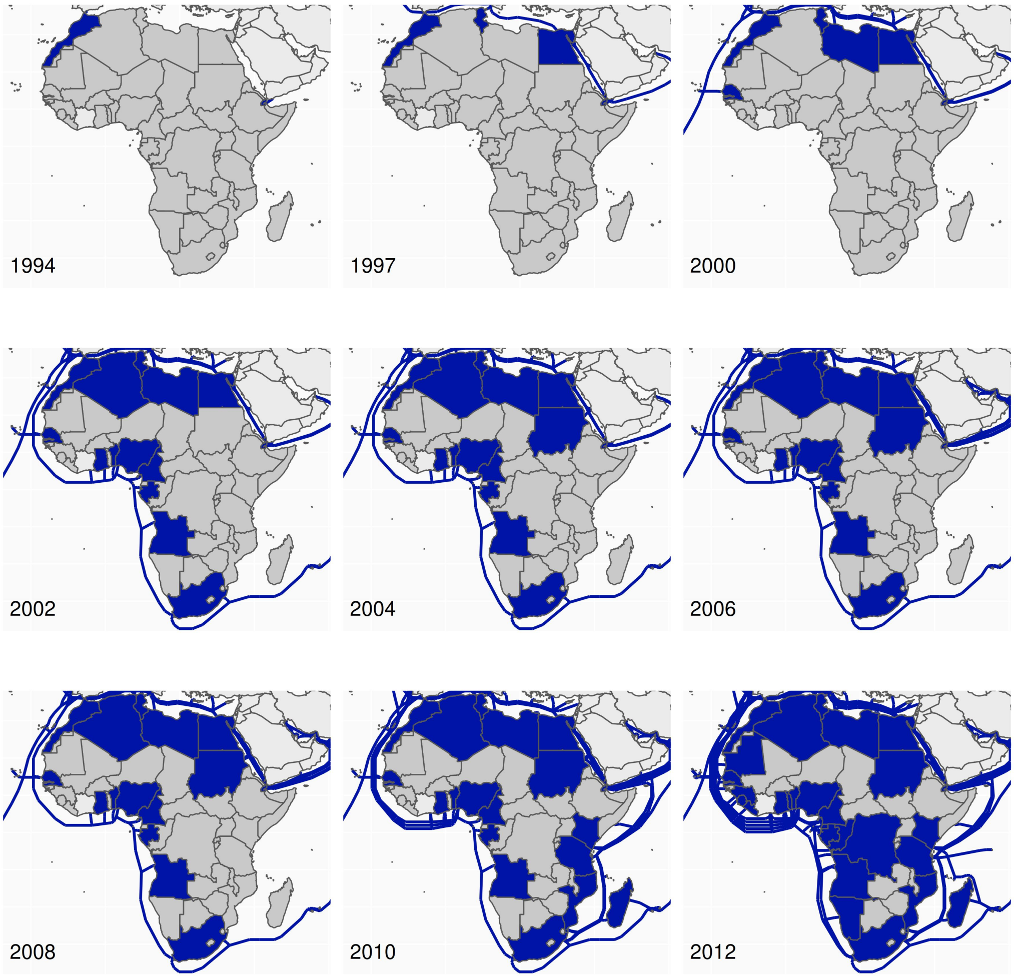
Notes: This figure shows the average unit cost per Mb/s capacity based on 2008 thousands US$. The source is Detecon (2013). The histogram in yellow refers to satellite transmission. That in gray refers to submarine fiber-optic transmission.

Figure 3: The Status of Discount Window Facilities in Africa



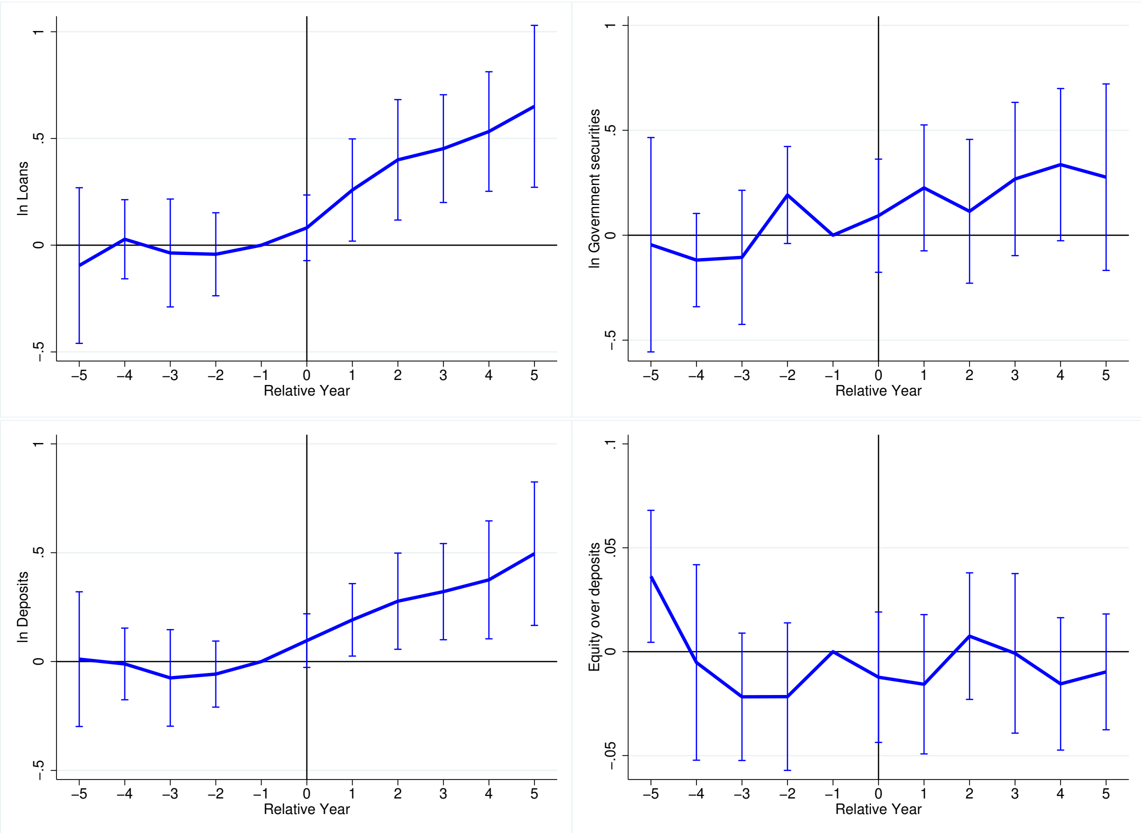
Notes: This figure shows the status of the discount window facilities in Africa. Data are from local sources, the International Monetary Fund and World Bank documentation. On the y axis is reported the number of Central banks in Africa. On the x axis there is a classification of the central banks based on how active they are in providing discount window facilities: active, not active and unstable. Active implies that access to the discount window is available given the policy rate. Not Active means that there is no facility dedicated to loans to commercial banks. Unstable indicates that the facility does not necessarily provide funds given a policy rate, see Choudhary and Limodio (2017) for more on this.

Figure 4: The Arrival of High-Speed Internet in Africa



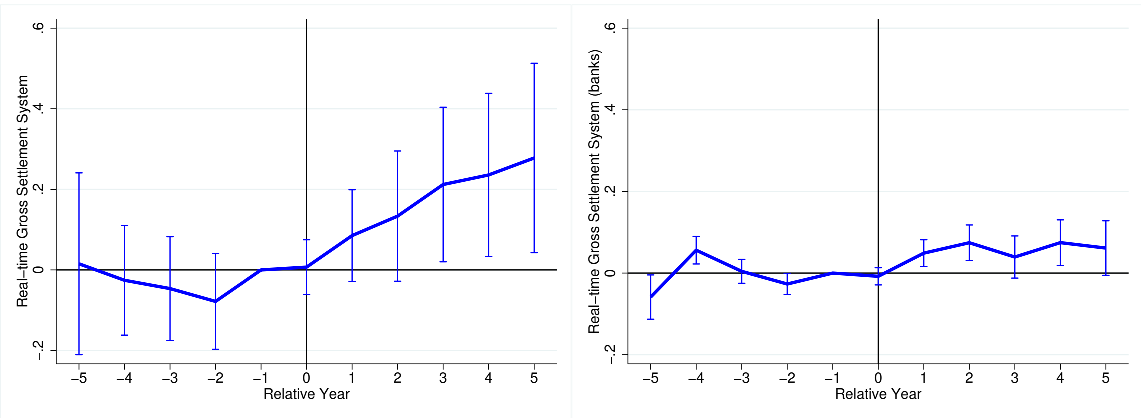
Notes: This figure shows the gradual arrival of submarine fiber-optic cables in Africa. In our analysis, we focus on first connections. Data are from Telegeography maps and they are available online. The top panel refers to cables ready to serve during the 90's, from 1994, when the first cables connected Morocco and Djibouti, to 2000. The panel in the middle shows the progressive connection to high-speed internet at the beginning of the XXI century. On the left, there are cables ready to serve in 2002. In the center, those ready to serve in 2004. Finally, on the right, cables ready to serve in 2006. The bottom panel refers to the last stage of first connections. On the left, there are cables ready to serve in 2008. In the middle, cables ready to serve in 2010. Finally, on the right, cables ready to serve in 2012, that is also the last year in which countries in our sample receive connection. The group of treated African coastal countries is reported in dark blue.

Figure 5: An Event Study on High-Speed Internet and Banking



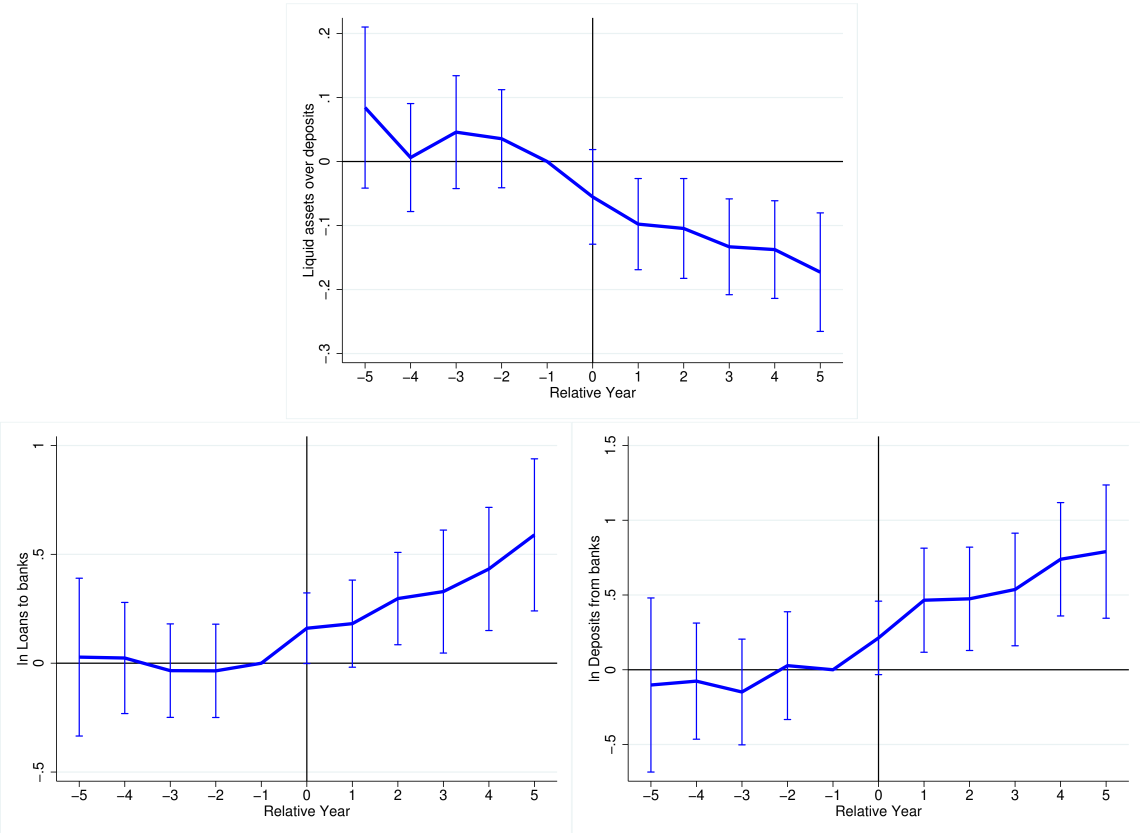
Notes: This figure shows the event study for the period 5 years before–5 years after the arrival of the first submarine cable in country c. The y axis reports coefficients for the dependent variables: ln Loans (the natural logarithm of net loans (in million of US dollars)); ln Government securities (the natural logarithm of government securities (in million of US dollars)); ln Deposits (the natural logarithm of deposits and short-term funding (in million of US dollars)); and Equity over DST (the share of total equity over deposits and short-term funding). The x axis refers to the relative time from the arrival of the cable. The blue (solid) line connects point estimates relative to the base year (-1). 95% confidence intervals are also reported, and standard errors are clustered at country level.

Figure 6: An Event Study on High-Speed Internet and RTGS Adoption



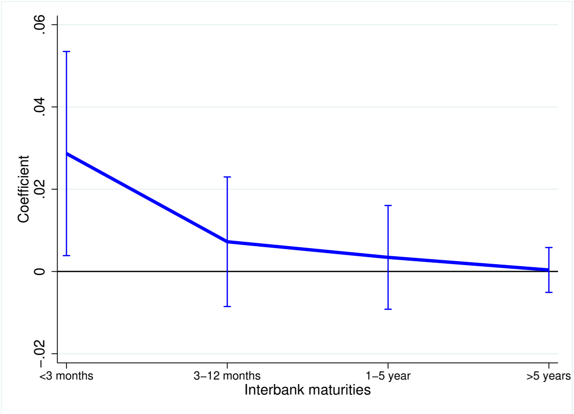
Notes: This figure shows the event study for the period 5 years before–5 years after the arrival of the first submarine cable in country c. The y axis reports coefficients for the dependent variables: RTGS (Real-time Gross Settlement System); and RTGS bank (RTGS adoption for the single bank). The x axis refers to the relative time from the arrival of the cable. The blue (solid) line connects point estimates relative to the base year (-1). 95% confidence intervals are also reported, and standard errors are clustered at country (bank) level.

Figure 7: An Event Study on High-Speed Internet and Interbank Markets



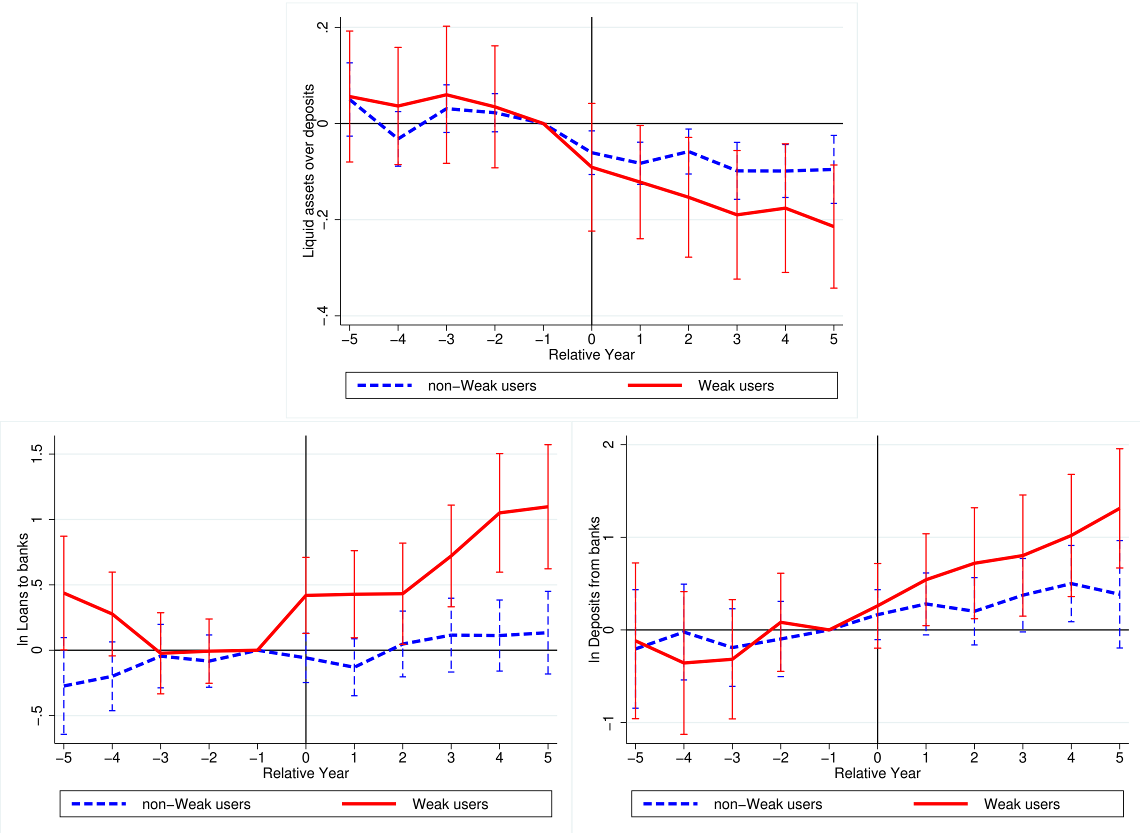
Notes: This figure shows the event study for the period 5 years before–5 years after the arrival of the first submarine cable in country c. The y axis reports coefficients for the dependent variables: Liquid Assets over DST (the share of liquid assets over deposits and short-term funding); ln Loans to Banks (the natural logarithm of loans to banks (in million of US dollars)); and ln Deposits from Banks (the natural logarithm of loans to banks (in million of US dollars)). The x axis refers to the relative time from the arrival of the cable. The blue (solid) line connects point estimates relative to the base year (-1). 95% confidence intervals are also reported, and standard errors are clustered at country level.

Figure 8: High-Speed Internet and Interbank Maturity



Notes: This plot reports the coefficients of the regression of interbank maturities on $Submarine_{ct}$. The blue (solid) line connects point estimates. 95% confidence intervals are also reported, and standard errors are clustered at country level.

Figure 9: An Event Study on High-Speed Internet, Interbank Markets and Weak Users



Notes: This figure shows the event study for the period 5 years before–5 years after the arrival of the first submarine cable in country c. The y axis reports coefficients for the dependent variables: Liquid Assets over DST (the share of liquid assets over deposits and short-term funding); ln Loans to Banks (the natural logarithm of loans to banks (in million of US dollars)); and ln Deposits from Banks (the natural logarithm of loans to banks (in million of US dollars)). The x axis refers to the relative time from the arrival of the cable. The blue (dashed) line connects point estimates relative to the base year (-1) for non-Weak users. The red (solid) line connects point estimates relative to the base year (-1) for the Weak users. 95% confidence intervals are also reported, and standard errors are clustered at bank level.

Tables

Table 1: Summary Statistics

	(1)	(2)	(3)	(4)	(5)	(6)
Variables	Obs.	Mean	Std. Dev.	Median	5th P.tile	95th P.tile
	Panel A - Dependent variables: banks (1997-2018)					
Loans	6652	5.00	2.08	4.93	1.78	8.35
Government	3680	3.89	2.33	3.90	0.24	7.57
Deposits	6702	5.61	1.99	5.54	2.57	8.91
Equity/dep. & ST	6552	0.20	0.23	0.14	0.05	0.54
RTGS bank	10569	0.53	0.50	1	0	1
Loans to Banks	6344	3.77	2.05	3.84	0.50	7.19
Deposits from Banks	4980	2.91	2.47	3.06	-1.40	6.72
Liquid Assets/dep. & ST	6674	0.46	0.46	0.37	0.09	0.99
	Panel B - Dependent variables: firms (2002-2018)					
Access to Finance	25389	0.64	0.48	1	0	1
Loans from Banks	27863	0.19	0.39	0	0	1
Loans Maturity	2694	3.25	1.07	3.58	1.10	4.68
Sales	23828	-1.73	2.60	-2.00	-5.57	3.07
Workforce	26981	3.11	1.28	2.83	1.61	5.66
	Panel C - Independent variables					
Sample of Banks:	6776					
Submarine	6776	0.72	0.45	1	0	1
Submarine × Weak user	4170	0.25	0.43	0	0	1
Sample of Firms:	32761					
Submarine	32761	0.84	0.36	1	0	1
Submarine × Weak Interbank	32495	0.26	0.44	0	0	1

Notes: This table reports the summary statistics. Panel A refers to dependent variables related to banks indicators. Panel B refers to dependent variables related to firms indicators. Panel C focuses on the main predictors. Column 1 refers to the number of observations. Columns 2 and 3 refer to the mean and the standard deviation. Columns 4, 5, and 6, refer to the 50th, 5th, and 95th percentiles respectively.

Table 2: High-Speed Internet and Banking

	(1)	(2)	(3)	(4)
Variables	Loans	Government	Deposits	Equity
	ln	ln	ln	over DST
$Submarine_{ct}$	0.363***	0.211	0.263***	-0.007
	(0.128)	(0.151)	(0.090)	(0.011)
Bank FE	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes
Obs.	6617	3630	6671	6518
Adj. R sq.	0.863	0.789	0.881	0.575
MDV	5.009	3.917	5.623	0.196
SDDV	2.068	2.321	1.985	0.225

Notes: This table reports estimates from the staggered diff-in-diff design presented in equation (2). The dependent variables are as follows: Loans (natural logarithm of net loans (in million of US dollars)); Government (natural logarithm of government securities (in million of US dollars)); Deposits (natural logarithm of deposits and short-term funding (in million of US dollars)); and Equity (share of total equity over deposits and short-term funding). The main predictor is $Submarine_{ct}$, a binary variable for the arrival of the first fiber-optic submarine cable in the country. This dummy takes value zero before the arrival of the cable and 1 from the time of the arrival on. Obs. refers to the number of observations; Adj.R^2 is the adjusted R^2; MDV refers to the mean of the dependent variable; and SDDV refers to the standard deviation of the dependent variable. Fixed effects are at the bank and year level. Standard errors in parentheses, clustered at country level. ***, ** and * indicate significance at the 1%, 5% and 10% level, respectively.

Table 3: High-Speed Internet and Banking - Demand and Supply

Variables	(1) Loans ln	(2) Government ln	(3) Deposits ln	(4) Equity over DST
$Submarine_{ct}$	0.184**	-0.053	0.118	-0.014
	(0.091)	(0.128)	(0.072)	(0.014)
$Submarine_{gt}$	0.329***	0.424***	0.267***	0.015
	(0.090)	(0.128)	(0.084)	(0.014)
Bank FE	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes
Obs.	6481	3600	6535	6397
Adj. R sq.	0.865	0.789	0.883	0.577
MDV	5.039	3.930	5.653	0.194
SDDV	2.061	2.319	1.978	0.222

Notes: This table reports estimates from the staggered diff-in-diff design presented in equation (2). The dependent variables are as follows: Loans (natural logarithm of net loans (in million of US dollars)); Government (natural logarithm of government securities (in million of US dollars)); Deposits (natural logarithm of deposits and short-term funding (in million of US dollars)); and Equity (share of total equity over deposits and short-term funding). The main predictors are: $Submarine_{ct}$, a binary variable for the arrival of the first fiber-optic submarine cable in the country, and $Submarine_{gt}$, a binary variable for the arrival of the first fiber-optic submarine cable in the country of the headquarter. These dummies take value zero before the arrival of the cable and 1 from the time of the arrival on. Obs. refers to the number of observations; Adj.R^2 is the adjusted R^2; MDV refers to the mean of the dependent variable; and SDDV refers to the standard deviation of the dependent variable. Fixed effects are at the bank and year level. Standard errors in parentheses, clustered at bank level. ***, ** and * indicate significance at the 1%, 5% and 10% level, respectively.

Table 4: High-Speed Internet and Banking - Supply

Variables	(1) Loans ln	(2) Government ln	(3) Deposits ln	(4) Equity over DST
$Submarine_{gt}$	0.289**	0.251	0.238**	0.010
	(0.113)	(0.161)	(0.113)	(0.017)
Bank FE	Yes	Yes	Yes	Yes
Country-Year FE	Yes	Yes	Yes	Yes
Obs.	6423	3513	6480	6345
Adj. R sq.	0.890	0.824	0.900	0.581
MDV	5.052	3.969	5.662	0.195
SDDV	2.060	2.308	1.976	0.222

Notes: This table reports estimates from the staggered diff-in-diff design presented in equation (2). The dependent variables are as follows: Loans (natural logarithm of net loans (in million of US dollars)); Government (natural logarithm of government securities (in million of US dollars)); Deposits (natural logarithm of deposits and short-term funding (in million of US dollars)); and Equity (share of total equity over deposits and short-term funding). The main predictor is $Submarine_{gt}$, a binary variable for the arrival of the first fiber-optic submarine cable in the country of the headquarter. This dummy takes value zero before the arrival of the cable and 1 from the time of the arrival on. Obs. refers to the number of observations; Adj.R^2 is the adjusted R^2; MDV refers to the mean of the dependent variable; and SDDV refers to the standard deviation of the dependent variable. Fixed effects are at the bank and country by year level. Standard errors in parentheses, clustered at bank level. ***, ** and * indicate significance at the 1%, 5% and 10% level, respectively.

Table 5: High-Speed Internet and RTGS Adoption

Variables	(1) RTGS Country-Level	(2) RTGS Bank-Level
$Submarine_{ct}$	0.160*	0.055***
	(0.084)	(0.020)
country FE	Yes	No
Bank FE	No	Yes
Year FE	Yes	Yes
Obs.	814	10568
Adj. R sq.	0.670	0.701
MDV	0.456	0.527
SDDV	0.498	0.499

Notes: This table reports estimates from the staggered diff-in-diff design presented in equation (2). The dependent variables are as follows: RTGS (a dummy variable equal to one if the country as adopted the RTGS); RTGS bank (a dummy variable that is equal to one if the bank participates to the RTGS). The main predictor is Submarine_ct, a binary variable for the arrival of the first fiber-optic submarine cable in the country. This dummy takes value zero before the arrival of the cable and 1 from the time of the arrival on. Obs. refers to the number of observations; Adj.R^2 is the adjusted R^2; MDV refers to the mean of the dependent variable; and SDDV refers to the standard deviation of the dependent variable. Fixed effects are at the country (or bank) and year level. Standard errors in parentheses, clustered at country (or bank) level. ***, ** and * indicate significance at the 1%, 5% and 10% level, respectively.

Table 6: High-Speed Internet and Interbank Markets

Variables	(1) Liquid Assets over DST	(2) Loans to Banks ln	(3) Deposits from Banks ln
$Submarine_{ct}$	-0.139***	0.264**	0.540***
	(0.030)	(0.117)	(0.192)
Bank FE	Yes	Yes	Yes
Year FE	Yes	Yes	Yes
Obs.	6641	6308	4940
Adj. R sq.	0.429	0.767	0.669
MDV	0.463	3.783	2.928
SDDV	0.441	2.053	2.466

Notes: This table reports estimates from the staggered diff-in-diff design presented in equation (2). The dependent variables are as follows: Liquid Assets over DST (share of liquid assets over deposits and short-term funding); Loans to Banks (natural logarithm of loans to banks (in million of US dollars)); and Deposits from Banks (natural logarithm of loans to banks (in million of US dollars)). The main predictor is $Submarine_{ct}$, a binary variable for the arrival of the first fiber-optic submarine cable in the country. This dummy takes value zero before the arrival of the cable and 1 from the time of the arrival on. Obs. refers to the number of observations; Adj.R^2 is the adjusted R^2; MDV refers to the mean of the dependent variable; and SDDV refers to the standard deviation of the dependent variable. Fixed effects are at the bank and year level. Standard errors in parentheses, clustered at country level. ***, ** and * indicate significance at the 1%, 5% and 10% level, respectively.

Table 7: High-Speed Internet, Interbank Markets and Weak Users

Variables	(1) Liquid Assets over DST	(2) Loans to Banks ln	(3) Deposits from Banks ln	(4) Loans ln
$Submarine_{ct}$	-0.077***	0.066	0.265	0.210*
	(0.024)	(0.120)	(0.170)	(0.113)
$Submarine_{ct}$	-0.105**	0.397**	0.586***	0.274*
$\times$ $Weak\ user_{ic}$	(0.043)	(0.162)	(0.224)	(0.140)
Bank FE	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes
Obs.	4121	3915	2932	4109
Adj. R sq.	0.474	0.735	0.634	0.841
MDV	0.466	3.500	2.547	4.766
SDDV	0.385	1.943	2.326	1.914

Notes: This table reports estimates from the staggered diff-in-diff design presented in equation (3). The dependent variables are as follows: Liquid Assets over DST (share of liquid assets over deposits and short-term funding); Loans to Banks (natural logarithm of loans to banks (in million of US dollars)); and Deposits from Banks (natural logarithm of loans to banks (in million of US dollars)). The main predictors are: $Submarine_{ct}$, a binary variable for the arrival of the first fiber-optic submarine cable in the country; and $Submarine_{ct} \times Weak\ user_{ic}$, the interaction between $Submarine_{ct}$ and a dummy that takes value one if the bank was below the median of interbank loans and deposits (with respect to its own country and the period before the arrival of high-speed internet). Obs. refers to the number of observations; Adj.R^2 is the adjusted R^2; MDV refers to the mean of the dependent variable; and SDDV refers to the standard deviation of the dependent variable. Fixed effects are at the bank and year level. Standard errors in parentheses, clustered at bank level. ***, ** and * indicate significance at the 1%, 5% and 10% level, respectively.

Table 8: High-Speed Internet, Interbank Markets and Weak Users - Country-Year FEs

Variables	(1) Liquid Assets over DST	(2) Loans to Banks ln	(3) Deposits from Banks ln	(4) Loans ln
$Submarine_{ct}$	-0.111***	0.449***	0.594**	0.364***
$\times Weak\ user_{ic}$	(0.039)	(0.158)	(0.239)	(0.129)
Bank FE	Yes	Yes	Yes	Yes
Country-Year FE	Yes	Yes	Yes	Yes
Obs.	4052	3824	2826	4037
Adj. R sq.	0.504	0.743	0.659	0.871
MDV	0.460	3.494	2.553	4.786
SDDV	0.377	1.943	2.317	1.915

Notes: This table reports estimates from the staggered diff-in-diff design presented in equation (3). The dependent variables are as follows: Liquid Assets over DST (share of liquid assets over deposits and short-term funding); Loans to Banks (natural logarithm of loans to banks (in million of US dollars)); and Deposits from Banks (natural logarithm of loans to banks (in million of US dollars)). The main predictors is $Submarine_{ct} \times Weak\ user_{ic}$, the interaction between $Submarine_{ct}$, a binary variable for the arrival of the first fiber-optic submarine cable in the country, and $Weak\ user_{ic}$, a dummy that takes value one if the bank was below the median of interbank loans and deposits (with respect to its own country and the period before the arrival of high-speed internet). Obs. refers to the number of observations; Adj.R^2 is the adjusted R^2; MDV refers to the mean of the dependent variable; and SDDV refers to the standard deviation of the dependent variable. Fixed effects are at the bank and country by year level. Standard errors in parentheses, clustered at bank level. ***, ** and * indicate significance at the 1%, 5% and 10% level, respectively.

Table 9: High-Speed Internet and Firms

Variables	(1) Access to finance dummy	(2) Loans from banks dummy	(3) Loans maturity ln	(4) Sales ln	(5) Workforce ln
$Submarine_{ct}$	0.143***	0.038	1.005**	2.153	-0.099
	(0.050)	(0.046)	(0.393)	(1.772)	(0.139)
Country FE	Yes	Yes	No	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes
Obs.	25389	27863	2694	23828	26706
Adj. R sq.	0.0945	0.109	0.201	0.253	0.0920
MDV	0.638	0.191	3.253	-1.734	3.068
SDDV	0.481	0.393	1.071	2.600	1.197

Notes: This table reports estimates from the staggered diff-in-diff design presented in equation (2). The dependent variables are: Access to finance (dummy variable, where 1 indicates easy access to finance); Loans from banks (dummy variable, where 1 indicates at least one loan from a commercial bank); Loans maturity (natural logarithm of the term, in months, of loans from banks); Sales (natural logarithm of total annual sales); and Workforce (natural logarithm of the number of permanent and temporary full-time employees). The main predictor is $Submarine_{ct}$, a binary variable for the arrival of the first fiber-optic submarine cable in the country. This dummy takes value zero before the arrival of the cable and 1 from the time of the arrival on. Obs. refers to the number of observations; Adj.R^2 is the adjusted R^2; MDV refers to the mean of the dependent variable; and SDDV refers to the standard deviation of the dependent variable. Fixed effects are at the country and year level. Standard errors in parentheses, clustered at country level. ***, ** and * indicate significance at the 1%, 5% and 10% level, respectively.

Table 10: High-Speed Internet, Firms and Weak Interbank Markets

Variables	(1) Access to finance dummy	(2) Loans from banks dummy	(3) Loans maturity ln	(4) Sales ln	(5) Workforce ln
$Submarine_{ct}$	0.059	0.007	0.456**	0.641	-0.236
	(0.072)	(0.060)	(0.192)	(1.276)	(0.172)
$Submarine_{ct}$	0.165**	0.064	0.743*	3.192**	0.275*
$\times$ $Weak\ user_c$	(0.075)	(0.051)	(0.425)	(1.298)	(0.143)
Country FE	Yes	Yes	No	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes
Obs.	25126	27597	2694	23590	26454
Adj. R sq.	0.094	0.110	0.257	0.269	0.093
MDV	0.636	0.191	3.253	-1.747	3.069
SDDV	0.481	0.393	1.071	2.605	1.199

Notes: This table reports estimates from the staggered diff-in-diff design presented in equation (4). The dependent variables are: Access to finance (dummy variable, where 1 indicates easy access to finance); Loans from banks (dummy variable, where 1 indicates at least one loan from a commercial bank); Loans maturity (natural logarithm of the term, in months, of loans from banks); Sales (natural logarithm of total annual sales); and Workforce (natural logarithm of the number of permanent and temporary full-time employees). The main predictors are: $Submarine_{ct}$, a binary variable for the arrival of the first fiber-optic submarine cable in the country; and $Submarine_{ct} \times Weak\ user_c$, the interaction between $Submarine_{ct}$ and a dummy that takes value one if the country was below the median of interbank loans and deposits (in the period before the arrival of high-speed internet). Obs. refers to the number of observations; Adj.R^2 is the adjusted R^2; MDV refers to the mean of the dependent variable; and SDDV refers to the standard deviation of the dependent variable. Fixed effects are at the country and year level. Standard errors in parentheses, clustered at country level. ***, ** and * indicate significance at the 1%, 5% and 10% level, respectively.

Chapter 3

Mayoral Wage and Public Procurement *

Angelo D'Andrea **

Abstract

In this paper I run a regression discontinuity design analysis to document the causal effect of mayoral wages on procurement outcomes in a large database of public procurement contracts in Italy. To identify the effect of mayoral wages, I use a unique characteristic of the Italian legislation, namely that mayoral remuneration varies at predetermined population thresholds. The main results are as follows. First, I show that higher mayoral wages (i.e., the treatment group) are not related to differences in aggregate measures of procurement, such as the number of tenders, the total procurement expenditure, and the mean value of the contract in the municipality. Second, I show that certain specific procurement outcomes are affected by higher mayoral wages: the number of admitted offers increases, as do the final rebates on the reserve price; the probability that the same firm is awarded a contract repeatedly decreases; and for a limited sample, cost overruns go down. Finally, I provide evidence that re-election incentives play a role in explaining the effect of mayoral wages. This paper is the first to document a direct relationship between mayoral salaries and public procurement.

JEL: H0, H5, H57, H7, H76

Keywords: Public procurement, Local governments, Public wages

* This paper has benefitted from discussions with and suggestions from Gianmarco Daniele, Francesco Decarolis, Tommaso Giommoni, Pamela Giustinelli, Paolo Pinotti, Edoardo Spina, Guido Tabellini, and all the participants in the Political Economy RG at Bocconi University. I also thank the Baffi-CAREFIN center at Bocconi University for supporting this research. All errors are my own.

** angelo.dandrea@unibocconi.it, Bocconi University, Department of Economics, Via Roentgen 1, 20136 Milan (Italy).

1 Introduction

In this paper I use a regression discontinuity (RD) design to explore the effect of mayoral wages on the provision of public goods. I analyze a large database of Italian public procurement contracts administered by local municipalities to estimate the local effect of mayoral wages on procurement outcomes. Procurement outcomes include aggregate measures on the amount of opened tenders; certain ex-ante procurement attributes (the number of bidders, the final rebate and the characteristics of the winning contracts); the identity of the winners; and, for a limited sample, time and cost overruns.

The study has three motivations, the first of which is the central role played by public procurement in the provision of public goods at the local level. In 2017, the aggregate amount of procurement contracts in Italy was about 140 billion euros, 7.2% of GDP, and municipalities were among the top procurers, with a relative share of 8.3%.[1] Second, public procurement is highly relevant for local municipalities, and the legislation grants a great deal of discretion to local mayors and their executive committees. Finally, while remuneration for political activity and its effect on public governance has attracted considerable attention in the literature, there is no study, to the best of my knowledge, that directly relates the salary of local politicians to procurement performance.[2]

The area of public procurement represents a good laboratory in which to investigate my research question. The evaluation of procurement reveals virtues and flaws of local institutions, highlighting their ability to raise public funds and to provide public goods to the community, as well as detecting misconduct. In Italy, almost 51% of administrations investigated for potential malfeasance or corruption are municipalities.[3] Moreover, procurement design and performance are used to define red flags and benchmarks for the evaluation of local government achievements.

The literature has emphasized three channels through which high salaries could lead local politicians to perform differently. First, there may be a selection effect. Higher wages reduce the opportunity cost of being elected and encourage more highly educated people to run for election. At the same time, high wages increase the value of being in office and may encourage all citizens, educated or not, to run for office. These two mechanisms work in opposite directions, and the final effect of selection on local

[1] Statistics are from the 2017 Annual Report, table 7.15, presented by the National Anti-corruption Authority (ANAC) to the Italian Senate in June 2018.

[2] The relationship between politicians' remuneration and their performance is an issue of perennial interest. In Italy, even in a situation of generalized mistrust towards the political spectrum, there are legislative proposals on the appropriateness of increasing the salaries of Members of the Parliament to incentivize the "best" citizens to run for public offices.

[3] ANAC, 2017 Annual Report.

government performance depends on which of the two mechanisms dominates.

Second, there may be a re-election incentive effect. Higher wages increase the monetary value of being in office, thereby increasing the value of re-election. As a consequence, mayors may favor the interests of special groups when those groups are considered key to their re-election. In the case where the pivotal group corresponds to the mass of citizens, local performance would be aligned with the community's interests and procurement procedures would be set properly. In the other case, when certain groups are pivotal, politicians may direct public investments toward their special interests, potentially generating a decrease in social surplus.

Third, and sometimes overlooked, there may be a motivational incentive (”morale”, to use the terminology of Akerlof (1982)), such that higher wages may increase feelings of job satisfaction, leading politicians to behave better and acting as a deterrent to corruption. This channel, that takes into account the increase in the opportunity cost of illegal practices when wages are higher, is of particular interest in the context of public procurement.

In this paper I contribute to the existing literature in two ways: first, I investigate the effects of higher mayoral remuneration on procurement outcomes; second, I document the mechanisms that are behind the main findings. In order to implement the analysis, I exploit one specific feature of the Italian legislation, namely that the level of mayoral wages depends on the population of the municipality. In particular, wages are determined based on nine thresholds that fully partition the distribution of inhabitants per municipality.

My identification strategy relies on the assumption that mayoral wages cannot be manipulated around the discontinuity threshold. For identification purposes, I use the threshold of 5,000 inhabitants, above which level mayors earn a higher wage. Since the wages are associated with the number of inhabitants in the municipality, and population data are from historical reports and are gathered by an independent third party,[4] mayors are unable to misreport population data in order to receive a higher wage. Nevertheless, there is the possibility that they might use domestic migration flows to inflate their municipality's population. For that reason, I check the continuity assumption around the threshold using graphical and statistical tools discussed by McCrary (2008) and Lee (2008).

The universe of contracts that I use in this paper is provided by the Italian Anti-corruption National Authority (A.N.A.C.)[5] and takes into account all procurement contracts (public works, provisions and

[4] Compensation for local councillors in Italy is regulated by art. 156, comma 2, TUEL, which establishes that compensation based on population criteria must be based on the population of two years before the current year, using data from ISTAT.

[5] Autorità Nazionale Anticorruzione.

services) with a value above $40,000$ €. A further selection is made by excluding special regions and contracts that are not issued through open auctions or negotiated procedures.[6] Finally, I select the sample using the procedure suggested by Imbens and Kalyanaraman (2012). Specifically, I focus on the contracts administered by those municipalities within the interval around the discontinuity threshold, selected via the optimal bandwidth method. In this quasi-experimental setup, municipalities that lie in a small interval around the threshold are likely to be identical in terms of observable and unobservable characteristics, and higher wages should therefore be quasi-randomly assigned across treated and control mayors.

The main results from my RD analysis are as follows. First, the data on aggregate measures of public procurement (number of opened tenders, total procurement expenditure, and the mean value of the contracts) do not show any discontinuity around the threshold. Second, higher mayoral wages influence the specific characteristics of tenders.

There is a positive and significant effect of mayoral wages on the number of admitted offers per tender and the final rebate on the reserve price; a significant decrease in the probability that a firm is awarded multiple contracts; and, for a restricted sample of contracts, a decrease in cost overruns. While these results may be driven by different combinations of the selection, re-election, and/or morale mechanisms mentioned above, I further document that re-election incentives are particularly influential. To the best of my knowledge, the present study is the first to identify the causal effect of mayoral wages on procurement outcomes and to suggest the mechanism behind this effect.

This paper is related to different strands of the literature in economics and political economy. Primarily, it relates to the literature studying the link between politicians' remuneration and their performance. According to the efficiency wage theory, workers' productivity increases in proportion to the wage they are paid. The parallel for politicians exists, but predictions about politicians' behavior are highly heterogeneous, given the unique interests that motivate political activity. On the one hand, some studies highlight a positive relationship between politicians' wages and performance. This may be due to a better selection of candidates, since higher remuneration reduces the opportunity costs of serving in a political position and induces more educated citizens to run for election. It can also be due to higher incentives for re-election, since higher remuneration increases the monetary value of being re-elected

[6] Other layers of selection can be based on the reserve price. For example, Baldi et al. (2016) suggest that very complicated projects are usually treated in a special way by local bureaucracies. By using reserve price, I can exclude projects whose value is extraordinarily high. Another possibility is to use the EU threshold for public works. For my sample period, the EU threshold is set at $5,186,000$ €, or at $1,000,000$ € if the tender includes multiple lots and the value of the lot represents more than 20% of the total value. I have repeated my analysis using these criteria, and the results remain unchanged.

and could lead incumbents to improve their performance to increase their probability of re-election. Finally, the positive relationship may obtain because higher salaries improve the morale of politicians and reduce the probability of (active) misconduct. Various models in the political economy realm predict a positive relationship between political return and average government quality (see Besley (2004) and Caselli and Morelli (2004)). On the other hand, some studies find a negative relationship between politicians' wages and performance. Higher compensation might have pernicious effects by encouraging low-quality candidates to run for election. Messner and Polborn (2004) and Mattozzi and Merlo (2008) support this hypothesis. Moreover, higher wages might induce politicians to favor special groups in order to achieve a consensus for re-election. Depending on which group is favored and how, the effects on social welfare can be negative.

The question of whether and how politicians' wages affect their performance remains an empirical one. This paper contributes to the literature on the efficiency wage theory by providing empirical evidence on the specific effect of mayoral wages on public procurement. Efficiency can be identified in those aspects of the procurement contracts that are related to the level of competition of the tender, the identity of the winner, and time and cost overruns.

Second, this paper is related to the literature that looks at the connection between public wages and misconduct in local bureaucracies. In a seminal paper, Becker and Stigler (1974) highlight a negative relationship between wages and corruption at the level of intermediate auditing intensity. Two subsequent studies, Di Tella and Schargrodsky (2003) and Bandiera et al. (2009), empirically identify corruption in local public procurement. The former uses a one-shot policy for a single event study in Buenos Aires during the 1990s and shows how higher wages for bureaucrats reduce public corruption. The latter makes the crucial distinction between active and passive waste and shows that most of the inefficiencies in public procurement characterizing their sample come from the low competence of bureaucrats rather than from corrupt practices.

The paper's contribution to this literature is two-fold. First, it shows how local political leaders make different use of procurement procedures depending upon their salary level. Above the threshold of mayoral wages, open auctions and the lowest price criterion become demonstrably more common. Secondly, the paper directly investigates the relationship between wages and corruption, making use of a novel source of data from Decarolis et al. (2020).

The paper is also related to the literature linking political activity to the use of discretion.[7] In my

[7] The empirical analysis contains panels that investigate changes in performance by mayors managing open auctions and negotiated procedures, with negotiated procedures further divided in those that are standard and non-standard. These typologies

empirical analysis, I make broad use of the heterogeneity coming from higher (or lower) discretion, to emphasize its impact on the different outcomes. For public procurement, higher discretion can be associated with manipulation of ex-ante competition and the establishment of prolonged relationships with a restricted number of firms, which can lead to a reduction in consumer surplus. Lalive and Schmutzler (2011) show empirical evidence on this occurrence. On the other hand, when contracts are incomplete, buyer discretion can have positive effects on public procurement outcomes through reputation and experience (see, for example, Manelli and Vincent (1995), Bajari and Tadelis (2001), Calzolari and Spagnolo (2009), and Coviello et al. (2018)).

My paper contributes to this debate by providing evidence on how discretion interacts with mayoral wages, using an empirical model whose main specifications are all augmented to distinguish between open auctions and negotiated procedures.

Regarding the empirical strategy, this article follows the strand of the literature that uses the exogeneity of the population threshold as a policy identifier. The most closely related paper is that of Coviello and Gagliarducci (2017), who study the impact of politicians' tenure in office on the outcomes of public procurement tenders. They use the introduction of a two-term limit on the mayoral office in Italy in March 1993 to set up a quasi-experiment that aims to identify the effect of tenure on procurement. In contrast, I concentrate on wages. I follow Ferraz and Finan (2009) and Gagliarducci and Nannicini (2013) in implementing my identification strategy based on exogenously determined wage differentials. While the latter focuses on aggregate measures of local government efficiency, the purpose of my paper is to explicitly analyze procurement contracts.

Finally, in regard to the features of the tender to include and their interpretation, this paper refers to a well-established literature in the field of the economics of public procurement (see Coviello and Gagliarducci (2017), Kang and Miller (2017), and Coviello et al. (2018)).

The rest of the paper is organized as follows. In section 2, I describe the institutional framework. In section 3, I present the data. Section 4 presents the identification strategy. Section 5 reports the empirical analysis and the main results, and section 6 describes my robustness checks. Finally, in section 7 I conclude.

are mostly distinguished by their different levels of discretion.

2 Institutional Background

The Italian framework for public procurement has several unique features. First, it is conducted at different levels. A small portion of the contracts are issued by national authorities, whereas the vast majority are administered at the regional, provincial, or municipal level. In this paper, I focus on municipalities.[8] Second, the legislation distinguishes three types of contracts: public works, provisions, and services, with the latter defined as a residual category. Most of the papers in the existing literature focus on public works, ignoring provisions and services. I exploit the breadth of my dataset to take into account this important distinction and to show the existence of heterogeneous wage effects on procurement outcomes.

The majority of the contracts are set by auctions.[9] The value of the project for each auction is estimated by internal technicians, sometimes supported by the expertise of a private figure or the assistance of the public authority in charge of setting prices at the national level, the Consip. Standardization of procurement contracts is critical, and it is the means by which the legislation attempts to limit discretion. In terms of public works, the office responsible for procurement evaluates the types and quantities of inputs needed to complete the project, and the final value is then computed using a menu of standardized costs provided by the national authority. I follow Decarolis (2014) in arguing that the public buyer running the auction is not in full control of its value. In terms of provisions and services, municipalities often make direct use of the Consip, in which case the margins for manoeuvring shrink substantially.

Notwithstanding these aspects, the law that regulates public procurement in Italy is extremely complex and rich in technicalities.[10] These technicalities lead to corresponding loopholes and generate mar-

[8] Among public buyers, municipalities are the top procurers.

[9] Also, negotiated procedures are generally implemented through a (restricted) auction.

[10] The complexity of the Italian law lies in two dimensions. On the one hand, the auction mechanism that is used is somewhat unconventional. The winner of the auction does not correspond to the one offering the higher rebate on the reserve price (the maximum amount that the buyer is willing to pay), but rather is determined by a mathematical algorithm. After a preliminary trimming of the top 10% of the collected rebates, the rebates that exceed the average by more than the average deviation (the ”anomaly threshold”) are also excluded. The winning rebate is the highest of the nonexcluded rebates (that are below the anomaly threshold) (see Coviello and Gagliarducci (2017) for further details). On the other hand, the number of exceptions to the basic cases is so great that the actual level of discretion in the hands of bureaucrats is overly high. Although legislators have made numerous attempts to make the procurement system more transparent (see, for example, the use of the trimming mean algorithm, the introduction of Consip and the introduction of the ”Centrali Uniche di Committenza” (CUCs) as central aggregators of procurement contracts, and the continuous modifications of the basic procurement law), there is still a lot of room for–and evidence of– manipulation.

gins for manipulation. It is these margins for manipulation of procurement contracts by people involved in the procurement process (directly or indirectly) that motivate my research question.

In this paper, I claim the role of the mayor to be pivotal, and I focus on the effects of the mayoral wage on procurement outcomes. Mayors are not directly involved in the design of the procurement contract, nor should they involved in its ongoing process. However, public procurement represents a core activity for municipalities, and in fact both anecdotal evidence and official documents show that mayors and their entourages actively participate in the different phases of the procurement process.[11] Mayors play a crucial role in the way that procurement contracts are offered and subsequently allocated to firms, and this is even more transparent in small municipalities, which constitute the unit of analysis of this paper. Importantly, the same arguments can be applied to the entire junta. Without loss of generality, the paper could focus on the local effects of higher executive committee wages on local public procurement.

A potential threat to the identification of the effect of mayoral wages on procurement outcomes is that the wages themselves might be endogenous. In order to overcome this major problem, I exploit a quasi-experimental framework along the lines of that of Gagliarducci and Nannicini (2013).

In Italy, since 1960 mayoral wages have depended on the size of the population resident in the municipality. The wage is set exogenously by the central government in accordance with a nine-thresholds scheme. The same rule applies to all the members of the municipal government, including the executive officers appointed by the mayor, although in their case the salary is significantly lower. A key characteristic is that wages do not increase monotonically with population, but sharply. Thus, since population size cannot be (in principle) manipulated, the Italian institutional setting provides a natural source of local exogenous variation for mayoral salaries.

Mayoral compensation is not the only policy that varies with the population size. In table 1, I present a summary of different policies that use population brackets to discriminate among Italian municipalities. Besides mayoral remuneration, population size also determines the compensation of the members of the executive committee and of the councilors; the size of the council; the size of the executive committee; the electoral rule; and whether or not a minimum number of signatures is required to run for office. Importantly, all the other policies that discriminate based on population (e.g., transfers from the central government) change proportionally according to the number of inhabitants.

[11] Public procurement contracts are used to evaluate the performance of local governments. The ANAC establishes benchmarks and red flags to monitor local political leaders. Moreover, among the public administrations that have been under investigation, 51% are municipalities, and manipulation of tenders is a recurring theme (ANAC, 2017 Annual Report).

In this paper, I make extensive use of one of the nine thresholds, specifically the threshold of 5,000 inhabitants. Although the identified local effect can be difficult to generalize to different population levels, several reasons motivate this choice. First, the 5,000-inhabitant threshold has unique characteristics that allow me to separately identify the effect of wages on procurement. Indeed, no other changes occur at this threshold apart from the variation in mayoral wages (and the related executive committee remuneration) (see table 1). Second, cities with fewer than 10,000 inhabitants account for about 85% of Italian municipalities and 30% of the country's population. Third, the sample size is suitable for the empirical analysis.[12] Interestingly, these municipalities are also ideal in the context of the Becker-Stigler theory, where the existence of a negative relationship between wages and corruption requires an intermediate level of audit activity. In addition, the monitoring of political activity in small municipalities is relatively easier.

The analysis is deliberately implemented over the period 2013–2017 to rule out the possibility of overlapping effects. In 1999 the Italian government approved a system of fiscal rules called the domestic stability pact (DSP) that applied to municipalities and limited their discretion to run public deficits.[13] From 2001 to 2013, the DSP used the same 5,000-inhabitant threshold to discriminate between municipalities, exempting those below the threshold from its application. In 2013 the fiscal regulations changed, and the threshold was moved to 1,000 inhabitants. I exploit this 2013 legislative change to properly identify the effect of mayoral remuneration on local procurement outcomes.[14]

As of 2017, the real gross mayoral wages ranges from $1,291$ € per month for municipalities with

---

[12] Another possibility would be to use the threshold of 50,000 inhabitants. However, the sample size in that case would be tiny, given the small number of cities in Italy.

[13] See also Grembi et al. (2016) and Coviello et al. (2017).

[14] It can be claimed that the inclusion of municipalities with fewer than 5,000 inhabitants in the sphere of application of the DSP following 2013 does constitute an effect per se. This interpretation is meaningful and deserves further comment.
First, if being under the scrutiny of the public authority for the first time can foster improved performance, we would expect to see mayors in small municipalities (those below the cutoff) performing better. This would act as an attenuation bias for my findings, and the overall effect should be the smoothness of procurement outcomes around the threshold. This interpretation strengthens my discontinuity outcomes reported in section 5.

On the other hand, if being exempted from the DSP application for a long time can generate persistent effects, one could argue that differences in procurement outcomes are due to the persistence of bad practices among municipalities below the cutoff, rather than to the effect of mayoral wages. Interestingly, I find that mayors below and above the threshold do not differ in the amount of money they spend on public procurement, nor they differ in the number of tenders they issue (aspects that are directly related to the DSP). Indeed, the only differences between those mayors are in the special characteristics of the tenders they issue, an aspect that is more difficult to reconcile with the objectives and features of the DSP rules. Section 5 and appendix B.2 help to clarify these issues.

fewer than 1,000 inhabitants up to $7,798$ € for those with more than 1,000,000 residents. At the 5,000-inhabitant threshold, the gross mayoral salary increases by 28.6%, from $2,169$ € to $2,789$ €. These wages are quite sizable compared to those of the rest of the population: in 2014, the average gross labor income in Italian cities with fewer than 5,000 inhabitants was $1,380$ € per month, while in cities with between 5,000 and 25,000 inhabitants it was about $1,500$ €. Serving as mayor provides a significant source of income for a large fraction of the population, especially for mayors of small cities.

There are several limitations to the use of the population threshold as a strategy to identify the effect of mayoral wages on public procurement outcomes. One major limitation relates to policy overlap, which has the potential to invalidate the identification strategy. In the sample period, 2013–2017, there is no overlapping policy at the 5,000-inhabitant threshold that prevent the isolation of the aforementioned effect.

Three minor limitations related to the Italian institutional framework might affect, to a certain extent, the interpretation of the results. First, the compensation of the municipal executive committee co-moves with that of the mayor at the 5,000-inhabitant threshold. Although I cannot distinguish between the two effects, this is a second-order problem for at least two reasons: the effect of an average increase in politicians' wages is of interest in and of itself, and the magnitude of the compensation of the executive committee is small. Second, mayors can have different sources of revenue, earning their money from different jobs. Even though this scenario is plausible, the possibility of earning outside income affects only the external validity of the results. Moreover, external income adds noise to the empirical estimates. In that regard, evidence on the effect of mayoral wages on public procurement is an indicator of a powerful relationship. Third, under specific and documented circumstances, the executive committee can grant up to an additional 15% salary increase to the mayor, conditional on the approval of the Ministry of Internal Affairs. If applied, this policy would change the (quantitative) interpretation of the estimated results. However, this is quite a rare event.

3 Data and Sample

Data on mayors refer to those in charge from 2013 to 2017 in all Italian municipalities. The dataset contains information about the identity, gender, age, highest educational attainment (self-reported), political affiliation, and previous job (self-reported) of the elected mayor, and the exact duration of service. Moreover, I collect yearly information at the municipality level about the size of the resident population and certain demographic and geographic characteristics. The individual-level data on mayors are gathered from the Anagrafe Amministratori Locali by the Italian Ministry of the Interior, while the municipal-level data are from the the Italian National Institute of Statistics (ISTAT). Details about the data sources are provided in appendix D.1.

I combine the above information with data on procurement contracts administered by each mayor between 2013 and 2017 that have a project value greater than or equal to $40,000$€. Data on procurement have been provided by the ANAC and include tender-level information about the following: the value of the tender, the number of lots into which the tender is divided, the value of each specific lot, the type of contract (public work, provision or service), the choice of contractual partner, the type of contract procedure (auction, granting, etc.), the award criterion, the number of bidding firms, the identity of the winning bidder, the reserve price, the price rebate and a series of dates that specify different stages of the procurement process. Data about ex-post renegotiation are available, but for only a limited sample of contracts.

Since the aim of the study is to evaluate whether mayoral wages generate distortion in the procurement setup, the dependent variables that I use in the empirical analysis are meant to capture some forms of possible (in)efficiency. I divide the outcome variables into three groups.
First, I include aggregate indicators that account for the *number of reported tenders*, the *total amount of procurement expenditure*, and the *mean value of the contracts* by mayor over the sample period. Second, I account for procurement features at the tender level. *Bidders* is a variable providing the number of admitted offers to a specific tender. Although this variable may have some limitations, it is usually seen as a proxy for competition in the adjudication mechanism. *Rebate* is the winning rebate over the initial value of the project. From the local government's point of view, the higher the rebate the better.[15] *Incumbent* is a dummy variable that measures the firm's incumbency. It is constructed using information about the winner's identity, and it indicates whether the winner of a particular contract has won at least one other contract with the same mayor.[16] It is not clear in advance whether this measure positively or negatively affects the procurement outcome. It is worth noting, however, that a relatively high occurrence of incumbent winners in the same municipality is regarded, at least in principle, with suspicion. *Sum incumbency* counts the number of contracts obtained by the same firm in its prolonged relationship with a mayor. Third, for a restricted sample, I generate indicators of *Time overruns* and *Cost overruns*. These

[15] There are situations in which a rebate that is too high can be symptomatic of bad practices. Indeed, very low contract prices can signal corrupt behaviors, where the difference between the actual rebate and the winning rebate is the amount of money that goes to the bureaucrat as a bribe.

[16] All the results presented in the next sections are robust to the following alternative specifications: winning repeated contracts with the same mayor in the same mandate, and winning repeated contracts that correspond to the same or different tenders.

measures are used extensively in the literature to proxy for the ex-post efficiency of the procurement contract. The lower the time and cost overruns, the higher the level of efficiency.

In the second part of the analysis, I also show how manipulation of the tenders can take place. To do this, I use different variables. *Tendering period* measures the duration, in days, of the call for tenders. This variable is one of the main indicators of transparency following the guidelines of the ANAC. As a rule of thumb, the higher the duration, the higher the probability that firms become aware of the tender, the higher the competition level, and the more transparent the procurement process. *Lots* indicates the order of partition of the tender. A high number of lots can be indicative of suspicious practices, given the different publicity requirements related to the value of the contract. Although this may be a minor issue for public works, where the selection of the sample makes contracts homogeneous, it can be an interesting point for provisions and services. Finally, *Open auction* and *Lowest price* are dummy variables indicating the procedure and the adjudication criterion through which the contractor is chosen.

My sample selection techniques aim to homogenize the sample. First, I focus only on those contracts that are tendered via an open auction or a negotiated procedure. This allows the contracts that I consider to be comparable to each other. Second, I remove regions with special status, which are subject to distinct procurement regulations. Third, I select the sample using the procedure suggested by Imbens and Kalyanaraman (2012) and Calonico et al. (2014, 2017). Specifically, I focus on the contracts administered by municipalities within the interval around the discontinuity threshold, selected using the optimal bandwidth method. In this quasi-experimental setup, municipalities that lie in a small interval around the threshold are likely to be identical in terms of observable and unobservable characteristics, and this allows increased wages to be quasi-randomly assigned across treated and control mayors.[17]

In terms of mayoral wages, I follow the system of population thresholds adopted by the Italian legislators and focus on the interval of 3,000–10,000 inhabitants around the 5,000-inhabitant cutoff.

3.1 Descriptive Statistics

In table 2, I present summary statistics for the municipalities that comprise my sample.

The majority of the Italian cities in my sample are located in the north, in particular in the northwest, which also corresponds to the more productive area of the country. The sample includes a number of southern cities as well, while the central region of the country is less well represented. The great ma-

[17] For the final version of this paper, I do not perform any further selection. However, as a control exercise, I removed contracts that are missing some of the aforementioned variables. The results (available upon request) are not affected by this additional layer of selection.

jority of the mayors are male, even though on average females slightly outnumber males in the overall population. The average number of years in office of the mayors is two and a half, meaning that, on average, during the period 2013–2017 these political leaders are in the middle of their mandate. In Italy, mayors of municipalities with 3,000 or more inhabitants are subject to term limits: these mayors cannot run for election after their second term. In the sample, 73% of the mayors are in their first term, while the remaining 27% are in their second and last term. The majority of the mayors are born in the same province as their municipality, and most of them belong to a civic list.

Table 3 describes the main features of the tenders included in the sample. Statistics refer to the sample after selection and focus on municipalities with a population in the interval 3,250–9,750 inhabitants around the relevant threshold. Data include a total of 11,866 tenders, with an average of 7 bidders per auction and a mean winning rebate of 15.74%. Open auctions are uncommon, and the lowest price is by far the most commonly used adjudication mechanism. Single firms dominate among the winners, whereas single participation is limited. Interestingly, almost 21% of the tenders are awarded to firms that have already won at least one other tender with the same mayor. The breadth of the sample allows for differentiation among procurement types. Table 3 clearly shows on which aspects procurement tenders differ. Public works are generally characterized by a higher number of admitted offers, while the participation of a single firm is a rare event. Public works tenders usually feature higher rebates on the reserve price, with the average rebate more than 8% higher than in the case of provisions and services. The reserve price is generally higher for public works contracts, but there are no relevant differences between public works and provisions and services in terms of the number of lots and the tender period. Incumbent firms appear to be more likely to win repeated contracts for provisions and services than for public works. The heterogeneity among procurement types is an important element that this study contributes to the extant literature. For that reason, I explicitly take it into account throughout the empirical analysis.

Table 4 reports descriptive statistics by procurement type and municipality size, distinguishing between cities below and above 5,000 inhabitants. For public works contracts, the number of bidders and the final rebate both increase above the threshold, possibly indicating an overall increase in the level of competition for public procurement. In that regard, my results follow the theoretical arguments and empirical findings of Conley and Decarolis (2016). The mean reserve price increases by 10%, while the likelihood of firms winning repeated public works tenders does not change, just like the other indicators in the table. Focusing on provisions and services, things are quite similar. In particular, the reserve price and the final rebate go up, whereas the other variables remain more or less unchanged around the threshold.

4 Regression Discontinuity Design

The specific features of the Italian legislation, notably the relationship between mayoral remuneration and population thresholds, create the ideal conditions for the use of RD design (Hahn et al., 2001; Imbens and Lemieux, 2008; Lee and Lemieux, 2010) to estimate the effect of mayoral wages on procurement outcomes.

The economic intuition behind the RD methodology is as follows. Estimates are obtained by comparing features of the tenders that are managed by mayors whose municipalities lie just above or just below the 5,000-inhabitant threshold. In these cases, the outcomes of the tenders differ because of the mayoral wage incentive but should otherwise be identical, taking into account observable and unobservable characteristics.

The key assumptions of RD design are the following: (1) the forcing variable (population) is continuously distributed around the cutoff (no sorting); (2) the probability of being in the treatment group (receiving a higher pay) changes discontinuously at the threshold; (3) in the absence of treatment, the expected outcome changes continuously around the threshold (the continuity assumption).

Hahn et al. (2001) show that the RD design nonparametrically identifies the LATE of compliers at the cutoff under these further assumptions: the treatment variable is a monotonic deterministic function of the forcing variable; the forcing variable crossing the discontinuity threshold only impacts the outcome through the treatment (validity of the exclusion restriction); and the random effect of treatment and treatment assignment function are jointly independent of the forcing variable around the threshold.

Denote Ω_i as the wage of the mayor. Specifically, $\Omega_i = H$ if the municipality has a population higher than 5,000 inhabitants and $\Omega_i = L$ if the population is lower. Let P_i represent the population level and $\bar{p}$ identify the relevant threshold. Finally, denote Y_i as the public procurement outcome (one outcome for each dependent variable). Then, the LATE of the mayoral remuneration on tender outcomes, at the population threshold, is defined by

$$\lim_{h \to 0} \mathbb{E}(Y_i | P_i = \bar{p} + h) - \mathbb{E}(Y_i | P_i = \bar{p} - h), \tag{1}$$

which identifies a sharp RD design.

Equation (1) provides the intention-to-treat effect (ITT). Under the continuity assumption of the population around the threshold (and of the unobservables), the ITT is an unbiased estimate of the average treatment effect (ATE) of the change in mayoral wages on procurement outcomes.

4.1 Estimation

The estimation of equation (1) is performed using a data-driven, local-polynomial-based robust inference procedure. Specifically, I follow Calonico et al. (2014) in performing a bias-corrected inference procedure that is robust to large bandwidths. In the context of sharp RD, this implementation offers robust bias-corrected confidence intervals for average treatment effects at the cutoff.

To construct nonparametric estimators and confidence intervals, I perform local linear regression in the neighborhood of the cutoff and select different data-driven bandwidths, following Imbens and Kalyanaraman (2012) and Calonico et al. (2014, 2017). For reviews on classical inference approaches in the RD design, see also Imbens and Lemieux (2008), Lee and Lemieux (2010), DiNardo and Lee (2011), and references therein.

The main specification of my model is defined as follows. The forcing variable is a "delta" population computed as the difference between the population in a municipality and 5,000. As a result, the discontinuity cutoff is set at zero. I use a first-order polynomial regression, that identifies local linear regression (LLR). Following Cattaneo et al. (2019), the local linear estimator is in practice both the preferred and the most common choice for RD design. The local polynomial used to construct the bias correction is a second-order, local quadratic regression (LQR). The kernel function is triangular, while results are robust to the application of the uniform or the Epanechnikov specification. The selection of the bandwidths is made using a second-generation plug-in selector in line with Calonico et al. (2017). In the main specification, I set the bandwidths to minimize the mean square error, imposing the constraint that they be equal around the cutoff. Finally, errors are clustered at the city level, specifying a heteroskedasticity-robust nearest-neighbor variance estimator with the minimum number of neighbors set at three. Robustness checks, with the inclusion of covariates (Calonico et al., 2017, 2019) and alternative specifications, are provided in section 6.

5 Empirical Analysis

5.1 Testing the continuity assumption

In the RD framework, the validity of the assumption of continuity around the threshold is a necessary condition for obtaining causal estimates. I test this assumption both graphically and statistically, following the suggestions of McCrary (2008) and Lee (2008). The graphical and statistical methodologies are, in some ways, complementary.

In figure 1 I present a histogram of the population values around the 5,000-inhabitant threshold.[18] The figure suggests no sorting occurs around the cutoff and provides preliminary evidence for the validity of the continuity assumption. To formally test this result, I perform a McCrary density test. This test implements local linear density estimates and overcomes problems arising from separately estimated kernel density functions.[19] First, the routine generates a very undersmoothed histogram. The bins of the histogram are defined such that none of them includes points both to the left and to the right of the point of discontinuity. Then, the routine performs local linear smoothing of the histogram. The midpoints of the histogram bins are treated as regressors, while the normalized counts of the number of observations falling into each bin are treated as outcome variables. To accommodate the potential discontinuity in the density, local linear smoothing is conducted separately for the bins to the right and to the left of the point of potential discontinuity.

In figure 2, the results from the McCrary test suggest that no sorting occurs around the threshold, meaning that mayors do not manipulate population statistics to obtain higher remuneration. In table 5, I report the parametric version of the McCrary test with a specification by year.

Lee (2008) suggests an alternative procedure that aims to check the continuity of predetermined variables at the threshold. A predetermined variable is one that does not experience a jump at the discontinuity point. The lack of jumps at the cutoff for predetermined variables is necessary to ensure full identification of the effect of mayoral wages on procurement outcomes.

I test this occurrence in figure 3, in which a set of predetermined characteristics are investigated: *Area*, that measures the area of the municipality in square kilometers; *Sea level*, a measure of the height above sea level of the city administrative center; *North* and *South*, which are dummy variables that take the value of one if the municipality is in the north or south, respectively, of Italy.

In figure 3 I plot nonparametric estimates of the examined variables against the delta population, which is the difference between the number of inhabitants in a municipality and 5,000. The graphs use dots to depict local sample means over nonoverlapping bins that partition a restricted support of P_i, together with a local linear regression estimate for the control and treatment units separately. The figure, in line with the density tests, suggests no evidence of discontinuity around the threshold.[20]

[18] The threshold takes a value of zero here, because I use the delta population in the main specification.

[19] These problems are related to the bias at the boundaries.

[20] The jump in the variable *South* could be suspicious. First, it is worth noting that this discontinuity is not significant at any standard value of significance. Second, in section 6 I replicate all the estimates, controlling for the macro-region the municipality

This evidence, taken as a whole, supports the validity of the RD assumption that there is no manipulation around the threshold. As a consequence, I conclude that the set of municipalities that compose my sample is quasi-experimentally assigned around the value of 5,000 inhabitants.

5.2 Procurement outcomes: Graphical analysis

In this section, I present graphical evidence on the effects of mayoral wages on public procurement variables. Following the main structure of the paper, I split the analysis into two parts.

First, I test whether (at the discontinuity point) municipalities with mayors earning higher wages tend to issue more (fewer) tenders, spend a larger (smaller) amount of money on public procurement, and set up tenders with a higher (lower) average value. I refer to these measures as aggregate procurement outcomes.[21] Results are reported in figure 4, where zero represents the 5,000-inhabitant threshold. As we can see, none of the considered measures displays discontinuity around the threshold. This fact is of particular interest. On the one hand, it urges us to see beyond aggregate statistics and look at the details of individual tenders. On the other hand, it seems to suggest that a decade of Italian legislative reforms led to a substantial alignment in the way that public procurement is handled.[22]

Second, I test for the impact of higher mayoral wages on the individual characteristics of the tenders. I focus specifically on the number of admitted offers, the final rebate over the reserve price (in percentage terms), the probability that the same firm is awarded multiple contracts, and the number of contracts that each firm wins with the same mayor. I refer to these measures as procurement outcomes. Figure 5 depicts the results graphically, using the standard discontinuity threshold of 5,000 inhabitants (figure A1 in appendix A replicates the same graph using quadratic approximations). Results are as follows. First, there is a significant jump upwards in the number of admitted offers. Second, there is a corresponding increase in the final rebate over the reserve price. Third, the probability that a firm is awarded multiple contracts, and the number of contracts awarded, both decrease significantly. These results are not conclusive, but they all support the same intuition: the increase in the number of admitted offers,

belongs. For all the considered specifications, estimates keep their sign and magnitude, whereas significance levels increase.

[21] Using the data provided by ANAC, I have also looked at the transparency with which municipalities report information. I have constructed an indicator of transparency as the ratio of the number of tenders with a minimum set of information and those actually assigned. Here, the minimum set of information includes: the reserve price of the project, the number of admitted offers, the final rebate, the tendering period, the adjudication mechanism, and the identity of the winner. The results shows no discontinuity of the transparency indicator around the threshold.

[22] These results differ slightly from those in Gagliarducci and Nannicini (2013). For the period 1993–2001, they show that municipalities below the threshold tend to spend significantly more money and in a more inefficient way.

together with the increase in the final rebate and the decrease in the probability that firms win repeated contracts, suggest (with some caveats) that tenders issued just above the threshold are characterized by a higher level of competition.

To summarize, the graphical analysis suggests two things: that mayoral wages have no impact on aggregate procurement outcomes, but that they do have an effect on the way that mayors manage the procurement process, with higher wages associated with more ex-ante competition and a lower probability that the same firm is awarded multiple contracts. In the following section, I use the same dependent variables but complement simple graphical analysis with a more rigorous empirical analysis based on state-of-the-art nonparametric techniques.

5.3 Procurement outcomes: Nonparametric analysis

In this section, I report local-polynomial-based estimates for the mean treatment effect of mayoral wages on procurement outcomes in the RD design.[23] I allow for fully data-driven inference with automatic selection of the optimal bandwidth.[24] The main specification of the model implies point estimates in an interval of the discontinuity of distance 1,750–4,750 inhabitants;[25] local linear regressions to construct the point estimator; local quadratic regressions to implement the bias correction; a triangular kernel to construct the local polynomial estimators; an MSE-optimal bandwidth selector (below and above the cutoff) for the RD treatment-effect estimator; and a cluster-robust NN variance estimator at the municipality level. The analysis is again divided into two parts.

In table 6 I report the aggregate procurement outcomes. In line with the graphical analysis, it shows no discontinuity around the threshold for any of the variables considered at the mayoral level: number of reported tenders, expenditure on public procurement, and average tender value.[26] Results are qualitatively in line with those of Gagliarducci and Nannicini (2013), who find that mayors above the threshold tend to spend less on procurement contracts. However, the lack of discontinuity around the threshold suggests the need to examine the details of the individual tenders.

The main results of the paper, then, relate to the second part of the analysis, which focuses on pro-

[23] My analysis follows the suggestions made by Cattaneo et al. (2019).

[24] Optimal bandwidth selectors have the ability to minimize the bias-variance trade-off. On the one hand, increasing the interval increases the bias of the point estimate. On the other hand, the larger the interval, the lower the variance. Full discretion in the choice of the interval can noticeably affect the results.

[25] The choice of the intervals is intended to avoid possible overlap with the 3,000- and the 10,000-inhabitant thresholds.

[26] All variables have been weighted by the number of days for which the mayor appears in the dataset. Estimates were repeated with a focus on raw data, and the results remained unchanged.

curement outcomes at the level of the individual tender. In table 7 I report the core findings for all the contracts in the sample, without distinguishing between public works and provisions and services.

The discontinuity in mayoral wages is associated with a statistically significant increase in the number of bidders (see table 7, col. 1). The ratio of the final rebate over the reserve price (col. 2) shows that the final rebate goes up at the threshold: the discount on the reserve price increases by more than 2% when mayors earn a higher wage. The probability that a firm is awarded multiple contracts decreases by 13%, an effect that is significant at the 99% level (col. 3). The number of contracts awarded by a mayor to the same firm decreases by 0.21 (col. 4). Taken together, these results suggest a specific (and consistent) interpretation. Column 1 provides (partial) evidence that the ex-ante level of competition increases discontinuously around the threshold. Column 2 reinforces this interpretation and shows that the final rebate benefits from the competitive environment. Finally, columns 3 and 4 show that increased competition leads to more firms being awarded public contracts.

To summarize, the findings from the nonparametric analysis confirm those from the graphical analysis. The 28% increase in (gross) mayoral wages of the mayor, which is due to the exogenous regulation of mayoral remuneration, is associated with more competitive tenders and a broader spectrum of firms participating in the public sphere.

In table 7 I exploit the first layer of heterogeneity in my dataset. The table reports estimates for the full sample first for all tenders, including public works and provisions and services (panel A), and then separately for all tenders sold in open auctions (panel B) and via negotiated procedures (panel C). Evidence from these panels offers new and useful insights. Notably, most of the findings in panel A are due to the open auctions. The increase in competition, identified by a large number of admitted offers and the resulting lower rebate on the reserve price, only characterizes open auctions. The magnitude of the effect is large: at the threshold, there is a 20-unit increase in the number of bidders and a 13% higher rebate. At the same time, the decrease in the probability that firms win repeated auctions (25%) and the decrease in the number of tenders won are large and significant at the 90% level. For negotiated procedures, the effect on the number of admitted offers and final rebate vanishes, while the effect on the probability that firms win repeated contracts, although lower in magnitude, remains significant. Indeed, much of the effect of mayoral wages on procurement outcomes seems to be related to the different management of open auctions, with the common denominator of the reduced probability that firms win repeated contracts. These findings, though not conclusive, are of particular interest.[27]

[27] This step of the investigation provides the first evidence that policy implications in public procurement can be highly heterogeneous. The difference between open auctions and negotiated procedures has not been adequately addressed in prior studies.

Negotiated procedures generally allow for a higher level of discretion. In that regard, it is not surprising that the number of admitted offers is lower and that it does not change discontinuously around the threshold. Open auctions, in contrast, allow for less discretion, and sometimes they are required (e.g., for important and high-value tenders). One way to make up for the lost discretion is through the setting of the auction features. As I show in the next section, there is a special discontinuity characterizing my data such that, for open auctions, mayors who are below the threshold tend to use a more flexible adjudication mechanism. The use of less-transparent criteria in the setup of the open auction can, indeed, be the main explanation behind the findings in panel B.

Given the granularity of my data, in appendix A.2 I further decompose the sample and investigate the effects of mayoral wages on procurement outcomes for different typologies of negotiated procedures. I distinguish between negotiated procedures with low discretion, which are mostly restricted auctions created by legislators to streamline the procurement process, and negotiated procedures with high discretion, characterized by less-stringent requirements, which require the following of specific procedures. Table A2 in appendix A.2 shows that contracts with high-discretion procedures are the one that drive the outcomes in panel C of table 7.[28]

To summarize, table 7 reports two main findings. First, the treatment effect of an increase in mayoral wages is associated with a higher level of competition in public procurement, especially when the tenders involve open auctions. Second, at the cutoff there is a significant decrease in the probability that firms win multiple contracts with the same mayor.

An interesting feature of my dataset is that, unlike the standard in the literature, it includes data on contracts for both public works and provisions and services. In tables 8 and 9, I exploit this layer of heterogeneity and examine the effects of mayoral wages on public procurement outcomes, distinguishing for the typology of the contract.

Table 8 refers to public works contracts[29] and is divided into three panels. Panel A includes all the tenders, panel B the open-auction tenders, and panel C the tenders involving negotiated procedures. Regarding the choice of the dependent variables, all the columns correspond to those in table 7. Results for the whole sample (panel A) are in line with those reported in table 7. The coefficient related to the number of admitted offers is positive and significant, as is the positive (but not significant) coefficient

[28] High-discretion procedures are the typical type of negotiated procedure for small municipalities: in my sample, more than 90% of the negotiated procedures belong to that category.

[29] In this paper, whenever I use the term *public works*, I refer specifically to construction. Construction contracts comprise 92% of all of the public works tenders included in the sample. I focus on construction in order to improve the overall level of comparison between different tenders.

associated with the final rebate. In terms of the identity of the winner, the probability of winning repeated tenders and the number of tenders obtained with the same mayor both decrease significantly at the threshold. For open auction (panel B), all the coefficients amplify dramatically and become significant. At the threshold, there is a 30-unit increase in the number of bidders and an 13% higher rebate, significant at the 95% level. At the same time, the decrease in the probability that firms win repeated auctions (19%) and the decrease in the number of tenders won are large and significant at the 90% level. For negotiated procedures (panel C), it is worth noting that the signs of the coefficients are the same as those in panel A, and significance is also preserved for admitted offers. Results in table 7 show no effect on competition for negotiated procedures. Here, I complement those findings and show that those effects are present if I focus on public works. The results for public works are relatively consistent, since public work contracts, unlike those for provisions and services, are less standardized and leave more room for maneuvering.

Table 9 focuses on provisions and services. The effects of mayoral wages on procurement outcomes at the cutoff are quite different from tables 7 and 8. The effect of mayoral wages on the number of admitted offers vanishes (col. 1), and there is a significant increase in the final rebate over the reserve price in open auctions, whereas that effect disappears (and the sign switches to negative) in the case of negotiated procedures (col. 2). Finally, the estimates in columns 3 and 4, though not significant, are qualitatively in line with the general finding that winning multiple contracts becomes more difficult at the threshold. I believe that the fact that provisions and services contracts follow more standardized rules, with a national authority (the Consip) directly involved in the procurement process, is behind these results.

Tables 8 and 9 provide preliminary evidence highlighting the heterogeneous effects of increasing mayoral wages on procurement outcomes, conditional on the type of procurement. For public works, the main results associated with the increase in mayoral remuneration are an increase in the number of admitted offers and final rebate on the reserve price (higher competition), and a concurrent reduction in the probability that firms win multiple contracts. Regarding provisions and services, in contrast, there is still a discontinuous increase in the value of the final rebate as well as a reduction in multiple wins, while the effect on the number of bidders vanishes.

Taken together, the findings in tables 7, 8, and 9 show how mayoral wages affect procurement tenders and how that effect can be heterogeneous, depending on the nature of the contract. The main result is that tenders above the threshold are subject to higher competition and more alternation of the winners.

However, competition does not necessarily mean efficiency. It can still be the case that lower rebates are compensated for by ex-post renegotiation, and/or that higher discounts are symptomatic of manipulation. At the same time, it may be the case that prolonged relationships with the same firms are driven by reputational concerns rather than by pure favoritism. To test whether increased competition is associated with better performance, I construct two measures of efficiency that are widely used in the literature: *Time overrun* and *Cost overrun*. *Time overrun* measures the difference, in days, between the actual date of delivery and the estimated date at the signature of the contract (weighted by the estimated time of execution). *Cost overrun*, similarly, measures the difference between the final cost of the project and the cost estimated at the signing date (weighted by the estimated cost).[30] I expect that both time and cost overruns decrease at the threshold, meaning that the increase in tender competition is accompanied by a parallel increase in overall efficiency.

The availability of data on time and cost overruns is limited to a restricted subsample of my dataset. For time overruns, the limited subsample mostly includes tenders for public works, whereas tenders for provisions and services tend to have very diverse arrangements. For cost overruns, the associated subsample features a more balanced distribution of the two typologies of procurement contracts. For these reasons, I present estimates for *Time overrun* and *Cost overrun* in two separate tables.

Table 10 examines cost overruns and focuses on all the contracts, without distinguishing between public works and provisions and services. Cost overruns go down at the threshold, meaning that the final cost of the project decreases (col. 1). The results for open auctions and negotiated procedures (cols. 2 and 3, respectively), confirm this finding and show that cost overruns decrease more for open auctions.

Table 11 focuses on public works and adds to the previous table a proxy for time overruns. The estimates for public works cost overruns replicate those in table 10, whereas those associated with public works time overruns are (still) negative but not significant.

These findings show that increasing the mayoral wage leads to more competitive procurement tenders. Moreover, the cost overrun results suggest that the increase in competition is, in turn, associated with gains in efficiency (i.e., a discontinuous reduction in cost overruns).

Limited availability of data reduces the power of the analysis. A possible concern about the above-mentioned relationship is that the sample for which I analyze cost overruns is not randomly extracted from the overall sample. To address this concern, I perform two exercises. First, I verify that the availability of information on cost overruns does not show any discontinuity around the threshold. Second, I repeat the analysis in table 7, restricting the sample to those observations for which I have data on cost overruns. Results, reported in Table A3 of appendix A.3, confirm the consistency of the main findings.

[30] I measure time overruns as $\frac{closing\ date - estimated\ closing\ date}{estimated\ time\ of\ execution}$. I measure cost overruns as $\frac{final\ cost - awarding\ price}{awarding\ price}$.

Notwithstanding these facts, I recommend to interpret the evidence on cost overruns with caution.

5.3.1 Mechanisms

In this section, I complement the analysis on the effects of mayoral wages on public procurement by providing evidence on the mechanisms that are behind these findings.

The first question that I seek to answer is how mayors and their executive committees can design the tenders to achieve their aims. Put differently: which are the features of the tenders, chosen by local political leaders (with all the caveats discussed in section 2), that can explain the findings in the previous section? The answer is that the difference is in the details.

Since the details are very difficult to observe, I confine my analysis here to some proxies for transparency that help explain how competition and efficiency may be fostered and/or limited by discretion.[31]

Table 12 is a first attempt in this direction. Panel A refers to all the tenders in the sample. Panel B focuses on open auctions. Regressions from column 1 employ a dummy variable for negotiated procedures, while the regressions for column 2 use a dummy that indicates whether the adjudication mechanism is the lowest price (rather than the most economically advantageous) tender. Column 3 measures the duration, in days, of the call for tenders. This variable is one of the indicators of transparency prompted by the ANAC guidelines. As a rule of thumb, the higher the duration, the higher the probability that firms become aware of the tender, and the more transparent the procurement process. Finally, Column 4 indicates the number of lots into which the tender is divided. A higher number of lots can be indicative of suspicious practices, given that low-value contracts can be subject to more flexible regulations.[32]

Point estimates from Panel A are qualitatively intuitive (although not statistically significant).
Above the threshold, mayors tend to issue more open-auction tenders, make more extensive use of the lowest-price mechanism, stretch the publicity period of the tender, and reduce the number of lots, thus increasing the level of transparency of the tender. These results are in line with the findings in the previous section and (partially) explain how the procurement process can be designed in order to shape its outcomes. Panel B completes the analysis and shows that, for open auctions, there is a noticeable

[31] I want to stress that it is not my intention to go beyond a prima facie analysis. In Italy, the National Anti-corruption Authority (ANAC) is in charge of monitoring the fairness of the procurement process. A call for tenders is a very complicated document, full of technicalities. Indeed, any misconduct is generally hidden in those technicalities. Since the details are difficult to observe, it is far from possible for me to provide an all-encompassing analysis here.

[32] The problem is not as great for public works, where the sample selection makes contracts relatively homogeneous.

increase in the use of the lowest-price mechanism at the threshold. This finding is of particular interest and can explain the striking results I found for open auctions. While mayors above the threshold use the lowest-price adjudication mechanism in association with open auctions, those below the threshold tend to adopt the most economically advantageous tender. The latter approach provides greater discretion in the setting of the participation criteria and can dramatically limit competition.[33]

The second question that I seek to answer is this: why do higher wages induce mayors and their executive committees to behave differently? The literature has proposed three possible reasons: selection, re-election incentives, and morale. In what follows, I investigate these mechanisms and link them with public procurement outcomes.

First, higher wages reduce the opportunity cost of mayoral service and may therefore encourage more highly educated people to run for election. If that is the case, the effects of higher wages on procurement outcomes–more competition and greater efficiency– can be explained by the election of highly educated mayors. To test this hypothesis, I code a dummy variable *High edu* that is equal to one if the mayor has at least a bachelor's degree and zero otherwise, and I test whether this variable changes discontinuously at the threshold. The results, reported in figure 6, highlight two things worth noting. First, mayors' educational attainments are generally high, with more than 50% of the mayors being highly educated. Second, no discontinuity can be detected, meaning that education is not a driver of the differential outcomes. These results differ from the ones in Ferraz and Finan (2009) and Gagliarducci and Nannicini (2013), where selection is crucial to explain the effect of wages on public performance.[34]

Second, higher wages increase the monetary value of being in office, therefore increasing the value of re-election. When the pivotal group whose support is required for re-election corresponds to the mass of citizens, local procurement performance aligns with the community's interests and procurement procedures will be set properly. Following this reasoning, the difference in outcomes that I have shown in the previous section may be due to the increased incentives, for mayors above the threshold, to be re-elected. To test this hypothesis, I exploit a unique feature of the Italian legislation in which the second term of the mayoral mandate is binding. I separate the sample into two subsamples. The first subsam-

[33] By itself, the "most economically advantageous tender" mechanism should not be discounted. Indeed, there are situations in which it can significantly improve performance. Here, I just want to highlight two facts: first, that it is a less transparent tool compared to the lowest-price mechanism; and second, that there is a nontrivial discontinuity in its usage for open auctions, and that this discontinuity is paralleled by a reduction in ex-ante competition and overall efficiency.

[34] I impute this divergence to differences in the sample composition. In my sample, 93% of the mayors have a high school diploma or higher degree. Of these, 53% have at least a bachelor's degree. The difference in these percentages can be explained by the increased access to education that has characterized advanced economies in the last fifty years.

ple includes mayors who are in their first term and have the opportunity to be re-elected. I expect those mayors to be affected by re-election incentives and procurement outcomes to jump discontinuously at the threshold. The second subsample includes mayors that are in their second (and final) term. For these mayors, I expect attenuated effects of higher wages on procurement outcomes. The results are reported in table 13. Panel A refers to mayors in their first term, and panel B to those in their second term. As we can see, the results are consistent with the hypothesis. Procurement outcomes change significantly at the threshold for mayors who are in their first term, whereas the wage effect vanishes for mayors in their second term. This is also in line with the findings of Coviello and Gagliarducci (2017), who show that tenure in office has an influence on procurement outcomes. My results enrich their framework by showing that mayoral wages can be the underlying cause of the tenure effect via re-election incentives.

Third, it is possible that higher wages have an effect on morale that can act as a deterrent to corruption. To test this hypothesis, which refers to the increase in the opportunity cost of illegal practices when wages are higher, I replicate the main exercise but exclude municipalities that have been investigated for corruption.[35] Information on investigated municipalities is very difficult to obtain. Here, I use confidential data from Decarolis et al. (2020)[36] provided by the Agenzia Informazioni e Sicurezza Interna (AISI), the Italian equivalent of the FBI. I code a dummy variable, *Investigated municipality*, that is equal to one if at least one of the public officials in charge of awarding the procurement contract has been flagged for suspected corruption. I use this dummy to screen the sample, and I replicate the analysis considering only the municipalities that are not under investigation. The results are reported in table 14. The main findings of the paper remain unaltered. Higher wages are related to more bidders at the tender, a higher final rebate, and a lower probability of winning multiple contracts, independent of the municipality being under the scrutiny of law enforcement. This provides evidence that morale incentives are not pivotal in explaining the effect of mayoral wages on public procurement performance.

To summarize, tenders issued just above the threshold are designed in a more "transparent" way. Moreover, re-election incentives seem to underpin the results in section 5.3, where higher wages are associated with more competition and higher efficiency.

[35] In a preliminary analysis, I also test whether there is a discontinuity around the threshold for municipalities that have been investigated. Evidence is provided in figure A4 of appendix A.4.

[36] I am extremely grateful to the authors for sharing their unique, very high-scale and high-quality dataset with me.

6 Robustness checks

In this section, I provide robustness checks to support the validity of the findings in section 5.3.

A first concern, related to the discontinuous increase in competition for tenders just above the threshold, is that municipalities that are above the cutoff have more firms competing for the same contract. If that is the case, the results might be mechanically driven by more competition among firms. To test this hypothesis, I collect data on the number of firms in each municipality from the 2011 Italian census of the firms (data are from the ISTAT) and check for continuity around the threshold. Results (reported in appendix B.1, figure B1) distinguish between total firms and firms operating in the construction sector.[37] As we can see, no discontinuity can be detected. Thus, the results on the increase in competition around the threshold are not driven by more competitors fighting for the public contract.

A second concern relates to the so-called domestic stability pact (DSP). In 1999 the Italian government approved a system of fiscal rules, the DSP, that limited the discretion of municipalities to run public deficits. From 2001 to 2013, the DSP used the 5,000-inhabitant threshold to discriminate between municipalities, with those below the threshold exempted from its application. In 2013 the fiscal regulations changed, and the threshold was moved to 1,000 inhabitants. In the main analysis, which focuses on data from 2013–2017, I exploit the uniqueness of the change in mayoral wages at the cutoff to identify its effect on local procurement outcomes and offer several arguments regarding the channels behind this effect. Still, there is an aspect that merits further investigation: the possibility that illegal or manipulative practices associated with exemption from the DSP persist over time and explain the main findings in tables 7 and following.[38] I use two arguments to address this concern. First, in the main analysis I show that mayors below and above the threshold do not significantly differ in the amount of money spent on public procurement, nor they differ in the number of tenders they issue (aspects that are directly related to the DSP application). Indeed, the only differences relate to the special characteristics of the tenders, an aspect that is more difficult to reconcile with the objectives and requirements of the DSP. Second, in appendix B.2, tables B2a and B2b, I replicate the main analysis on aggregate and individual tender outcomes for the universe of procurement contracts issued by Italian municipalities during the period 2010–2012 (a period in which both mayoral wages and the DSP applicability changed discontinuously at the threshold). Two things are worth noting. Aggregate outcomes show a significant discontinuity at the threshold, in line with municipalities on the two sides of the cutoff being subject to different DSP

[37] The use of the construction sector is very important, as it aligns with the public works data that I have extensively analyzed throughout the paper.

[38] I thank two anonymous referees for proposing this argument.

regimes. Mayors below the threshold open more tenders and spend significantly more on public procurement contracts. Individual outcomes are smoother than those presented in table 7 (and figure 5). In particular, there is no discontinuity in the number of bidders per tender or the final rebate on the reserve price (even if the estimates are qualitatively similar). I interpret these findings as follows. Before 2013, the effect of mayoral wages on individual procurement outcomes was attenuated by the difference in DSP compliance. Mayors below the threshold were allowed to extensively use the leverage of public procurement, and tenders' specific features were of second order. After 2013, those mayors had to adapt to the DSP regulations and, lacking the lever of higher spending, the wage effect was expressed through the tenders' specific characteristics.

A third issue is related to the reform of the public procurement code enacted in Italy in 2016.[39] This reform was the result of a long and troubled process that had been long expected by the professions. The complexity of the new code, however, has been a significant limitation. From its official publication in April 2016, the code has been surrounded by uncertainty and subject to a lack of implementation. It is common knowledge that the Italian system took years to adapt to the requirements of the new regulation. Since my sample ends in 2017, I expect the effects of the reform to be of second order. To test this hypothesis, I proceed in two ways. First, in table 15 I add year fixed effects to the main specification. Second, I directly control for the reform by restricting the sample to those tenders opened before the approval of the code. Results, reported in appendix B.3, table B3, show that the coefficients associated with the main dependent variables are in line with those in table 7. This evidence confirms that the reform does not alter the way mayoral wages affect procurement outcomes.

Another interesting aspect that I investigate in my setting is related to the variability in local purchasing power. Specifically, I am able to exploit variability in the purchasing power of different geographic locations in Italy to obtain a measure of the heterogeneous intensity of the mayoral wage effect. If higher wages matter, I should find stronger differences in procurement outcomes at the threshold in those municipalities where purchasing power is higher.[40] I test this hypothesis in appendix B.4, table B4, where

[39] On April 18, 2016, the Italian government definitively approved Legislative Decree no. 50, implementing Directives 2014/23/EU, 2014/24/EU and 2014/25/EU of the European Parliament and European Council of 26 February 2014 "on public procurement and awarding concession contracts, procurement by entities operating in the water, energy, transport and postal services sectors and on the reorganization of the Public Procurement Regulation" (new code). The new code came into force on April 19, 2016. However, it created many uncertainties for economic operators, requiring a massive revision by legislators through Correction notice no. 164 of 15 July 2016 (with 170 corrections in 220 articles). For that reason, I consider August 1, 2016, as the first day of full application of the new code.

[40] I thank one of the anonymous referees for this suggestion.

I restrict the sample to municipalities that are located in the southern region of Italy, where purchasing power is greater. Estimates from table B4 support this assumption and show that the effect of mayoral wages on public procurement is particularly relevant in places where the increase in mayoral wage is substantial.

In what follows, I consider six possible concerns regarding the discontinuous relationship between procurement outcomes and mayoral wages. I focus specifically on the full sample, without distinguishing between public works and provisions and services. As a consequence, panel A of table 7 and its related estimates identify my preferred benchmark.

First, in table 15 I repeat the main analysis, adding to the basic specification a set of covariates that control for time (year fixed effects), geography (indicators for Italian regions and macro-regions), tender characteristics (type of tender, reserve price, adjudication mechanism, publicity duration), and mayoral characteristics (education, gender, province of birth, number of years in office). Local polynomials methods can easily accommodate additional covariates, but the latter must satisfy the important condition that they be balanced at the cutoff. When the empirical evidence shows that predetermined covariates differ systematically at the cutoff, the assumption of continuity of the potential outcomes is implausible, and thus the nonparametric continuity-based RD framework is no longer appropriate without further (restrictive) assumptions about the data-generating process. Consistent with that, none of the covariates included in table 15 experiences discontinuity around the threshold. Results are consistent with those obtained in the main specification. Ex-ante competition, proxied by the number of admitted offers, significantly increases, and the final rebate over the reserve price increases as well. The probability that a firm wins multiple contracts with the same mayor decreases by almost 12.5%. Finally, for the restricted sample of contracts for which I have data, cost overruns also decrease (the coefficient is almost significant at the 90%). Point estimates are similar to those in table 7, whereas confidence intervals substantially decrease, reflecting the improvement in overall efficiency brought by the inclusion of control variables.

Second, I test whether there are significant effects at simulated cutoff values. The key RD identifying assumption is the continuity of the regression functions for treatment and control units at the cutoff in the absence of the treatment. Hence, I test whether the estimable regression functions for the control and treatment units are continuous at points other than the true cutoff. Evidence of discontinuities away from the threshold can signal potential issues in the identification strategy. For this test, I generate a time series of simulated thresholds ranging from 4,000 to 6,000 inhabitants in steps of 100. I then make statistical estimations and inferences for RD treatment effects at artificial cutoff points separately for the control and treatment groups. To avoid contamination due to real treatment effects, I use only treated observations for cutoffs above the actual cutoff, and I only use control observations for cutoffs below the actual cutoff. Restricting the observations in this way guarantees that the analysis of each fake threshold uses only observations with the same treatment status. Results are reported appendix C.1, figure C1. Significance away from the true cutoff is rarely detected, and in the cases in which it appears it is close to the region of nonsignificance.

Third, I investigate the sensitivity of the results to the response of units that are located very close to the cutoff. The idea behind this approach is to exclude such units and then repeat the estimation and inference analysis using the remaining sample. This idea is sometimes referred to as a "donut hole" approach. Even though systematic manipulation of score values close to the threshold is not a great concern in my setting, this strategy is still useful in assessing the sensitivity of the results to the unavoidable extrapolation involved in local polynomial estimation, as the few observations closest to the cutoff are likely to be the most influential when fitting the local polynomials. In appendix C.2, figure C2, I show the results for holes up to 100 in steps of 10 units. The coefficients associated with all the examined variables preserve their signs and (in general) their levels of significance, even with the introduction of donut holes.

Fourth, I check for asymmetric bandwidths. Results, reported in appendix C.3, table C3, largely confirm the local effects of increased mayoral wages on the number of admitted offers, the final rebate, and the probability that firms win multiple contracts.

Fifth, I check for different clustering of the standard errors. The results, with standard errors clustered by mayor, are consistent with those reported in table 7 (see appendix C.4, table C4).

Sixth, I repeat the main analysis using a higher-order polynomial approximation. Results are presented in appendix C.5, table C5a, for a second-order approximation and in appendix C.5, table C5b, for a fourth-order approximation. Point estimates are similar to those in the main specification, even if statistical significance is reduced[41].

Seventh, I analyze the sensitivity of the results in table 7 to the bandwidth choice. In contrast to the donut hole approach, which investigates sensitivity as units from the center of the neighborhood around the cutoff are removed, here I investigate sensitivity as units are added or removed at the endpoints of the neighborhood. In the continuity-based approach, this falsification test is implemented by changing the bandwidth used for local polynomial estimation. In Appendix C.6, figure C6, I present the results for

[41] I refer the reader to Gelman and Imbens (2019) to substantiate the appropriateness of choosing local first-order polynomials to empirically test my research question in a RD framework.

the number of admitted offers, the final rebate on the reserve price, and *Incumbent*, using a bandwidth window that ranges from 0 to 1,500. The estimated effects are significant across different bandwidths below and above the optimal bandwidth.

7 Conclusion

In this paper, I run a regression discontinuity design analysis to document the causal effect of mayoral wages on procurement outcomes in a large data set of public procurement contracts in Italy. To perform the analysis, I make use of a specific feature of the Italian legislation in which mayoral wages are determined using a threshold scheme based on the number of inhabitants of a municipality. I focus on municipalities with a population between 3,250 and 9,750 inhabitants to exploit the unique properties of the 5,000-inhabitant cutoff that allow for the identification of the mayoral wages effect. Mayors of municipalities with a population above 5,000 earn higher wages, while mayors of municipalities with fewer than 5,000 inhabitants earn lower wages. My identification strategy relies on the assumption that mayoral remuneration cannot be manipulated around the threshold. Since the mayoral wage is associated with the number of inhabitants of the municipality, this prevents mayors from misreporting population data in order to obtain a higher remuneration. I focus on contracts administered by municipalities within the interval around the discontinuity threshold, selected via the optimal bandwidth method. In this quasi-experimental setup, municipalities that lie in a small interval around the cutoff are likely to be identical in terms of observable and unobservable characteristics, and increased wage should be as if quasi-randomly assigned across treated and control mayors.

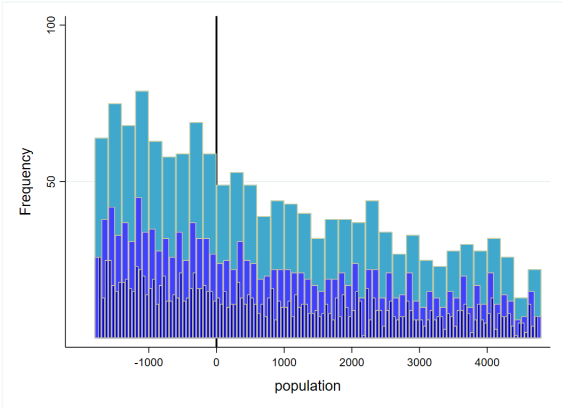
The main results of the paper are as follows. First, aggregate measures of public procurement-the number of tenders, the total procurement expenditure, and the mean value of the contract- do not show any discontinuity around the threshold. The average treatment effect of higher mayoral remuneration is not statistically different from zero. Second, higher mayoral wages positively affect some of the specific tender characteristics. The number of admitted offers and the final rebate on the reserve price experience a significant jump upwards (a signal of higher competition). Interestingly, these findings are mostly associated with open auctions. Moreover, an increase in the mayoral wage causes a significant decrease in the probability that a firm is awarded multiple contracts by the same mayor. This holds true both for open auctions and negotiated procedures, for both public works and provisions and services tenders. Third, for a restricted sample, I also show that the increase in competition is paralleled by an increase in the efficiency of the contract, evidenced by a reduction in cost overruns. Importantly, all the findings persist when I repeat the analysis controlling for the time period, the geographic location of the municipality, contract features, and mayoral characteristics.

While the above results may be driven by different combinations of the mechanisms described in the introduction, I provide evidence that re-election incentives are particularly influential in procurement outcomes.

This study has two main motivations: first, the central role played by public procurement in the provision of public goods at the local level, and second, the lack of empirical evidence on the direct relationship between politicians' monetary incentives and local procurement performance. I contribute to the literature through this attempt to fill this gap. Specifically, I provide the first set of evidence on the effects of increased mayoral wage on procurement outcomes. I also propose several preliminary considerations regarding the existence of heterogeneous effects generated by monetary incentives, distinguishing between public works and provisions and services as well as between open auctions and negotiated procedures. Finally, I outline the possible mechanism behind these effects. To the best of my knowledge, this is the first study that has identified a causal effect of mayoral wages on procurement outcomes with some degree of precision and robustness.

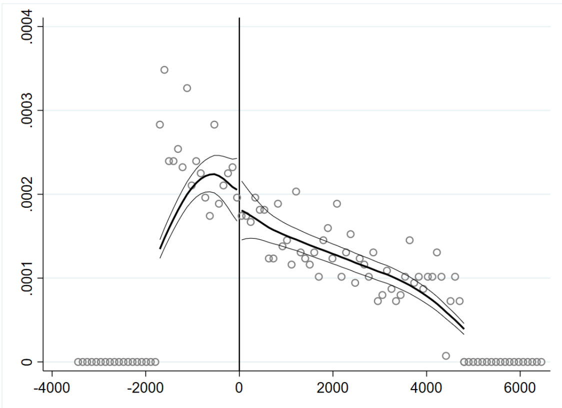
Figures

Figure 1. Population density around 5,000 inhabitants



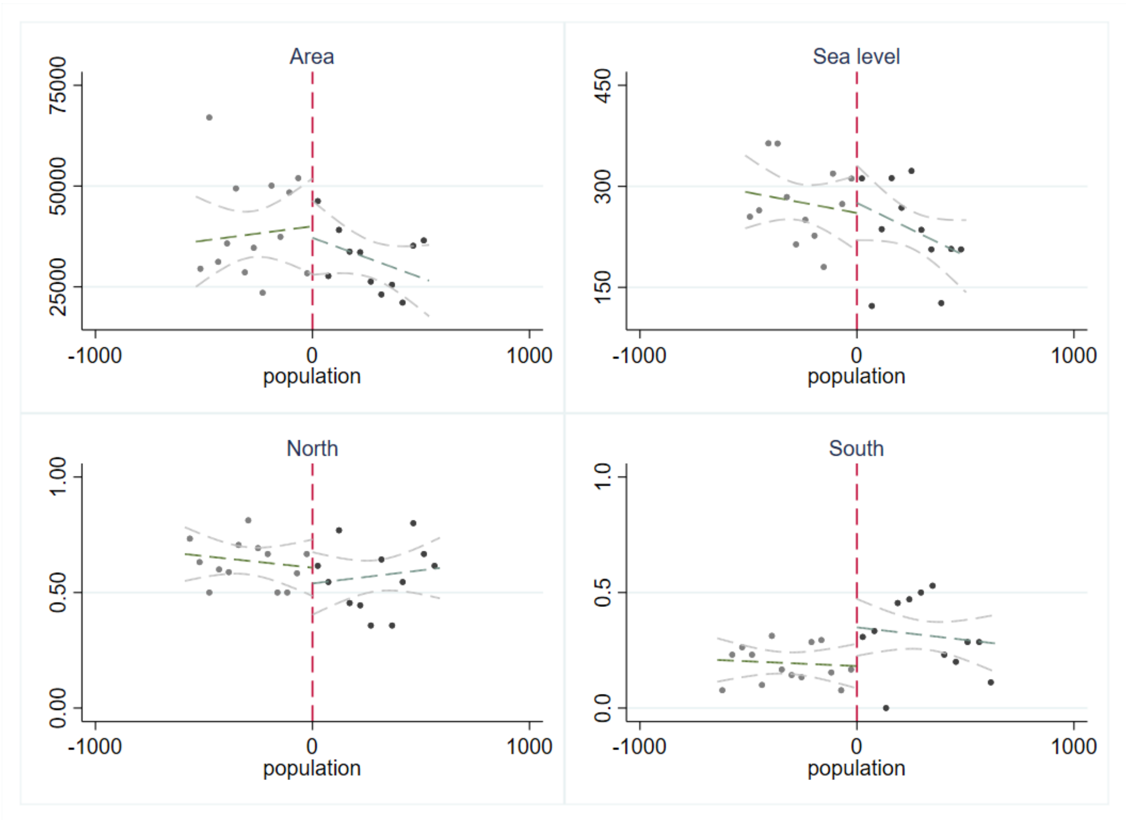
Notes: This figure depicts the frequency of cities around the delta = 0 population threshold (vertical line).

Figure 2. McCrary discontinuity density test results



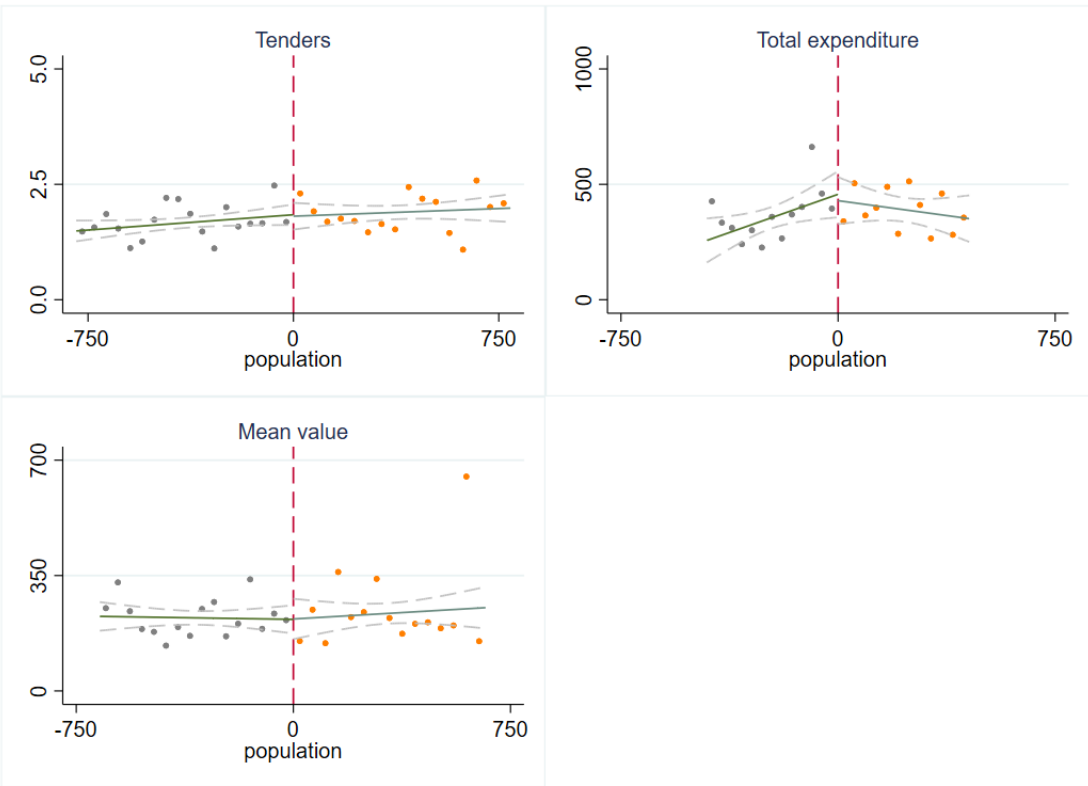
Notes: For population values below 5,000, mayors earn a lower wage. For values above 5,000, mayors earn a higher wage. Circles represent average observed values, the bold solid line is a kernel estimate (see McCrary (2008)), and the two thin lines are 95 percent confidence intervals. The vertical line represents the delta = 0 population threshold.

Figure 3. Predetermined characteristics



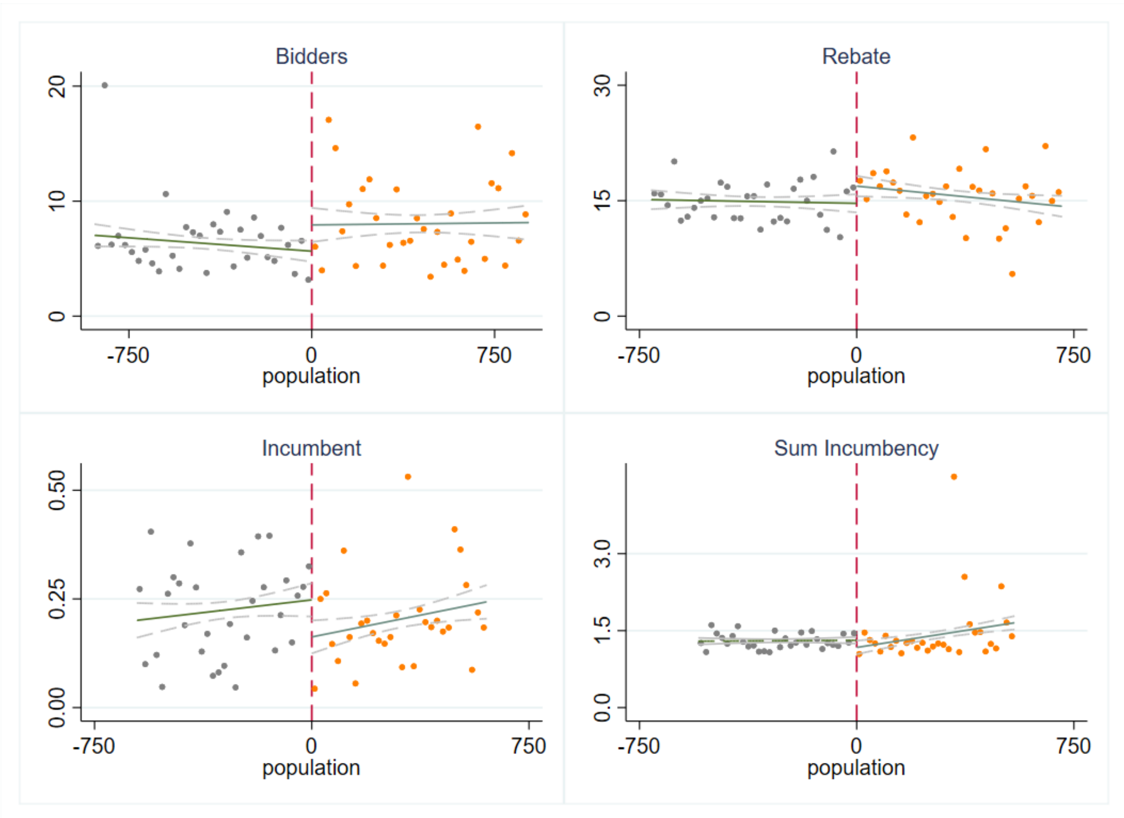
Notes: These graphs report characteristics for pre-determined variables at the city level, for sample municipalities with a population between 3,250 and 9,750 inhabitants. Data are from ISTAT and the Anagrafe Amministratori Locali (Ministry of the Interior). Area corresponds to the area of the municipality in square kilometers. Sea level, in meters, refers to the height above sea level of the city administrative center. North and South are dummy variables identifying the macro-region of each municipality. Local linear approximations are provided together with confidence intervals at 90%. The vertical line represents the delta = 0 population threshold.

Figure 4. Graphical analysis: The effect of mayoral wages on aggregate procurement outcomes



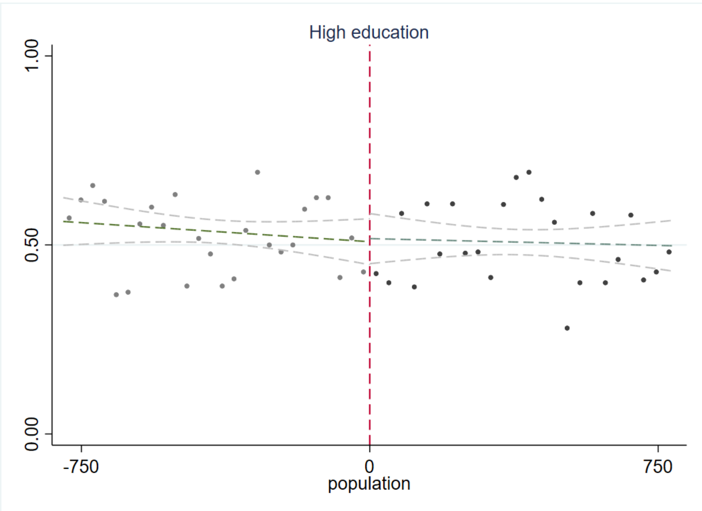
Notes: Data are from the Anti-corruption National Authority (ANAC) for all procurement contracts above 40,000€ offered by municipalities during the period 2013–2017. Sample selection follows the process described in sections 3 and 4. Inferences are made for a close interval of the population size around the threshold (interval is 3,250–9,750 inhabitants). Estimates are made following the procedure described in section 5.3. The top-left panel refers to the number of tenders issued by each mayor, weighted by the mayor's number of years in office. The top-right panel refers to the amount of money (in euros) spent by each mayor on public procurement, weighted by the mayor's number of years in office. The bottom panel indicates the mean value of the tenders, by mayor. The solid lines are local linear approximations, and the dashed lines are their associated 90% confidence intervals. The dashed vertical lines denote the discontinuity, normalized to zero (delta population).

Figure 5. Graphical analysis: The effect of mayoral wages on tender-specific procurement outcomes



Notes: Data are from the Anti-corruption National Authority (ANAC) for all procurement contracts above 40,000€ offered by municipalities during the period 2013–2017. Sample selection follows the process described in sections 3 and 4. Inferences are made for a close interval of the population size around the threshold (interval is 3,250–9,750 inhabitants). Estimates are made following the procedure described in section 5.3. The top-left panel refers to *Bidders*, the number of admitted offers. The top-right panel refers to *Rebate*, the final discount on the reserve price. The bottom-left panels refers to *Incumbent*, and indicates whether the winning firm has won at least one other contract with the same mayor. The bottom-right panel refers to the number of contracts won by a firm during the mayor's terms. The solid lines are local linear approximations, and the dashed lines are their associated 90% confidence intervals. The dashed vertical lines denote the discontinuity, normalized to zero (delta population).

Figure 6. Channels of the mayoral wage effect: Selection



Notes: Data are from the Anagrafe Amministratori Locali (Ministry of the Interior). Inferences are made for close interval of the population size around the threshold (interval is 3,250–9,750 inhabitants). *High edu* refers to mayors that have at least a bachelor's degree. The dashed lines are local linear approximation and the corresponding 90% confidence intervals. The vertical dashed line denotes the discontinuity, normalized to be zero (delta population).

Tables

Table 1. Legislative thresholds for the Italian municipalities

Population	Mayoral Wage	Executive Comm. Wage	Council Fee	Executive Comm. Size	Council Size	Electoral Rule	Election Signatures
Below 1,000	1,290	15%	18	3	12	Single	NO
1,000-3,000	1,450	15%	18	3	12	Single	YES
3,000-5,000	2,170	15%	18	5	16	Single	YES
5,000-10,000	2,790	45%	18	5	16	Single	YES
10,000-15,000	3,100	45%	22	6	20	Single	YES
15,000-20,000	3,100	45%	22	6	20	Runoff	YES
20,000-30,000	3,100	45%	22	6	20	Runoff	YES
30,000-50,000	3,460	45%	36	8	30	Runoff	YES
50,000-60,000	4,130	60%	36	8	30	Runoff	YES
60,000-100,000	4,130	60%	36	8	30	Runoff	YES
100,000-250,000	5,010	60%	36	10	40	Runoff	YES
250,000-500,000	5,780	65%	60	11	46	Runoff	YES
Above 500,000	7,800	65%	104	12-13	50-64	Runoff	YES

Notes: Population is the number of resident inhabitants. Mayoral Wage and Executive Comm. Wage refer to the monthly gross wages of the mayor and the members of the executive committee, respectively. The latter is expressed as a percentage of the former. Council Fee is the reimbursement per quarterly City Council session paid to councilors, measured in euros. The wage thresholds at 1,000 and 10,000 inhabitants were introduced in 2000; all of the others date back to 1960. Executive Comm. Size is the maximum allowed number of executive committee members appointed by the mayor, as set by the law April the 7th, 2014, n. 56. Council Size is the number of seats in the City Council. All of the size thresholds were set in 1960. Since 1993, Electoral Rule can be either single round (60% premium) or runoff (66% premium) voting. Finally, Election Signatures refers to the requirement for candidates to present a minimum number of signatures in order to run for the mayoral elections.

Table 2. Descriptive statistics: Municipal and mayoral characteristics

Variables	Mean	s.d.	p50	N
	Municipal characteristics			
South	0.23	0.42	0.00	1421
Center	0.15	0.36	0.00	1421
North	0.62	0.49	1.00	1421
Total population	5769.87	1807.56	5428.00	1421
Male population	2864.71	908.37	2683.00	1421
Female population	2962.40	944.76	2779.00	1421
	Mayoral characteristics			
Low edu	0.07	0.25	0.00	1874
Medium edu	0.41	0.49	0.00	1874
High edu	0.53	0.50	1.00	1874
Female	0.14	0.35	0.00	1945
Age	52.93	10.37	53.00	1945
Years in office	2.55	1.59	3.00	1945
Second term	0.27	0.45	0.00	1938
Born in province	0.80	0.40	1.00	1945
Party list	0.73	0.44	1.00	1945

Notes: Data are for municipalities with a population between 3,250 and 9,750 inhabitants that are part of the sample. Data sources are listed in appendix D. *South*, *Center*, and *North* are dummy variables identifying the macro-region of the municipality. *Total population* refers to the average total population over the sample period. The table also shows the population by gender. In terms of mayoral characteristics, three dummies indicate the level of education of the mayor, where *Low edu* represents the elementary level, *Med edu* represents the secondary level, and *High edu* represents the attainment of at least a bachelor's degree. *Female* is a dummy variable indicating that the mayor is a woman. *Age* is the age of the mayor. *Years in office* captures the number of consecutive years the mayor has been in office. *Second term* indicates that the mayor is in his or her second and final term. *Born in province* indicates that the mayor was born in the same province as the municipality he or she governs. Finally, *Party list* is a dummy indicating that the mayor ran in the election as the candidate of a civic list.

Table 3. Descriptive statistics: Tender characteristics across procurement types

	Mean	s.d.	p50	N	Mean	s.d.	p50	N	Mean	s.d.	p50	N
	All auctions				Public works				Provisions and services			
Bidders	7.18	17.46	3.00	11866	9.73	21.29	4.00	7374	2.54	3.01	2.00	3941
Rebate	15.74	14.38	12.65	10998	18.47	13.41	18.30	7067	9.87	14.47	3.08	3398
Tend. period	22.98	26.85	19.00	11866	22.65	28.20	19.00	7374	23.50	23.56	19.00	3941
Lots	1.11	0.77	1.00	10245	1.02	0.25	1.00	6429	1.27	1.27	1.00	3298
Reserve price	221.7	462.0	114.0	11866	228.3	433.4	118.54	7374	215.1	528.7	108.8	3941
Incumbent	0.21	0.41	0.00	11866	0.19	0.39	0.00	7374	0.24	0.43	0.00	3941
Single firm	0.94	0.24	1.00	11816	0.95	0.22	1.00	7346	0.92	0.27	1.00	3919
Single part.	0.24	0.43	0.00	11866	0.13	0.34	0.00	7374	0.46	0.50	0.00	3941
Lowest price	0.74	0.44	1.00	11490	0.86	0.35	1.00	7228	0.51	0.50	1.00	3724
Open auction	0.20	0.40	0.00	11866	0.15	0.36	0.00	7374	0.29	0.45	0.00	3941

Notes: Data are from the ANAC, for all public procurement contracts above 40,000€ offered by municipalities during the period 2013–2017. Statistics refer to the sample selection as described in sections 3 and 4. The data explicitly refer to a close interval of the population size around the threshold (interval is 3,250–9,750 inhabitants). *Bidders* is the number of admitted offers. *Rebate* is the final discount on the reserve price. *Tend. period* is the amount of time, in days, during which offers can be submitted. *Lots* is the number of lots into which the tender is split. *Reserve price* is the starting value set by the contracting authority, in 1,000 €. *Incumbent* indicates whether the winning firm has won at least one other contract with the same mayor. *Single firm* indicates whether the winner has acted as a single firm (rather than as part of a group). *Single part.* indicates that the firm has participated alone. *Lowest price* is the adjudication criterion. *Open auction* indicates that the contractors was chosen via an open auction. I provide statistics for the overall sample and also separately for public works and for provisions and services.

Table 4. Descriptive statistics: Tender characteristics around the threshold

Variables	Mean	s.d.	p50	N	Mean	s.d.	p50	N
	Below 5,000 inhabitants							
	Public works				Provisions and services			
Bidders	8.33	15.89	4.00	2484	2.33	2.38	2.00	1156
Rebate	17.48	13.34	16.80	2379	8.54	13.24	2.24	1008
Tend. period	22.57	31.61	19.00	2484	23.92	28.02	19.00	1156
Lots	1.03	0.31	1.00	2149	1.18	0.96	1.00	964
Reserve price	212.9	405.2	115.0	2484	195.3	379.3	105.2	1156
Incumbent	0.19	0.39	0.00	2484	0.23	0.42	0.00	1156
Single firm	0.96	0.20	1.00	2478	0.93	0.26	1.00	1149
Single part.	0.13	0.34	0.00	2484	0.49	0.50	0.00	1156
Lowest price	0.86	0.35	1.00	2442	0.53	0.50	1.00	1089
Open auction	0.15	0.36	0.00	2484	0.29	0.45	0.00	1156
	Above 5,000 inhabitants							
	Public works				Provisions and services			
Bidders	10.43	23.52	5.00	4895	2.63	3.23	2.00	2789
Rebate	18.98	13.41	18.95	4693	10.42	14.91	3.56	2394
Tend. period	22.71	26.37	19.00	4895	23.34	21.44	19.00	2789
Lots	1.02	0.21	1.00	4283	1.31	1.38	1.00	2338
Reserve price	236.1	446.7	120.3	4895	223.9	580.3	109.7	2789
Incumbent	0.19	0.39	0.00	4895	0.24	0.43	0.00	2789
Single firm	0.95	0.23	1.00	4873	0.92	0.27	1.00	2774
Single part.	0.13	0.34	0.00	4895	0.44	0.50	0.00	2789
Lowest price	0.86	0.35	1.00	4791	0.50	0.50	0.00	2639
Open auction	0.15	0.36	0.00	4895	0.29	0.46	0.00	2789

Notes: Data are from the ANAC, for all public procurement contracts above 40,000€ offered by municipalities during the period 2013–2017. Statistics refer to the sample selection as described in sections 3 and 4. The data explicitly refer to a close interval of the population size around the threshold (interval is 3,250–9,750 inhabitants). *Bidders* is the number of admitted offers. *Rebate* is the final discount on the reserve price. *Tend. period* is the amount of time, in days, during which offers can be submitted. *Lots* is the number of lots into which the tender is split. *Reserve price* is the starting value set by the contracting authority, in 1,000 €. *Incumbent* indicates whether the winning firm has won at least one other contract with the same mayor. *Single firm* indicates whether the winner has acted as a single firm (rather than as part of a group). *Single part.* indicates that the firm has participated alone. *Lowest price* is the adjudication criterion. *Open auction* indicates that the contractors was chosen via an open auction. I provide statistics for public works and for provisions and services, distinguishing between tenders issued by municipalities below and above the 5,000-inhabitant threshold.

Table 5. McCrary discontinuity test

	All	2013	2014	2015	2016	2017
Disc.	-0.116	-0.105	-0.002	-0.024	-0.160	-0.108
s.e.	0.146	0.192	0.158	0.202	0.213	0.243

Notes: The running variable is the difference between the municipality population and 5,000. The rows report the coefficients and standard errors of the McCrary discontinuity test. The columns report values for each sample year.

Table 6. The effect of mayoral wages on aggregate procurement outcomes: Full sample

	(1)	(2)	(3)
	Tenders	Total expenditure	Mean value
		All tenders	
Conv.	0.0609	-99.93	-17.77
	(0.254)	(112.7)	(31.87)
Robust	0.0853	-141.3	-26.64
	(0.306)	(128.4)	(36.54)
Avg.	1.769	370.2	227.6
N	1945	1945	1945
h	793.4	457.2	670

Notes: Data are from the Anti-corruption National Authority (ANAC) for all procurement contracts above 40,000€ offered by municipalities during the period 2013–2017. Sample selection follows the process described in sections 3 and 4. Inferences are made for a close interval of the population size around the threshold (interval is 3,250–9,750 inhabitants). Estimates are made following the procedure described in section 5.3. Column (1) refers to the number of tenders issued by each mayor, weighted by the number of mayor's years in office. Column (2) refers to the amount of money (in euros) spent by each mayor on public procurement, weighted by the number of mayor's years in office. Finally, column (3) indicates the mean value of the tender, by mayor. The table reports point estimates and standard errors following conventional and robust procedures. It also provides the average value of the dependent variable, the number of observations, and the selected bandwidths for point estimation. Standard errors are in parentheses. * $p < .10$, ** $p < .05$, *** $p < .01$.

Table 7. The effect of mayoral wage on procurement outcomes: Full sample

	(1) Bidders	(2) Rebate	(3) Incumbent	(4) Sum Incumbency
	PANEL A: All tenders			
Conv.	3.805**	2.327*	-0.125***	-0.206***
	(1.831)	(1.374)	(0.0377)	(0.0709)
Robust	4.453**	2.232	-0.136***	-0.215**
	(2.120)	(1.674)	(0.0447)	(0.0863)
Avg.	7.182	15.74	0.213	1.345
N	11866	10998	11866	11866
h	891.6	710.8	605.5	547
	PANEL B: Open auctions			
Conv.	17.83***	12.56***	-0.230*	-0.277
	(6.694)	(3.118)	(0.128)	(0.252)
Robust	20.09***	13.75***	-0.256*	-0.341
	(7.549)	(3.474)	(0.149)	(0.291)
Avg.	18.43	15.67	0.150	1.228
N	2355	2145	2355	2355
h	523.6	397.6	416.5	317.3
	PANEL C: Negotiated procedures			
Conv.	0.337	0.345	-0.0914***	-0.0841
	(0.425)	(1.536)	(0.0340)	(0.0638)
Robust	0.514	-0.111	-0.0939**	-0.0670
	(0.474)	(1.826)	(0.0407)	(0.0752)
Avg.	4.397	15.76	0.228	1.374
N	9511	8853	9511	9511
h	432.9	492.4	488.4	333.6

Notes: Data are from the Anti-corruption National Authority (ANAC) for all procurement contracts above 40,000€ offered by municipalities during the period 2013–2017. Sample selection follows the process described in sections 3 and 4. Inferences are made for a close interval of the population size around the threshold (interval is 3,250–9,750 inhabitants). Estimates are made following the procedure described in section 5.3. Column (1) refers to *Bidders*, the number of admitted offers. Column (2) reports *Rebate*, the final discount rate on the reserve price. Column (3), *Incumbent*, shows whether the winning firm has won at least one other contract with the same mayor. Finally, column (4), *Sum Incumbency*, refers to the number of contracts won by a firm during the mayor's term(s). The table reports point estimates and standard errors following conventional and robust procedures. It also provides the average value of the dependent variable, the number of observations, and the selected bandwidths for point estimation. Panel A refers to the full sample. Panel B focuses on open auctions. Panel C focuses on negotiated procedures. Standard errors are in parentheses. * $p < .10$, ** $p < .05$, *** $p < .01$.

Table 8. The effect of mayor's wage on procurement outcomes: public works

	(1) Bidders	(2) Rebate	(3) Incumbent	(4) Sum Incumbency
		PANEL A: All tenders		
Conv.	6.033***	1.788	-0.133***	-0.122*
	(2.119)	(1.479)	(0.0401)	(0.0713)
Robust	6.692***	2.235	-0.146***	-0.139
	(2.414)	(1.726)	(0.0474)	(0.0856)
Avg.	9.726	18.47	0.191	1.280
N	7374	7067	7374	7374
h	672.2	597.7	411.3	373.8
		PANEL B: Open auctions		
Conv.	26.79*	11.83**	-0.161*	-0.164*
	(14.90)	(5.359)	(0.0929)	(0.0986)
Robust	30.58*	13.83**	-0.190*	-0.195*
	(17.29)	(6.081)	(0.108)	(0.116)
Avg.	33.93	20	0.110	1.146
N	1127	1091	1127	1127
h	672.6	392.1	582.2	514
		PANEL C: Negotiated procedures		
Conv.	0.864*	1.031	-0.116***	-0.102
	(0.508)	(1.441)	(0.0434)	(0.0748)
Robust	1.081*	1.406	-0.127**	-0.111
	(0.567)	(1.700)	(0.0517)	(0.0908)
Avg.	5.359	18.19	0.206	1.305
N	6247	5976	6247	6247
h	500.7	750.3	428.3	404.6

Notes: Data are from the Anti-corruption National Authority (ANAC) for all procurement contracts above 40,000€ offered by municipalities during the period 2013–2017. Sample selection follows the process described in sections 3 and 4. Inferences are made for a close interval of the population size around the threshold (interval is 3,250–9,750 inhabitants). Estimates are made following the procedure described in section 5.3. Column (1) refers to *Bidders*, the number of admitted offers. Column (2) reports *Rebate*, the final discount rate on the reserve price. Column (3), *Incumbent*, shows whether the winning firm has won at least one other contract with the same mayor. Finally, column (4), *Sum Incumbency*, refers to the number of contracts won by a firm during the mayor's term(s). The table reports point estimates and standard errors following conventional and robust procedures. It also provides the average value of the dependent variable, the number of observations, and the selected bandwidths for point estimation. Panel A refers to the full sample. Panel B focuses on open auctions. Panel C focuses on negotiated procedures. Standard errors are in parentheses. * $p < .10$, ** $p < .05$, *** $p < .01$.

Table 9. The effect of mayor's wage on procurement outcomes: provisions and services

	(1) Bidders	(2) Rebate	(3) Incumbent	(4) Sum Incumbency
	PANEL A: All tenders			
Conv.	-0.0936	4.101*	-0.116	-0.286**
	(0.256)	(2.228)	(0.0746)	(0.137)
Robust	-0.182	3.745	-0.134	-0.271*
	(0.294)	(2.620)	(0.0882)	(0.164)
Avg.	2.541	9.874	0.237	1.401
N	3941	3398	3941	3941
h	573.8	586.3	462.3	378.1
	PANEL B: Open auctions			
Conv.	0.795	9.300**	-0.152	-0.489
	(0.651)	(3.912)	(0.162)	(0.324)
Robust	0.935	10.39**	-0.178	-0.558
	(0.742)	(4.486)	(0.190)	(0.391)
Avg.	3.198	10.67	0.184	1.297
N	1150	978	1150	1150
h	742.7	569.8	485.3	377.8
	PANEL C: Negotiated procedures			
Conv.	-0.438	-0.428	-0.0793	-0.0911
	(0.268)	(2.498)	(0.0692)	(0.127)
Robust	-0.571*	-1.790	-0.0865	-0.139
	(0.308)	(2.743)	(0.0819)	(0.141)
Avg.	2.270	9.554	0.259	1.444
N	2791	2420	2791	2791
h	526.1	417.9	704.5	363.6

Notes: Data are from the Anti-corruption National Authority (ANAC) for all procurement contracts above 40,000€ offered by municipalities during the period 2013–2017. Sample selection follows the process described in sections 3 and 4. Inferences are made for a close interval of the population size around the threshold (interval is 3,250–9,750 inhabitants). Estimates are made following the procedure described in section 5.3. Column (1) refers to *Bidders*, the number of admitted offers. Column (2) reports *Rebate*, the final discount rate on the reserve price. Column (3), *Incumbent*, shows whether the winning firm has won at least one other contract with the same mayor. Finally, column (4), *Sum Incumbency*, refers to the number of contracts won by a firm during the mayor's term(s). The table reports point estimates and standard errors following conventional and robust procedures. It also provides the average value of the dependent variable, the number of observations, and the selected bandwidths for point estimation. Panel A refers to the full sample. Panel B focuses on open auctions. Panel C focuses on negotiated procedures. Standard errors are in parentheses. * $p < .10$, ** $p < .05$, *** $p < .01$.

Table 10. The effect of mayoral wages on cost overruns

	(1)	(2)	(3)
	Cost overrun	Cost overrun	Cost overrun
		OA	NP
		All tenders	
Conv.	-0.0577	-0.438	-0.0892***
	(0.0408)	(0.522)	(0.0330)
Robust	-0.0764*	-0.530	-0.105***
	(0.0435)	(0.654)	(0.0373)
Avg.	0.0899	0.117	0.0857
N	5492	726	4766
h	577.4	655.3	372.6

Notes: Data are from the Anti-corruption National Authority (ANAC) for all procurement contracts above 40,000€ offered by municipalities during the period 2013–2017. Sample selection follows the process described in sections 3 and 4. Inferences are made for a close interval of the population size around the threshold (interval is 3,250–9,750 inhabitants). Estimates are made following the procedure described in section 5.3. Column (1) refers to cost overruns. Column (2) refers to cost overruns for open-auction tenders only. Column (3) refers to cost overruns for negotiated-procedure tenders only. The table reports point estimates and standard errors following conventional and robust procedures. It also provides the average value of the dependent variable, the number of observations, and the selected bandwidths for point estimation. Standard errors are in parentheses. * $p < .10$, ** $p < .05$, *** $p < .01$.

Table 11. The effect of mayoral wages on time and cost overruns: Public works

	(1)	(2)	(3)	(4)
	Time overrun	Cost overrun	Cost overrun	Cost overrun
			OA	NP
	All tenders			
Conv.	-0.415	-0.0717*	-0.417	-0.112***
	(1.036)	(0.0415)	(0.498)	(0.0380)
Robust	-0.652	-0.0900**	-0.519	-0.128***
	(1.319)	(0.0450)	(0.627)	(0.0429)
Avg.	0.843	0.102	0.136	0.0968
N	2565	4477	555	3922
h	439.5	566.6	814.1	345.3

Notes: Data are from the Anti-corruption National Authority (ANAC) for all public works contracts above 40,000€ offered by municipalities during the period 2013–2017. Sample selection follows the process described in sections 3 and 4. Inferences are made for a close interval of the population size around the threshold (interval is 3,250–9,750 inhabitants). Estimates are made following the procedure described in section 5.3. Column (1) refers to time overruns, with values trimmed below the 1st percentile and above the 99th percentile. Column (2) refers to cost overruns. Column (3) refers to cost overruns for open-auction tenders only. Column (4) refers to cost overruns for negotiated-procedure tenders only. The table reports point estimates and standard errors following conventional and robust procedures. It also provides the average value of the dependent variable, the number of observations, and the selected bandwidths for point estimation. Standard errors are in parentheses. * $p < .10$, ** $p < .05$, *** $p < .01$.

Table 12. Procurement tender design (means)

	(1) Negotiated procedure	(2) Lowest price	(3) Publicity period	(4) Lots
	PANEL A: All tenders			
Conv.	-0.0207	0.0567	1.204	-0.0199**
	(0.0414)	(0.0420)	(2.688)	(0.00897)
Robust	-0.0117	0.0704	1.184	-0.0128
	(0.0497)	(0.0495)	(3.401)	(0.0107)
Avg.	0.802	0.744	22.98	1.091
N	11866	11490	11866	11817
h	916	773.8	1102	555.7
	PANEL B: Open auctions			
Conv.		0.398***	-0.998	-0.103
		(0.122)	(7.610)	(0.120)
Robust		0.440***	-1.742	-0.0862
		(0.137)	(9.556)	(0.139)
Avg.		0.442	32.33	1.233
N		2192	2355	2334
h		797	1201	781.9

Notes: Data are from the Anti-corruption National Authority (ANAC) for all procurement contracts above 40,000€ offered by municipalities during the period 2013–2017. Sample selection follows the process described in sections 3 and 4. Inferences are made for a close interval of the population size around the threshold (interval is 3,250–9,750 inhabitants). Estimates are made following the procedure described in section 5.3. Column (1) refers to negotiated procedures. Column (2) indicates the lowest-price adjudication mechanism. Column (3) is the publicity period for the tender. Column (4) refers to the number of lots. The table reports point estimates and standard errors following conventional and robust procedures. It also provides the average value of the dependent variable, the number of observations, and the selected bandwidths for point estimation. Panel A refers to the full sample. Panel B focuses on open auctions. Standard errors are in parentheses. * $p < .10$, ** $p < .05$, *** $p < .01$.

Table 13. Channels of the mayoral wage effect: Re-election

	(1)	(2)	(3)	(4)
	Bidders	Rebate	Incumbent	Sum Incumbency
	PANEL A: All tenders - First term			
Conv.	3.938*	2.882*	-0.178***	-0.230**
	(2.131)	(1.656)	(0.0427)	(0.0895)
Robust	4.606*	2.831	-0.201***	-0.273***
	(2.496)	(2.024)	(0.0478)	(0.101)
Avg.	7.393	15.85	0.207	1.334
N	8671	8065	8671	8671
h	699.2	511.1	324.3	336.5
	PANEL B: All tenders - Second term			
Conv.	3.390	-0.413	-0.0160	-0.0232
	(2.408)	(1.988)	(0.0665)	(0.0704)
Robust	4.069	0.296	0.00879	-0.0315
	(2.788)	(2.293)	(0.0766)	(0.0812)
Avg.	6.621	15.38	0.227	1.374
N	3133	2874	3133	3133
h	565.9	784	574	596.7

Notes: Data are from the Anti-corruption National Authority (ANAC) for all procurement contracts above 40,000€ offered by municipalities during the period 2013–2017. Sample selection follows the process described in sections 3 and 4. Inferences are made for a close interval of the population size around the threshold (interval is 3,250–9,750 inhabitants). Estimates are made following the procedure described in section 5.3. Column (1) refers to *Bidders*, the number of admitted offers. Column (2) reports *Rebate*, the final discount rate on the reserve price. Column (3), *Incumbent*, shows whether the winning firm has won at least one other contract with the same mayor. Finally, column (4), *Sum Incumbency*, refers to the number of contracts won by a firm during the mayor's term(s). The table reports point estimates and standard errors following conventional and robust procedures. It also provides the average value of the dependent variable, the number of observations, and the selected bandwidths for point estimation. Panel A refers to the full sample - First term. Panel B refers to to the full sample - Second term. Standard errors are in parentheses. * $p < .10$, ** $p < .05$, *** $p < .01$.

Table 14. Channels of the wage effect: Morale

	(1)	(2)	(3)	(4)
	Bidders	Rebate	Incumbent	Sum Incumbency
	PANEL A: All tenders			
Conv.	4.277*	3.703***	-0.107***	-0.164***
	(2.280)	(1.404)	(0.0344)	(0.0609)
Robust	4.820*	4.184**	-0.122***	-0.181**
	(2.733)	(1.628)	(0.0386)	(0.0707)
Avg.	6.780	15.54	0.207	1.332
N	8717	8078	8717	8717
h	910.6	960	998.1	602.8

Notes: Data are from the Anti-corruption National Authority (ANAC) for all procurement contracts above 40,000€ offered by municipalities during the period 2013–2017 (excluding investigated municipalities). Sample selection follows the process described in sections 3 and 4. Inferences are made for a close interval of the population size around the threshold (interval is 3,250–9,750 inhabitants). Estimates are made following the procedure described in section 5.3. Column (1) refers to *Bidders*, the number of admitted offers. Column (2) reports *Rebate*, the final discount rate on the reserve price. Column (3), *Incumbent*, shows whether the winning firm has won at least one other contract with the same mayor. Finally, column (4), *Sum Incumbency*, refers to the number of contracts won by a firm during the mayor's term(s). The table reports point estimates and standard errors following conventional and robust procedures. It also provides the average value of the dependent variable, the number of observations, and the selected bandwidths for point estimation. Standard errors are in parentheses. * $p < .10$, ** $p < .05$, *** $p < .01$.

Table 15. The effect of mayoral wages on procurement outcomes: Inclusion of covariates

	(1)	(2)	(3)	(4)	(5)
	Bidders	Rebate	Incumbent	Sum Inc.	Cost over
	All tenders				
Conv.	3.391**	3.194***	-0.120***	-0.206***	-0.0445
	(1.613)	(0.919)	(0.0316)	(0.0484)	(0.0380)
Robust	3.220*	3.560***	-0.124***	-0.220***	-0.0648
	(1.927)	(1.064)	(0.0373)	(0.0582)	(0.0404)
Municipality controls	Yes	Yes	Yes	Yes	Yes
Region dummies	Yes	Yes	Yes	Yes	Yes
Year dummies	Yes	Yes	Yes	Yes	Yes
Tender controls	Yes	Yes	Yes	Yes	Yes
Mayor controls	Yes	Yes	Yes	Yes	Yes
Avg.	7.182	15.74	0.213	1.345	0.0899
N	11109	10384	11109	11109	5255
h	656.5	643.2	471.4	418.6	506.5

Notes: Data are from the Anti-corruption National Authority (ANAC) for all procurement contracts above 40,000€ offered by municipalities during the period 2013–2017. Sample selection follows the process described in sections 3 and 4. Inferences are made for a close interval of the population size around the threshold (interval is 3,250–9,750 inhabitants). Estimates are made following the procedure described in section 5.3. Column (1) refers to *Bidders*, the number of admitted offers. Column (2) reports *Rebate*, the final discount rate on the reserve price. Column (3), *Incumbent*, shows whether the winning firm has won at least one other contract with the same mayor. Column (4), *Sum Incumbency*, refers to the number of contracts won by a firm during the mayor's term(s). Finally, column (5) refers to Cost overruns as defined in section 5. The table reports point estimates and standard errors following conventional and robust procedures. It also provides the set of covariates included, the average value of the dependent variable, the number of observations, and the selected bandwidths for point estimation. Standard errors are in parentheses. * $p < .10$, ** $p < .05$, *** $p < .01$.

References

ABI. Italy. bank's initiatives in b2b digital marketplace. 2001.

S. Abis. Man vs. machine: Quantitative and discretionary equity management. *Working Paper*, 2017.

D. Acemoglu and F. Zilibotti. Was prometheus unbound by chance? risk, diversification, and growth. *Journal of political economy*, 105(4):709–751, 1997.

African Development Fund. Multinational-project to reform system and means of payment in uemoa countries. *African Development Fund - Project Completion Report*, 2002.

AGCOM. Annual report. 2011.

T. Ahnert, S. Doerr, N. Pierri, and Y. Timmer. Does it help? information technology in banking and entrepreneurship. Technical report, International Monetary Fund, 2021.

G. A. Akerlof. Labor contracts as partial gift exchange. *The quarterly journal of economics*, 97(4):543–569, 1982.

A. Akerman, I. Gaarder, and M. Mogstad. The skill complementarity of broadband internet. *The Quarterly Journal of Economics*, 130(4):1781–1824, 2015.

M. Ali Choudhary and N. Limodio. Liquidity risk and long-term finance: Evidence from a natural experiment. *The Review of Economic Studies*, 2021.

F. Allen, X. Gu, and O. Kowalewski. The interbank market puzzle. *Working Paper*, 2018.

P. Angelini and N. Cetorelli. The effects of regulatory reform on competition in the banking industry. *journal of Money, credit and banking*, pages 663–684, 2003.

J. D. Angrist and G. W. Imbens. Two-stage least squares estimation of average causal effects in models with variable treatment intensity. *Journal of the American statistical Association*, 90(430):431–442, 1995.

J. Atack, F. Bateman, M. Haines, and R. A. Margo. Did railroads induce or follow economic growth?: Urbanization and population growth in the american midwest, 1850–1860. *Social Science History*, 34 (2):171–197, 2010.

A. Augereau and S. Greenstein. The need for speed in emerging communications markets: upgrades to advanced technology at internet service providers. *International Journal of Industrial Organization*, 19 (7):1085–1102, 2001.

P. Bajari and S. Tadelis. Incentives versus transaction costs: A theory of procurement contracts. *Rand journal of Economics*, pages 387–407, 2001.

S. Baldi, A. Bottasso, M. Conti, and C. Piccardo. To bid or not to bid: That is the question: Public procurement, project complexity and corruption. *European Journal of Political Economy*, 43:89–106, 2016.

O. Bandiera, A. Prat, and T. Valletti. Active and passive waste in government spending: evidence from a policy experiment. *American Economic Review*, 99(4):1278–1308, 2009.

A. Banerjee, E. Duflo, and N. Qian. On the road: Access to transportation infrastructure and economic growth in china. Technical report, National Bureau of Economic Research, 2012.

Bank of Italy. The italian financial system: trends and perspectives. 2003.

Bank of Italy. Annual report. 2008.

R. Bartlett, A. Morse, R. Stanton, and N. Wallace. Consumer-lending discrimination in the era of fintech. *Unpublished working paper. University of California, Berkeley*, 2018.

N. Baum-Snow. Did highways cause suburbanization? *The quarterly journal of economics*, 122(2):775–805, 2007.

M. L. Bech and B. Hobijn. Technology diffusion within central banking: the case of real-time gross settlement. *FRB of New York Staff Report*, (260), 2006.

T. Beck and R. Levine. *Handbook of Finance and Development*. Edward Elgar Publishing, 2018.

T. Beck, H. Pamuk, R. Ramrattan, and B. R. Uras. Payment instruments, finance and development. *Journal of Development Economics*, 133:162–186, 2018.

G. S. Becker and G. J. Stigler. Law enforcement, malfeasance, and compensation of enforcers. *The Journal of Legal Studies*, 3(1):1–18, 1974.

V. R. Bencivenga and B. Smith. Financial intermediation and endogenous growth. *Review of Economic Studies*, 58(2):195–209, 1991.

M. Benetton, G. Compiani, and A. Morse. Cryptomining: Local evidence from china and the us. Technical report, Working paper, 2019.

T. Berg, V. Burg, A. Gombović, and M. Puri. On the rise of fintechs–credit scoring using digital footprints. Technical report, National Bureau of Economic Research, 2018.

A. N. Berger. The economic effects of technological progress: evidence from the banking industry. *Journal of Money, credit and Banking*, pages 141–176, 2003.

A. N. Berger and R. DeYoung. The effects of geographic expansion on bank efficiency. *Journal of financial services research*, 19(2):163–184, 2001.

A. N. Berger and G. F. Udell. The institutional memory hypothesis and the procyclicality of bank lending behavior. *Journal of financial intermediation*, 13(4):458–495, 2004.

M. Bertrand, A. Schoar, and D. Thesmar. Banking deregulation and industry structure: Evidence from the french banking reforms of 1985. *The Journal of Finance*, 62(2):597–628, 2007.

T. Besley. Paying politicians: theory and evidence. *Journal of the European Economic Association*, 2(2-3): 193–215, 2004.

S. Bharath, S. Dahiya, A. Saunders, and A. Srinivasan. So what do i get? the bank's view of lending relationships. *Journal of financial Economics*, 85(2):368–419, 2007.

S. Bhattacharya and G. Chiesa. Proprietary information, financial intermediation, and research incentives. *Journal of financial Intermediation*, 4(4):328–357, 1995.

Ç. Bircan and R. De Haas. The limits of lending? banks and technology adoption across russia. *The Review of Financial Studies*, 33(2):536–609, 2020.

N. Bloom, C. Genakos, R. Sadun, and J. Van Reenen. Management practices across firms and countries. *The Academy of Management Perspectives*, 26(1):12–33, 2012.

P. Bolton, T. Santos, and J. A. Scheinkman. Outside and inside liquidity. *The Quarterly Journal of Economics*, 126(1):259–321, 2011.

P. Bolton, Y. Li, N. Wang, and J. Yang. Dynamic banking and the value of deposits. Technical report, National Bureau of Economic Research, 2021.

K. Borusyak and X. Jaravel. Revisiting event study designs. *Available at SSRN 2826228*, 2017.

F. Braggion, A. Manconi, N. Pavanini, and H. Zhu. The value of. *European Corporate Governance Institute–Finance Working Paper*, 719, 2020.

E. Breza and C. Kinnan. Measuring the equilibrium impacts of credit: Evidence from the indian microfinance crisis. *NBER Working Paper No. 24329*, 2018.

E. Breza and A. Liberman. Financial contracting and organizational form: Evidence from the regulation of trade credit. *The Journal of Finance*, 72(1):291–324, 2017.

E. Breza, M. Kanz, and L. Klapper. Workplace signaling and financial commitment: Evidence from a field experiment. *AEA Papers and Proceedings*, 108:438–43, 2018.

J. R. Brown and A. Goolsbee. Does the internet make markets more competitive? evidence from the life insurance industry. *Journal of political economy*, 110(3):481–507, 2002.

G. Buchak, G. Matvos, T. Piskorski, and A. Seru. Fintech, regulatory arbitrage, and the rise of shadow banks. *Journal of Financial Economics*, 130(3):453–483, 2018.

F. J. Buera, J. P. Kaboski, and Y. Shin. Finance and development: A tale of two sectors. *The American Economic Review*, 101(5):1964–2002, 2011.

M. Bugamelli, F. Lotti, M. Amici, E. Ciapanna, F. Colonna, F. D'Amuri, S. Giacomelli, A. Linarello, F. Manaresi, G. Palumbo, et al. Productivity growth in italy: a tale of a slow-motion change. *Bank of Italy Occasional Paper*, (422), 2018.

D. R. Burnett, R. Beckman, and T. M. Davenport. *Submarine Cables: the handbook of Law and Policy*. Martinus Nijhoff Publishers, 2013.

P. Bustos, G. Garber, and J. Ponticelli. Capital accumulation and structural transformation. *The Quarterly Journal of Economics*, 135(2):1037–1094, 2020.

C. W. Calomiris and M. Carlson. Corporate governance and risk management at unprotected banks: National banks in the 1890s. *Journal of Financial Economics*, 119(3):512–532, 2016.

S. Calonico, M. D. Cattaneo, and R. Titiunik. Robust nonparametric confidence intervals for regression-discontinuity designs. *Econometrica*, 82(6):2295–2326, 2014.

S. Calonico, M. D. Cattaneo, M. H. Farrell, and R. Titiunik. rdrobust: Software for regression-discontinuity designs. *The Stata Journal*, 17(2):372–404, 2017.

S. Calonico, M. D. Cattaneo, M. H. Farrell, and R. Titiunik. Regression discontinuity designs using covariates. *Review of Economics and Statistics*, 101(3):442–451, 2019.

G. Calzolari and G. Spagnolo. Relational contracts and competitive screening. 2009.

C. Cambini and L. Sabatino. Digital highways and firm turnover. 2021.

F. Campante, R. Durante, and F. Sobbrio. Politics 2.0: The multifaceted effect of broadband internet on political participation. *Journal of the European Economic Association*, 16(4):1094–1136, 2018.

T. S. Campbell and W. A. Kracaw. Information production, market signalling, and the theory of financial intermediation. *The Journal of Finance*, 35(4):863–882, 1980.

G. Caprio, L. Laeven, and R. Levine. Governance and bank valuation. *Journal of Financial Intermediation*, 16(4):584–617, 2007.

L. Carter. *Submarine Cables and the Oceans: Connecting the World*. Number 31. UNEP/Earthprint, 2010.

F. Caselli and M. Morelli. Bad politicians. *Journal of public economics*, 88(3-4):759–782, 2004.

L. Casolaro and G. Gobbi. Information technology and productivity changes in the banking industry. *Economic Notes*, 36(1):43–76, 2007.

M. D. Cattaneo, N. Idrobo, and R. Titiunik. *A practical introduction to regression discontinuity designs: Foundations*. Cambridge University Press, 2019.

N. Cetorelli and L. S. Goldberg. Liquidity management of us global banks: Internal capital markets in the great recession. *Journal of International Economics*, 88(2):299–311, 2012a.

N. Cetorelli and L. S. Goldberg. Banking globalization and monetary transmission. *The Journal of Finance*, 67(5):1811–1843, 2012b.

N. Cetorelli and P. E. Strahan. Finance as a barrier to entry: Bank competition and industry structure in local us markets. *The Journal of Finance*, 61(1):437–461, 2006.

A. Chandra and E. Thompson. Does public infrastructure affect economic activity?: Evidence from the rural interstate highway system. *Regional science and urban economics*, 30(4):457–490, 2000.

M. A. Choudhary and N. Limodio. Deposit volatility, liquidity and long-term investment: Evidence from a natural experiment in pakistan. *IGIER Working Paper*, 2017.

E. Ciapanna and D. Sabbatini. Broadband in italy: an overview. *Bank of Italy Occasional Paper*, (34), 2008.

R. Ciciretti, I. Hasan, and C. Zazzara. Do internet activities add value? evidence from the traditional banks. *Journal of financial services research*, 35(1):81–98, 2009.

S. Claessens and N. Van Horen. The impact of the global financial crisis on banking globalization. *IMF Economic Review*, 63(4):868–918, 2015.

P. Coen and J. Coen. A structural model of interbank network formation and contagion. 2019.

D. Comin and B. Hobijn. Cross-country technology adoption: making the theories face the facts. *Journal of monetary Economics*, 51(1):39–83, 2004.

T. G. Conley and F. Decarolis. Detecting bidders groups in collusive auctions. *American Economic Journal: Microeconomics*, 8(2):1–38, 2016.

D. Coviello and S. Gagliarducci. Tenure in office and public procurement. *American Economic Journal: Economic Policy*, 9(3):59–105, 2017.

D. Coviello, I. Marino, T. Nannicini, N. Persico, et al. Direct propagation of a fiscal shock: Evidence from italy's stability pact. In *CSEF Working Paper No 484*. 2017.

D. Coviello, A. Guglielmo, and G. Spagnolo. The effect of discretion on procurement performance. *Management Science*, 64(2):715–738, 2018.

B. R. Craig and Y. Ma. Intermediation in the interbank lending market. *Working Paper.*, 2018.

G. S. Crawford, N. Pavanini, and F. Schivardi. Asymmetric information and imperfect competition in lending markets. *American Economic Review*, 108(7):1659–1701, 2018.

N. Crouzet, A. Gupta, and F. Mezzanotti. Shocks and technology adoption: Evidence from electronic payment systems. *Kellogg School of Management Working Paper*, 2019.

N. Czernich, O. Falck, T. Kretschmer, and L. Woessmann. Broadband infrastructure and economic growth. *The Economic Journal*, 121(552):505–532, 2011.

F. D'Acunto, N. Prabhala, and A. G. Rossi. The promises and pitfalls of robo-advising. *The Review of Financial Studies*, 32(5):1983–2020, 2019.

F. D'Acunto, P. Ghosh, R. Jain, and A. G. Rossi. How costly are cultural biases? *Available at SSRN 3736117*, 2020.

A. D'Andrea and N. Limodio. High-speed internet, financial technology and banking in africa. *BAFFI CAREFIN Centre Research Paper*, (2019-124), 2019.

C. De Chaisemartin and X. d'Haultfoeuille. Two-way fixed effects estimators with heterogeneous treatment effects. *American Economic Review*, 110(9):2964–96, 2020.

C. De Roure, L. Pelizzon, and P. Tasca. How does p2p lending fit into the consumer credit market? 2016.

F. Decarolis. Awarding price, contract performance, and bids screening: Evidence from procurement auctions. *American Economic Journal: Applied Economics*, 6(1):108–32, 2014.

F. Decarolis and G. Rovigatti. From mad men to maths men: Concentration and buyer power in online advertising. *American Economic Review*.

F. Decarolis, R. Fisman, P. Pinotti, and S. Vannutelli. Rules, discretion, and corruption in procurement: Evidence from italian government contracting. Technical report, National Bureau of Economic Research, 2020.

H. Degryse and S. Ongena. The impact of technology and regulation on the geographical scope of banking. *Oxford review of economic policy*, 20(4):571–590, 2004.

H. Degryse and S. Ongena. Distance, lending relationships, and competition. *The Journal of Finance*, 60 (1):231–266, 2005.

G. Dell'Ariccia and R. Marquez. Information and bank credit allocation. *Journal of financial Economics*, 72(1):185–214, 2004.

E. Denbee, C. Julliard, Y. Li, and K. Yuan. Network risk and key players: A structural analysis of interbank liquidity. *Journal of Financial Economics*, 2021.

T. DeStefano, R. Kneller, and J. Timmis. Broadband infrastructure, ict use and firm performance: Evidence for uk firms. *Journal of Economic Behavior & Organization*, 155:110–139, 2018.

Detecon. Economic impact of submarine cables disruptions. *Asia-Pacific Economic Cooperation - Policy Support Unit*, 2013.

E. Detragiache, T. Tressel, and P. Gupta. Foreign banks in poor countries: theory and evidence. *The Journal of Finance*, 63(5):2123–2160, 2008.

R. DeYoung. The performance of internet-based business models: Evidence from the banking industry. *The Journal of Business*, 78(3):893–948, 2005.

M. Di Maggio and V. Yao. Fintech borrowers: lax-screening or cream-skimming? Technical report, National Bureau of Economic Research, 2020.

R. Di Tella and E. Schargrodsky. The role of wages and auditing during a crackdown on corruption in the city of buenos aires. *The Journal of Law and Economics*, 46(1):269–292, 2003.

D. W. Diamond and P. H. Dybvig. Bank runs, deposit insurance, and liquidity. *Journal of political economy*, 91(3):401–419, 1983.

J. DiNardo and D. S. Lee. Program evaluation and research designs. In *Handbook of labor economics*, volume 4, pages 463–536. Elsevier, 2011.

S. Djankov, C. McLiesh, and A. Shleifer. Private credit in 129 countries. *Journal of financial Economics*, 84 (2):299–329, 2007.

D. Donaldson. Railroads of the raj: Estimating the impact of transportation infrastructure. *American Economic Review*, 108(4-5):899–934, 2018.

M. Draca, R. Sadun, and J. Van Reenen. Productivity and icts: A review of the evidence. *The Oxford handbook of information and communication technologies*, 2007.

T. Duprey and M. Lé. Bankscope dataset: getting started. *Available at SSRN 2191449*, 2016.

G. Duranton and M. A. Turner. Urban growth and transportation. *Review of Economic Studies*, 79(4): 1407–1440, 2012.

F. D'Acunto, M. Prokopczuk, and M. Weber. Historical antisemitism, ethnic specialization, and financial development. *The Review of Economic Studies*, 86(3):1170–1206, 2019.

B. Eichengreen, R. Lafarguette, and A. Mehl. Cables, sharks and servers: Technology and the geography of the foreign exchange market. Technical report, National Bureau of Economic Research, 2016.

L. Einav, M. Jenkins, and J. Levin. The impact of credit scoring on consumer lending. *The RAND Journal of Economics*, 44(2):249–274, 2013.

R. Felici and M. Pagnini. Distance, bank heterogeneity and entry in local banking markets. *The Journal of Industrial Economics*, 56(3):500–534, 2008.

C. Ferraz and F. Finan. Motivating politicians: The impacts of monetary incentives on quality and performance. Technical report, National Bureau of Economic Research, 2009.

R. Fisman, D. Paravisini, and V. Vig. Cultural proximity and loan outcomes. *American Economic Review*, 107(2):457–92, 2017.

D. Foos, L. Norden, and M. Weber. Loan growth and riskiness of banks. *Journal of Banking & Finance*, 34 (12):2929–2940, 2010.

C. Forman, A. Goldfarb, and S. Greenstein. The internet and local wages: Convergence or divergence? Technical report, National Bureau of Economic Research, 2009.

W. S. Frame, L. D. Wall, and L. J. White. Technological change and financial innovation in banking: Some implications for fintech. 2018.

A. Fuster, M. Plosser, P. Schnabl, and J. Vickery. The role of technology in mortgage lending. *The Review of Financial Studies*, 32(5):1854–1899, 2019.

S. Gagliarducci and T. Nannicini. Do better paid politicians perform better? disentangling incentives from selection. *Journal of the European Economic Association*, 11(2):369–398, 2013.

A. Gavazza, M. Nardotto, and T. Valletti. Internet and politics: Evidence from uk local elections and local government policies. *The Review of Economic Studies*, 86(5):2092–2135, 2019.

A. Gelman and G. Imbens. Why high-order polynomials should not be used in regression discontinuity designs. *Journal of Business & Economic Statistics*, 37(3):447–456, 2019.

G. Gobbi and E. Sette. Do firms benefit from concentrating their borrowing? evidence from the great recession. *Review of Finance*, 18(2):527–560, 2014.

I. Goldstein and A. Pauzner. Demand–deposit contracts and the probability of bank runs. *the Journal of Finance*, 60(3):1293–1327, 2005.

I. Goldstein, W. Jiang, and G. A. Karolyi. To fintech and beyond. *The Review of Financial Studies*, 32(5): 1647–1661, 2019.

V. Grembi, T. Nannicini, and U. Troiano. Do fiscal rules matter? *American Economic Journal: Applied Economics*, pages 1–30, 2016.

R. E. Gropp, S. Corvoisier, et al. Contestability, technology and banking. Technical report, ZEW-Leibniz Centre for European Economic Research, 2009.

P. Grosjean. The institutional legacy of the ottoman empire: Islamic rule and financial development in south eastern europe. *Journal of Comparative Economics*, 39(1):1–16, 2011.

V. Guerrieri and G. Lorenzoni. Liquidity and trading dynamics. *Econometrica*, 77(6):1751–1790, 2009.

S. Guriev, N. Melnikov, and E. Zhuravskaya. 3g internet and confidence in government. *The Quarterly Journal of Economics*, 2021.

J. Hahn, P. Todd, and W. Van der Klaauw. Identification and estimation of treatment effects with a regression-discontinuity design. *Econometrica*, 69(1):201–209, 2001.

T. H. Hannan and J. M. McDowell. The determinants of technology adoption: The case of the banking firm. *The RAND Journal of Economics*, pages 328–335, 1984.

R. Hauswald and R. Marquez. Information technology and financial services competition. *The Review of Financial Studies*, 16(3):921–948, 2003.

Z. He, S. Jiang, D. Xu, and X. Yin. Investing in lending technology: It spending in banking. *University of Chicago, Becker Friedman Institute for Economics Working Paper*, (2021-116), 2021.

F. Heider, M. Hoerova, and C. Holthausen. Liquidity hoarding and interbank market rates: The role of counterparty risk. *Journal of Financial Economics*, 118(2):336–354, 2015.

A. Hertzberg, J. M. Liberti, and D. Paravisini. Information and incentives inside the firm: Evidence from loan officer rotation. *The Journal of Finance*, 65(3):795–828, 2010.

A. Hertzberg, J. M. Liberti, and D. Paravisini. Public information and coordination: evidence from a credit registry expansion. *The Journal of Finance*, 66(2):379–412, 2011.

A. Hertzberg, A. Liberman, and D. Paravisini. Screening on loan terms: evidence from maturity choice in consumer credit. *The Review of Financial Studies*, 31(9):3532–3567, 2018.

S. Higgins. Financial technology adoption. *Working Paper*, 2019.

J. Hjort and J. Poulsen. The arrival of fast internet and employment in africa. *American Economic Review*, 109(3):1032–79, 2019.

G. Imbens and K. Kalyanaraman. Optimal bandwidth choice for the regression discontinuity estimator. *The Review of economic studies*, 79(3):933–959, 2012.

G. W. Imbens and T. Lemieux. Regression discontinuity designs: A guide to practice. *Journal of econometrics*, 142(2):615–635, 2008.

G. W. Imbens and J. M. Wooldridge. Recent developments in the econometrics of program evaluation. *Journal of economic literature*, 47(1):5–86, 2009.

P. Impiglia, F. Laurenti, L. Magnone, and R. Persico. La rete in rame di telecom italia: caratteristiche e potenzialità per lo sviluppo delle tecnologie xds. *Notiziario Tecnico Telecom Italia*, 13(1):74–89, 2004.

V. Ioannidou and S. Ongena. "time for a change": loan conditions and bank behavior when firms switch banks. *The Journal of Finance*, 65(5):1847–1877, 2010.

W. Jack and T. Suri. Risk sharing and transactions costs: Evidence from kenya's mobile money revolution. *American Economic Review*, 104(1):183–223, January 2014.

J. P. Kaboski and R. M. Townsend. The impact of credit on village economies. *American Economic Journal: Applied Economics*, 4(2):98–133, 2012.

K. Kang and R. A. Miller. Winning by default: Why is there so little competition in government procurement? *Unpublished working paper*, 2017.

J. Keil and S. Ongena. It's the end of bank branching as we know it (and we feel fine). *September*, 29: 20–83, 2020.

A. I. Khwaja and A. Mian. Tracing the impact of bank liquidity shocks: Evidence from an emerging market. *American Economic Review*, 98(4):1413–42, 2008.

J. Kolko. Broadband and local growth. *Journal of Urban Economics*, 71(1):100–113, 2012.

R. S. Kroszner and P. E. Strahan. What drives deregulation? economics and politics of the relaxation of bank branching restrictions. *The Quarterly Journal of Economics*, 114(4):1437–1467, 1999.

R. Lalive and A. Schmutzler. Auctions vs negotiations in public procurement: which works better? *Available at SSRN 1919531*, 2011.

D. S. Lee. Randomized experiments from non-random selection in us house elections. *Journal of Econometrics*, 142(2):675–697, 2008.

D. S. Lee and T. Lemieux. Regression discontinuity designs in economics. *Journal of economic literature*, 48(2):281–355, 2010.

H. E. Leland and D. H. Pyle. Informational asymmetries, financial structure, and financial intermediation. *The journal of Finance*, 32(2):371–387, 1977.

R. Levine. Financial development and economic growth: views and agenda. *Journal of economic literature*, 35(2):688–726, 1997.

R. Levine. Finance and growth: theory and evidence. *Handbook of economic growth*, 1:865–934, 2005.

J. M. Liberti and M. A. Petersen. Information: Hard and soft. *Review of Corporate Finance Studies*, 8(1): 1–41, 2019.

J. M. Liberti, A. Seru, and V. Vig. Information, credit, and organization. *Credit, and Organization (June 2016)*, 2016.

N. Limodio and F. Strobbe. Bank deposits and liquidity regulation: Evidence from ethiopia. *Mimeo*, 2017.

C. Lin, C. Ma, Y. Sun, and Y. Xu. The telegraph and modern banking development, 1881–1936. *Journal of Financial Economics*, 141(2):730–749, 2021.

M. Manacorda and A. Tesei. Liberation technology: Mobile phones and political mobilization in africa. *Econometrica*, 88(2):533–567, 2020.

A. M. Manelli and D. R. Vincent. Optimal procurement mechanisms. *Econometrica: Journal of the Econometric Society*, pages 591–620, 1995.

A. Mattozzi and A. Merlo. Political careers or career politicians? *Journal of Public Economics*, 92(3-4): 597–608, 2008.

J. McCrary. Manipulation of the running variable in the regression discontinuity design: A density test. *Journal of econometrics*, 142(2):698–714, 2008.

McKinsey and Company. Roaring to life: Growth and innovation in african retail banking. 2018.

M. Messner and M. K. Polborn. Paying politicians. *Journal of Public Economics*, 88(12):2423–2445, 2004.

G. Michaels. The effect of trade on the demand for skill: Evidence from the interstate highway system. *The Review of Economics and Statistics*, 90(4):683–701, 2008.

F. S. Mishkin and P. E. Strahan. What will technology do to financial structure? *NBER working paper*, (w6892), 1999.

P. E. Mistrulli, L. Antelmo, M. Galardo, I. Garrì, D. Pellegrino, D. Revelli, V. Savino, et al. Why do banks close? the geography of branch pruning. Technical report, Bank of Italy, Economic Research and International Relations Area, 2019.

A. J. Padilla and M. Pagano. Endogenous communication among lenders and entrepreneurial incentives. *The review of financial studies*, 10(1):205–236, 1997.

D. Paravisini, V. Rappoport, P. Schnabl, and D. Wolfenzon. Dissecting the effect of credit supply on trade: Evidence from matched credit-export data. *The Review of Economic Studies*, 82(1):333–359, 2015.

L. Pascali. Banks and development: Jewish communities in the italian renaissance and current economic performance. *Review of Economics and Statistics*, 98(1):140–158, 2016.

L. Pascali. The wind of change: Maritime technology, trade, and economic development. *American Economic Review*, 107(9):2821–54, 2017.

M. A. Petersen and R. G. Rajan. Does distance still matter? the information revolution in small business lending. *The journal of Finance*, 57(6):2533–2570, 2002.

M. N. Pierri and M. Y. Timmer. *Tech in fin before fintech: Blessing or curse for financial stability?* International Monetary Fund, 2020.

T. Rephann and A. Isserman. New highways as economic development tools: An evaluation using quasi-experimental matching methods. *Regional Science and Urban Economics*, 24(6):723–751, 1994.

G. Rodano, N. Serrano-Velarde, and E. Tarantino. Bank financing in times of crisis: Evidence from rdd time series estimates. 2013.

L.-H. Roller and L. Waverman. Telecommunications infrastructure and economic development: A simultaneous approach. *American economic review*, 91(4):909–923, 2001.

G. Saint-Paul. Technological choice, financial markets and economic development. *European Economic Review*, 36(4):763–781, 1992.

C. Schenone. Lending relationships and information rents: Do banks exploit their information advantages? *The Review of Financial Studies*, 23(3):1149–1199, 2010.

S. A. Sharpe. Asymmetric information, bank lending, and implicit contracts: A stylized model of customer relationships. *The journal of finance*, 45(4):1069–1087, 1990.

C. Steinwender. Real effects of information frictions: When the states and the kingdom became united. *American Economic Review*, 108(3):657–96, 2018.

J. E. Stiglitz and A. Weiss. Asymmetric information in credit markets and its implications for macroeconomics. *Oxford Economic Papers*, 44(4):694–724, 1992.

H. Tang. Peer-to-peer lenders versus banks: substitutes or complements? *The Review of Financial Studies*, 32(5):1900–1938, 2019.

R. M. Townsend. Intermediation with costly bilateral exchange. *The Review of Economic Studies*, 45(3): 417–425, 1978.

J. Vesala. *Technological transformation and retail banking competition: Implications and measurement*. 2000.

X. Vives and Z. Ye. Information technology and bank competition. *Available at SSRN 3863988*, 2021.

W. J. Wilhelm, Jr. The internet and financial market structure. *Oxford Review of Economic Policy*, 17(2): 235–247, 2001.

World Bank. Global financial development report 2015/2016 : Long-term finance. *World Bank*, 2015.

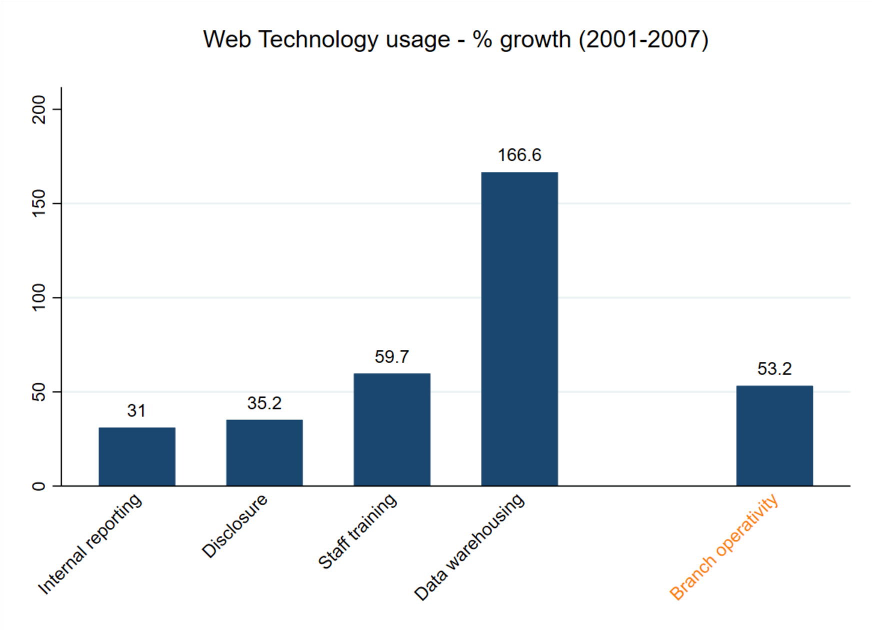
F. Zilibotti. Endogenous Growth and Intermediation in an 'Archipelago' Economy. *Economic Journal*, 104(423):462–473, March 1994.

APPENDICES

Appendix: Broadband and Bank Intermediation

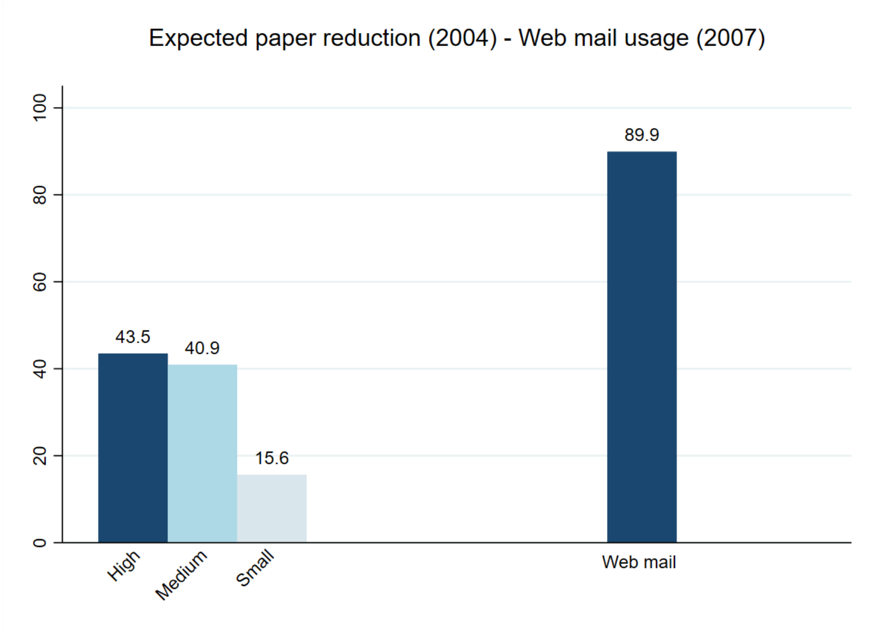
Additional Figures

Figure 5: Usage of web technologies



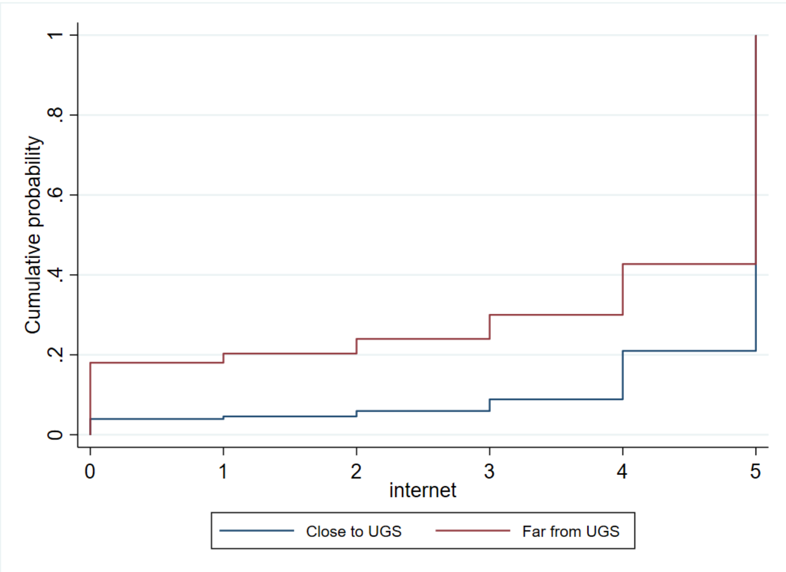
Notes: This figure reports the % growth of the use of web technologies, during the period 2001-2007, for different functions within Italian banks. The source is the Economic Analysis, 2002-2008, from the Italian Banking Association (ABI). Notice that for *Branch operativity*, we do not have the % growth but simply the % usage of web technology in 2007.

Figure 6: Web e-mails



Notes: This figure reports expectations about paper reduction, in 2004, within bank branches. Much of the banks included in the survey expected a high or medium reduction of paper. At the same time and for the same unit of analysis, the figure reports the % usage of web e-mails in 2007. The source is the Economic Analysis, 2002-2008, from the Italian Banking Association (ABI).

Figure 7: Instrument Monotonicity Test



Notes: Angrist and Imbens (1995) instrument's monotonicity test. The assumption of monotonicity of the LATE is not verifiable, but has testable implications on the CDFs of internet for municipalities close or far from the UGS. That is, they should never cross. Here, we plot the CDFs of internet for banks close to the UGS (blue solid line) and far from the UGS (red solid line). In order to separate the two groups, we use a dummy variable below/above the median of UGS distance, that proxy well for our continuous instrument. Values of the CDFs refer to the post broadband period. Since the two CDFs never cross, the instrument passes the test.

Additional Tables

Table B1: Balance Table

	Close	Far	Norm. diff.	N.
Surface	35.31	55.06	(-.24)	4400
Altitude	205.08	336.16	**(-.37)**	4400
North	.69	.55	(.21)	4528
SL per capita	.24	.43	**(-.39)**	4528
Dist. province capital	2.67	3.23	**(-.61)**	4492
Pop. growth	.07	.04	**(.3)**	4361
Adults growth	.05	.02	(.21)	4528
Graduate growth	.88	.8	(.15)	4528
Foreigners growth	2.53	2.83	(-.09)	4524
Buildings growth	.13	.1	(.13)	4528
Firms growth	.1	.04	**(.31)**	4528
Employees growth	.05	.05	(.02)	4528
Income p.c. growth	.19	.17	(.19)	4361

Notes: balance table. This table compares several geographical and socioeconomic indicators, for municipalities that are at different distances from the necessary infrastructure for broadband. We distinguish between municipalities close (below the median of distance) and far (above the median of distance) from the closest UGS. Column 1 reports the average value of each variable for municipalities close to the UGS. Column 2 reports the average value of each variable for municipalities far from the UGS. Column 3 reports the normalized difference as in Imbens and Wooldridge (2009). Values above 0.25 can be considered problematic. Finally, column 4 reports the total number of observations (municipalities).

Table B2: Regressions of Internet on Banks' Number of Loans and Extended Credit with Controls

	(1) Ln (N. loans)	(2) Ln (Ext. credit)	(3) Ln (N. loans)	(4) Ln (Ext. credit)
Internet	0.006**	0.007	0.029*	0.064***
	(0.002)	(0.004)	(0.016)	(0.023)
Controls	X	X	X	X
Bank-Mun FE	X	X	X	X
Year FE	X	X	X	X
Method	OLS	OLS	IV	IV
F-statistic			142.5	141.2
Mean	28.9	30240.062	28.9	30240.062
R-squared	0.902	0.861		
N	123350	122869	123350	122869

Notes: This table reports estimates from OLS and 2SLS as presented in equations (1) and (3). The dataset is at the bank-municipality-year level. The dependent variables are: *Ln(N. loans)*, the natural logarithm of the number of loans issued by bank b in year t; and *Ln(Ext. Credit)*, the natural logarithm of the amount of loans granted by bank b in year t. The main predictor is *Internet*, a measure of ADSL coverage in the municipality, based on a six-point asymmetric scale. Columns 1 and 2 refer to basic OLS estimates. Columns 3 and 4 refer to 2SLS estimates, where the variable *Internet* is instrumented by the interaction between *Distance from UGS* and a dummy variable *post2001*. *Controls* refers to municipality-level variables: the natural logarithm of total population; elderly population, number of private firms, number of employees, distance from the provincial capital, interacted with a second-order polynomial-time trend. *Method* reports the used estimator; *F-statistic* reports the Sanderson-Windmeijer multivariate F-statistic, when the 2SLS methodology is adopted; *Mean* refers to the mean of the dependent variable; *R-squared* is the adjusted R^2; and *N.* refers to the number of observations. Fixed effects are at the bank-municipality and year level. Standard errors, in parentheses, are clustered at the municipality level. ***, ** and * indicate significance at the 1%, 5% and 10% level, respectively.

Table B3: Regressions of Internet on Banks' Number of Loans

	Ln (N. loans)	Ln (N. loans)	Ln (N. loans)	Ln (N. loans)
Internet	0.039**			
	(0.016)			
Years Since Good Internet		0.063**		
		(0.026)		
Good access			0.226**	
			(0.095)	
Some access				0.254**
				(0.107)
Bank-Mun FE	X	X	X	X
Year FE	X	X	X	X
Method	IV	IV	IV	IV
F-statistic	134.9	121.7	117.9	118.3
Mean	28.79	28.79	28.79	28.79
R-squared				
N	124243	124243	124243	124243

Notes: This table reports estimates from 2SLS as presented in equation (3). The dataset is at the bank-municipality-year level. The dependent variable is *Ln(N. loans)*, the natural logarithm of the number of loans issued by bank *b* in year *t*. The main predictors are: *Internet*, a measure of ADSL coverage in the municipality, based on a six-point asymmetric scale; *Years since good internet*, a variable that counts the number of years since the percentage of households with access to the ADSL was above 50%; *Good access*, a dummy variable that takes value 1 if broadband access is above 50%, and zero otherwise; and *Some access*, a dummy variable that takes value 1 if broadband access is above 0%, and zero otherwise. All our predictors are instrumented by the interaction between *Distance from UGS* and a dummy variable *post2001*. *Method* reports the used estimator; *F-statistic* reports the Sanderson-Windmeijer multivariate F-statistic; *Mean* refers to the mean of the dependent variable; *R-squared* is the adjusted R^2; and *N.* refers to the number of observations. Fixed effects are at the bank-municipality and year level. Standard errors, in parentheses, are clustered at the municipality level. ***, ** and * indicate significance at the 1%, 5% and 10% level, respectively.

Table B4: Regressions of Internet Placebo on Banks' Number of Lines and Firm's Extended Credit

	(1) Ln (N. loans)	(2) Ln (Ext. Credit)	(3) Ln (N. loans)	(4) Ln (Ext. Credit)
Internet placebo	0.000	-0.005	0.008	0.001
	(0.003)	(0.005)	(0.009)	(0.015)
Bank-Mun FE	X	X	X	X
Year FE	X	X	X	X
Method	OLS	OLS	IV	IV
F-statistic			100.4	99.6
Mean	26.12	24395.638	26.12	24395.638
R-squared	0.932	0.906		
N	72277	71905	72277	71905

Notes: This table reports estimates from OLS and 2SLS as presented in equations (1) and (3). The dataset is at the bank-municipality-year level. The dependent variables are: *Ln(N. loans)*, the natural logarithm of the number of loans issued by bank *b* in year *t*; and *Ln(Ext. Credit)*, the natural logarithm of the amount of loans granted by bank *b* in year *t*. The main predictor is *Internet placebo*, a fake measure of ADSL coverage in the municipality, based on a six-point asymmetric scale. The sample includes years from 1998 to 2003, where we assign ADSL data of 2006 to years from 2001 to 2003. Columns 1 and 2 refer to basic OLS estimates. Columns 3 and 4 refer to 2SLS estimates, where the variable *Internet* is instrumented by the interaction between *Distance from UGS* and a dummy variable *post2001*. *Method* reports the used estimator; *F-statistic* reports the Sanderson-Windmeijer multivariate F-statistic, when the 2SLS methodology is adopted; *Mean* refers to the mean of the dependent variable; *R-squared* is the adjusted R^2; and *N.* refers to the number of observations. Fixed effects are at the bank-municipality and year level. Standard errors, in parentheses, are clustered at the municipality level. ***, ** and * indicate significance at the 1%, 5% and 10% level, respectively.

Table B5: Regressions of Internet on the Share of Connected Firms out of the municipality

	(1) Share (Connected firms)	(2) Share (Connected firms)
Internet	0.021***	0.116***
	(0.002)	(0.010)
Bank-Mun FE	X	X
Year FE	X	X
Method	OLS	IV
F-statistic		132.2
Mean	0.31	0.31
R-squared	0.515	
N	92654	92654

Notes: This table reports estimates from OLS and 2SLS as presented in equations (1) and (3). The dataset is at the bank-municipality-year level, and focuses on new loans. The dependent variables is the *Share(Connected firms)*, the share of the loans originated with firms connected to broadband. The main predictor is *Internet*, a measure of ADSL coverage in the municipality, based on a six-point asymmetric scale. Column 1 refers to basic OLS estimates. Column 2 refers to 2SLS estimates, where the variable *Internet* is instrumented by the interaction between *Distance from UGS* and a dummy variable *post2001*. *Method* reports the used estimator; *F-statistic* reports the Sanderson-Windmeijer multivariate F-statistic, when the 2SLS methodology is adopted; *Mean* refers to the mean of the dependent variable; *R-squared* is the adjusted R^2; and *N.* refers to the number of observations. Fixed effects are at the bank-municipality and year level. Standard errors, in parentheses, are clustered at the municipality level. ***, ** and * indicate significance at the 1%, 5% and 10% level, respectively.

Table B6: Regressions of Internet on Loan Geography

	(1) Dummy (Diff. Province)	(2) Asinh (Distance)	(3) Dummy (Diff. Province)	(4) Asinh (Distance)
Internet	0.001	0.003	0.033***	0.074**
	(0.002)	(0.007)	(0.011)	(0.037)
Bank-Year FE	X	X	X	X
Bank-Mun FE	X	X	X	X
Firm FE	X	X	X	X
Method	OLS	OLS	IV	IV
F-statistic			191.5	215.4
Mean	0.42	77.35	0.42	77.35
R-squared	0.672	0.756		
N	633732	567594	633732	567594

Notes: This table reports estimates from OLS and 2SLS as presented in equations (1) and (3). The dataset is at the firm-bank-municipality-year level, and focuses on new loans. The dependent variables are: *Dummy(Diff. Province)*, a dummy variable for the loan being originated outside the province of the bank; and *Asinh(Distance)*, the inverse hyperbolic sine of the geodesic distance between the centroid of the municipality of the bank and the location of the firm. The main predictor is *Internet*, a measure of ADSL coverage in the municipality, based on a six-point asymmetric scale. Columns 1 and 2 refer to basic OLS estimates. Columns 3 and 4 refer to 2SLS estimates, where the variable *Internet* is instrumented by the interaction between *Distance from UGS* and a dummy variable *post2001*. *Method* reports the used estimator; *F-statistic* reports the Sanderson-Windmeijer multivariate F-statistic, when the 2SLS methodology is adopted; *Mean* refers to the mean of the dependent variable; *R-squared* is the adjusted R^2; and *N.* refers to the number of observations. Fixed effects are at the bank-municipality, firm and bank-year level. Standard errors, in parentheses, are clustered at the municipality level. ***, ** and * indicate significance at the 1%, 5% and 10% level, respectively.

Table B7: Regressions of Internet on Average Interest Rates - Small Municipalities

	(1) Average Rate	(2) Average Rate
Internet	0.009	0.309**
	(0.012)	(0.139)
Bank-Mun FE	X	X
Year FE	X	X
Method	OLS	IV
F-statistic		43.0
Mean	6.25	6.25
R-squared	0.546	
N	16637	16637

Notes: This table reports estimates from OLS and 2SLS as presented in equations (1) and (3). The dataset is at the bank-municipality-year level, and includes information on small municipalities (below the median of population) only. The dependent variable is *Average Rate*, the (weighted) average interest rate on loans issued by bank *b* in year *t*. The main predictor is *Internet*, a measure of ADSL coverage in the municipality, based on a six-point asymmetric scale. Column 1 refers to the basic OLS estimate. Column 2 and refers to the 2SLS estimate, where the variable *Internet* is instrumented by the interaction between *Distance from UGS* and a dummy variable *post2001*. *Method* reports the used estimator; *F-statistic* reports the Sanderson-Windmeijer multivariate F-statistic, when the 2SLS methodology is adopted; *Mean* refers to the mean of the dependent variable; *R-squared* is the adjusted R^2; and *N.* refers to the number of observations. Fixed effects are at the bank-municipality and year level. Standard errors, in parentheses, are clustered at the municipality level. ***, ** and * indicate significance at the 1%, 5% and 10% level, respectively.

Table B8: Regressions of Internet on Banks' Internal efficiency (productivity and quality) - Small Municipalities

	(1)	(2)	(3)	(4)
	Ln	Asinh	Ln	Asinh
	(Ext./Empl.)	(NPLs/N. loans)	(Ext./Empl.)	(NPLs/N. loans)
Internet	-0.014**	0.000	-0.165***	-0.002
	(0.006)	(0.000)	(0.055)	(0.003)
Bank-Mun FE	X	X	X	X
Year FE	X	X	X	X
Method	OLS	OLS	IV	IV
F-statistic			81.6	90.6
Mean	750.624	0.01	750.624	0.01
R-squared	0.720	0.337		
N	29465	31045	29465	31045

Notes: This table reports estimates from OLS and 2SLS as presented in equations (1) and (3). The dataset is at the bank-municipality-year level, and includes information on small municipalities (below the median of population) only. The dependent variables are: *Ln(Ext./Empl.)*, the natural logarithm of the amount of credit issued by bank employee; and *Asinh(NPLs/N. loans)*, the inverse hyperbolic sine of the share of non performing loans on total loans. The main predictor is *Internet*, a measure of ADSL coverage in the municipality, based on a six-point asymmetric scale. Columns 1 and 2 refer to basic OLS estimates. Columns 3 and 4 refer to 2SLS estimates, where the variable *Internet* is instrumented by the interaction between *Distance from UGS* and a dummy variable *post2001*. *Method* reports the used estimator; *F-statistic* reports the Sanderson-Windmeijer multivariate F-statistic, when the 2SLS methodology is adopted; *Mean* refers to the mean of the dependent variable; *R-squared* is the adjusted R^2; and *N.* refers to the number of observations. Fixed effects are at the bank-municipality and year level. Standard errors, in parentheses, are clustered at the municipality level. ***, ** and * indicate significance at the 1%, 5% and 10% level, respectively.

Table B9: Regressions of Internet on Bank Competitors in the municipality - Small Municipalities

	(1)	(2)
	Ln (Competitors)	Ln (Competitors)
Internet	0.002**	-0.007
	(0.001)	(0.010)
Mun FE	X	X
Year FE	X	X
Method	OLS	IV
F statistic		125.4
Mean	2.03	2.03
R-squared	0.908	
N	20839	20839

Notes: This table reports estimates from OLS and 2SLS as presented in equations (1) and (3). The dataset is at the municipality-year level, and includes information on small municipalities (below the median of population) only. The dependent variable is *Ln(competitors)*, the natural logarithm of the number of bank (physical) competitors in municipality m, in year t. The main predictor is *Internet*, a measure of ADSL coverage in the municipality, based on a six-point asymmetric scale. Column 1 refers to basic OLS estimates. Column 2 refers to 2SLS estimates, where the variable *Internet* is instrumented by the interaction between *Distance from UGS* and a dummy variable *post2001*. *Method* reports the used estimator; *F-statistic* reports the Sanderson-Windmeijer multivariate F-statistic, when the 2SLS methodology is adopted; *Mean* refers to the mean of the dependent variable; *R-squared* is the adjusted R^2; and *N.* refers to the number of observations. Fixed effects are at the municipality and year level. Standard errors, in parentheses, are clustered at the municipality level. ***, ** and * indicate significance at the 1%, 5% and 10% level, respectively.

Table B10: Regressions of Internet on Competition (Deposits) - Small Municipalities

	(1) HHI	(2) Top 3 Share	(3) Top 5 Share	(4) HHI	(5) Top 3 Share	(6) Top 5 Share
Internet	-0.003***	-0.002***	-0.002***	-0.032***	-0.023***	-0.023***
	(0.001)	(0.000)	(0.000)	(0.005)	(0.003)	(0.003)
Mun FE	X	X	X	X	X	X
Year FE	X	X	X	X	X	X
Method	OLS	OLS	OLS	IV	IV	IV
F-statistic				215.0	215.0	215.0
Mean	0.84	0.99	0.99	0.84	0.99	0.99
R-squared	0.824	0.192	0.180			
N	28161	28161	28161	28161	28161	28161

Notes: This table reports estimates from OLS and 2SLS as presented in equations (1) and (3). The dataset is at the municipality-year level, and includes information on small municipalities (below the median of population) only. The dependent variables are: *HHI*, the Herfindahl–Hirschman Index of bank deposits in municipality m and year t; *Top 3 share*, the share of deposits owned by top 3 banks in the municipality; and *Top 5 share*, the share of deposits owned by top 5 banks in the municipality. The main predictor is *Internet*, a measure of ADSL coverage in the municipality, based on a six-point asymmetric scale. Columns 1 to 3 refer to basic OLS estimates. Columns 4 to 6 refer to 2SLS estimates, where the variable *Internet* is instrumented by the interaction between *Distance from UGS* and a dummy variable *post2001*. *Method* reports the used estimator; *F-statistic* reports the Sanderson-Windmeijer multivariate F-statistic, when the 2SLS methodology is adopted; *Mean* refers to the mean of the dependent variable; *R-squared* is the adjusted R^2; and *N.* refers to the number of observations. Fixed effects are at the municipality and year level. Standard errors, in parentheses, are clustered at the municipality level. ***, ** and * indicate significance at the 1%, 5% and 10% level, respectively.

Table B11: Regressions of Internet on Loan Geography - Small Municipalities

	(1) Dummy (Diff. Province)	(2) Asinh (Distance)	(3) Dummy (Diff. Province)	(4) Asinh (Distance)
Internet	0.004**	-0.008	0.019	0.039
	(0.002)	(0.009)	(0.021)	(0.068)
Bank-Year FE	X	X	X	X
Bank-Mun FE	X	X	X	X
Firm FE	X	X	X	X
Method	OLS	OLS	IV	IV
F-statistic			4.7	4.0
Mean	0.15	14.61	0.15	14.61
R-squared	0.948	0.955		
N	9572	8120	9572	8120

Notes: This table reports estimates from OLS and 2SLS as presented in equations (1) and (3). The dataset is at the firm-bank-municipality-year level. It focuses on new loans, and includes information on small municipalities (below the median of population) only. The dependent variables are: *Dummy(Diff. Province)*, a dummy variable for the loan being originated outside the province of the bank; and *Asinh(Distance)*, the inverse hyperbolic sine of the geodesic distance between the centroid of the municipality of the bank and the location of the firm. The main predictor is *Internet*, a measure of ADSL coverage in the municipality, based on a six-point asymmetric scale. Columns 1 and 2 refer to basic OLS estimates. Columns 3 and 4 refer to 2SLS estimates, where the variable *Internet* is instrumented by the interaction between *Distance from UGS* and a dummy variable *post2001*. *Method* reports the used estimator; *F-statistic* reports the Sanderson-Windmeijer multivariate F-statistic, when the 2SLS methodology is adopted; *Mean* refers to the mean of the dependent variable; *R-squared* is the adjusted R^2; and *N.* refers to the number of observations. Fixed effects are at the bank-municipality, firm and bank-year level. Standard errors, in parentheses, are clustered at the municipality level. ***, ** and * indicate significance at the 1%, 5% and 10% level, respectively.

Table B12: Regressions of Internet on Existing Relationships - Small Municipalities

	(1)	(2)	(3)	(4)
	Dummy (Multiple)	Dummy (Multiple Bank)	Dummy (Multiple)	Dummy (Multiple Bank)
Internet	0.011*	0.002	-0.014	-0.059
	(0.006)	(0.005)	(0.057)	(0.047)
Bank-Year FE	X	X	X	X
Bank-Mun FE	X	X	X	X
Firm FE	X	X	X	X
Method	OLS	OLS	IV	IV
F-statistic			4.7	4.7
Mean	0.86	0.1	0.86	0.1
R-squared	0.836	0.793		
N	9572	9572	9572	9572

Notes: This table reports estimates from OLS and 2SLS as presented in equations (1) and (3). The dataset is at the firm-bank-municipality-year level. It focuses on new loans, and includes information on small municipalities (below the median of population) only. The dependent variables are: *Dummy(Multiple)*, a dummy variable for the loan issued to a firm already having a credit relationship; and *Dummy(Multiple Bank)*, a dummy variable for the loan issued to a firm already having a credit relationship with the same bank (in a different place). The main predictor is *Internet*, a measure of ADSL coverage in the municipality, based on a six-point asymmetric scale. Columns 1 and 2 refer to basic OLS estimates. Columns 3 and 4 refer to 2SLS estimates, where the variable *Internet* is instrumented by the interaction between *Distance from UGS* and a dummy variable *post2001*. *Method* reports the used estimator; *F-statistic* reports the Sanderson-Windmeijer multivariate F-statistic, when the 2SLS methodology is adopted; *Mean* refers to the mean of the dependent variable; *R-squared* is the adjusted R^2; and *N.* refers to the number of observations. Fixed effects are at the bank-municipality, firm and bank-year level. Standard errors, in parentheses, are clustered at the municipality level. ***, ** and * indicate significance at the 1%, 5% and 10% level, respectively.

Appendix: High-speed Internet, Financial Technology and Banking

Appendix A

Table A1: African Countries and First Fiber-optic Submarine Cable

Country	Location	Cable	RTS
Algeria	Coast	ALPAL2	jul 2002
Angola	Coast	SAT3/WASC	apr 2002
Benin	Coast	SAT3/WASC	apr 2002
Cameroon	Coast	SAT3/WASC	apr 2002
Cape Verde	Coast	ATLANTIS	feb 2000
Comoros	Coast	EASSy	jul 2010
Congo	Coast	WACS	may 2012
Cote D'Ivoire	Coast	SAT3/WASC	apr 2002
Democratic Republic Of Congo	Coast	WACS	may 2012
Djibouti	Coast	Aden-Djibouti	1994
Egypt	Coast	Aletar	apr 1997
Equatorial Guinea	Coast	ACE	dec 2012
Gabon	Coast	SAT3/WASC	apr 2002
Gambia	Coast	ACE	dec 2012
Ghana	Coast	SAT3/WASC	apr 2002
Guinea	Coast	ACE	dec 2012
Kenya	Coast	TEAMS	jul 2009
Liberia	Coast	ACE	dec 2012
Libya	Coast	ITALY-LIBYA	1998
Madagascar	Coast	EASSy	nov 2009
Mauritania	Coast	ACE	dec 2012
Mauritius	Coast	SAFE	apr 2002
Morocco	Coast	Estepona-Tetouan	jul 1994
Mozambique	Coast	SEACOM	jul 2009
Namibia	Coast	WACS	may 2012
Nigeria	Coast	SAT3/WASC	apr 2002
Sao Tome and Principe	Coast	ACE	dec 2012
Senegal	Coast	ATLANTIS	feb 2000
Seychelles	Coast	SEAS	aug 2012
Sierra Leone	Coast	ACE	dec 2012
South Africa	Coast	SAT3/WASC	apr 2002
Sudan	Coast	SAS-1	apr 2003
Togo	Coast	WACS	may 2012
Tunisia	Coast	Trapani-Kelibia	nov 1995
United Republic Of Tanzania	Coast	SEACOM	jul 2009

Appendix B

In table B1 we show the dependent and independent variables, with sources and description.

Table B1: Data Sources and Description

Variables	Source	Description
		Banking variables
Loans	Bankscope	Net loans,
	and BankFocus	in millions of US$.
Government	Bankscope	Government securities,
	and BankFocus	in millions of US$.
Deposits	Bankscope	Deposits and short-term funding,
	and BankFocus	in millions of US$.
Equity/dep. & ST	Bankscope	Share of total equity
	and BankFocus	over deposits and short-term funding.
RTGS bank	Central Banks websites	Whether the bank adopts
	and bank's reports	the RTGS system.
Loans to Banks	Bankscope	Loans and advances to banks,
	and BankFocus	in millions of US$.
Deposits from Banks	Bankscope	Deposits from banks,
	and BankFocus	in millions of US$.
Liquid Assets/dep. & ST	Bankscope	Share of liquid assets
	and BankFocus	over deposits and short-term funding.
		Firms variables
Access to Finance	World Bank Enterprise Survey	Whether access to finance
		is a minor obstacle.
Loans from Banks	World Bank Enterprise Survey	Whether the activity
		is financed by loans from banks.
Loans Maturity	World Bank Enterprise Survey	The maturity of loans,
		in months.
Sales	World Bank Enterprise Survey	Total annual sales,
		in millions of US$.
Workforce	World Bank Enterprise Survey	Number of permanent and temporary
		full-time employees.
		Independent variables
Submarine	Telegeography maps	Whether the country
		is reached by a fiber-optic cable.
Submarine HQ	Telegeography maps,	Whether the country of the headquarter
	CvH database, and bank's reports	is reached by a fiber-optic cable.

Appendix C

In table C1 we provide empirical evidence to support the exogeneity of the timing of connection. Here, we propose cross-country regressions where the year of arrival of the submarine cable is regressed on three indicators that proxy for the economic outlook of the country and its banking sector before the arrival of the cable. *GDP growth* indicates the growth of GDP at constant US$; *Private credit to GDP* refers to the amount of private credit in the economy as a share of GDP; and *Central bank assets to GDP* refers to the amount of assets held by the central bank as a share of GDP. All the data come from the WB GFD database.

Table C2 complements table C1. It shows panel regressions where submarine connectedness is regressed over the current values of the aforementioned economic and banking indicators, and their 10–years average.

Table C1: Submarine Cables Arrival - Cross Sectional Regressions

Variables	(1) Submarine Year	(2) Submarine Year	(3) Submarine Year
GDP growth (pre)	-5.720		
	(17.84)		
Private credit to GDP (pre)		-0.099	
		(0.060)	
Central bank assets to GDP (pre)			-0.087
			(0.118)
Obs.	34	31	31
Adj. R sq.	-0.028	0.049	-0.018
MDV	2006	2006	2005
SDDV	6.273	6.290	6.246

Notes: This table reports estimates from the regression of the year of arrival of the submarine cable on different economic proxies (computed before the arrival of the cable). GDP growth indicates the growth of GDP at constant US$; Private credit to GDP refers to the amount of private credit in the economy as a share of GDP; and Central bank assets to GDP refers to the amount of assets held by the central bank as a share of GDP. Obs. refers to the number of observations; Adj.R^2 is the adjusted R^2; M.D.V. refers to the mean of the dependent variable. Fixed effects are at the country and year level. Standard errors in parentheses, clustered at country level. ***, ** and * indicate significance at the 1%, 5% and 10% level, respectively.

Table C2: Submarine Cables Arrival - Panel Regressions

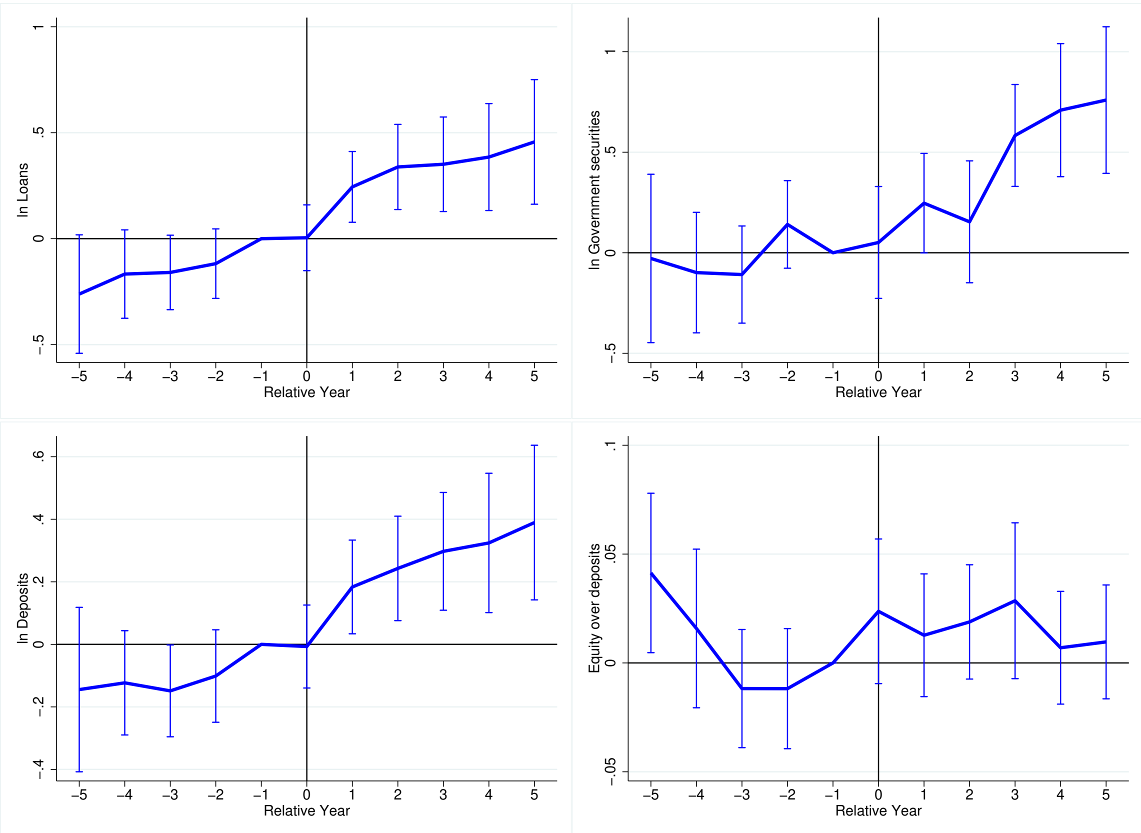
	(1)	(2)	(3)
Variables	Submarine	Submarine	Submarine
GDP growth	0.002		
	(0.046)		
GDP growth (avg10)	-0.458		
	(0.380)		
Private credit to GDP		-0.001	
		(0.001)	
Private credit to GDP (avg10)		0.001	
		(0.002)	
Central bank assets to GDP			-0.002
			(0.003)
Central bank assets to GDP (avg10)			-0.011*
			(0.006)
Country FE	Yes	Yes	Yes
Year FE	Yes	Yes	Yes
Obs.	1392	1262	1103
Adj. R sq.	0.701	0.719	0.730
MDV	0.304	0.316	0.347
SDDV	0.460	0.465	0.476

Notes: This table reports estimates from the regression of the dummy Submarine on different economic proxies and their averages. GDP growth indicates the growth of GDP at constant US$; Private credit to GDP refers to the amount of private credit in the economy as a share of GDP; and Central bank assets to GDP refers to the amount of assets held by the central bank as a share of GDP. All these variables are also considered with their 10-years average. Obs. refers to the number of observations; Adj.R^2 is the adjusted R^2; M.D.V. refers to the mean of the dependent variable. Fixed effects are at the country and year level. Standard errors in parentheses, clustered at country level. ***, ** and * indicate significance at the 1%, 5% and 10% level, respectively.

Appendix D

In figure D1 and table D1, we perform the event study and the diff-in-diff regression when connection of the headquarter, $Submarine_{gt}$, is used as the main predictor. Our outcome variables are the same as in the main analysis: *Loans*, the natural logarithm of net loans, that proxies for bank credit to the private sector; *Government*, the natural logarithm of government securities, that quantifies the investments of banks in government bonds; *Deposits*, the natural logarithm of bank deposits and short-term funding; and *Equity/Deposits and short-term funding*, that measures the amount of shareholders' equity over bank deposits.

Figure D1: An Event Study on High-Speed Internet and Banking - Headquarter



Notes: This figure shows the event study for the period 5 years before–5 years after the arrival of the first submarine cable in the country of the headquarter, g. The y axis reports coefficients for the dependent variables: ln Loans (the natural logarithm of net loans (in million of US dollars)); ln Government securities (the natural logarithm of government securities (in million of US dollars)); ln Deposits (the natural logarithm of deposits and short-term funding (in million of US dollars)); and Equity over DST (the share of total equity over deposits and short-term funding). The x axis refers to the relative time from the arrival of the cable. The blue (solid) line connects point estimates relative to the base year (-1). 95% confidence intervals are also reported, and standard errors are clustered at country level.

Table D1: High-Speed Internet and Banking - Headquarter

	(1)	(2)	(3)	(4)
Variables	Loans	Government	Deposits	Equity
	ln	ln	ln	over DST
$Submarine_{gt}$	0.442***	0.388***	0.339***	0.006
	(0.080)	(0.120)	(0.080)	(0.012)
Bank FE	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes
Obs.	6481	3600	6535	6397
Adj. R sq.	0.865	0.790	0.883	0.577
MDV	5.039	3.930	5.653	0.194
SDDV	2.061	2.319	1.978	0.222

Notes: This table reports estimates from the staggered diff-in-diff design presented in equation (2). The dependent variables are as follows: Loans (natural logarithm of net loans (in million of US dollars)); Government (natural logarithm of government securities (in million of US dollars)); Deposits (natural logarithm of deposits and short-term funding (in million of US dollars)); and Equity (share of total equity over deposits and short-term funding). The main predictor is $Submarine_{gt}$, a binary variable for the arrival of the first fiber-optic submarine cable in the country of the headquarter. This dummy takes value zero before the arrival of the cable and 1 from the time of the arrival on. Obs. refers to the number of observations; Adj.R^2 is the adjusted R^2; MDV refers to the mean of the dependent variable; and SDDV refers to the standard deviation of the dependent variable. Fixed effects are at the bank and year level. Standard errors in parentheses, clustered at parent level. ***, ** and * indicate significance at the 1%, 5% and 10% level, respectively.

Appendix E

Table E1 is the balance table for the variable *Weak user*, as defined in section 3.2.1. The first two rows refer to *loans to banks* and *deposits from banks,* that we use to generate this variable. The other rows refer to: the log of total assets; the log of total deposits; assets over deposits; loans over assets; government securities over assets; equity over assets; and liquid assets over total assets. The last column provides the difference between the group of *Weak users* and that of *non-Weak users,* with the respective levels of significance.

Figure E1 presents a scatter plot where weak users and non-weak users are compared on the basis of their log loans/ log deposits composition.

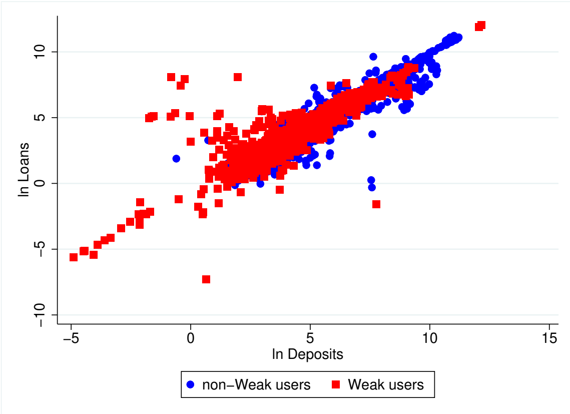
Table E2 replicates the analysis in table 2 on the restricted sample of banks for which the variable *Weak user* is non-missing.

Table E1: Balance Table - Weak users

	(1)	(2)	(3)
	non-Weak user	Weak user	difference
log (LtB)	3.960	1.823	-2.174***
	(1.243)	(1.501)	(0.229)
log (DfB)	2.439	0.788	-1.621***
	(1.905)	(1.844)	(0.275)
log (Ass)	5.675	3.924	-1.767***
	(1.312)	(1.325)	(0.173)
log (DST)	5.406	3.470	-1.946***
	(1.376)	(1.504)	(0.227)
Ass/dep	1.698	1.937	0.207
	(5.051)	(3.634)	(0.444)
Loans/ass	0.436	0.432	-0.003
	(0.194)	(0.191)	(0.015)
Gov/ass	0.151	0.140	-0.002
	(0.139)	(0.121)	(0.018)
Equ/ass	0.115	0.199	0.086***
	(0.115)	(0.194)	(0.016)
Liq/ass	0.393	0.370	-0.026
	(0.194)	(0.191)	(0.020)
Obs.	171	153	324

Notes: This table is the balance table for the variable Weak user. The reduced number of banks in this sample stems from the definition of Weak user: banks with no data before the cable arrival are automatically excluded. log (LtB) is the natural logarithm of loans to banks (in millions US$); log (DfB) is the natural logarithm of deposits from banks (in millions US$); log (Ass) is the natural logarithm of total assets (in millions US$); log (DST) is the natural logarithm of deposits and short-term funding (in millions US$); Ass/dep is the ratio between assets and deposits, excluded loans to banks and deposits from banks; Loans/ass is the ratio between net loans and total assets; Gov/ass is the ratio between government securities and total assets; Equ/ass is the ratio between equity and total assets; finally, Liq/ass is the share of liquid assets over total assets. Column 1 refers to non-Weak user. Column 2 refers to Weak user. Finally, Column 3 reports their difference. ***, ** and * indicate significance at the 1%, 5% and 10% level, respectively.

Figure E1: Scatter Plot - Weak vs non-Weak users



Notes: This plot shows the correlation between ln Loans (the natural log of net loans (in mil. of US dollars)) and ln Deposits (the natural log of deposits and s-t funds (in mil. of US dollars)), for Weak users (in square red) and non-Weak users (in circle blue).

Table E2: High-Speed Internet and Banking - Restricted Sample

Variables	(1) Loans ln	(2) Government ln	(3) Deposits ln	(4) Equity over DST
$Submarine_{ct}$	0.329***	0.126	0.270***	-0.008
	(0.097)	(0.141)	(0.079)	(0.017)
Bank FE	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes
Obs.	4109	2175	4137	4056
Adj. R sq.	0.840	0.709	0.846	0.532
MDV	4.766	3.430	5.318	0.197
SDDV	1.914	2.123	1.855	0.214

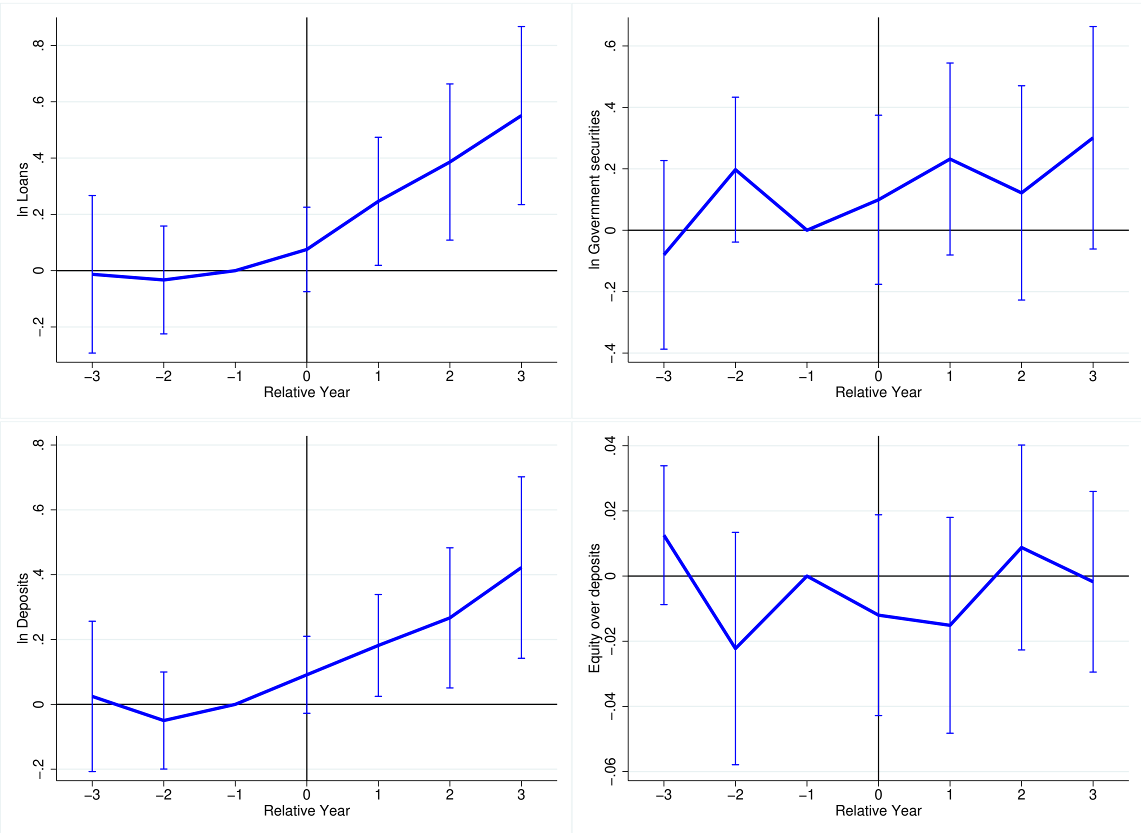
Notes: This table reports estimates from the staggered diff-in-diff design presented in equation (2). Differently from table 2, here the analysis focuses on the restricted sample of banks for which Weak user is non missing. The dependent variables are as follows: Loans (natural log of net loans (in mil. of US dollars)); Government (natural log of government securities (in mil. of US dollars)); Deposits (natural log of deposits and short-term funding (in mil. of US dollars)); and Equity (share of total equity over deposits and short-term funding). The main predictor is $Submarine_{ct}$, a binary variable for the arrival of the first fiber-optic submarine cable in the country. This dummy takes value zero before the arrival of the cable and 1 from the time of the arrival on. Obs. refers to the number of observations; Adj.R^2 is the adjusted R^2; MDV refers to the mean of the dependent variable; and SDDV refers to the standard deviation of the dependent variable. Fixed effects are at the bank and year level. Standard errors in parentheses, clustered at country level. ***, ** and * indicate significance at the 1%, 5% and 10% level, respectively.

Appendix F

F.1

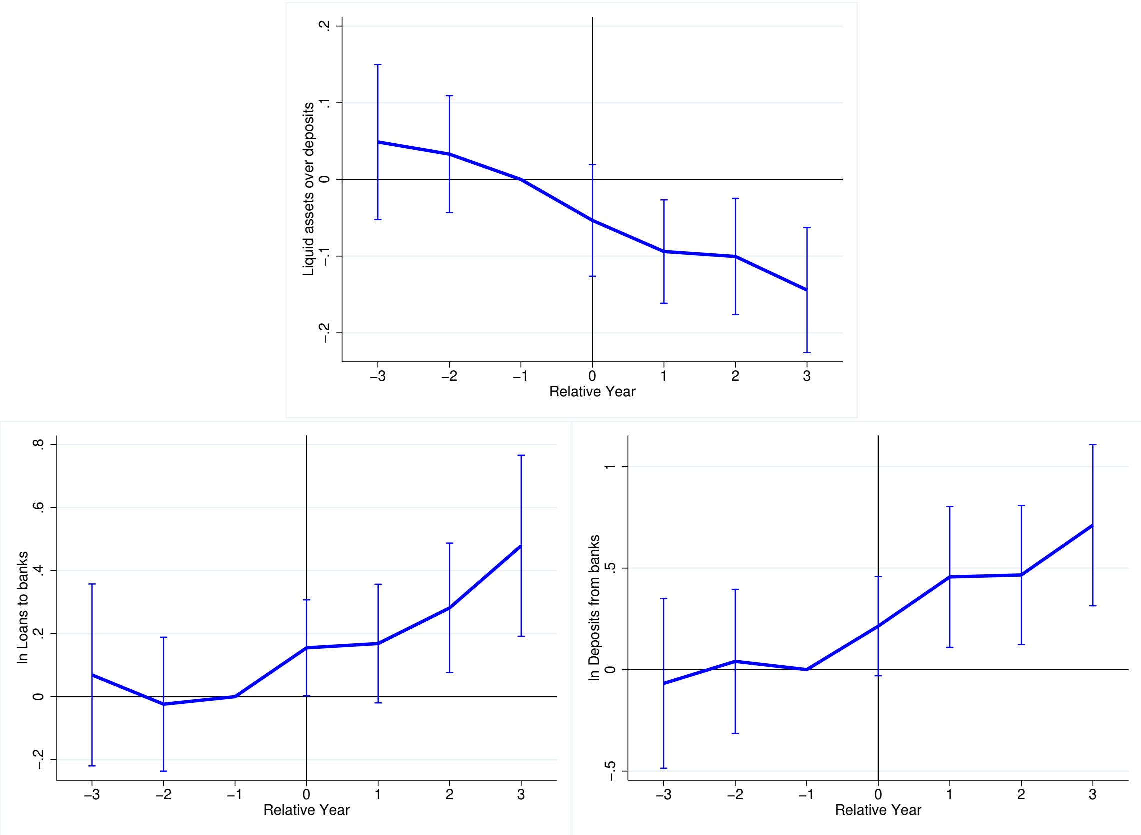
Figures F1a and F1b, replicate the event studies in the main analysis but focusing on a 3-years event window. Figure F1 refers to banking, while figure F2 focuses on interbank outcomes.

Figure F1a: An Event Study on High-Speed Internet and Banking (3-years)



Notes: This figure shows the event study for the period 3 years before–3 years after the arrival of the first submarine cable in country c. The y axis reports coefficients for the dependent variables: ln Loans (the natural logarithm of net loans (in million of US dollars)); ln Government securities (the natural logarithm of government securities (in million of US dollars)); ln Deposits (the natural logarithm of deposits and short-term funding (in million of US dollars)); and Equity over DST (the share of total equity over deposits and short-term funding). The x axis refers to the relative time from the arrival of the cable. The blue (solid) line connects point estimates relative to the base year (-1). 95% confidence intervals are also reported, and standard errors are clustered at country level.

Figure F1b: An Event Study on High-Speed Internet and Interbank Outcomes (3-years)



Notes: This figure shows the event study for the period 3 years before–3 years after the arrival of the first submarine cable in country c. The y axis reports coefficients for the dependent variables: Liquid Assets over DST (the share of liquid assets over deposits and short-term funding);ln Loans to Banks (the natural logarithm of loans to banks (in million of US dollars)); and ln Deposits from Banks (the natural logarithm of loans to banks (in million of US dollars)). The x axis refers to the relative time from the arrival of the cable. The blue (solid) line connects point estimates relative to the base year (-1). 95% confidence intervals are also reported, and standard errors are clustered at country level.

F.2

Tables F2a and F2b control for the (possible) confounding effects of changes in the consolidation procedure in bank's balance sheet. Here, we substitute bank fixed effects, that we use in the main specifications of table 2 and table 6, with bank by consolidation code fixed effects.

Table F2a: High-Speed Internet and Banking - Consolidation

	(1)	(2)	(3)	(4)
Variables	Loans	Government	Deposits	Equity
	ln	ln	ln	over DST
$Submarine_{ct}$	0.358***	0.286**	0.235***	0.003
	(0.113)	(0.129)	(0.082)	(0.014)
Bank-Consol FE	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes
Obs.	6603	3611	6658	6506
Adj. R sq.	0.898	0.812	0.912	0.610
MDV	5.011	3.917	5.624	0.196
SDDV	2.062	2.320	1.981	0.225

Notes: This table reports estimates from the staggered diff-in-diff design presented in equation (2). The dependent variables are as follows: Loans (natural logarithm of net loans (in million of US dollars)); Government (natural logarithm of government securities (in million of US dollars)); Deposits (natural logarithm of deposits and short-term funding (in million of US dollars)); and Equity (share of total equity over deposits and short-term funding). The main predictor is $Submarine_{ct}$, a binary variable for the arrival of the first fiber-optic submarine cable in the country. This dummy takes value zero before the arrival of the cable and 1 from the time of the arrival on. Obs. refers to the number of observations; Adj.R^2 is the adjusted R^2; MDV refers to the mean of the dependent variable; and SDDV refers to the standard deviation of the dependent variable. Fixed effects are at the bank-consolidation and year level. Standard errors in parentheses, clustered at country level. ***, ** and * indicate significance at the 1%, 5% and 10% level, respectively.

Table F2b: High-Speed Internet and Interbank Outcomes - Consolidation

Variables	(1) Liquid Assets over DST	(2) Loans to Banks ln	(3) Deposits from Banks ln
$Submarine_{ct}$	-0.139***	0.230**	0.489**
	(0.030)	(0.110)	(0.183)
Bank-Consol FE	Yes	Yes	Yes
Year FE	Yes	Yes	Yes
Obs.	6629	6292	4913
Adj. R sq.	0.459	0.797	0.699
MDV	0.462	3.783	2.929
SDDV	0.441	2.053	2.466

Notes: This table reports estimates from the staggered diff-in-diff design presented in equation (2). The dependent variables are as follows: Liquid Assets over DST (share of liquid assets over deposits and short-term funding); Loans to Banks (natural logarithm of loans to banks (in million of US dollars)); and Deposits from Banks (natural logarithm of loans to banks (in million of US dollars)). The main predictor is $Submarine_{ct}$, a binary variable for the arrival of the first fiber-optic submarine cable in the country. This dummy takes value zero before the arrival of the cable and 1 from the time of the arrival on. Obs. refers to the number of observations; Adj.R^2 is the adjusted R^2; MDV refers to the mean of the dependent variable; and SDDV refers to the standard deviation of the dependent variable. Fixed effects are at the bank-consolidation and year level. Standard errors in parentheses, clustered at country level. ***, ** and * indicate significance at the 1%, 5% and 10% level, respectively.

F.3

Tables F3a and F3b add to the main specifications control for the quality of the institutions (following Allen et al. (2018)). *Regulatory quality,* captures perceptions of the ability of the government to formulate and implement sound policies and regulations that permit and promote private sector development; *Rule of law*, captures perceptions of the extent to which agents have confidence in and abide by the rules of society, and in particular the quality of contract enforcement, property rights, and the courts; *Control of corruption*, captures perceptions of the extent to which public power is exercised for private gain, including both petty and grand forms of corruption, as well as capture of the state by elites and private interests. Data are from the WB WGI database.

Table F3a: High-Speed Internet and Banking - WGI Controls

Variables	(1) Loans ln	(2) Government ln	(3) Deposits ln	(4) Equity over DST
$Submarine_{ct}$	0.270**	0.200	0.182*	-0.009
	(0.108)	(0.168)	(0.091)	(0.011)
Regulatory quality	0.638***	0.587*	0.506**	-0.026
	(0.233)	(0.344)	(0.200)	(0.026)
Rule of law	0.186	-0.266	0.079	0.001
	(0.210)	(0.262)	(0.163)	(0.023)
Control corruption	0.248	0.285	0.169	-0.014
	(0.172)	(0.369)	(0.153)	(0.029)
Bank FE	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes
Obs.	5592	3087	5628	5497
Adj. R sq.	0.876	0.808	0.887	0.602
MDV	5.074	3.949	5.696	0.194
SDDV	2.052	2.349	1.971	0.223

Notes: This table reports estimates from the staggered diff-in-diff design presented in equation (2). The dependent variables are as follows: Loans (natural logarithm of net loans (in million of US dollars)); Government (natural logarithm of government securities (in million of US dollars)); Deposits (natural logarithm of deposits and short-term funding (in million of US dollars)); and Equity (share of total equity over deposits and short-term funding). The main predictors are: $Submarine_{ct}$, a binary variable for the arrival of the first fiber-optic submarine cable in the country. This dummy takes value zero before the arrival of the cable and 1 from the time of the arrival on; Regulatory quality captures perceptions of the ability of the government to formulate and implement sound policies and regulations that permit and promote private sector development; Rule of law captures perceptions of the extent to which agents have confidence in and abide by the rules of society, and in particular the quality of contract enforcement, property rights, the police, and the courts, as well as the likelihood of crime and violence; Control corruption captures perceptions of the extent to which public power is exercised for private gain, including both petty and grand forms of corruption, as well as capture of the state by elites and private interests. Obs. refers to the number of observations; Adj.R^2 is the adjusted R^2; MDV refers to the mean of the dependent variable; and SDDV refers to the standard deviation of the dependent variable. Fixed effects are at the bank and year level. Standard errors in parentheses, clustered at country level. ***, ** and * indicate significance at the 1%, 5% and 10% level, respectively.

Table F3b: High-Speed Internet and Interbank Outcomes - WGI Controls

Variables	(1) Liquid Assets over DST	(2) Loans to Banks ln	(3) Deposits from Banks ln
$Submarine_{ct}$	-0.137***	0.196*	0.494***
	(0.033)	(0.101)	(0.165)
Regulatory quality	-0.066	0.289	0.708
	(0.059)	(0.267)	(0.432)
Rule of law	-0.079	0.117	0.307
	(0.059)	(0.336)	(0.353)
Control corruption	-0.042	-0.356	0.365
	(0.051)	(0.243)	(0.256)
Bank FE	Yes	Yes	Yes
Year FE	Yes	Yes	Yes
Obs.	5599	5311	4207
Adj. R sq.	0.445	0.777	0.686
MDV	0.453	3.827	2.951
SDDV	0.438	2.064	2.479

Notes: This table reports estimates from the staggered diff-in-diff design presented in equation (2). The dependent variables are as follows: Liquid Assets over DST (share of liquid assets over deposits and short-term funding); Loans to Banks (natural logarithm of loans to banks (in million of US dollars)); and Deposits from Banks (natural logarithm of loans to banks (in million of US dollars)). The main predictors are: $Submarine_{ct}$, a binary variable for the arrival of the first fiber-optic submarine cable in the country. This dummy takes value zero before the arrival of the cable and 1 from the time of the arrival on; Regulatory quality captures perceptions of the ability of the government to formulate and implement sound policies and regulations that permit and promote private sector development; Rule of law captures perceptions of the extent to which agents have confidence in and abide by the rules of society, and in particular the quality of contract enforcement, property rights, the police, and the courts, as well as the likelihood of crime and violence; Control corruption captures perceptions of the extent to which public power is exercised for private gain, including both petty and grand forms of corruption, as well as capture of the state by elites and private interests. Obs. refers to the number of observations; Adj.R^2 is the adjusted R^2; MDV refers to the mean of the dependent variable; and SDDV refers to the standard deviation of the dependent variable. Fixed effects are at the bank and year level. Standard errors in parentheses, clustered at country level. ***, ** and * indicate significance at the 1%, 5% and 10% level, respectively.

F.4

Tables F4a and F4b replicate the main analysis on an extended sample, where missing values are imputed using multivariate normal regressions.

Table F4a: High-Speed Internet and Banking - Imputed Values

	(1)	(2)	(3)	(4)
Variables	Loans	Government	Deposits	Equity
	ln	ln	ln	over DST
$Submarine_{ct}$	0.356**	0.262	0.257**	-0.019*
	(0.127)	(0.169)	(0.087)	(0.012)
Bank FE	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes
Obs.	6746	6746	6746	6746
MDV	4.977	5.611	3.587	0.209
SDDV	2.089	1.994	2.369	0.262

Notes: This table reports estimates from the staggered diff-in-diff design presented in equation (2). Missing values of the dependent variables have been imputed using multivariate normal regressions. The dependent variables are as follows: Loans (natural logarithm of net loans (in million of US dollars)); Government (natural logarithm of government securities (in million of US dollars)); Deposits (natural logarithm of deposits and short-term funding (in million of US dollars)); and Equity (share of total equity over deposits and short-term funding). The main predictor is $Submarine_{ct}$, a binary variable for the arrival of the first fiber-optic submarine cable in the country. This dummy takes value zero before the arrival of the cable and 1 from the time of the arrival on. Obs. refers to the number of observations; MDV refers to the mean of the dependent variable; and SDDV refers to the standard deviation of the dependent variable. Fixed effects are at the bank and year level. Standard errors in parentheses, clustered at country level. ***, ** and * indicate significance at the 1%, 5% and 10% level, respectively.

Table F4b: High-Speed Internet and Interbank Outcomes - Imputed Values

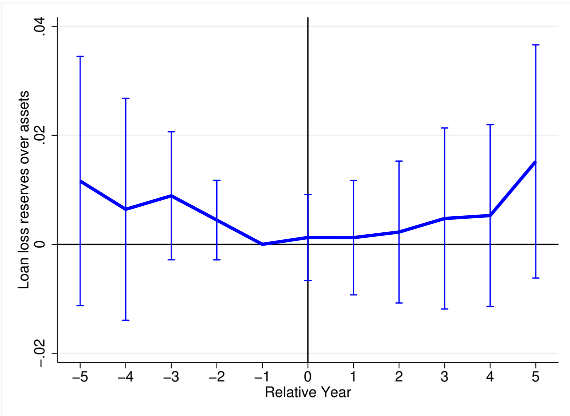
	(1)	(2)	(3)
Variables	Liquid Assets over DST	Loans to Banks ln	Deposits from Banks ln
$Submarine_{ct}$	-0.141***	0.249**	0.413**
	(0.030)	(0.111)	(0.171)
Bank FE	Yes	Yes	Yes
Year FE	Yes	Yes	Yes
Obs.	6746	6746	6746
MDV	0.468	3.737	2.563
SDDV	0.452	2.073	2.557

Notes: This table reports estimates from the staggered diff-in-diff design presented in equation (2). Missing values of the dependent variables have been imputed using multivariate normal regressions. The dependent variables are as follows: Liquid Assets over DST (share of liquid assets over deposits and short-term funding); Loans to Banks (natural logarithm of loans to banks (in million of US dollars)); and Deposits from Banks (natural logarithm of loans to banks (in million of US dollars)). The main predictor is $Submarine_{ct}$, a binary variable for the arrival of the first fiber-optic submarine cable in the country. This dummy takes value zero before the arrival of the cable and 1 from the time of the arrival on. Obs. refers to the number of observations; MDV refers to the mean of the dependent variable; and SDDV refers to the standard deviation of the dependent variable. Fixed effects are at the bank and year level. Standard errors in parentheses, clustered at country level. ***, ** and * indicate significance at the 1%, 5% and 10% level, respectively.

F.5

Figure F5 performs our standard event-study for loan loss reserves. We use this graph to detect possible changes in risk attitudes following the arrival of high-speed internet.

Figure F5: An Event Study on High-Speed Internet and Loan loss reserves



Notes: This figure shows the event study for the period 5 years before–5 years after the arrival of the first submarine cable in country c. The y axis reports coefficients for the dependent variable: Loan loss reserves over assets (the share of loan loss reserves over total assets). The x axis refers to the relative time from the arrival of the cable. The blue (solid) line connects point estimates relative to the base year (-1). 95% confidence intervals are also reported, and standard errors are clustered at country level.

F.6

Tables F6a and F6b replicate tables 9 and 10 in the main text, but focusing on expenses for electricity and raw materials. The latter provides evidence on the changes in inputs usage following the arrival of the submarine cable.

Table F6a: High-Speed Internet and Firms Inputs

	(1)	(2)
Variables	Electricity	Raw materials
	ln	ln
$Submarine_{ct}$	2.038	3.262*
	(1.869)	(1.716)
Country FE	Yes	Yes
Year FE	Yes	Yes
Obs.	23270	12304
Adj. R sq.	0.276	0.359
MDV	-6.154	-3.017
SDDV	2.380	2.918

Notes: This table reports estimates from the staggered diff-in-diff design presented in equation (2). The dependent variables are: Electricity (natural logarithm of total amount of expenses in electricity); and Raw materials (natural logarithm of total amount of expenses in raw materials). The main predictor is $Submarine_{ct}$, a binary variable for the arrival of the first fiber-optic submarine cable in the country. This dummy takes value zero before the arrival of the cable and 1 from the time of the arrival on. Obs. refers to the number of observations; Adj.R^2 is the adjusted R^2; MDV refers to the mean of the dependent variable; and SDDV refers to the standard deviation of the dependent variable. Fixed effects are at the country and year level. Standard errors in parentheses, clustered at country level. ***, ** and * indicate significance at the 1%, 5% and 10% level, respectively.

Table F6b: High-Speed Internet and Firms Inputs - Weak interbank

Variables	(1) Electricity ln	(2) Raw materials ln
Submarine_ct	0.748	2.441
	(1.392)	(1.717)
$Submarine_{ct}$	2.566*	1.411
$\times$ $Weak\ user_c$	(1.446)	(1.706)
Country FE	Yes	Yes
Year FE	Yes	Yes
Obs.	23188	12293
Adj. R sq.	0.288	0.361
MDV	-6.160	-3.018
SDDV	2.380	2.919

Notes: This table reports estimates from the staggered diff-in-diff design presented in equation (4). The dependent variables are: Electricity (natural logarithm of total amount of expenses in electricity); and Raw materials (natural logarithm of total amount of expenses in raw materials). The main predictors are: $Submarine_{ct}$, a binary variable for the arrival of the first fiber-optic submarine cable in the country; and $Submarine_{ct}$ $\times$ $Weak\ user_c$, the interaction between $Submarine_{ct}$ and a dummy that takes value one if the country was below the median of interbank loans and deposits (in the period before the arrival of high-speed internet). Obs. refers to the number of observations; Adj.R^2 is the adjusted R^2; MDV refers to the mean of the dependent variable; and SDDV refers to the standard deviation of the dependent variable. Fixed effects are at the country and year level. Standard errors in parentheses, clustered at country level. ***, ** and * indicate significance at the 1%, 5% and 10% level, respectively.

Appendix: Mayoral Wage and Public Procurement

A.1

Results depicted in figure A.1 complement the evidence in figure 5. The figure shows the behavior of the dependent variables around the 5,000-inhabitant threshold, using quadratic local approximations.

Figure A1. Graphical analysis: Tender-specific procurement outcomes

Notes: Data are from the Anti-corruption National Authority (ANAC) for all procurement contracts above 40,000€ offered by municipalities during the period 2013–2017. Sample selection follows the process described in sections 3 and 4. Inferences are made for a close interval of the population size around the threshold (interval is 3,250–9,750 inhabitants). Estimates are made following the procedure described in section 5.3. The top-left panel refers to *Bidders*, the number of admitted offers. The top-right panel refers to *Rebate*, the final discount on the reserve price. The bottom-left panels refers to *Incumbent*, and indicates whether the winning firm has won at least one other contract with the same mayor. The bottom-right panel refers to the number of contracts won by a firm during the mayor's terms. The solid lines are local quadratic approximations, and the dashed lines are their associated 90% confidence intervals. The dashed vertical lines denote the discontinuity, normalized to zero (delta population).

A.2

Table A2 presents a further decomposition of table 7. Given the granularity of my data, I investigate the effects of mayoral wages on procurement outcomes for different typologies of negotiated procedures. I distinguish between negotiated procedures with low levels of discretion, which are mostly restricted auctions created by legislators to streamline the procurement process, and negotiated procedures with high levels of discretion, which are characterized by less stringent requirements and have specific conditions that must be met in order to be utilized. Results in table A2 show that high-discretion contracts are the ones that drive the outcomes in panel C of table 7.

Table A2. The effect of mayoral wages on procurement outcomes: Negotiated procedures (full sample)

	(1)	(2)	(3)	(4)
	Bidders	Rebate	Incumbent	Sum Incumbency
	Negotiated procedures - high discretion			
Conv.	0.308	0.375	-0.0975**	-0.123*
	(0.453)	(1.534)	(0.0389)	(0.0682)
Robust	0.431	-0.167	-0.103**	-0.117
	(0.508)	(1.789)	(0.0458)	(0.0810)
Avg.	4.305	15.63	0.231	1.378
N	8486	7915	8486	8486
h	511.1	473	487.9	396.7

Notes: Data are from the Anti-corruption National Authority (ANAC) for all procurement contracts above 40,000€ offered by municipalities during the period 2013–2017. Sample selection follows the process described in sections 3 and 4. Inferences are made for a close interval of the population size around the threshold (interval is 3,250–9,750 inhabitants). Estimates are made following the procedure described in section 5.3. Column (1) refers to *Bidders*, the number of admitted offers. Column (2) reports *Rebate*, the final discount rate on the reserve price. Column (3), *Incumbent*, shows whether the winning firm has won at least one other contract with the same mayor. Finally, column (4), *Sum Incumbency*, refers to the number of contracts won by a firm during the mayor's term(s). The table reports point estimates and standard errors following conventional and robust procedures. It also provides the average value of the dependent variable, the number of observations, and the selected bandwidths for point estimation. This table further decomposes the negotiated procedures in table 7 and refers to high-discretion procedures. Standard errors are in parentheses. * $p < .10$, ** $p < .05$, *** $p < .01$.

A.3

In table A3 I report the results from replicating the analysis in table 7, but restricting the sample to observations for which cost overrun data are available. Results support the randomness of the restricted sample and show coefficients similar to those reported in table 7.

Table A3. The effect of mayoral wages on procurement outcomes: Restricted sample

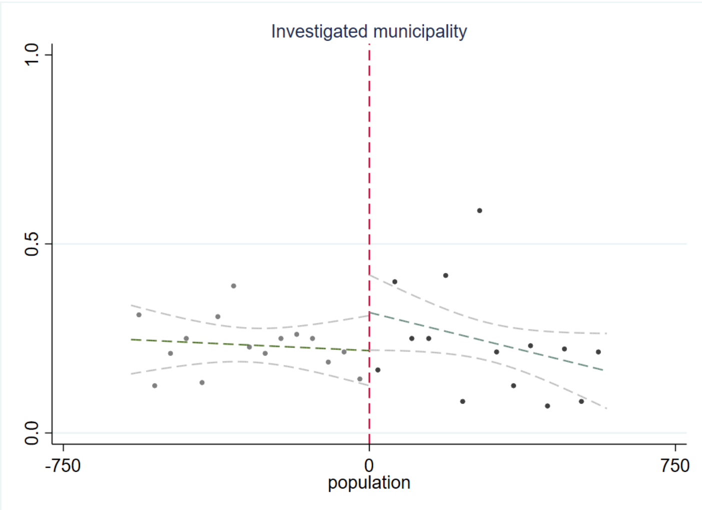
	(1)	(2)	(3)	(4)
	Bidders	Rebate	Incumbent	Sum Incumbency
			All tenders	
Conv.	3.760*	1.272	-0.125*	-0.105
	(2.259)	(2.329)	(0.0645)	(0.108)
Robust	4.262*	1.566	-0.133*	-0.112
	(2.482)	(2.825)	(0.0785)	(0.130)
Avg.	8.512	17.07	0.230	1.367
N	5492	5492	5492	5492
h	513.2	446.5	511.7	376.4

Notes: Data are from the Anti-corruption National Authority (ANAC) for all procurement contracts above 40,000€ offered by municipalities during the period 2013–2017 (the sample is restricted to contracts that have data on cost overruns). Sample selection follows the process described in sections 3 and 4. Inferences are made for a close interval of the population size around the threshold (interval is 3,250–9,750 inhabitants). Estimates are made following the procedure described in section 5.3. Column (1) refers to *Bidders*, the number of admitted offers. Column (2) reports *Rebate*, the final discount rate on the reserve price. Column (3), *Incumbent*, shows whether the winning firm has won at least one other contract with the same mayor. Finally, column (4), *Sum Incumbency*, refers to the number of contracts won by a firm during the mayor's term(s). The table reports point estimates and standard errors following conventional and robust procedures. It also provides the average value of the dependent variable, the number of observations, and the selected bandwidths for point estimation. Standard errors are in parentheses. * $p < .10$, ** $p < .05$, *** $p < .01$.

A.4

Figure A4 complements the analysis in table 14. I focus on investigated municipalities, following Decarolis et al. (2020). Although the probability of having an investigated municipality is higher above the threshold, we observe no (significant) discontinuity.

Figure A4. Investigated municipalities

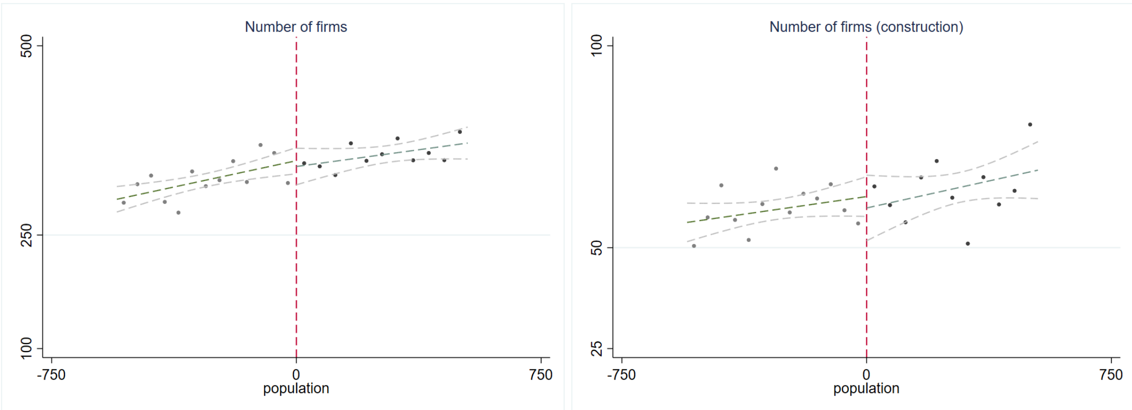


Notes: Data are from Decarolis et al. (2020). Inferences are made for a close interval of the population size around the threshold (interval is 3,250–9,750 inhabitants). Investigated municipality refers to a municipality that has been under the scrutiny of the law enforcement. The dashed lines are a local linear approximation and its corresponding 90% confidence intervals. The vertical dashed line denotes the discontinuity, normalized to zero (delta population).

B.1

In figure B1 I test whether the main results from section 5, specifically regarding the increase in tender competition around the threshold, are mechanically driven by having more firms competing for the same contract. I collect data on the number of firms in each municipality from the 2011 Italian census of the firms (data are from the ISTAT) and check for continuity around the threshold. The results are reported in figure B1. The left panel refers to the total number of firms. The right panel focuses on firms operating in the construction sector. In both cases no discontinuity can be detected.

Figure B1. Distribution of firms, by municipality



Notes: Data come from the ISTAT, 2011 firms census. Inferences are made for a close interval of the population size around the threshold (interval is 3,250–9,750 inhabitants). The left panel depicts data for all firms in the sample. The right panel focuses on firms in the construction sector. The dashed lines are a local linear approximation and its corresponding 90% confidence intervals. The vertical dashed line denotes the discontinuity, normalized to zero (delta population).

B.2

In tables B2a and B2b I replicate the main analysis on aggregate and individual tender outcomes for the universe of procurement contracts issued by Italian municipalities during the period 2010-2012.

Table B2a. The effect of mayoral wages on aggregate procurement outcomes: 2010–2012

	(1)	(2)	(3)
	Tenders	Total expenditure	Mean value
		All tenders	
Conv.	-1.260***	-164.9**	-4.163
	(0.393)	(82.44)	(43.23)
Robust	-1.418***	-171.8*	-18.07
	(0.457)	(99.48)	(50.40)
Avg.	2.237	375.5	188.1
N	1693	1693	1693
h	482.2	696.7	823.3

Notes: Data are from the Anti-corruption National Authority (ANAC) for all procurement contracts above 40,000€ offered by municipalities during the period 2010–2012. Sample selection follows the process described in sections 3 and 4. Inferences are made for a close interval of the population size around the threshold (interval is 3,250–9,750 inhabitants). Estimates are made following the procedure described in section 5.3. Column (1) refers to the number of tenders issued by each mayor, weighted by the number of mayor's years in office. Column (2) refers to the amount of money (in euros) spent by each mayor on public procurement, weighted by the number of mayor's years in office. Finally, column (3) indicates the mean value of the tender, by mayor. The table reports point estimates and standard errors following conventional and robust procedures. It also provides the average value of the dependent variable, the number of observations, and the selected bandwidths for point estimation. Standard errors are in parentheses. * $p < .10$, ** $p < .05$, *** $p < .01$.

Table B2b. The effect of mayoral wages on procurement outcomes: 2010–2012

	(1)	(2)	(3)	(4)
	Bidders	Rebate	Incumbent	Sum Incumbency
			All tenders	
Conv.	0.465	1.972	-0.131***	-0.128***
	(1.935)	(1.715)	(0.0382)	(0.0428)
Robust	0.564	1.534	-0.152***	-0.156***
	(2.378)	(2.041)	(0.0439)	(0.0490)
Avg.	7.404	14.52	0.192	1.266
N	8130	7869	8133	8133
h	1150	730.2	596.3	622.3

Notes: Data are from the Anti-corruption National Authority (ANAC) for all procurement contracts above 40,000€ offered by municipalities during the period 2010–2012. Sample selection follows the process described in sections 3 and 4. Inferences are made for a close interval of the population size around the threshold (interval is 3,250–9,750 inhabitants). Estimates are made following the procedure described in section 5.3. Column (1) refers to *Bidders*, the number of admitted offers. Column (2) reports *Rebate*, the final discount rate on the reserve price. Column (3), *Incumbent*, shows whether the winning firm has won at least one other contract with the same mayor. Finally, column (4), *Sum Incumbency*, refers to the number of contracts won by a firm during the mayor's term(s). The table reports point estimates and standard errors following conventional and robust procedures. It also provides the average value of the dependent variable, the number of observations, and the selected bandwidths for point estimation. Standard errors are in parentheses. * $p < .10$, ** $p < .05$, *** $p < .01$.

B.3

Table B3 reports results from the main specification, removing tenders that were opened after the reform of 2016. As we can see, the local effects of increased mayoral wages on the number of admitted offers, the final rebate, and the probability that firms win multiple contracts are in line with those in table 7.

Table B3. The effect of mayoral wages on procurement outcomes: Pre–2016 reform

	(1)	(2)	(3)	(4)
	Bidders	Rebate	Incumbent	Sum Incumbency
			All tenders	
Conv.	3.875***	1.644	-0.0760*	-0.164*
	(1.417)	(1.367)	(0.0402)	(0.0889)
Robust	4.105**	1.669	-0.0802*	-0.182*
	(1.691)	(1.673)	(0.0484)	(0.106)
Avg.	7.472	15.84	0.220	1.355
N	9203	8569	9203	9203
h	477.2	588.1	492.8	393.1

Notes: Data are from the Anti-corruption National Authority (ANAC) for all procurement contracts above 40,000€ offered by municipalities during the period 2013–mid2016. Sample selection follows the process described in sections 3 and 4. Inferences are made for a close interval of the population size around the threshold (interval is 3,250–9,750 inhabitants). Estimates are made following the procedure described in section 5.3. Column (1) refers to *Bidders*, the number of admitted offers. Column (2) reports *Rebate*, the final discount rate on the reserve price. Column (3), *Incumbent*, shows whether the winning firm has won at least one other contract with the same mayor. Finally, column (4), *Sum Incumbency*, refers to the number of contracts won by a firm during the mayor's term(s). The table reports point estimates and standard errors following conventional and robust procedures. It also provides the average value of the dependent variable, the number of observations, and the selected bandwidths for point estimation. Standard errors are in parentheses. * $p < .10$, ** $p < .05$, *** $p < .01$.

B.4

In table B4 I replicate the analysis in table 7, panel A, restricting the sample to municipalities that are located in the south of Italy, where purchasing power is relatively higher. The coefficients of all estimates are magnified with respect to the estimates in table 7 and provide evidence that local mayoral wage effects are heterogeneous and depend on purchasing power.

Table B4. The effect of mayoral wages on procurement outcomes: Purchasing power

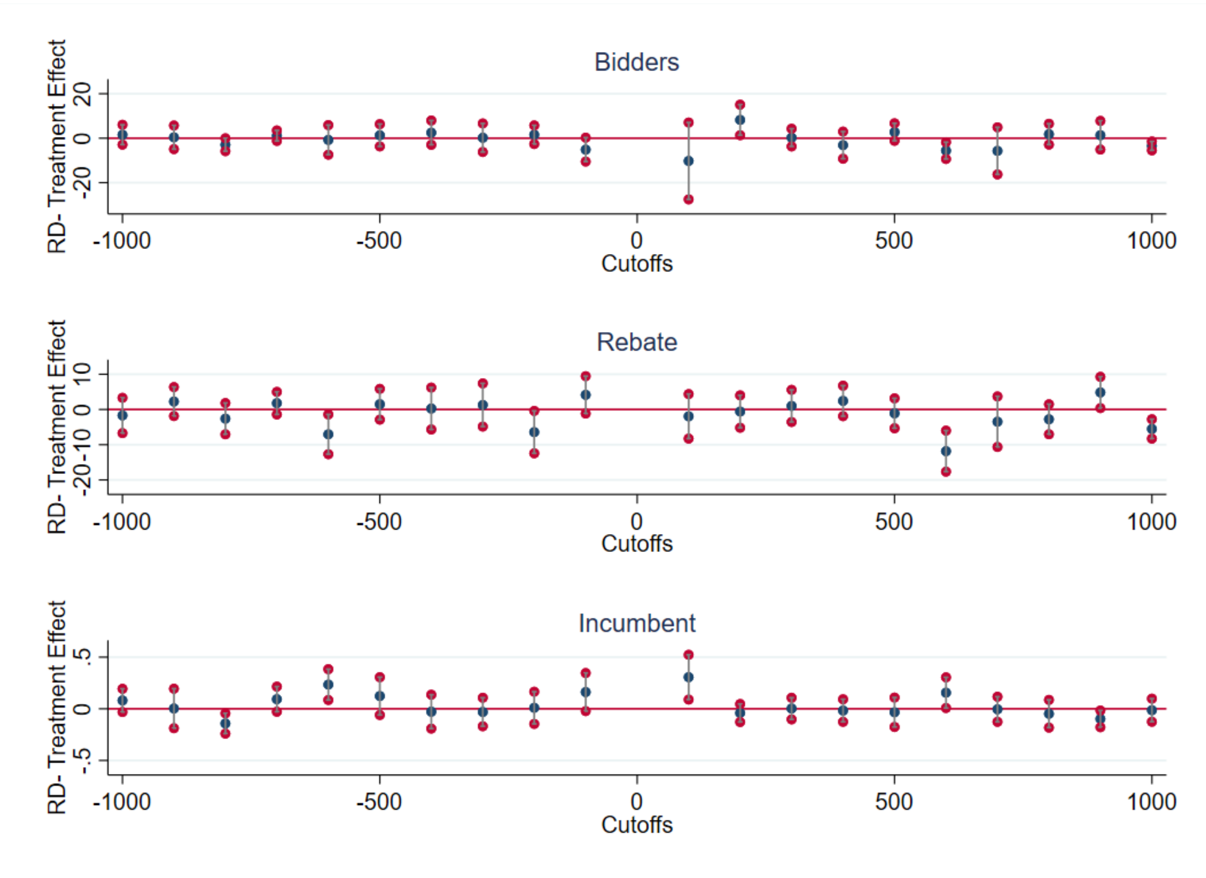
	(1)	(2)	(3)	(4)
	Bidders	Rebate	Incumbent	Sum Incumbency
			All tenders	
Conv.	12.48***	10.95***	-0.251***	-0.618***
	(4.211)	(2.302)	(0.0647)	(0.125)
Robust	14.29***	11.94***	-0.277***	-0.548***
	(4.644)	(2.530)	(0.0701)	(0.135)
Avg.	11.83	17.52	0.174	1.310
N	2600	2450	2600	2600
h	920.6	695.5	578.5	413.5

Notes: Data are from the Anti-corruption National Authority (ANAC) for all procurement contracts above 40,000€ offered by municipalities during the period 2013–2017. Sample selection follows the process described in sections 3 and 4. Inferences are made for a close interval of the population size around the threshold (interval is 3,250–9,750 inhabitants). Estimates are made following the procedure described in section 5.3. Column (1) refers to *Bidders*, the number of admitted offers. Column (2) reports *Rebate*, the final discount rate on the reserve price. Column (3), *Incumbent*, shows whether the winning firm has won at least one other contract with the same mayor. Finally, column (4), *Sum Incumbency*, refers to the number of contracts won by a firm during the mayor's term(s). The table reports point estimates and standard errors following conventional and robust procedures. It also provides the average value of the dependent variable, the number of observations, and the selected bandwidths for point estimation. Standard errors are in parentheses. * $p < .10$, ** $p < .05$, *** $p < .01$.

C.1

Figure C1 presents findings from tests of the significance of the effect of mayoral wages on procurement outcomes at simulated cutoff values. For this test, I generate a time series of simulated thresholds ranging from 4,000 to 6,000 inhabitants in steps of 100. I conduct statistical estimation and make inferences for RD treatment effects at artificial cutoff points, for control and treatment units separately. To avoid contamination due to real treatment effects, I use only treated observations for cutoffs above the actual cutoff, and I use only control observations for cutoffs below the actual cutoff. Restricting the observations in this way guarantees that the analysis of each simulated threshold uses only observations with the same treatment status. Results show that significance away from the true cutoff is rarely detected, and when it does occur, it appears close to the region of nonsignificance.

Figure C1. Time series of simulated cutoffs (4,000 to 6,000 inhabitants)

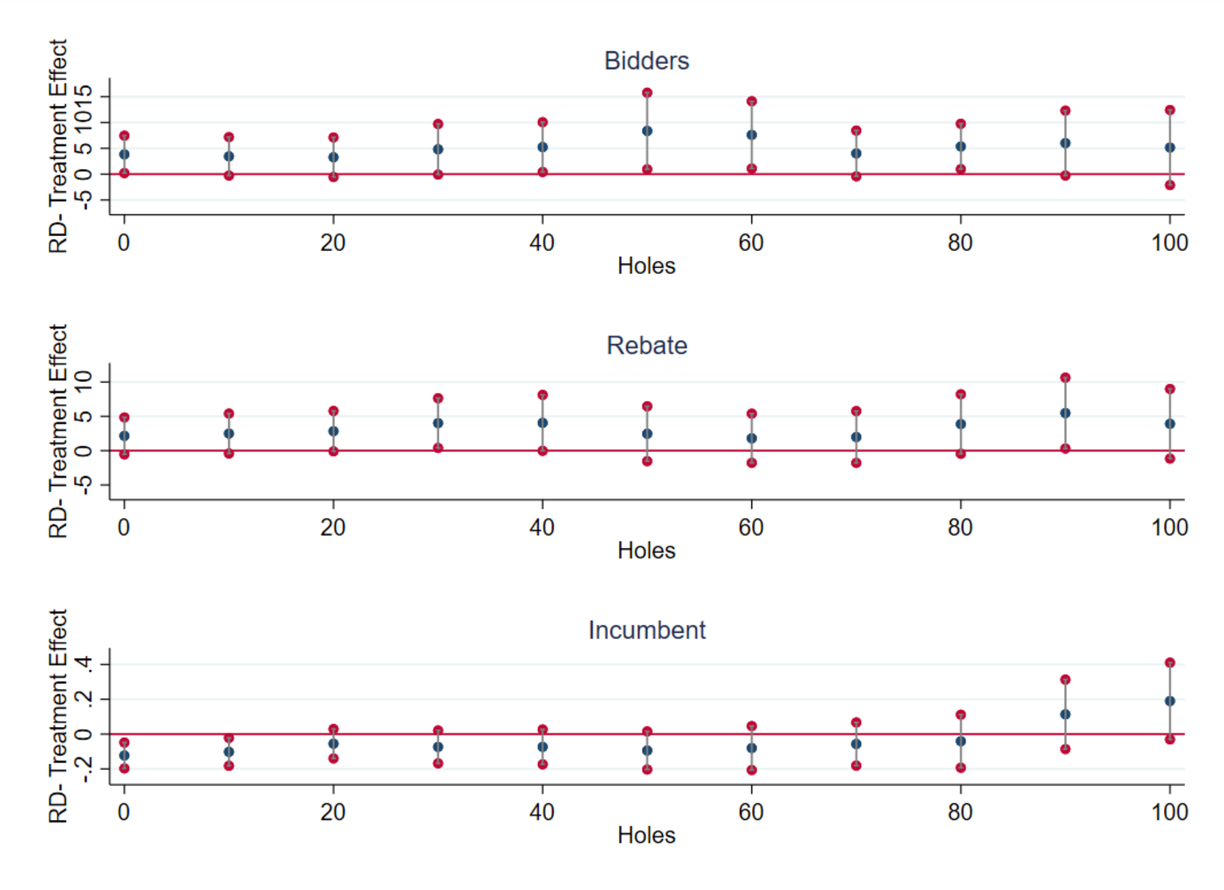


Notes: This figure reports point estimates with corresponding confidence intervals, for the number of admitted offers, the final rebate, and incumbent winners.

C.2

Figure C2 reports the sensitivity of the results to the response of units located very close to the cutoff. This is sometimes referred to as a "donut hole" approach. Although systematic manipulation of score values close to the threshold is not a great concern in my setting, this strategy is still useful in assessing the sensitivity of the results to the unavoidable extrapolation involved in local polynomial estimation, as the few observations closest to the cutoff are likely to be the most influential when fitting the local polynomials. The coefficients associated with all the examined variables preserve their signs and (in general) their levels of significance, even when donut holes are introduced.

Figure C2. "Donut holes" sensitivity test



Notes: This figure reports the results of a "donut holes" sensitivity test for 10-unit holes around the 5,000-inhabitant threshold, up to 100. The figure depicts point estimates with corresponding confidence intervals, for the number of admitted offers, the final rebate, and incumbent winners.

C.3

Table C3 reports the results from the main specification, checking for asymmetric bandwidths. The local effects of increased mayoral wages on the number of admitted offers, the final rebate, and the probability that firms win multiple contracts are in line with those in table 7.

Table C3. The effect of mayoral wages on procurement outcomes: Asymmetric bandwidths

	(1)	(2)	(3)	(4)
	Bidders	Rebate	Incumbent	Sum Incumbency
			All tenders	
Conv.	3.440***	2.024*	-0.113***	-0.194***
	(1.199)	(1.196)	(0.0355)	(0.0721)
Robust	3.520**	2.290	-0.125***	-0.229***
	(1.477)	(1.431)	(0.0416)	(0.0848)
Avg.	7.182	15.74	0.213	1.345
N	11866	10998	11866	11866
h_l	409.5	487.4	345.0	310.9
h_u	1746.0	945.6	1051.9	663.3

Notes: Data are from the Anti-corruption National Authority (ANAC) for all procurement contracts above 40,000€ offered by municipalities during the period 2013–2017. Sample selection follows the process described in sections 3 and 4. Inferences are made for a close interval of the population size around the threshold (interval is 3,250–9,750 inhabitants). Estimates are made following the procedure described in section 5.3. Column (1) refers to *Bidders*, the number of admitted offers. Column (2) reports *Rebate*, the final discount rate on the reserve price. Column (3), *Incumbent*, shows whether the winning firm has won at least one other contract with the same mayor. Finally, column (4), *Sum Incumbency*, refers to the number of contracts won by a firm during the mayor's term(s). The table reports point estimates and standard errors following conventional and robust procedures. It also provides the average value of the dependent variable, the number of observations, and the selected (upper and lower) bandwidths for point estimation. Standard errors are in parentheses. * $p < .10$, ** $p < .05$, *** $p < .01$.

C.4

Table C4 presents estimates of the main specification with a clustering of standard errors by mayors (standard errors in tables 7-15 are clustered by municipality). Findings are completely consistent with those presented in table 7.

Table C4. The effect of mayoral wages on procurement outcomes: Clustering of standard errors by mayor

	(1)	(2)	(3)	(4)
	Bidders	Rebate	Incumbent	Sum Incumbency
			All tenders	
Conv.	3.805**	2.324	-0.125***	-0.204***
	(1.806)	(1.422)	(0.0374)	(0.0728)
Robust	4.454**	2.237	-0.136***	-0.216**
	(2.095)	(1.741)	(0.0444)	(0.0884)
Avg.	7.182	15.74	0.213	1.345
N	11866	10998	11866	11866
h	892.4	715.8	605.1	556.6

Notes: Data are from the Anti-corruption National Authority (ANAC) for all procurement contracts above 40,000€ offered by municipalities during the period 2013–2017. Sample selection follows the process described in sections 3 and 4. Inferences are made for a close interval of the population size around the threshold (interval is 3,250–9,750 inhabitants). Estimates are made following the procedure described in section 5.3. Column (1) refers to *Bidders*, the number of admitted offers. Column (2) reports *Rebate*, the final discount rate on the reserve price. Column (3), *Incumbent*, shows whether the winning firm has won at least one other contract with the same mayor. Finally, column (4), *Sum Incumbency*, refers to the number of contracts won by a firm during the mayor's term(s). The table reports point estimates and standard errors following conventional and robust procedures. It also provides the average value of the dependent variable, the number of observations, and the selected bandwidths for point estimation. Robust standard errors, in parentheses, are clustered by mayor. * $p < .10$, ** $p < .05$, *** $p < .01$.

C.5

Tables C5a and C5b show the results from the main specification using second-order and fourth-order polynomial approximations, respectively. The findings are consistent with those in table 7.

Table C5a. The effect of mayoral wages on procurement outcomes: Second-order polynomial approximation

	(1)	(2)	(3)	(4)
	Bidders	Rebate	Incumbent	Sum Incumbency
			All tenders	
Conv.	3.246	2.271	-0.127***	-0.211**
	(2.120)	(1.746)	(0.0418)	(0.0856)
Robust	2.505	2.007	-0.125***	-0.186**
	(2.113)	(1.960)	(0.0471)	(0.0942)
Avg.	7.182	15.74	0.213	1.345
N	11866	10998	11866	11866
h	761.9	839.2	908.9	785.8

Notes: Data are from the Anti-corruption National Authority (ANAC) for all procurement contracts above 40,000€ offered by municipalities during the period 2013–2017. Sample selection follows the process described in sections 3 and 4. Inferences are made for a close interval of the population size around the threshold (interval is 3,250–9,750 inhabitants). Estimates are made following the procedure described in section 5.3 but using a second-order polynomial. Column (1) refers to *Bidders*, the number of admitted offers. Column (2) reports *Rebate*, the final discount rate on the reserve price. Column (3), *Incumbent*, shows whether the winning firm has won at least one other contract with the same mayor. Finally, column (4), *Sum Incumbency*, refers to the number of contracts won by a firm during the mayor's term(s). The table reports point estimates and standard errors following conventional and robust procedures. It also provides the average value of the dependent variable, the number of observations, and the selected bandwidths for point estimation. Standard errors are in parentheses. * $p < .10$, ** $p < .05$, *** $p < .01$.

Table C5b. The effect of mayoral wages on procurement outcomes: Fourth-order polynomial approximation

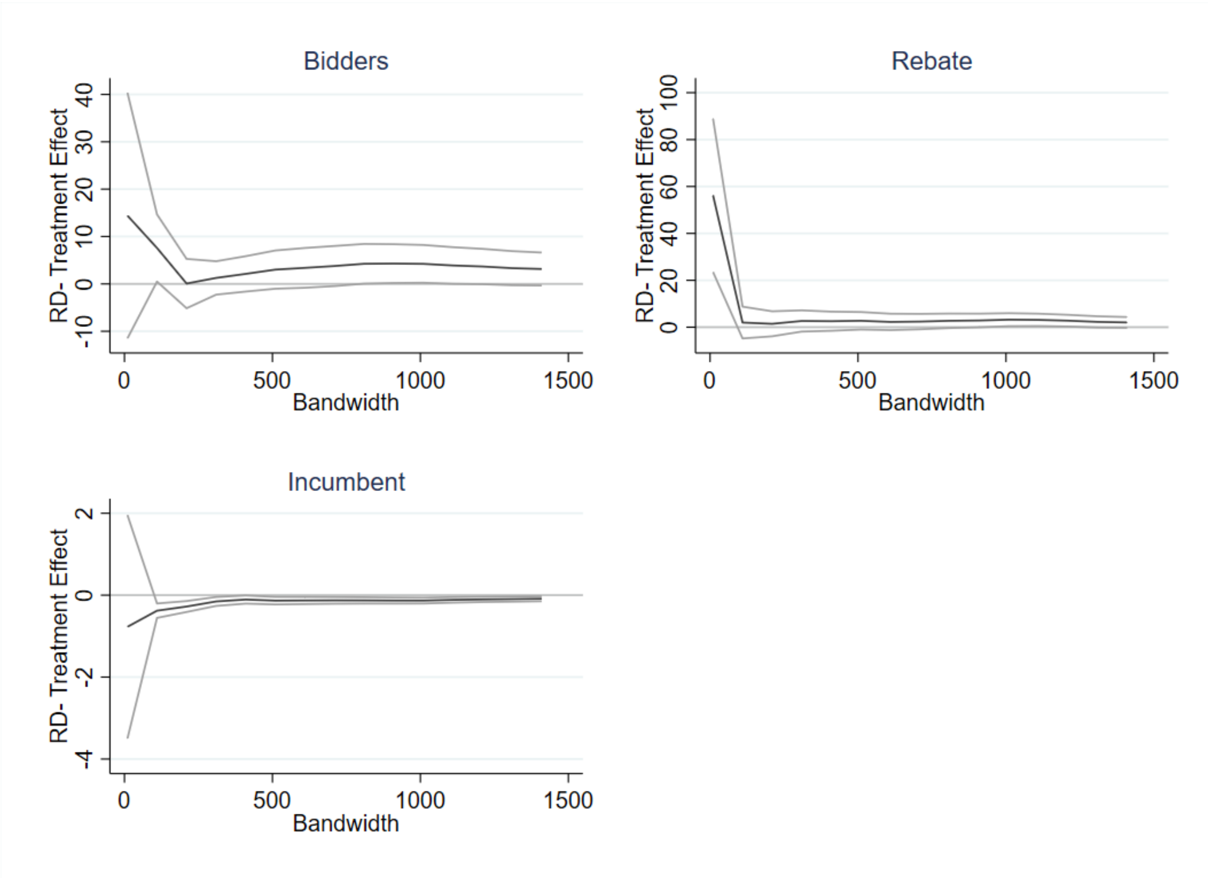
	(1)	(2)	(3)	(4)
	Bidders	Rebate	Incumbent	Sum Incumbency
			All tenders	
Conv.	1.426	2.962	-0.128**	-0.102
	(1.905)	(2.365)	(0.0531)	(0.107)
Robust	1.185	3.056	-0.119**	-0.117
	(1.974)	(2.550)	(0.0564)	(0.112)
Avg.	7.182	15.74	0.213	1.345
N	11866	10998	11866	11866
h	1083	947.8	1237	864.3

Notes: Data are from the Anti-corruption National Authority (ANAC) for all procurement contracts above 40,000€ offered by municipalities during the period 2013–2017. Sample selection follows the process described in sections 3 and 4. Inferences are made for a close interval of the population size around the threshold (interval is 3,250–9,750 inhabitants). Estimates are made following the procedure described in section 5.3, but using a fourth-order polynomial. Column (1) refers to *Bidders*, the number of admitted offers. Column (2) reports *Rebate*, the final discount rate on the reserve price. Column (3), *Incumbent*, shows whether the winning firm has won at least one other contract with the same mayor. Finally, column (4), *Sum Incumbency*, refers to the number of contracts won by a firm during the mayor's term(s). The table reports point estimates and standard errors following conventional and robust procedures. It also provides the average value of the dependent variable, the number of observations, and the selected bandwidths for point estimation. Standard errors are in parentheses. * $p < .10$, ** $p < .05$, *** $p < .01$.

C.6

Figure C6 analyzes the sensitivity of the results presented in table 7 to the bandwidth choice. It depicts estimate sensitivity as units are added or removed at the endpoints of the neighborhood. In the continuity-based approach, this falsification test is implemented by changing the bandwidth used for local polynomial estimation. Results are presented for the number of admitted offers, the final rebate on the reserve price, and *Incumbent*, using a bandwidth window ranging from 0 to 1,500. The estimated effects are significant across different bandwidths below and above the optimal bandwidth.

Figure C6. Sensitivity of the results to different bandwidth choice



Notes: The figure depicts point estimates with corresponding confidence intervals, for the number of admitted offers, the final rebate, and incumbent winners.

D.1

Table D1 reports the main sources of data used in this paper.

Table D1. Sources of data

Data	Source
Procurement variables	ANAC (Autorità Nazionale Anticorruzione)
Mayoral information	Anagrafe Amministratori Locali (Ministry of the Interior)
Data on municipalities	ISTAT (Istituto Nazionale di Statistica)
Data on firms	ISTAT (Istituto Nazionale di Statistica)